

5. СХЕМЫ РАЗМНОЖЕНИЯ ОСОБЕЙ

5.1. Рекомбинация генов

Пусть особи a_k^t и a_l^t с различающимися между собой генотипами $E(a_k^t)$ и $E(a_l^t)$ $[|E(a_k^t) \oplus E(a_l^t)| > 0]$ являются "родительской" парой, которая образована из особей популяции P^t по одной из рассмотренных в разделе 4 систем скрещивания.

Под *рекомбинацией генов* будем понимать схему размножения особей, которая моделирует акты синаммии и мейоза, удовлетворяющие законам наследственности Менделя.

По своей сути рекомбинация генов ведет к появлению новых сочетаний "родительских" генов, так как аллель любого гена "родительской" гомологичной хромосомы, согласно первого закона Менделя, целиком передается "потомку" по наследству. При этом гомологичные хромосомы "родителей" сравниваются по содержанию каждого гена. Если аллели в i -ом локусе одинаковы у "отцовской" и "материнской" хромосом $[e_{\theta}^{OT}(i) = e_{\theta}^M(i) = e_{\theta}(i)]$, то аллель $e_{\theta}(i)$ сохраняется в i -ом гене "потомка". В противном случае $[e_{\theta}^{OT}(i) \neq e_{\theta}^M(i)]$ в i -ый локус гамет "потомка" заносится с вероятностью $(1/2)$ либо аллель $e_{\theta}^{OT}(i)$, либо аллель $e_{\theta}^M(i)$. Эта операция случайного расхождения "родительских" генов по гаметам "потомков", согласно второго закона Менделя, проводится для всех "родительских" генов, для которых аллели не совпадают между собой. Рекомбинация генов позволяет воспроизвести два типа "потомков":

- копию одной из "родительских" гамет ("отца" a_k^t или "матери" a_l^t);
- гамету одного из "родителей", в которой некоторые гены имеют аллелеформы другого "родителя".

На рис. 5.1. приведен пример воспроизводства двух гибридных гамет "потомков" (b_1^t, b_2^t) , образованных из "родительской" пары $(a_k^t, a_l^t) \in P^t$ с помощью рекомбинации генов.

$$\mathbf{E}(\mathbf{a}_k^t): \quad \boxed{\mathbf{e}_\theta(1)} \quad \boxed{\mathbf{e}_\theta^{\text{OT}}(2)} \quad \boxed{\mathbf{e}_\theta(3)} \quad \boxed{\mathbf{e}_\theta^{\text{OT}}(4)} \quad \dots$$
$$\mathbf{E}(\mathbf{b}_1^t): \quad \begin{array}{|c|c|c|c|} \hline e_\theta(1) & \mathbf{e}_\theta^{\text{OT}}(2) & e_\theta(3) & \mathbf{e}_\theta^{\text{M}}(4) \\ \hline \end{array}$$
$$E(\mathbf{b}_2^t): \quad \boxed{e_0(1)} \quad \boxed{e_\theta^M(2)} \quad e_0(3) \quad \boxed{e_\theta^{\text{OT}}(4)} \quad \dots$$

В некоторых случаях какие-то аллели могут оказывать более сильное влияние на соответствующий признак особи и в процессе рекомбинации генов им необходимо отдавать большее предпочтение при формировании генов в гаметах "потомков".

Будем называть *рецессивным аллелем* аллельную форму a , которая проявляется лишь в гомозиготе (aa), когда "родители" имеют одинаковые аллели в рассматриваемом локусе аллельного гена $[e_0^{\text{OT}}(i) = e_0^M(i) = a]$, а *доминантным аллелем* - аллельную форму A , которая проявляется не только в гомозиготе (AA), но и в гетерозиготах (Aa или aA).

Взаимоотношение двух введенных аллелей a и A в конкретном локусе гомологичных хромосом зиготы обладает свойством *доминантности*

(доминирования), которое заключается в том, что доминантный аллель A всегда передается "потомку" независимо от того, принадлежит ли он "материнской" или "отцовской" гамете (рис.5.2), т.е. доминантный аллель A оказывает более сильное влияние на соответствующий признак "потомка" по сравнению с рецессивным аллелем a .

"Отцовская" гамета

□

$E(a_k^t):$



a



"Материнская" гамета □

□

$E(a_i^t):$



A



1



Гамета "потомка"

□ $E(b_i^t):$



A



а) доминантный аллель A
принадлежит "отцовской"
гамете

б) доминантный аллель A
принадлежит "материнской"
гамете

Рис.5.2. Доминирование доминантного аллеля A над рецессивным аллелем a .

Для учета доминантности некоторых из аллельных форм $e_\theta(k)$, $k = \overline{1, m_i}$, можно воспользоваться частотой аллели $e_\theta(k)$, находящейся в i -м локусе хромосомного набора популяции P^t :

$$P(e_\theta(k), i) = v_i / v, \quad k = \overline{1, m_i}, \quad (5.1)$$

где

v - численность популяции P^t ;

v_i - число генотипов в хромосомном наборе популяции P^t , в которых i -ый локус содержит аллельную форму $e_\theta(k)$;

m_i - число форм аллелей в i -м локусе ($1 \leq m_i \leq v$).

Тогда, если $P(e_\theta(k), i) > P(e_\theta(j), i)$, то аллель $e_\theta(k)$ считается доминантным аллелем; это приводит к тому, что при наличии в i -м аллельном гене двух аллелей $e_\theta(k)$ и $e_\theta(j)$, в i -ый локус гаметы "потомка" заносится доминантный аллель $e_\theta(k)$ с вероятностью 1 вместо вероятности $(1/2)$, принятой для рекомбинации генов без доминирования.

В качестве иллюстрации приведем реализацию схемы размножения особей путем рекомбинации генов для задачи оптимального разбиения графа $G(X, V, W)$ порядка n на два подграфа $G_1(X_1, V_1, W_1)$ и $G_2(X_2, V_2, W_2)$ порядка n_1 и n_2 , соответственно.

Обозначим символами e_i^{OT} , e_i^M и $e_i^П$ - аллельные формы i -го гена гамет "отца", "матери" и "потомка".

Алгоритм рекомбинации генов

1. $I := \{1, 2, \dots, n\}$; $I_1 := I_2 := \emptyset$; $n_1(t) := n_2(t) := 0$.
2. Для всех $i \in I$ формируются гомозиготные гены "потомка":
 - 2.1. Если $e_i^{OT} = e_i^M = 1$, то $\{e_i^П := 1\}$; $I_1 := I_1 \cup \{i\}$; $n_1(t) := n_1(t) + 1$;
 - 2.2. Если $e_i^{OT} = e_i^M = 0$, то $\{e_i^П := 0\}$; $I_2 := I_2 \cup \{i\}$; $n_2(t) := n_2(t) + 1$;
3. $I := I \setminus \{I_1 \cup I_2\}$ - множество гетерозиготных генов.
4. Случайным образом с вероятностью $(1/|I|)$ выбирается j -ый гетерозиготный ген ($j \in I$).
- ☐ 5. Случайным образом с вероятностью $(1/2)$ в j -ый локус гаметы "потомка" заносится e_j^{OT} или e_j^M ("1" или "0");
- ☐ 6. Если $e_j^П = 1$, то $n_1(t) := n_1(t) + 1$ иначе $n_2(t) := n_2(t) + 1$.
- ☐ 7. $I := I \setminus \{j\}$.
- нет

 8. $[n_1(t) = n_1 \text{ или } \forall n_2(t) = n_2]$?

↓ да
 9. Если $n_1(t) = n_1$, то $e_k^П := 0$ для всех $k \in I$ иначе $e_k^П := 1$ для всех $k \in I$ (гамета "потомка" $\square E(b_i^t)$ сформирована).

Для учета доминантности аллельных форм п.5 рассмотренного алгоритма должен быть заменен следующей процедурой:

5.1. Случайным образом с вероятностью (v_j / v) в j -ый локус "потомка" заносится "1" и с вероятностью $(1 - v_j / v)$ заносится "0" (Здесь v_j - число единиц в j -м локусе хромосомного набора популяции P^t численностью v).

Введенная процедура позволяет определить доминантный аллель A с помощью частот аллелей хромосомного набора текущей популяции P^t , т.к. каждый ген имеет всего две формы аллелей "1" или "0".

Пример 5.1.

Пусть для графа, приведенного на рис.1.1. задано два дихотомических разбиения (X'_1, X'_2) и (X''_1, X''_2) :

$$X'_1 = \{x_2, x_4, x_5, x_7, x_9\}, \quad X'_2 = \{x_1, x_3, x_6, x_8, x_{10}, x_{11}, x_{12}\} \quad \text{и}$$

$$X''_1 = \{x_1, x_2, x_5, x_8, x_{10}\} \quad X''_2 = \{x_3, x_4, x_6, x_7, x_9, x_{10}, x_{12}\}.$$

Будем считать, что (X'_1, X'_2) соответствует особи a_k^t , а разбиение (X''_1, X''_2) - особи a_l^t , которые образуют "родительскую" пару с генотипами, удовлетворяющими условию, что число "1" в каждом из них равно $n_1=5^*$:

□

"отцовская" гамета $E(a_k^t)$:	0	1	0	1	1	0	1	0	1	0	0	0	(13)		
	1	2	3	4	5	6	7	8	9	10	11	12			
"материнская" гамета $E(a_l^t)$:	1	1	0	0	1	0	0	1	0	1	0	0	(19)		
	1	2	3	4	5	6	7	8	9	10	11	12			

□

Для приведенных "родительских" гамет гомозиготные гены находятся во 2, 3, 5, 6, 11 и 12 локусах; локусы, которые занимают гетерозиготные гены, помечены крестиками:

□

x	1	0	x	1	0	x	x	x	x	0	0				
1	2	3	4	5	6	7	8	9	10	11	12				

□

В результате рекомбинации генов может быть получена, например, следующая гамета "потомка":

	e_1^M								e_4^{OT}		e_7^M	e_8^{OT}	e_9^M	e_{10}^M		
$E(b_1^t):$	1															
	1	2	3	4	5	6	7	8	9	10	11	12				

□

которой соответствует разбиение (X_1, X_2) :

$$X_1 = \{x_1, x_2, x_4, x_5, x_{10}\}, \quad X_2 = \{x_3, x_6, x_7, x_8, x_9, x_{11}, x_{12}\}.$$

В генотипе "потомка" $E(b_1^t)$ гены 2, 3, 5, 6, 11 и 12 (заштрихованные биты) совпадают с гомозиготными генами "родителей", сохраняя их аллели по наследству;

*) В скобках указаны степени приспособленности μ особей, имеющих данный генотип □ п.

аллели генов "потомка" 1, 7, 9 и 10 получены от соответствующих генов "материнского" генотипа (эти биты помечены символом e_i^M); гены "потомки" 4 и 8 получены от соответствующих генов "отцовского" генотипа (эти биты помечены символом e_i^{OT}). □

5.2. Простой кроссинговер

Размножение особей путем рекомбинации генов основывалось на независимом случайном обмене аллелями в аллельных генах гомологических хромосом и последующем расхождении последних по гаметам "потомков". Однако, это не означает, что гены, расположенные в одной гомологической хромосоме, всегда переходят в результате акта мейоза к "потомству" совместно в полном составе. Они могут расходиться в дочерние гаметы "потомков" также отдельными участками гомологических хромосом, состоящими из нескольких сцепленных между собой генов. Такой взаимный обмен участками гомологических хромосом называется *кроссинговером*.

В результате кроссинговера "родительские" гены передаются "потомству" не по одиночке, а в сочетаниях, отличных от тех, в которых они имели место в "родительских" гаметах. При этом сочетания генов и направление их изменений носят случайный характер.

При *простом кроссинговере* гомологические хромосомы зиготы, прежде чем разойтись по дочерним гаметам "потомков", разрываются в одном и том же месте на два участка, а затем обмениваются соответствующими участками сцепленных генов или восстанавливаются в исходном виде.

Точка разрыва гомологичных хромосом, которая может находиться в любом месте хромосомы между какой-либо парой ее генов, называется *точкой кроссинговера* α . Она делит "материнскую" и "отцовскую" хромосомы случайным образом на два участка: L_1^M, L_1^{OT} , содержащие гены от 1 до α , и L_2^M, L_2^{OT} , содержащие гены от $(\alpha+1)$ до n .

В результате простого кроссинговера воспроизводятся два типа "потомков" (рис.5.3):

- копия одной из "родительских" гамет ("отца" a_k^t или "матери" a_l^t) длиной L ;
- *рекомбинантная гамета*, содержащая сочетания генов, которые отсутствуют у каждого из "родителей" в отдельности.

α -точка кроссинговера

*

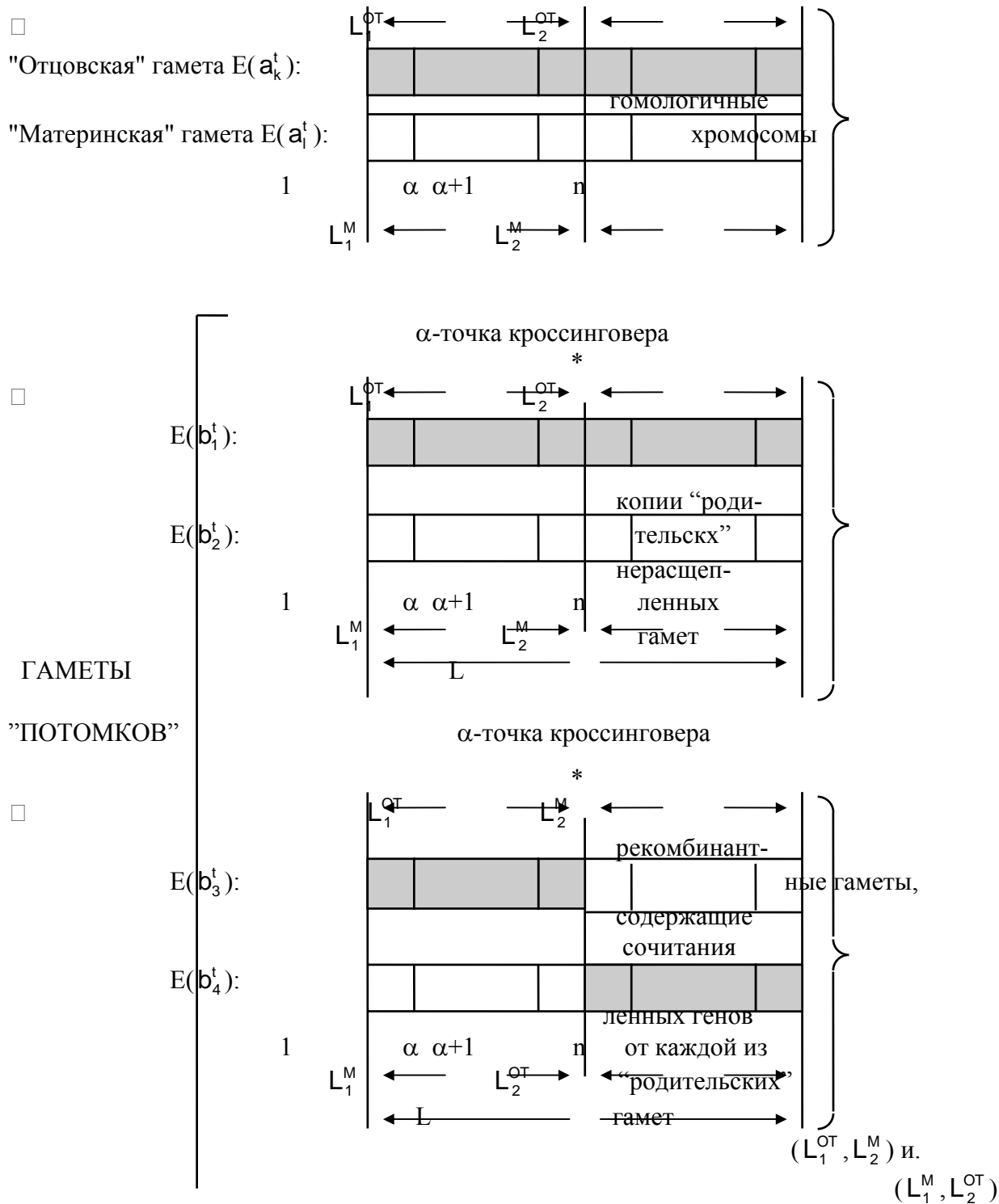


Рис. 5.3. Гаметы "потомков", получаемые в результате простого кроссинговера.

При построении символьной модели для задачи оптимального дихотомического разбиения графа G на структуру генотипа любой особи a_k^t было наложено требование, чтобы число "1" в битах бинарной комбинации $E(a_k^t)$ равнялось мощности множества вершин X_1 подграфа G_1 ($|X_1| = n_1$). В связи с этим после разрыва гомологичных хромосом и объединения различных их участков в одну хромосому

необходимо проверять указанное выше требование для полученной гаметы "потомка". Если число "1" в гамете "потомка" не равно n_1 , то необходимо "подкорректировать" содержание гаметы рекомбинантного типа. Один из способов такой корректировки, называемый *методом случайного перемешивания*, связан со случайной заменой по равновероятностному закону лишних "1" на "0" или лишних "0" на "1".

В качестве иллюстрации приведем реализацию схемы размножения с помощью простого кроссинговера для задачи оптимального дихотомического разбиения.

Алгоритм простого кроссинговера

(для особей $a_k^t, a_l^t \in P^t$ с различающимися генотипами)

1. Случайным образом с вероятностью $(1/(n-1))$ определяется точка кроссинговера α [$\alpha \in [1, n-1]$].
2. Каждая гомологическая хромосома разрывается на два участка в одной и той же точке α :

$$(L_1^{OT}, L_2^{OT}); (L_1^M, L_2^M).$$

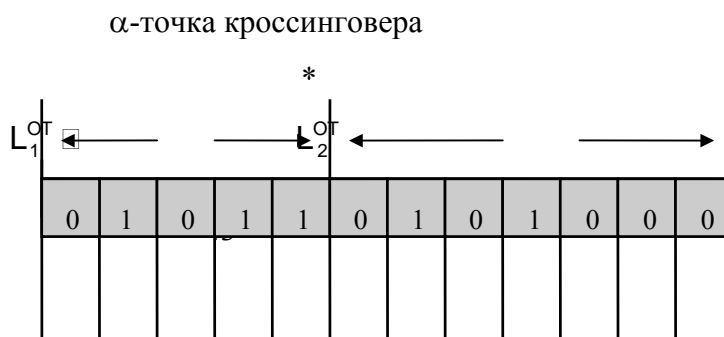
3. Формируются две рекомбинантные гаметы "потомков" b_1^t и b_2^t , объединяющие участок L_1^{OT} с участком L_2^M и участок L_1^M с участком L_2^{OT} :

$$E(b_1^t) = (L_1^{OT}, L_2^M) \text{ и } E(b_2^t) = (L_1^M, L_2^{OT}).$$

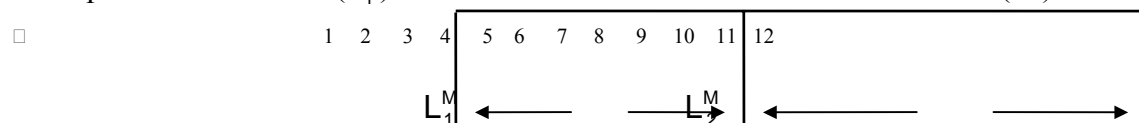
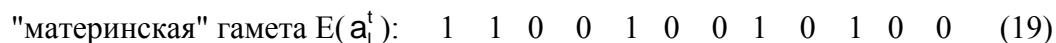
4. Если число "1" в $E(b_1^t)$ и $E(b_2^t)$ равно n_1 , то генотипы "потомков" b_1^t и b_2^t сформированы.

Пример 5.2.

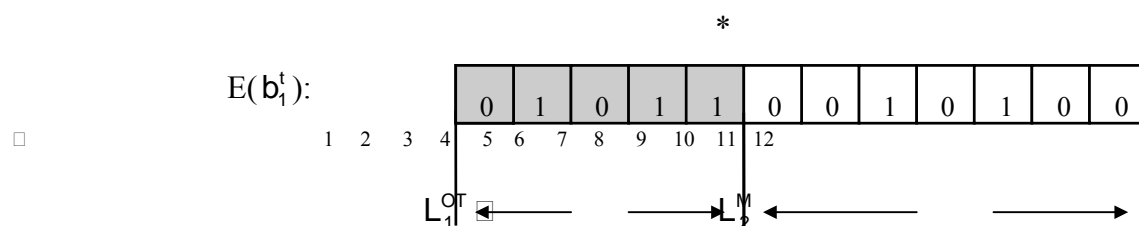
Пусть генотипы "родительской" пары соответствуют разбиениям графа G , приведенным в примере 5.1.:



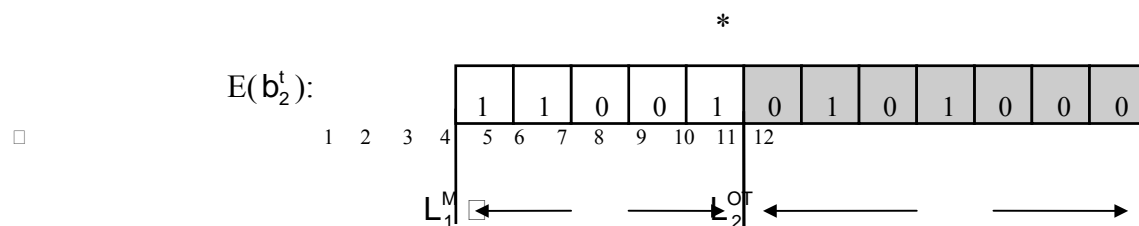
(13)



α -точка кроссинговера



α -точка кроссинговера


$$X'_1 = \{ x_2, x_4, x_5, x_8, x_{10} \},$$

$$X'_2 = \{x_1, x_3, x_6, x_7, x_9, x_{11}, x_{12}\},$$

$$X_2'' = \{x_1, x_2, x_5, x_7, x_9\},$$

$$X_2'' = \{X_3, X_4, X_6, X_8, X_{10}, X_{11}, X_{12}\}.$$

1. $I := \{1, 2, \dots, n\}$; $I_1 := I_2 := 0$.

2. Для всех $i \in I$ формируются гомозиготные гены "потомка":

2.1. Если $e_i^{OT} = e_i^M = 1$, то $\{e_i^\Pi := 1; \bar{l}_1 := \bar{l}_1 \cup \{i\}\}$;

2.2. Если $\mathbf{e}_i^{\text{OT}} = \mathbf{e}_i^{\text{M}} = 0$, то $\{\mathbf{e}_i^{\Pi} := 0; \bar{\mathbf{l}}_2 := \bar{\mathbf{l}}_2 \cup \{i\}\}$;

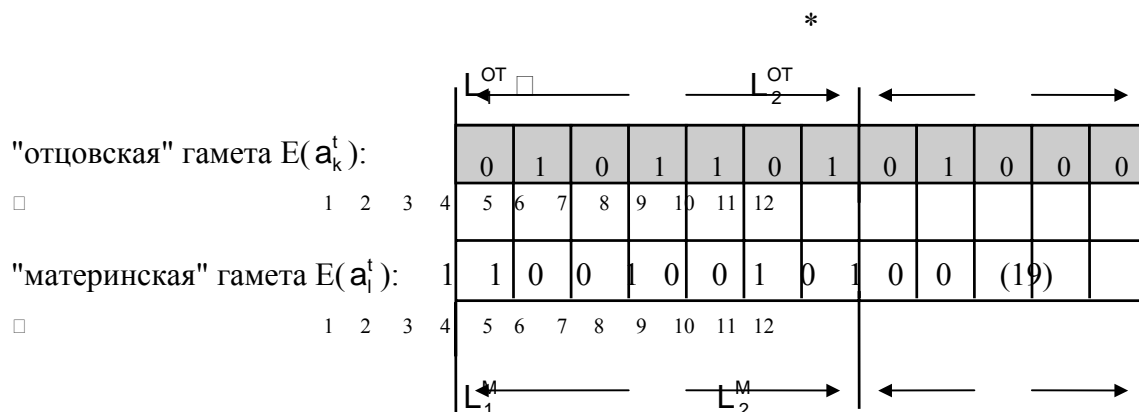
$$3. I_1 := I_1 \setminus \bar{I}_1; I_2 := I_2 \setminus \bar{I}_2.$$

4. Если $n_1(t) < n_2(t)$, то $e_i^\Pi := 1$ для всех $i \in I_1$, иначе $e_i^\Pi := 0$ для всех $i \in I_2$.

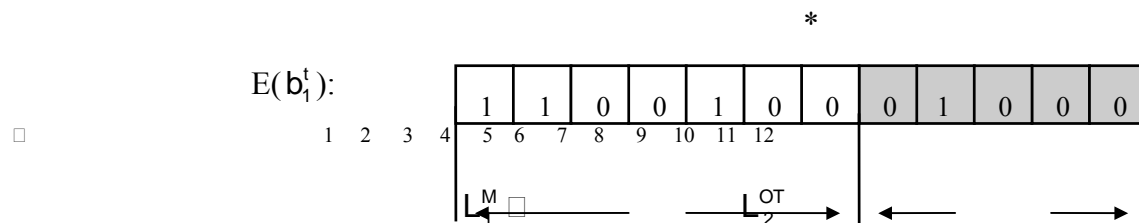
5.2. Если $n_1(t) \geq n_2(t)$, то {случайным образом с равной вероятностью из множества I_1 берется подмножество I_1' , содержащее $S_2' = (n_2 - n_2(t))$ локусов; принимаем $e_i^\Pi := 0$ для $i \in I_1'$; $e_i^\Pi := 1$ для всех $i \in I_1 \setminus I_1'$ } .

Пример 5.3.

α -точка кроссинговера



α -точка кроссинговера


$$I_2 = \{3, 4, 6, 7, 8, 10, 11, 12\}, \quad n_2(t) = 8.$$

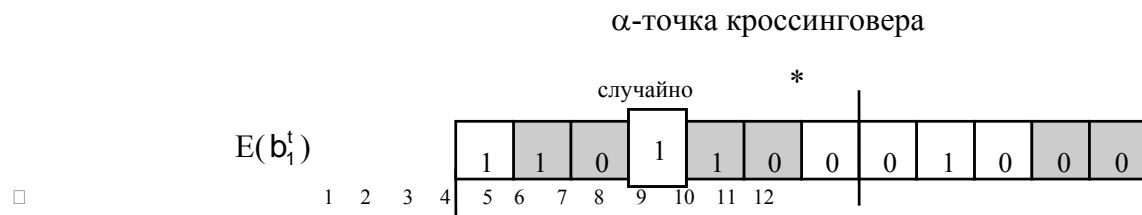
Гомозиготные гены, содержащие "1", составляют множество $I_1' = \{2, 5\}$; а содержащие "0" - множество $I_2' = \{3, 6, 11, 12\}$. Тогда $e_i^\Pi = 1$ для $i \in \{2, 5\}$ и $e_k^\Pi = 0$ для $k \in \{3, 6, 11, 12\}$.

Так как $n_1(t) < n_2(t)$, принимаем, что $e_i^\Pi := 1$ для $i \in I_1 \setminus I_1' := \{1, 9\}$.

Вычисляем $S_1 = (n_1 - n_1(t)) = 1$. Следовательно, из множества $I_2 \setminus I_2' = \{4, 7, 8, 10\}$ необходимо случайным образом выбрать один локус

(пусть это будет локус 4). Помещаем в этот локус "1": $e_4^\Pi := 1$. Окончательно принимаем $e_i^\Pi := 0$ для всех $i \in \{7, 8, 10\}$.

Откорректированная гамета $E(b_1^t)$ является генотипом "потомка" b_1^t :



Построенный генотип соответствует разбиению:

$$X_1 = \{x_1, x_2, x_4, x_5, x_9\}, \quad X_2 = \{x_3, x_6, x_7, x_8, x_{10}, x_{11}, x_{12}\}.$$

При инбридинге, когда генотипы "родителей" совпадают $[E(a_k^t) = E(a_l^t) = \dot{A}]$, простой кроссинговер применяется к одной из "родительских" гамет и позволяет воспроизвести двух "потомков".

Алгоритм простого кроссинговера

(для особей $a_k^t, a_l^t \in P^t$ с одинаковыми генотипами)

1. Случайным образом с вероятностью $(1/(n-1))$ определяется точка кроссинговера $\alpha (\alpha \in [1, n-1])$.
2. "Родительская" гамета E разрывается на два участка в точке α :
 L_1 с генами от 1 до α и L_2 с генами от $(\alpha+1)$ до n .
3. Генетическая информация, содержащаяся в сцепленных генах участка L_1 "родительской" гаметы, передается "потомку" полностью без каких-либо изменений: $e_i^\Pi := e_i, i \in [1, \alpha]$.
4. Определяются параметры $S_1 = (n_1 - n_1(t))$ и $S_2 = (n_2 - n_2(t))$, где
 $n_1(t)$ - число локусов на участке L_1 , содержащих "1";
 $n_2(t)$ - число локусов на участке L_2 , содержащих "0".
5. Если $S_1 = 0$, то $e_i^\Pi := 0$ для всех $i \in [\alpha+1, n]$ (Генотип "потомка" b_1^t сформирован).
6. Если $S_1 < S_2$, то {случайным образом с равной вероятностью из множества локусов участка L_2 выбирается S_1 локус; в выбранные локусы генотипа "потомка" заносятся "1", а

в оставшиеся свободные локусы генотипа "потомка" заносятся "0"}, иначе {случайным образом с равной вероятностью из множества локусов участка L_2 выбираются S_2 локусов; в выбранные локусы генотипа "потомка" заносятся "0", а в оставшиеся свободные локусы генотипа "потомки" заносятся "1" }.

Аналогичным образом формируется генотип "потомка" b_2^t при условии, что генетическая информация, содержащаяся в сцепленных генах участка L_2 "родительской" гаметы, передается "потомку" b_2^t полностью без каких-либо изменений.

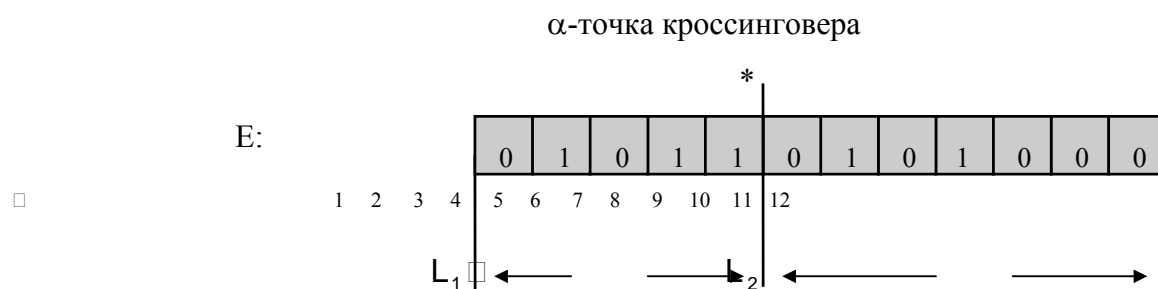
Пример 5.4.

Пусть "родительская" гамета E, соответствующая распределению:

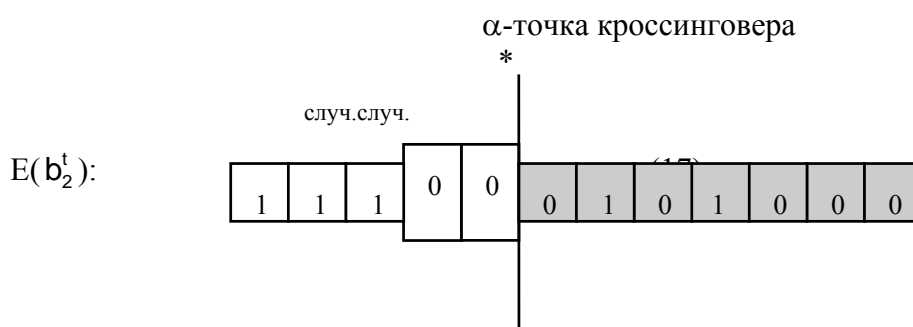
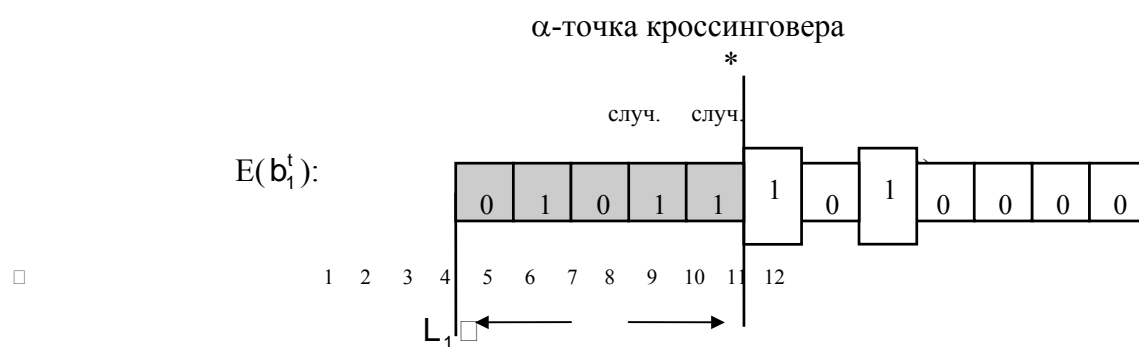
$$X_1 = \{ x_2, x_4, x_5, x_7, x_9 \},$$

$$X_2 = \{ x_1, x_3, x_6, x_8, x_{10}, x_{11}, x_{12} \}$$

разрывается в точке $\alpha=5$:



Генотипы "потомков" b_1^t и b_2^t , полученные с помощью простого кроссинговера, имеют следующий вид:



□

1 2 3 4 5 6 7 8 9 10 11 12

 L_2 □

Здесь локусы 6 и 7 в генотипе $E(b_1^t)$ и локусы 4 и 5 в генотипе $E(b_2^t)$ выбраны случайным образом с равной вероятностью, соответственно, из локусов участка L_2 и участка L_1 .

Генотипу $E(b_1^t)$ соответствует разбиение:

$$X'_1 = \{x_2, x_4, x_5, x_6, x_8\}, \text{ и } X'_2 = \{x_1, x_3, x_7, x_9, x_{10}, x_{11}, x_{12}\},$$

а генотипу $E(b_2^t)$ соответствует разбиение:

$$X''_1 = \{x_1, x_2, x_3, x_7, x_9\}, \text{ и } X''_2 = \{x_4, x_5, x_6, x_8, x_{10}, x_{11}, x_{12}\}.$$