

А. С. СЕРЕБРОВСКИЙ, член-корреспондент Академии Наук СССР

**АНАЛИЗ АЛЛОБАЛАНСА ХРОМОСОМЫ ДВУМЯ СИГНАЛЯМИ.
ИСПОЛЬЗОВАНИЕ НЕСКОЛЬКИХ КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ
ОДНОВРЕМЕННО**

В другом месте мы указывали уже, что при генетическом анализе количественных признаков новые перспективы могут быть открыты одновременным анализом нескольких различных количественных признаков. В настоящем сообщении мы рассматриваем этот вопрос применительно к анализу хромосомы двумя сигналами S_1 и S_2 . Анализируется не полный баланс хромосомы, но различие между балансами двух гомологичных хромосом, находящаяся в гетерозиготной особи, т. е. то, что мы называем аллобалансом.

Обозначим гены, влияющие на признак А, через A_1, A_2, \dots и гены, влияющие на признак В, через B_1, B_2, \dots . Пусть например мы имеем гетерозиготную самку дрозофилы или самца шелкопряда и скрещиваем ее с гомозиготным самцом (соответственно самкой) анализатором. Над каждым геном напишем силу его действия, взяв произвольный пример:

$$\begin{array}{c} \text{♀} \\ \frac{S_1 \quad B_1 \quad B_2 \quad A_1 \quad B_3 \quad A_2 \quad B_4 \quad S_2}{s_1 \quad b_1 \quad b_2 \quad a_1 \quad b_3 \quad a_2 \quad b_4 \quad s_2} \end{array} \times \begin{array}{c} \text{♂} \\ \frac{s_1 \quad b_1 \quad b_2 \quad a_1 \quad b_3 \quad a_2 \quad b_4 \quad s_2}{s_1 \quad b_1 \quad b_2 \quad a_1 \quad b_3 \quad a_2 \quad b_4 \quad s_2} \end{array}$$

Выберем из потомства от такого скрещивания группу S_1s_2 — особей, отличимых по сигналам S_1 и s_2 . Эти особи будут либо одиночными перекрестниками, либо тройными на участке S_1-S_2 . Запишем только одиночные, возможные в этой структуре, и определим у каждого величину признака А и В.

Место перекреста	Структура	Величина признака	
		А	В
1	$S_1b_1b_2a_1b_3a_2b_4s_2$	0	0
2	$S_1B_1b_2a_1b_3a_2b_4s_2$	0	2
3	$S_1B_1B_2a_1b_3a_2b_4s_2$	0	6
4	$S_1B_1B_2A_1b_3a_2b_4s_2$	3	6
5	$S_1B_1B_2A_1B_3a_2b_4s_2$	3	4
6	$S_1B_1B_2A_1B_3A_2b_4s_2$	7	4
7	$S_1B_1B_2A_1B_3A_2B_4s_2$	7	7

Нанеся 7 полученных точек на поле **A, B**, получим фиг. 1. Через 7 нанесенных точек мы протянули ломаную полоску. При нанесении этой полоски мы могли и не знать порядковых номеров 7 точек, но проведенная полоска единственно возможная, если при ее нанесении руководствоваться: 1) тем, чтобы углы (перегибы) полоски приходились на точку, а не на пустом месте, и 2) тем, чтобы отрезки полоски были параллельны либо оси **A**, либо оси **B**.

Обратим внимание на следующие свойства этой полоски.

1. Расстояние между точками равно силам генов: расстояние между точками 1 и 2, равное 2, есть сила гена $B_1 (= +2)$; расстояние между точками 2 и 3 ($= +4$) есть сила гена B_2 и т. д.

2. Параллелизм отрезков между соседними точками с осью **A** указывает, что соответствующий ген влияет на признак **A**, параллелизм отрезка с осью **B** указывает на влияние соответствующего гена на признак **B**. Например



Фиг. 1

участок 1—2 параллелен оси **B** и отвечает гену B_1 , участок 3—4 параллелен оси **A** и отвечает гену A_1 и т. д.

3. Порядок отрезков отвечает порядку расположения генов в хромосомах.

4. Число углов на полоске отвечает числу «смен» **A** и **B** в порядке их расположения по длине хромосомы. В нашем случае число и места «смен» следующие:



5. Частоты точек отражают частоту перекрестов между генами. Если например в эмпирическом материале точка 5 встречается вдвое чаще, чем точка 4, то, следовательно, расстояние между 3-м и 4-м генами (B_2 и A_1) меньше, чем расстояние между 4-м и 5-м генами (A_1 и B_3), хотя и не точно в 2 раза, а в пропорции $c_3(1-c_4) : (1-c_3)c_4$, где c_3 и c_4 — величины перекреста в долях единицы на участке 3 и 4.

Рассмотрим, какие выгоды представляет одновременное исследование двух признаков. Если бы мы наблюдали только за признаком **A**, то среди S_1s_2 особей имели бы величину признака 0, 3 и 7. Отсюда мы могли бы

заклЮчить, что на участке $S_1—s_2$ имеются два гена, влияющие на признак **A**, и что эти три величины отвечают генотипам $a_1 a_2 A_1 a_2$ и $A_1 A_2$ (1). Но о порядке расположения генов заключение было бы невозможно. По признаку **B** мы имели бы 5 величин 0, 2, 4, 6 и 7 и могли бы заключить о 4 генах, но о порядке их расположения заключение тоже было бы невозможно. Но, развернув их на плоскость, мы получаем возможность для такого суждения. Особое значение здесь приобретает расположение генов **B** между генами **A** и, обратно, генов **A** между генами **B**. С этой точки зрения тем яснее будет получающаяся картина, чем больше независимых признаков будем мы исследовать одновременно. Если в хромосомах расположены гены A_1 и A_2 , то между ними может не оказаться гена, влияющего на признак **B** (по крайней мере в гетерозиготном состоянии, необходимом для данного анализа). Но зато здесь может оказаться расположенным ген C_k , влияющий на признак **C**, или ген D_k , влияющий на признак **D**, и т. д.

Настоящий прием анализа имеет следующие осложнения.

Среди особей $S_1 s_2$ кроме одиночных перекрестников могут встретиться и тройные перекрестники, каждый из которых будет иметь свою величину признака **A** и **B**. Поэтому на нашем поле появятся еще дополнительные точки, которые могут осложнить чтение рисунка и проведение «полоски». Правда, тройные перекресты будут встречаться значительно реже одиночных. Например, при расстоянии между генами в 10% одиночные перекрестники будут встречаться более, чем в 100 раз чаще, чем тройные. Но если некоторые из промежутков между генами будут очень малы, а другие велики, то некоторые из одиночных перекрестов будут встречаться так же часто, как некоторые из тройных, или даже реже.

При небольшом расстоянии между сигналами тройные перекресты могут быть сведены к нулю. Но тогда возникает второе осложнение. Могут оказаться гены, влияющие на **A** и **B** и расположенные левее левой сигналы и правее правой. Тогда в группе особей $S_1 s_2$ могут оказаться особи с двойным перекрестом, один из которых прошел между сигналами, а второй где-либо вне осигналенного участка $S_1—S_2$. Это приведет к тому, что каждая «точка» нашего поля окажется вдвоенной, встроенной и т. д. Тем самым удвоится, встроится и т. д. вся полоска, а это осложнит и самое ее проведение. Против этого затруднения сможет помочь использование трех сигналов и более, что однако выходит за пределы темы данного сообщения.

Осложнение может быть вызвано и генами, расположенными в других хромосомах, которые тоже будут удваивать (и т. д.) наши точки. С этим осложнением можно бороться однако определением величины признака не по одной особи, а по группе их. Например, от описанного в начале скрещивания мы можем получить самцов дрозофилы (соответственно самок шелкопряда) и каждого из этих самцов скрестить с самками (соответственно самцами) анализаторами. Поскольку у самцов дрозофилы (и самок шелкопряда) перекреста не происходит, постольку все особи $S_1 s_2$, полученные от отдельного такого самца, будут сохранять $S_1 s_2$ -хромосому неизменной. На большом материале можно будет изучать точно величину признаков **A** и **B**, освободившись от влияния других хромосом, которые в этом или следующем поколении смогут быть элиминированы.

В качестве иллюстрации мы рассчитали частоты различных одиночных и тройных перекрестов для одного случая, при котором две сигналы расположены по концам хромосом на расстоянии 100 морганид друг от друга. Для простоты коинциденция принята всюду равной единице, чем повышена частота тройных перекрестов. На фиг. 2 кружками различных диаметров

яния 5, 10, 20 и 25 морганид. Более точные расстояния могут быть найдены решением системы уравнений:

$$\begin{aligned} c_1(1 - c_2)(1 - c_3)(1 - c_4)(1 - c_5)(1 - c_6)(1 - c_7) &= 0.37, \\ (1 - c_1)c_2(1 - c_3)(1 - c_4)(1 - c_5)(1 - c_6)(1 - c_7) &= 0.83, \\ (1 - c_1)(1 - c_2)c_3(1 - c_4)(1 - c_5)(1 - c_6)(1 - c_7) &= 0.175 \text{ и т. д.} \end{aligned}$$

Неизвестные c_1, c_2, \dots, c_7 —расстояния между генами или длины участков в величине перекреста (выраженных в долях единицы).

Вернемся еще раз к фиг. 2. Мы выбрали для решения 7 крупных точек. Но при (0; 9) имеется еще одна, самая крупная из мелких. Попробуем включить ее в решение, проведя полосу через точки (0; 0), (0; 4), (0; 9) (4; 9), (4; 2), (2; 2), (2; 7) и (7; 7). Это решение запишем так:

Участки (точки)	Гены	Сила	Площадь круга	Приблизит. расстояние в морганидах
1	S_1	+4	37	9.0
2	B_1	+5	83	20.1
3	B_1	+4	7	1.7
4	A_1	-5	1	0.3
5	B_3	-2	17.5	4.2
6	B_4	-2	37	9.0
7	A_2	+5	110	26.6
8	B_5	+5	83	20.1
9	A_3	+5	37	9.0
	S_2		412.5	100.0

Сравнивая это решение с предыдущим, мы находим в нем два лишних гена $B_2(+5)$ и $B_3(-5)$, разделенных друг от друга расстоянием в 2 морганиды. Такое совпадение, что тесно рядом сидят два взаимно нейтрализующихся гена, позволяет считать эти гены нереальными. Следует иметь в виду, что допущение пары рядом локализованных и взаимно нейтрализующихся генов возможно для любой точки хромосомы и не может быть ни доказано, ни опровергнуто. Поэтому допускать такое решение следует лишь тогда, когда к нему имеются веские доказательства. В нашем случае веских доказательств существования этой пары генов нет, и можно эту мелкую точку (0; 9) отнести к точкам тройного перекреста. Это допущение легко проверить. Если наше первое решение верно, то наиболее часто встречающимся тройным перекрестом будет перекрест 2, 5, 6. Он должен будет отвечать генотипу $S_1B_1a_1b_2a_2B_3a_3s_2$ и иметь величины признака $A=0$ и $B=9$.

Частота этой точки (0; 9) должна быть

$$(1 - c_1)c_2(1 - c_3)(1 - c_4)c_5c_6(1 - c_7),$$

или при нашем приближенном решении:

$$0.909 \cdot 0.205 \cdot 0.957 \cdot 0.909 \cdot 0.274 \cdot 0.205 \cdot 0.909 = 0.0083, \text{ или } 0.83\%,$$

тогда как самый редкий одиночный перекрест (3) должен встречаться с частотой

$$(1 - c_1)(1 - c_2)c_3(1 - c_4)(1 - c_5)(1 - c_6)(1 - c_7)$$

или при нашем приближенном решении

$$0.909 \cdot 0.195 \cdot 0.043 \cdot 0.909 \cdot 0.726 \cdot 0.795 \cdot 0.909 = 0.0148, \text{ или } 1.48\%.$$

Таким образом площадь кружка частоты тройного перекреста 2, 5, 6 должна быть примерно в два раза менее площади кружка перекреста 3. Действительно, в точке (0; 9) мы и находим самую крупную из мелких точек, с площадью около 7, т. е. в $2^{1/2}$ раза меньшую, чем площадь кружка в точке (4; 4). Расчеты совпадают, следовательно эту точку (0; 9) мы должны признать точкой 2, 5, 6-перекреста и нашу «полоску» через нее не проводить, т. е. остановиться на первом решении. Последнее действительно отвечает принятому при составлении задачи условию:

S_1	B_1^{+4}		A_1^{+4}	B_2^{-2}	A_2^{-2}		B_3^{+5}		A_3^{+5}	S_2
10	20		5	10		25		20		10

Лаборатория генетики
Московского университета.

Поступило
5 III 1937.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ М. Л. К а р п, Зоологический журнал, XV, вып. 1, 26—47 (1936).