

Дорогие ребята!

Поздравляем вас с участием в заключительном теоретическом этапе Московской олимпиады школьников по генетике! Отвечая на вопросы и выполняя задания, не спешите, так как ответы не всегда очевидны и требуют применения не только биологических знаний, но и общей эрудиции, логики и творческого подхода. Успеха Вам в работе!

Максимальное количество баллов: 100

Время на выполнение работы: 4 часа.

Задание 1. Окраска львиного зева (8 баллов).

Для анализа наследования цвета лепестков львиного зева Вы отобрали 3 отдельных растения и провели различные скрещивания. Лепестки растений №1 и №2 были розового цвета, а лепестки растения №3 были белого цвета. Известно, что все гены находятся на разных хромосомах. Результаты скрещиваний представлены в таблице:

Скрещивание	F1	F2
№1 x №2	все красные	9:7 = красные: розовые
№1 x №3	все красные	16:21:27 = белые: розовые: красные
№2 x №3	все розовые	1:3 = белые: розовые

1.1. Как наследуется признак? Напишите генотипы всех отображенных особей (4 балла), а также генотипы потомков F1.

Решение:

Из первого скрещивания видно, что по двум генам наблюдается комплементарное взаимодействие, где растения с генотипом A_V_ красного цвета, а все остальные – розового (**0.5 балла**). По следующему скрещиванию видно, что есть ещё один ген, который в рецессивном состоянии эпистатически подавляет два первых гена (**0.5 балла**).

___ cc - белый

A_bbC_, aaV_C_, aabbC_ - розовый

A_V_C_ - красный

Генотипы родительских особей (1.5 балла по 0.5 за генотип):

№1 - AAbbCC

№2 - aaBBCC

№3 - aaVVcc

Таблица скрещиваний (1.5 балла по 0.25 балла за каждую, верно заполненную клетку в первых двух столбцах):

Скрещивание	F1	F2
-------------	----	----

$AAbbCC \times aaBBCC$	$AaBbCC$	9:7 = красные : розовые
$AAbbCC \times aaBBcc$	$AaBbCc$	16:21:27 = белые : розовые : красные
$aaBBCC \times aaBBcc$	$aaBBCc$	1:3 = белые : розовые

Если, верно, указаны генотипы родителей и они пронумерованы в соответствии с условиями задачи – первый столбец считается верным.

1.2. Из скрещивания №1 x №2 из F2 вы отобрали наугад красное растение. С какой вероятностью это растение будет гетерозиготным хотя бы по 2 генам? (4 балла).

Решение:

У организмов из F2 скрещивания №1 x №2 по гену С наблюдается у всех гомозиготное доминантное состояние, соответственно для ответа нам нужны только растения с генотипом $AaBbCC$. Так как мы берём изначально красное растение, то оно не может иметь генотипы $Aabb$, $aaBB$, $aaBb$, $aabb$. Нарисуем решётку Пеннета, игнорируя ген С (по 0.125 за каждую верно заполненную клетку):

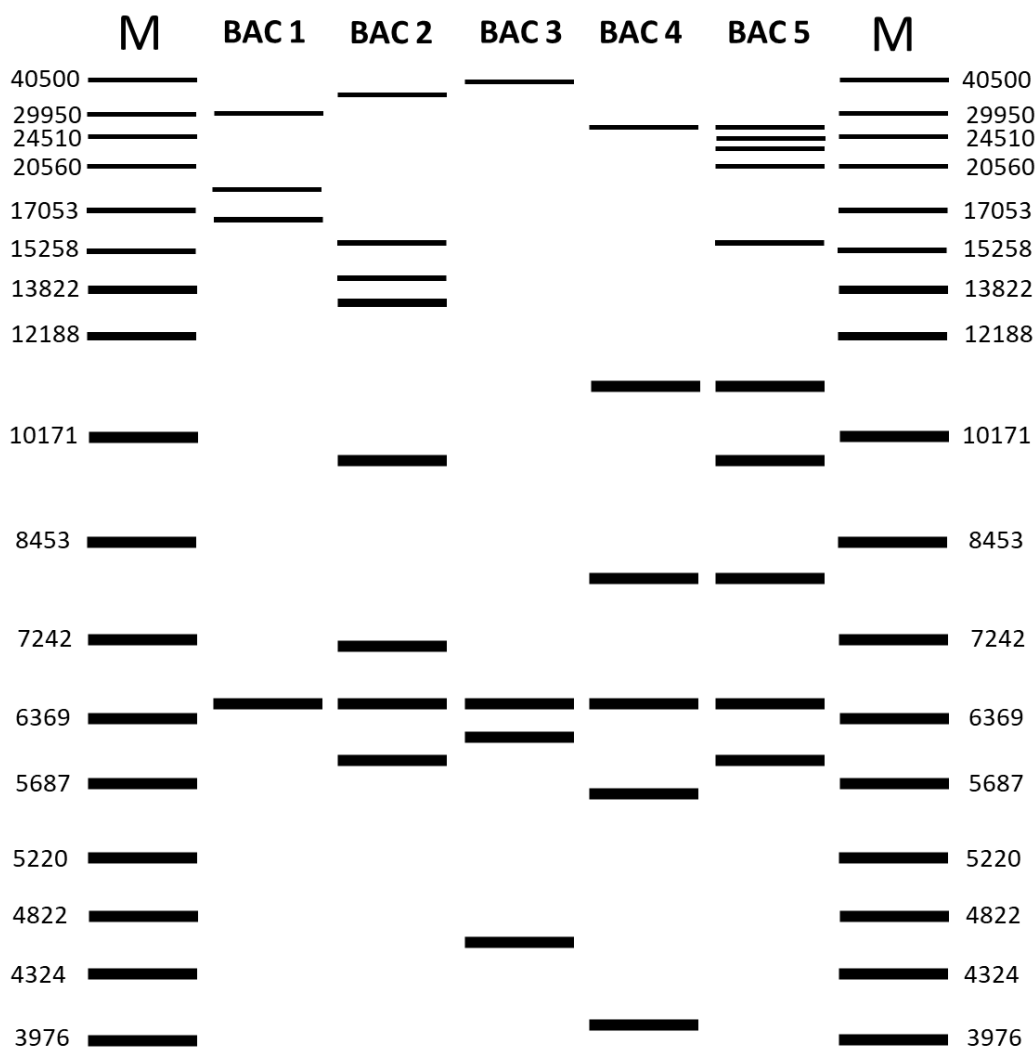
	AB	Ab	aB	ab
AB	AABB	AABb	AaBB	AaBb
Ab	AABb	AAbb	AaBb	Aabb
aB	AaBB	AaBb	aaBB	aaBb
ab	AaBb	Aabb	aaBb	aabb

Нам нужен генотип $AaBb$, их 4 ячейки, а всего вариантов 9, так как мы считаем только красные ячейки (1 балл). Искомая вероятность = 4/9 (1 балла)

Задание 2. Проект геном человека (10 баллов).

В ходе проекта «Геном человека» геномную ДНК из клеток человека вначале фрагментировали на множество больших перекрывающихся фрагментов, размером от примерно 50 до 300 тысяч пар нуклеотидов, после чего клонировали такие фрагменты в векторы, называемые бактериальными искусственными хромосомами (BAC, bacterial artificial chromosome). Затем каждую из таких BAC с вставкой уникального участка генома дробили на перекрывающиеся фрагменты меньшей длины, определяли последовательности нуклеотидов небольших фрагментов секвенированием по Сэнгеру, после чего по перекрытию последовательностей этих небольших фрагментов восстанавливали полную последовательность каждой уникальной BAC. Для того, чтобы собрать полную последовательность генома, необходимо было узнать порядок (взаимное расположение) BAC на хромосомах, то есть получить BAC-карту генома. Было предложено несколько подходов для установления взаимного расположения BAC. Один из таких подходов основан на использовании эндонуклеаз рестрикции (рестриктаз) – рестрикционный анализ. В этом случае, каждую BAC обрабатывают одной и той же рестриктазой (или комбинацией

рестриктаз), после чего устанавливают перекрытие ВАС по наличию общих фрагментов-продуктов рестрикции. Другой подход основан на использовании STS (sequence-tagged sites) – коротких уникальных последовательностей генома, присутствие которых можно проверить в каждой ВАС, например, с помощью полимеразной цепной реакции (ПЦР). Предположим, что перед вами стоит задача – определить взаимное расположение нескольких ВАС (ВАС 1-5) с перекрывающимися фрагментами определенного участка генома, используя два описанных выше подхода. На рисунке ниже показан результат электрофореза продуктов полного расщепления каждой из пяти ВАС смесью рестриктаз NotI, FseI и SrfI; М – маркер молекулярных длин (размер фрагментов подписан в парах нуклеотидов). Учтите, что выявляемый на всех дорожках фрагмент размером примерно 6500 п.н. – это фрагмент самого ВАС-вектора, а не вставки.



В таблице ниже знаком «+» отмечено наличие, а знаком «-» – отсутствие каждого из использованных в анализе STS-маркеров (А – G) в каждой из ВАС.

	STS						
	A	B	C	D	E	F	G
ВАС 1	-	-	-	+	-	+	+
ВАС 2	+	-	-	-	+	-	-
ВАС 3	+	-	+	-	-	-	-
ВАС 4	-	-	-	-	-	+	-
ВАС 5	-	+	-	+	-	+	-

2.1. Заполните таблицу, отметив знаком «+» перекрывание фрагментов генома в соответствующих ВАС, которое следует только из данных рестрикционного анализа. Если перекрывание фрагментов генома нельзя установить по результатам рестрикционного анализа, оставьте соответствующие ячейки пустыми (1.25 балла):

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					
ВАС 2					
ВАС 3					
ВАС 4					
ВАС 5					

Решение (по 0.125 за каждую верную ячейку):

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					
ВАС 2					
ВАС 3					
ВАС 4					
ВАС 5		+		+	

2.2. Заполните таблицу, отметив знаком «+» перекрывание фрагментов генома в соответствующих ВАС, которое следует только из данных STR-анализа. Если перекрывание фрагментов генома нельзя установить по результатам STR-анализа, оставьте соответствующие ячейки пустыми (1.25 балла):

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					

ВАС 2					
ВАС 3					
ВАС 4					
ВАС 5					

Решение (по 0.125 за каждую верную ячейку):

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					
ВАС 2					
ВАС 3		+			
ВАС 4	+				
ВАС 5	+			+	

2.3. Заполните таблицу, отметив, какие из фрагментов генома перекрываются друг с другом, учитывая данные как рестрикционного анализа, так и STR-анализа (поставьте знак «+» в ячейку, если последовательности в соответствующих ВАС перекрываются; при отсутствии перекрытия поставьте в соответствующую ячейку знак «-»; в заполненной таблице не должно остаться пустых ячеек) (2.5 балла)

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					
ВАС 2					
ВАС 3					
ВАС 4					
ВАС 5					

Решение (по 0.25 за каждую верную ячейку):

	ВАС 1	ВАС 2	ВАС 3	ВАС 4	ВАС 5
ВАС 1					
ВАС 2	-				
ВАС 3	-	+			
ВАС 4	+	-	-		
ВАС 5	+	+	-	+	

2.4. Укажите, какие из ВАС находятся по краям анализируемого участка генома? (3 балла)

Решение:

ВАС1 и ВАС3 (или просто 1 и 3) (по 1.5 балла за каждый верный ответ).

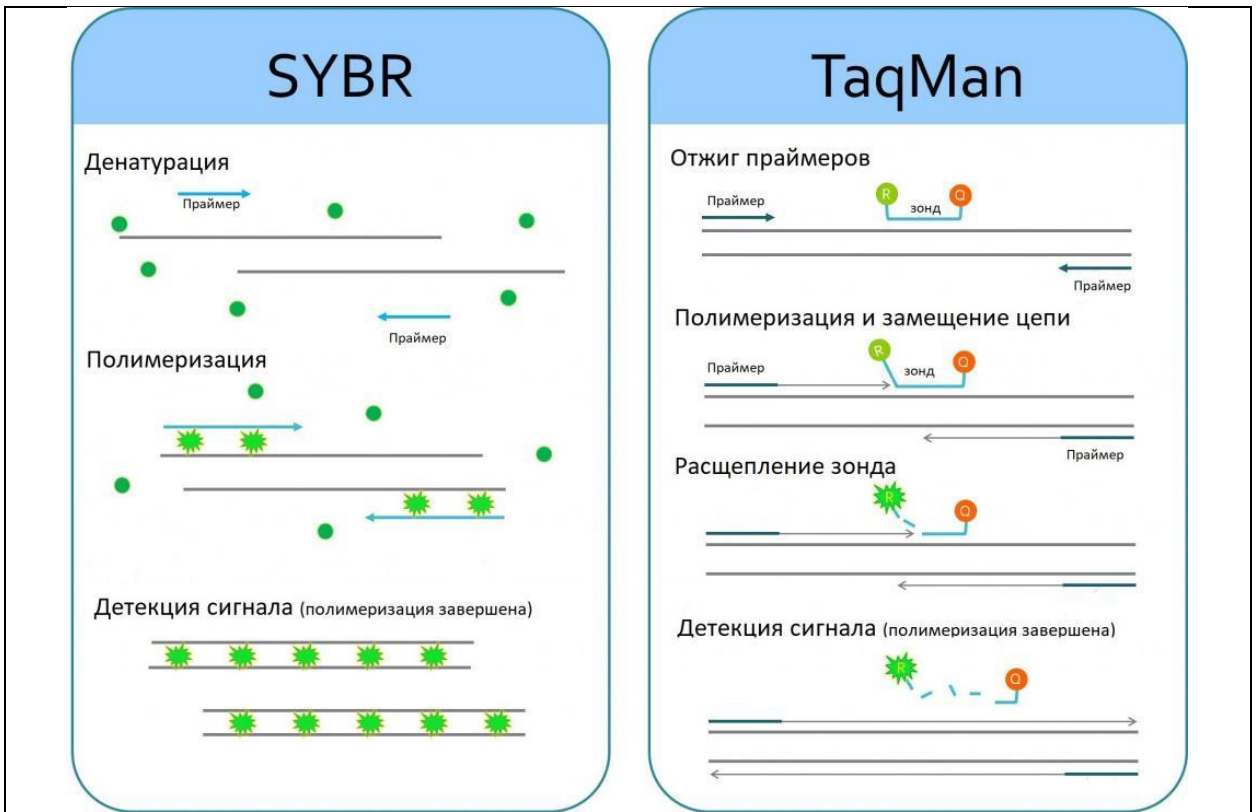
2.5. В какую из пяти ВАС встроен фрагмент наибольшего размера? (2 балла)

Решение (2 балла):

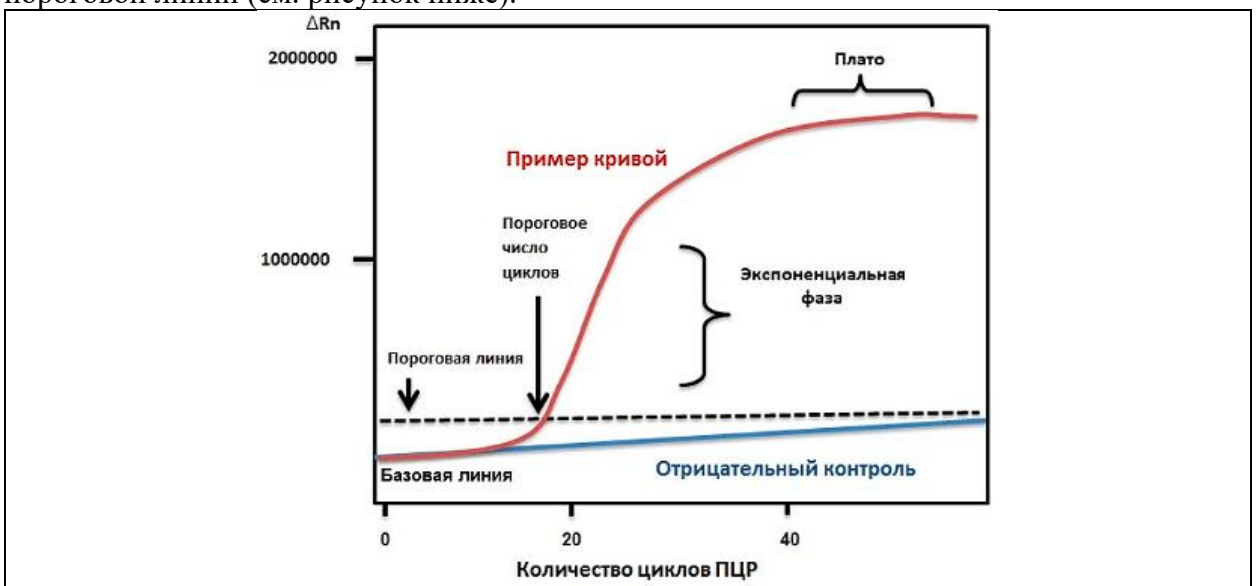
ВАС5

Задание 3. ПЦР в реальном времени (6 баллов).

Молекулярными биологами часто используется метод ПЦР в реальном времени. ПЦР в реальном времени основана на измерении интенсивности сигнала флуоресцентного красителя во время амплификации. Существуют разные варианты систем с красителями, две из которых (SYBR и TaqMan) показаны на схеме ниже.



Значение флуоресцентного сигнала в каждом образце измеряется на каждом цикле амплификации. Интенсивность сигнала пропорциональна количеству накопленного продукта ПЦР. В методе используются специфичные праймеры. Размер получаемого в результате амплификации фрагмента обычно не превышает 200 пар нуклеотидов. Для расчета начального числа копий матрицы необходимо определить тот этап реакции, когда амплификация находится в экспоненциальной фазе и при этом в компоненты реакции находятся в достаточном количестве. Один из подходов к анализу результатов называется «пороговый метод», когда для каждого образца определяют значение порогового цикла реакции C_t , threshold cycle, как точки пересечения графика накопления продукта и пороговой линии (см. рисунок ниже).



Метод ПЦР в реальном времени в целом довольно чувствительный и точный, тем не менее, возможно получение ложноположительных результатов. Для вычленения таковых можно

использовать кривые плавления (денатурации). Кривая плавления — зависимость скорости падения флуоресценции от температуры плавления. Она позволяет детектировать по изменению сигнала флуоресценции интеркалирующего красителя наличие конкретных фрагментов по соответствующим температурам плавления.

ПЦР в реальном времени используется для оценки вирусной нагрузки (число копий вируса на единицу объема анализируемого образца). Для этого готовят контрольные образцы — ряд разведений матрицы с известной концентрацией. Для определения количества молекул матрицы в анализируемом образце необходимо использовать калибровочную кривую зависимости C_t от логарифма количества молекул матрицы на старте реакции.

3.1. Почему при использовании системы TaqMan размер зонда не должен превышать 40 нуклеотидов? Ответ поясните (1 балл).

Решение:

Не будет эффективного гашения флуоресценции (FRET) (1 балл)

3.2. Почему массу матрицы в случае, если используется тотальная ДНК, на одну реакцию для получения достоверных результатов нужно выбирать из диапазона 1 пг – 100 нг? Ответ поясните (1.5 балла).

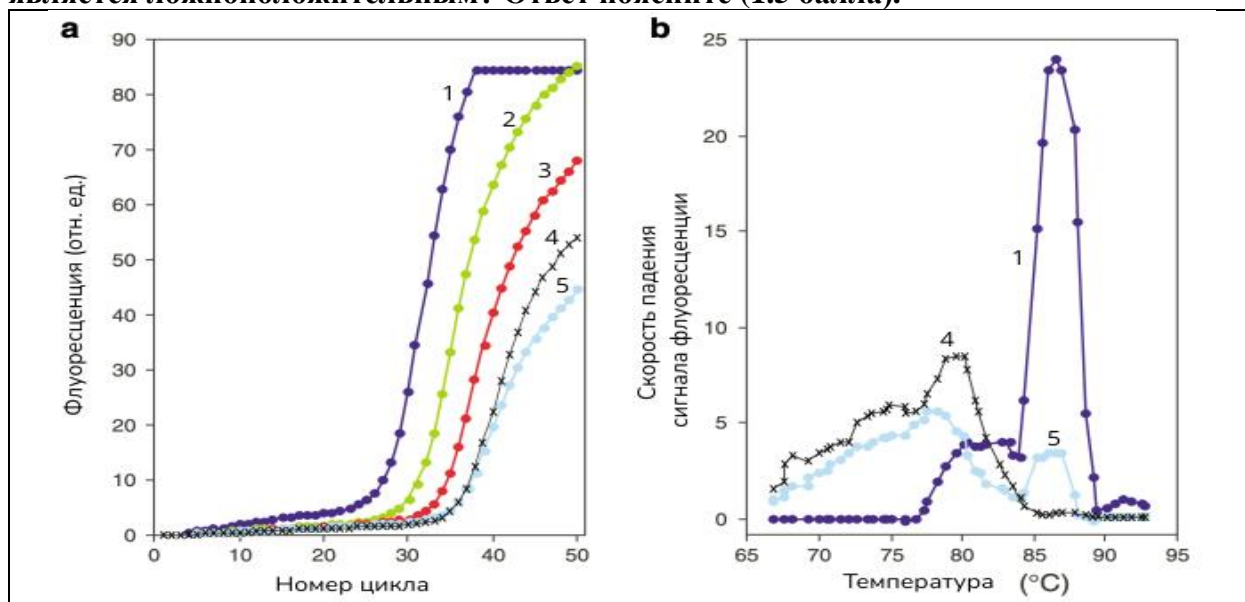
Решение:

(1) снижение эффективности ПЦР,

(2) невозможность достоверной детекции увеличения флуоресцентного сигнала

Полный балл за любое логичное объяснение из перечисленных

3.3. Ниже на рисунке а представлены кривые накопления продуктов для пяти образцов (1-5). Используются одни и те же условия ПЦР и праймеры. На рисунке б – кривые плавления для образцов 1, 4 и 5. Для какого из образцов результат ПЦР является ложноположительным? Ответ поясните (1.5 балла).



Решение:

Образец 4 – вода (на кривой плавления температура плавления продукта для образца 4 отличается, отсутствует пик в области 85–90 градусов, что указывает на то, что это результат неспецифичной амплификации) (0.5 баллов за ответ, 1 за пояснение)

0.5 баллов за ответ – ложноположительный результат

0.5 баллов за ответ – отрицательный контроль

3.4. Для проведения анализа на количество вирусных частиц SARS-CoV-2 в соскобе можно использовать метод ПЦР в реальном времени, когда сначала из образца выделяют матрицу (наследственный материал вируса). После этого матрицу помещают в пробирку, в которой происходят все необходимые для анализа процессы. Перечислите обязательные компоненты, которые нужно добавить в такую пробирку (кроме выделенной матрицы), чтобы в итоге получить данные о количестве вируса в соскобе при описанном подходе (не используйте аббревиатуры, пишите названия компонентов полностью и по-русски) (2 балла).

Решение (за все остальные пункты, кроме ОТ по 0.25 балла):

Обратная транскриптаза (0.5 балла)

Буфер

Дезоксирибонуклеозидтрифосфаты

Прямой праймер

Обратный праймер

Полимераза

Зонд/интеркалирующий краситель

Задание 4. Анализ вирусной нагрузки (15 баллов).

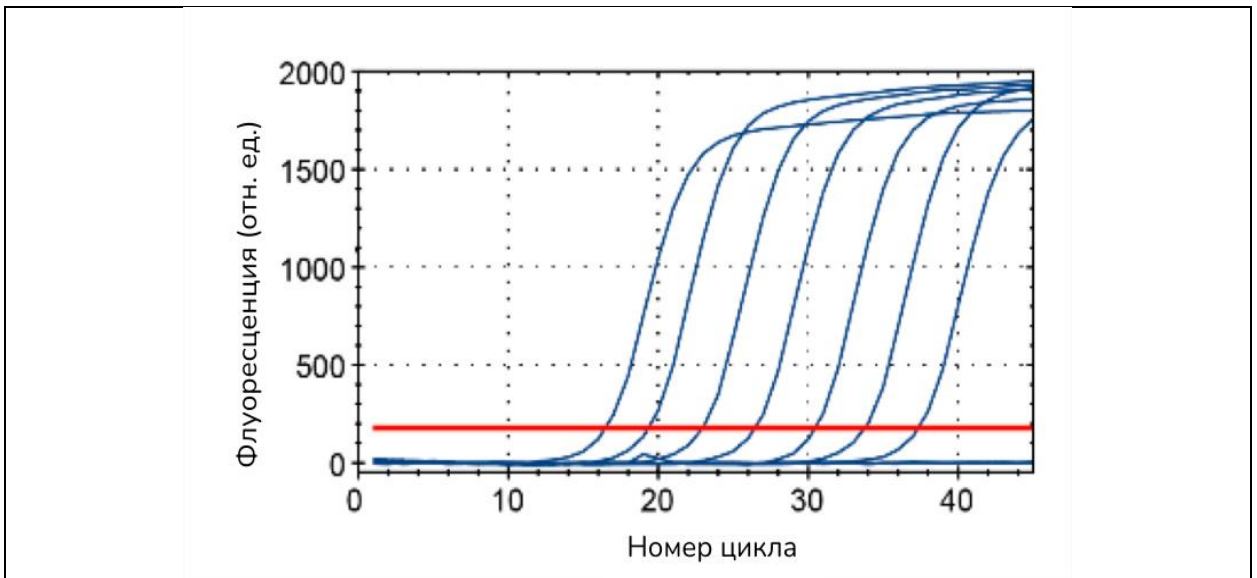
При работе с заданием используйте линейку и программируемый калькулятор!

Для определения вирусной нагрузки гепатита С при оценке эффективности проводимой противовирусной терапии у пациента X и прогнозировании развития хронического процесса провели ПЦР в реальном времени. Если концентрация вируса менее 8×10^5 МЕ/мл (для вируса гепатита С 1 МЕ/мл эквивалентна 4 копиям вируса на 1 мл цельной крови), то проводимое противовирусное лечение эффективно, если выше, то необходимо рассмотреть альтернативные варианты терапии.

При выделении наследственного материала вируса из 10 мл крови пациента X потери составили примерно 10%, конечный объем рабочего раствора образца составил 100 мкл. Для построения калибровочной кривой использовали ряд из семи последовательных разведений контрольной матрицы с известной концентрацией. Исходная концентрация контрольной матрицы составляла 20 нг/мкл. Объем растворов матриц, добавлявшийся в пробирку для ПЦР, составлял 5 мкл (и для контрольных, и для образца). Конечный объем смеси для ПЦР составлял 20 мкл. Средняя молекулярная масса нуклеотида составляет 340 Да. Размер генома вируса гепатита С составляет примерно 9600 нуклеотидов.

Результат анализа представлен на рисунке ниже. Семь кривых слева направо отражают результат для семи контрольных образцов с последовательными разведениями контрольной матрицы (самая левая линия – исходная неразведенная матрица, далее – разведение в 10 раз, далее – разведение в 100 раз и т.д.).

Постройте калибровочную кривую зависимости C_t от десятичного логарифма количества молекул контрольной матрицы на старте реакции. При построении калибровочной кривой количество циклов округляйте до целых значений. Значение C_t для образца из крови пациента X составляет 32 (на рисунке кривая для этого образца не представлена). Определите с использованием построенной Вами калибровочной кривой, является ли проводимое для пациента X противовирусное лечение эффективным. Ответ обоснуйте расчетами.



Решение:

Масса генома вируса (г/моль) = $9600 \cdot 340 = 3264000$ (1 балл)

Количество моль контрольной матрицы на пробирку рассчитывается, как масса контрольной матрицы на пробирку/молярную массу

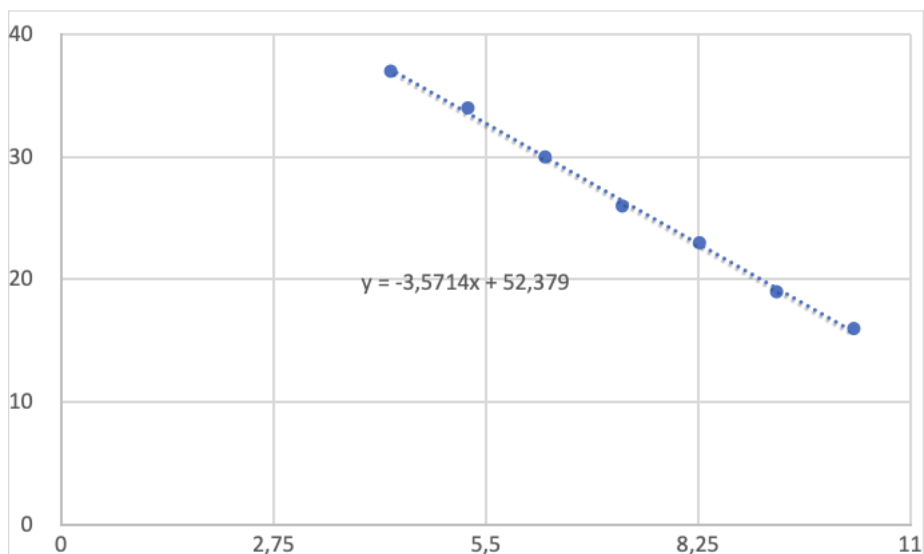
Количество копий рассчитывается как полученное количество вещества, умноженное на число авогадро ($6 \cdot 10^{23}$)

Затем получаем десятичный логарифм от данного значения (0.5 баллов за каждое верное значение)

Значения C_t получаем согласно графику (по 0.5 баллов за каждое верное значение)

Масса контрольной матрицы на пробирку, нг	Масса генома, г/моль	Количество моль контрольной матрицы на пробирку	Количество копий контрольной матрицы на пробирку	Lg кол-ва копий матрицы на пробирку	C_t (по графику из условия)
1100	3264000	0,0000000000000000306372549019608	18449754901,9608	10,2659906010949	16
210	3264000	0,0000000000000000306372549019608	1844975490,19608	9,26599060109492	19
31	3264000	0,0000000000000000306372549019608	184497549,019608	8,26599060109492	23
40,1	3264000	0,0000000000000000306372549019608	18449754,9019608	7,26599060109492	26
50,01	3264000	0,0000000000000000306372549019608	1844975,49019608	6,26599060109492	30
60,001	3264000	0,0000000000000000306372549019608	184497,549019608	5,26599060109492	34
70,0001	3264000	0,0000000000000000306372549019608	18449,7549019608	4,26599060109492	37

Построен график в логарифмических координатах (3 балла):



Используя полученное уравнение кривой:

Определено количество копий для образца с Ct = 32 (1 балл)

Верно рассчитано значение МЕ/мл у пациента (2 балла)

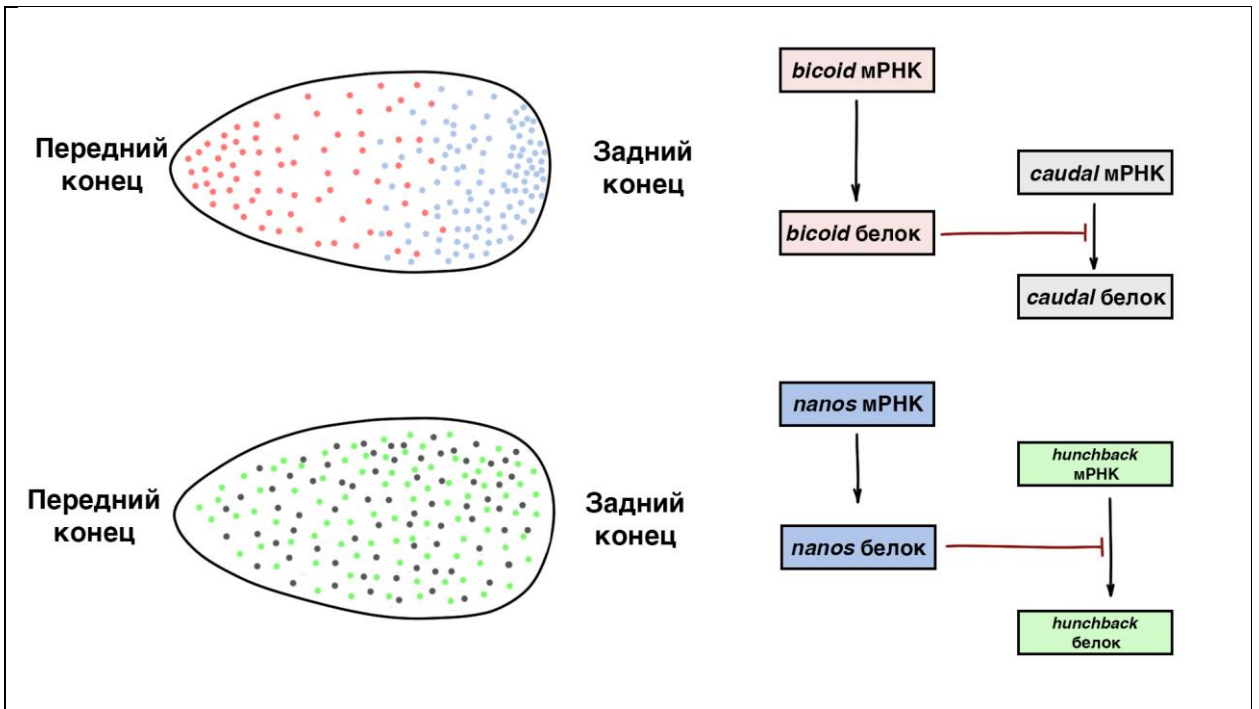
Lg кол-ва копий для образца с Ct=32 (по уравнению линии тренда)	Кол-во копий на одну ПЦР пробирку для образца пациента X (то есть, на 5 мкл рабочего раствора)	Кол-во копий на весь объем рабочего раствора после выделения из крови (на 100 мкл)	Учет потери 10% при выделении РНК	Количество копий на 1 мл крови	МЕ/мл у пациента X
5,70616565	508353,303340899	10167066,066818	11183772,6734998	1118377,26734998	279594,316837495

Вывод: терапия неэффективна (1 балл)

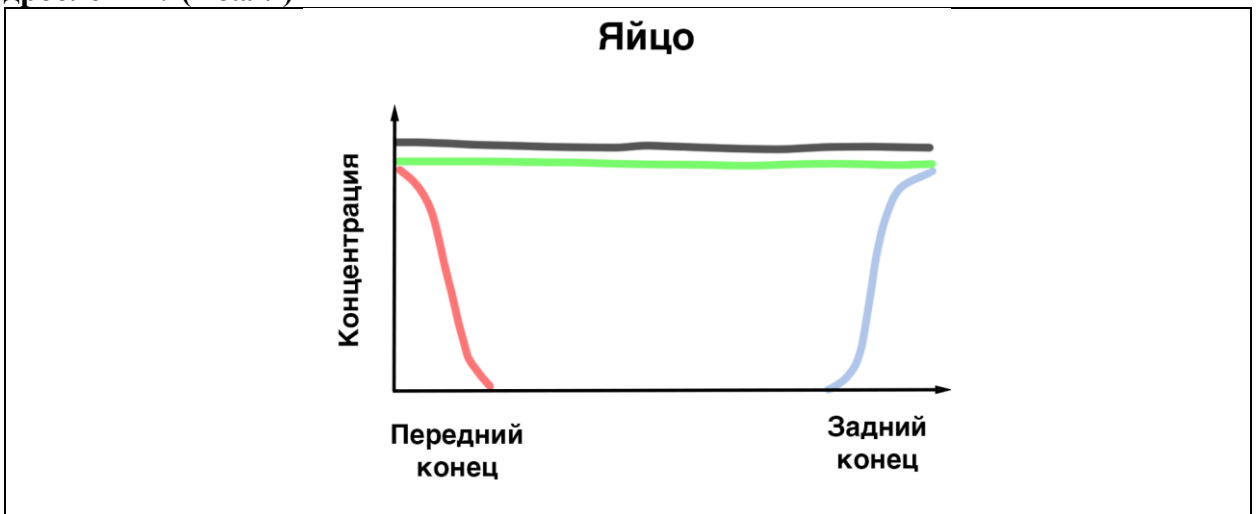
Задание 5. Развитие плодовой мушки (6 баллов).

Большинство животных имеют ярко выраженную полярность. Так, например, выделяют головной и хвостовой конец тела, спинную и брюшную сторону.

На рисунке показано схематично расположение мРНК двух генов *bicoid*, *nanos*, *hunchback* и *caudal* которые будут образовывать одноименные белки. На схеме регуляции показано, как эти белки взаимодействуют.

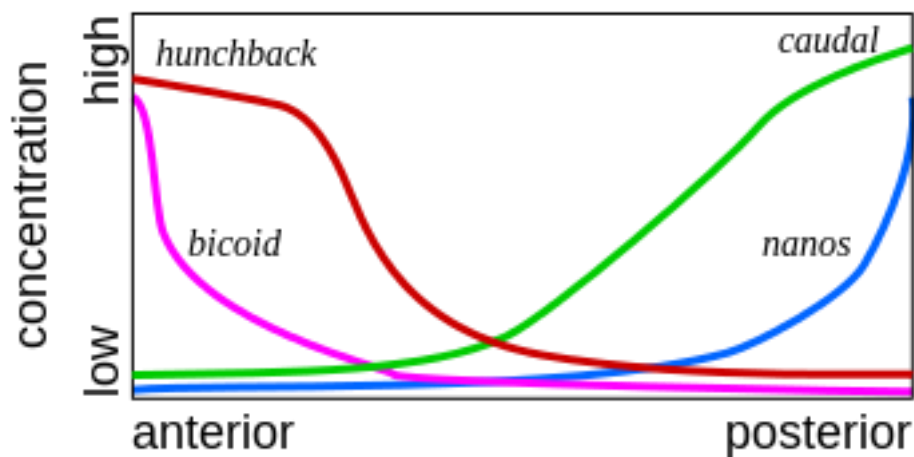


5.1. Если известно расположение мРНК генов в яйце, предположите (**нарисуйте график**), как будет выглядеть градиент *caudal* и *hunchback* в момент начала дробления? (1 балл)

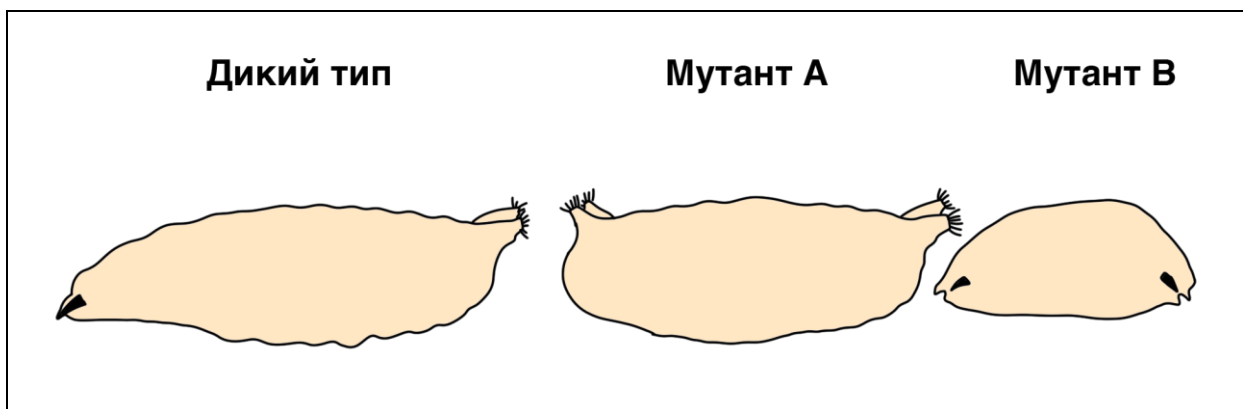


Решение:

Можно заметить, что согласно схеме белок *bicoid* ингибирует формирование белка *caudal*, а белок *nanos* ингибирует формирование белка *hunchback*. Таким образом (**по 0.5 балла за каждый верный градиент**):



5.2. На рисунках изображена личинка дрозофилы, которая развивается нормально (дикий тип), а также мутанты с потерей функции по генам (*bicoid* и *nanos*). Предположите, какие мутации произошли в каждом мутантном организме (2 балла).

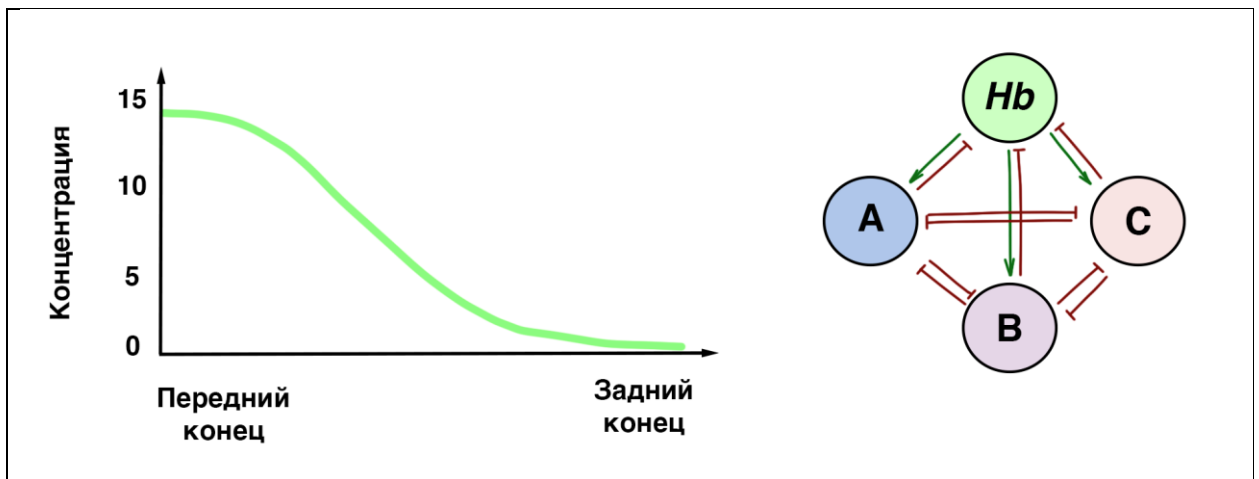


Решение:

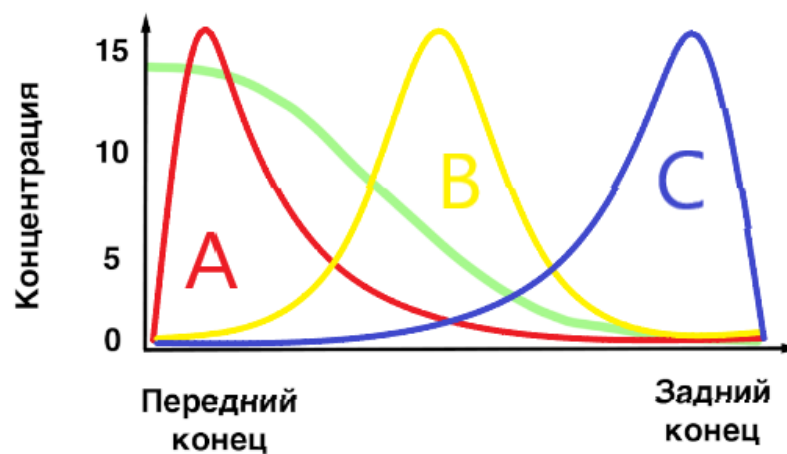
А – мутант по гену *bicoid* – нормальный передний конец не формируется (у личинки 2 задних конца) (1 балл)

В – мутант по гену *nanos* – нормальный задний конец не формируется (у личинки 2 передних конца) (1 балл)

5.3. Дрозофила является сегментированным животным. Ген *hunchback* запускает другие гены развития, которые разделяют развивающийся зародыш на большее число сегментов. Известно, что ген А активируется только при высокой концентрации *hunchback* (выше 10 нмоль/л). Ген В активируется при средней концентрации *hunchback* (ниже 10 нмоль/л, но выше 5 нмоль/л), а ген С запускается при низкой концентрации *hunchback* (ниже 5 нмоль/л). На схеме показана схема взаиморегуляции генов. На графике показано распределение белка *hunchback* в цитоплазме дробящейся яйцеклетки. Предположите, как будет выглядеть распределение белков А, В, С на стадии позднего эмбриона (нарисуйте график аналогичный первому заданию) (3 балла).



Решение (по 1 баллу за каждый верный градиент):



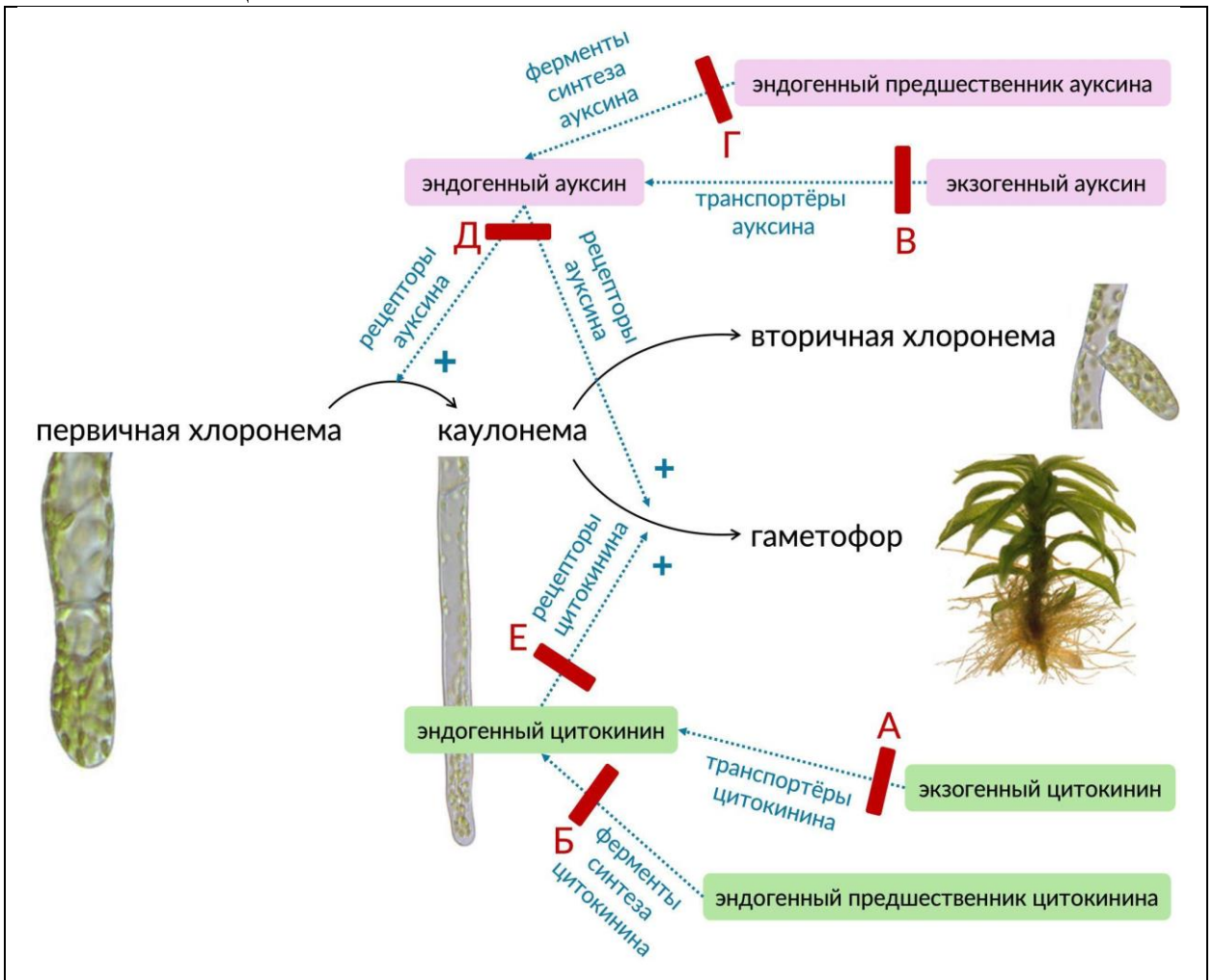
Задание 6. Гаметофиты мхов! (15 баллов)

Регуляция развития гаметофитов ауксинами и цитокининами активно исследуется на мохообразных, например, на *Physcomitrium patens*. Главная цель исследований состоит в том, чтобы выявить молекулярных игроков на разных этапах онтогенеза гаплоидного поколения.

Развитие гаметофита *P. patens* начинается с прорастания споры в первичную хлоронему - разветвлённую зелёную нить, клетки которой имеют много хлоропластов. На хлоронеме развиваются нити каулонемы, клетки которой крупнее, содержат меньше хлоропластов и со временем накапливают в пластидах красные каротиноиды. дальнейшая судьба каулонемы зависит от факторов среды: либо появится вторичная хлоронема, либо разовьются листостебельные гаметофоры. Появление каулонемы стимулируется ауксинами, в то время как развитие гаметофоров возможно только в случае совместного действия ауксинов и цитокининов. На рисунке 1 показаны этапы развития гаметофита, на которые указанные фитогормоны оказывают влияние.

Для раскрытия механизма регуляции, показанного на рисунке 1, использованы методы химического мутагенеза и анализа фенотипов растений на среде с добавлением растворов гормонов (экспериментальные группы) и без них (контрольные группы) - НУК, нафтилуксусная кислота, ауксин, и БАП, бензиламинопурин, цитокинин. Экспериментальные группы подвергались обработке сериями концентраций растворов НУК и БАП - от 500 нМ до 12.5 мкМ и от 50 нМ до 50 мкМ соответственно. Мутанты были получены при обработке спор растений дикого типа мутагеном и последующем отборе

мутантных фенотипов. Все мутанты были разделены на 8 категорий в зависимости от набора наблюдаемых отличий по фенотипу от дикого типа. У гаметофитов были изучены признаки размера нитей протонемы, количество нитей первичной хлоронемы и каулонемы, количество гаметофоров и наличие аномалий развития гаметофоров. Результаты анализа показаны в таблицах 1 и 2.



Регуляция развития гаметофита *Physcomitrium patens* ауксинами и цитокининами. Сплошные стрелки указывают на переход между стадиями онтогенеза гаметофита. Пунктирные стрелки обозначают процессы, вовлечённые в передачу сигналов гормонов, над стрелками подписаны участники этих процессов. Красными чертами с цифрами 1-5 отмечены участки схемы регуляции развития, где обнаружены мутации в рамках данного исследования.

Категория мутанта	Размер протонемы в сравнении с диким типом ¹	Положительное ●, отрицательное ● или отсутствие воздействия ● гормона на размер протонемы в сравнении с исходным размером протонемы мутанта на среде без гормонов		Положительное ●, отрицательное ● или отсутствие воздействия ● гормона на количество нитей хлоронемы и каулонемы у мутантов в сравнении с количеством нитей мутанта на среде без гормонов					
				Количество нитей первичной хлоронемы			Количество нитей каулонемы		
		Воздействие НУК 500 нМ – 12.5 μМ ³	Воздействие БАП 50 нМ – 50 μМ	Среда без гормонов ²	Воздействие НУК 500 нМ – 12.5 μМ ³	Воздействие БАП 50 нМ – 50 μМ	Среда без гормонов ²	Воздействие НУК 500 нМ – 12.5 μМ ³	Воздействие БАП 50 нМ – 50 μМ
1	▼	●	●	▲	●	●	0	0	0
2	▼	●	●	▲	●	●	▶	●	●
3	▶	●	●	▶	●	●	▶	●	●
4	▼	●	●	▲	●	●	▼	●●	●
5	▶	●●	●	▶	●●	●	▶	●●	●
6	▼	●	●	▲	●	●	▼	●●	●
7	▼	●	●	▲	●	●	▼	●	●
8	▼	●●	●	▶	●●	●	▶	●●	●
Дикий тип		●●	●		●●	●		●●	●

¹
▼ среднее значение размера протонемы у мутанта меньше, чем у дикого типа
▶ среднее значение размера протонемы у мутанта и дикого типа не отличаются

²
▼ среднее количество нитей у мутанта меньше, чем у дикого типа
▶ среднее количество нитей у мутанта и дикого типа не отличаются
▲ среднее количество нитей у мутанта больше, чем у дикого типа

³
●● низкие концентрации гормона не оказывают влияние на признак, а высокие влияют отрицательно
●● низкие концентрации гормона не оказывают влияние на признак, а высокие влияют положительно

Воздействие гормонов на размер и количество нитей протонемы *Physcomitrium patens*

Категория мутанта	Воздействие гормона на развитие и количество гаметофоров у дикого типа и мутантов					
	Количество гаметофоров на гаметофите			Развитие гаметофоров		
	Положительное ●, отрицательное ● или отсутствие воздействия ● гормона в сравнении с контролем на среде без гормонов			● нормальное развитие ● аномалии развития / арест развития ● стимуляция при низких конц., ингибирование при высоких конц. 0 гаметофоров не заложилась		
	Среда без гормонов в сравнении с диким типом	Воздействие НУК 500 нМ - 12.5 μМ	Воздействие БАП 50 нМ - 50 μМ	Среда без гормонов	Воздействие НУК 500 нМ - 12.5 μМ	Воздействие БАП 50 нМ - 50 μМ
1	0	0	0	0	0	0
2	0	0	●	●	●	●
3	▶	●	●	●	●	●●
4	▼	●●	●	●	●	●●
5	▶	●	●	●	●	●
6	▼	●●	●	●	●	●●
7	▼	●	●●	●	●	●●
8	0	0	0	0	0	0
Дикий тип		●	●	●	●	●●

Воздействие гормонов на количество гаметофоров и на наличие аномалий развития гаметофоров *Physcomitrium patens*.

6.1. Внимательно изучите результаты исследования, показанные в таблицах 1 и 2. Соотнесите категории мутантов (1-8) с предположительным местом мутации на схеме регуляции развития гаметофита *Physcomitrium patens* (А-Д) на рисунке 1. Несколько мутантов могут иметь мутацию в одном и том же месте схемы регуляции (8 баллов).

Категория мутантов	1	2	3	4	5	6	7	8
Место мутации на схеме регуляции								

Решение (по 1 баллу за каждую верную клетку таблицы):
В скобах указана возможный альтернативный ответ

Категория мутантов	1	2	3	4	5	6	7	8
Место мутации на схеме регуляции	Д	Б(В)	В	Г(Б)	А	Г(Б)	Г	Е

6.2. По мере развития гаметофита происходит накопление красных каротиноидов в пластидах клеток каулономы. Известно, что для стимуляции синтеза каротиноидов необходим ауксин, который активизирует экспрессию генов синтеза пигментов. Дополнительно было выяснено, что все изученные в этих экспериментах мутации рецессивны по отношению к аллелям дикого типа и не образуют групп сцепления. Было изучено наследование окраски каулономы в следующей серии скрещиваний. Мутант из категории 1 с частичным нарушением функции гена (достаточным для формирования гаметаангиев и развития зиготы) гибридизировали с мутантом 4. От спорофита было получено 1000 спор, которые были высеяны на питательную среду. 50 выросших гаметофитов были скрещены с мутантом из категории 3, от полученных спорофитов было собрано 10 000 спор. Эти споры посадили на среду, содержащую НУК. При условии, что все гаметофиты и спорофиты выживают и фертильны, а окраска зависит только от присутствия ауксина, оцените вероятность получения неокрашенных каулоном в последнем поколении. Ответ запишите в процентах, округлив десятичное число долей до второго знака после запятой. Аргументируйте ответ схемами скрещиваний (7 баллов).

Решение:

Пусть А и а - нормальный и мутантный аллель по рецептору ауксина; В и в - нормальный и мутантный аллель по ферменту синтеза ауксина; С и с - нормальный и мутантный аллель по транспортёру ауксина. Тогда первое скрещивание аВС х АдС (1 балл). Спорофиты единообразны АаВвСС (1 балл). В спорах и гаметофитах будет расщепление по генотипу $\frac{1}{4} ABC$, $\frac{1}{4} AbC$, $\frac{1}{4} aBC$, $\frac{1}{4} abC$ (1 балл).

При скрещивании этих гибридов с мутантом категории 3 АВс отсутствие окраски каулономы на среде с НУК будет наблюдаться только у генотипов гаметофитов с рецессивным а (нет рецептора ауксина) (1 балл), с генотипом АвС (есть рецептор, нарушен синтез, есть транспортёры ауксина) (1 балл).. Таким образом, фенотип отсутствия окраски на среде с НУК будет у $\frac{13}{32}$ потомков, то есть у 41%. (2 балла)

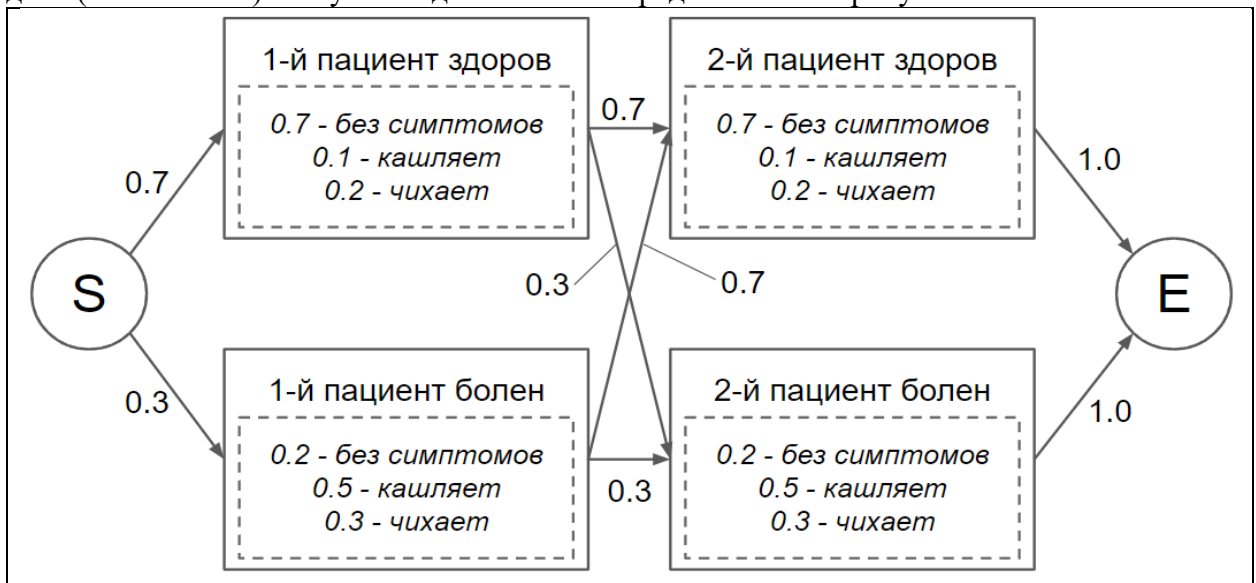
Задание 7. Скрытые марковские модели (17 баллов).

Скрытые марковские цепи (Hidden Markov Models - HMMs) - мощный алгоритмический инструмент огромного разнообразия биоинформатических программ. В основе данных

моделей данных лежит предположение, что каждое новое событие моделируемого процесса не зависит от предыдущего.

НММ может находиться в одном из состояний s_i , причем пользователю неизвестно, в каком именно. О текущем состоянии мы можем судить косвенно, так как каждое скрытое состояние s_i может с некоторой эмиссионной вероятностью $b_i(o_j)$ произвести наблюдаемой событие o_j . После того, как НММ произвела событие (не произвести никакое событие нельзя, т.е. сумма всех эмиссионных вероятностей равна 1.0) в каком-то состоянии, она переходит в новое состояние или остается в исходном, что определяется переходными вероятностями a_i для этого состояния (сумма всех переходных вероятностей для состояния всегда равняется 1.0).

В качестве примера давайте рассмотрим простую Марковскую модель для ситуации врачебного осмотра. Врач начинает свою работу (состояние S) и всегда принимает в день 2 пациента. Он может наблюдать у них максимум только 1 симптом: кашель или чихание. Исходя из наблюдаемого симптома (или его отсутствия) врач делает заключение о состоянии пациента. Когда врач закончил осмотр 2-х пациентов он завершает свой рабочий день (состояние E). Визуально данная НММ представлена на рисунке ниже.



7.1. Рассчитайте полную вероятность того, что за один день своей работы врач примет первым здорового пациента, который кашляет, и вторым - больного пациента, который не проявляет никаких симптомов (2 балла).

Решение (2 балла):

Врач всегда начинает осмотр из состояния S .

Нам надо рассчитать вероятность пройти по пути “ $S \rightarrow$ 1-й здоров \rightarrow 2-й болен $\rightarrow E$ ” и проявить симптомы “кашель \rightarrow без симптомов” в соответствующих состояниях. Так как события последовательные, то вероятности перемножаются и полная вероятность равна:

$$P = P(S \rightarrow \text{1-й здоров}) * P_{\text{1-й здоров}}(\text{кашляет}) * P(\text{1-й здоров} \rightarrow \text{2-й болен}) * P_{\text{2-й болен}}(\text{без симптомов}) * P(\text{2-й болен} \rightarrow E) = 0.7 * 0.1 * 0.3 * 0.2 * 1.0 = 0.0042.$$

7.2. Определите, какое наиболее правдоподобное заключение о здоровье пациентов сделает врач, если на осмотре он установил, что оба пациента чихают. Подтвердите ответ расчетами (6 баллов).

Решение:

Для определения наиболее правдоподобного диагноза нам надо рассчитать вероятности для всех возможных комбинаций состояний с условием, что оба наблюдаемых симптома были чиханием.

- а) оба здоровы: $P = 0.7 * 0.2 * 0.7 * 0.2 * 1.0 = 0.0196$
 б) 1-й здоров, 2-й болен: $P = 0.7 * 0.2 * 0.3 * 0.3 * 1.0 = 0.0126$
 в) 1-й болен, 2-й здоров: $P = 0.3 * 0.3 * 0.7 * 0.2 * 1.0 = 0.0126$
 г) оба больны: $P = 0.3 * 0.3 * 0.3 * 0.3 * 1.0 = 0.0081$

Исходя из наблюдаемых вероятностей, наиболее правдоподобным (с наибольшей вероятностью) будет заключение, что оба пациента здоровы. (по 1 баллу за верные вероятности => всего 4 балла; 2 балла за правильный вывод)

7.3. В биоинформатике скрытые марковские модели очень часто используются при работе с последовательностями нуклеиновых кислот или белков для предсказания значимых участков или даже генерации последовательностей с заданными свойствами. По образцу НММ для осмотра врача изобразите как будет выглядеть НММ для поиска CpG-островков в геноме с GC-составом 60%, если их ожидаемая частота встречаемости составляет 15%. Примите, что CpG-островки могут располагаться как в самом начале, так и в самом конце последовательности, но не могут идти друг за другом. Вероятности перехода из любого состояния в конец (E) примите равными 0.05 (9 баллов).

Подсказка: вам потребуется создать отдельное состояние для “несмысловых” участков, которое будет генерировать последовательность с нужным GC-составом.

Решение:

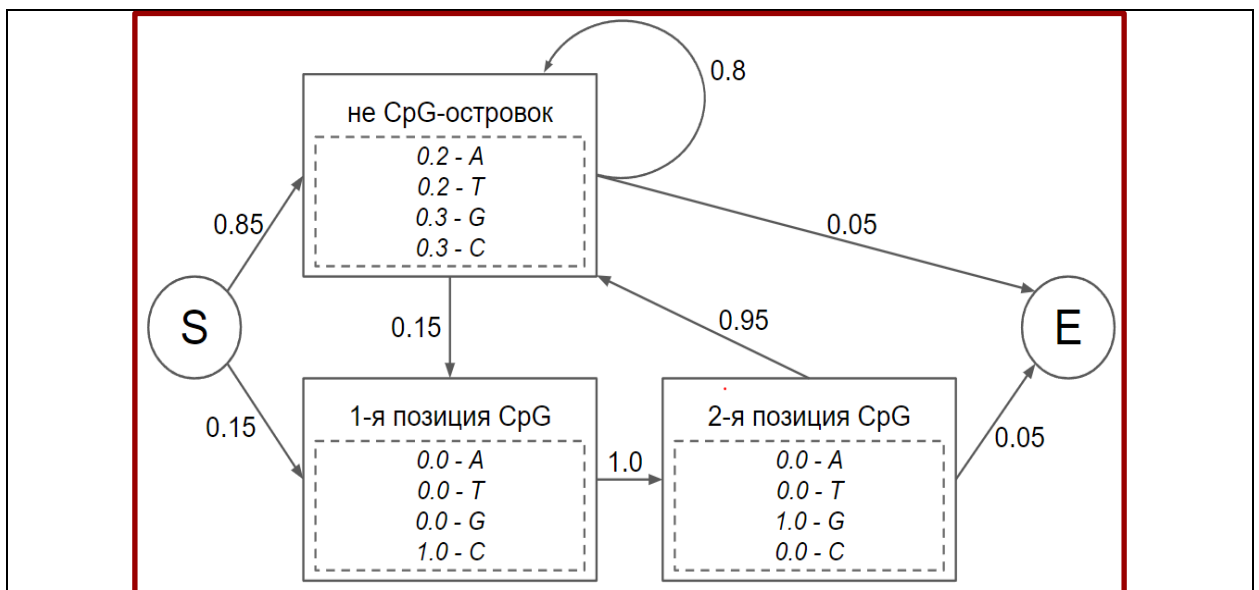
Известно: $P(S \rightarrow \text{1-я позиция CpG}) = P(\text{несмысловая} \rightarrow \text{1-я позиция CpG}) = 0.15$

$P(\text{1-я позиция CpG} \rightarrow \text{2-я позиция CpG}) = 1.0$ (т.к. CpG-островок не прерывается)

Для восстановления неизвестных переходных вероятностей необходимо помнить, что сумма всех переходных вероятностей, выходящих из каждого состояния равна 1.0.

Эмиссионные вероятности можно восстановить из условий задачи. В CpG-островке всегда на 1-й позиции генерируется C, а на 2-й генерируется G. В несмысловом участке вероятность генерации каждого нуклеотида рассчитывается из GC-состава генома по правилам Чаргаффа ($P(A) = P(T) = 0.2$; $P(G) = P(C) = 0.3$).

Соберем полученные вероятности в одну схему.



9 баллов за верную схему из них:

- 4 балла за комбинацию состояний, генерирующих CpG-островов (с вероятностями перехода и эмиссии)
- 1 балл за состояние, генерирующее не CpG-островов (со всеми вероятностями)
- 1 балл за состояния S и E с правильными переходными вероятностями
- 3 балла за все правильные переходные вероятности между состояниями

Задание 8. Полидактилия (15 баллов).

Полидактилия определяется аутосомно доминантным геном, частота которого в человеческой популяции составляет 0.5%. Известно, что признак имеет пенетрантность 80%.

8.1. Рассчитайте какая доля людей в популяции будет иметь фенотипические проявления полидактилии, если она находится в равновесии Харди-Вайнберга (2 балла).

Решение:

Необходимо рассчитать долю доминантных гомозигот (AA) и гетерозигот (Aa).

$$1 - q^2 = 1 - 0.995^2 = 0.009975 \text{ (1 балл)}$$

Из них только у 80% признак будет проявляться, значит:

$$0.009975 \cdot 0.8 = 0.00798 - \text{искомая доля людей, имеющих мутантный фенотип (1 балл)}$$

8.2. Полидактилия также может возникать в результате мутаций. Частота мутаций, приводящих к появлению гена полидактилии, составляет 3 на 1000 новых рождений за поколение (считайте, что вероятностью двойных мутаций можно пренебречь и все новые случаи – гетерозиготы). Пусть имеется популяция из 100.000 человек, находящаяся в равновесии Харди Вайнберга. В следующем поколении родится дополнительно 50.000 новых людей. Рассчитайте частоту людей с фенотипическим проявлением полидактилии в новой популяции (2 балла).

Решение:

Если популяция находится в равновесии, значит частоты аллелей и генотипов не изменяются, а значит в следующем поколении доля людей с унаследованным фенотипическим проявлением составит $0.00798 \cdot 150\,000 = 1197$ (1 балл). Помимо этого, возникнет 150 новых случаев полидактилии из которых проявятся только 120 (0.5 балла). Таким образом, суммарное количество людей, имеющих фенотипическое проявление заболевания составит 1317 (частота 0.00878) (0.5 балла)

При полностью верных расчетах, но использовании общей численности популяции равной 100 000 человек баллы снижаются вдвое.

8.3. В результате генетического скрининга было установлено, что в семье X мать является рецессивной гомозиготой по гену полидактилии. Известно, что в семье родился ребенок с полидактилией. Рассчитайте вероятность того, что отец был гомозиготным по гену полидактилии. Учитывайте, что полидактилия имеет неполную пенетрантность и может возникать в результате мутаций (8 баллов).

Решение:

Если генотип отца неизвестен, то необходимо рассмотреть три варианта: AA, Aa, aa.

Если генотип отца aa (частота 0.990025) то среди потомков не окажется детей с унаследованной полидактилией. Она может возникнуть в результате новой мутации с вероятностью 0.003. Такое заболевание проявляется в 80% случаев, поэтому итоговая вероятность: 0.00237606 (1 балл)

Если отец является гетерозиготой (0.00995), то среди потомков в 50% случае будут носители гена полидактилии и проявится она в 80% случаев. Во втором случае может возникнуть мутация, которая проявится с вероятностью 80%. В первом случае учитывать мутацию нет необходимости, так как у гетерозигот и доминантных гомозигот по условию одинаковая пенетрантность. Итоговая вероятность составит: 0.00399194 (2 балла)

Наконец, если отец был гомозиготным по гену полидактилии, то все потомки (независимо от мутаций) будут иметь фенотипические проявления с вероятностью 80%. Итоговая вероятность составит (1 балла):

Генотипы родителей	$aa * aa$	$Aa * aa$	$AA * aa$
Вероятность	0.990025	0.00995	0.000025
Генотипы потомков	aa	$Aa - 0.5$ $aa - 0.5$	Aa
Итоговая вероятность полидактилии	0.00237606	$0.00398 +$ $0.00001194 =$ 0.00399194	0.00002

Вероятность того, что отец был гомозиготен по доминантному гену можно подсчитать с помощью теоремы Байеса:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$P(A)$ – вероятность того, что отец гомозиготен по гену полидактилии (AA).

$P(B)$ – вероятность ребёнка иметь фенотипическое проявление полидактилии.

$P(B|A)$ – вероятность того, что ребенок будет иметь фенотипические проявления полидактилии, при условии того, что отец гомозиготен по гену полидактилии.

$P(A|B)$ – вероятность того, что отец гомозиготен по гену полидактилии, при условии, что ребенок имеет фенотипические проявления полидактилии – искомая вероятность!

$$P = \frac{0.00002}{0.00002 + 0.00237606 + 0.00399194} = \frac{0.00002}{0.006388} = 0.00313087 = 0.3\%$$

4 балла за верный расчет, 2 балла за правильное обозначение вероятностей.

Если вероятности в задаче не обозначены, но подсчет верный, то пункт оценивается в максимум баллов.

8.4. В результате генетического скрининга было установлено, что в семье X мать является рецессивной гомозиготой по гену полидактилии. Известно, что в семье родился ребенок с полидактилией. Какой генотип отца является самым вероятным? Ответ подтвердите расчетами (3 балла).

Решение:

Вероятность иметь генотип (AA)

$$P = \frac{0.00002}{0.00002 + 0.00237606 + 0.00399194} = \frac{0.00002}{0.006388} = 0.00313087 = 0.3\%$$

Рассчитаем вероятность иметь генотип (Aa) (1 балл):

$$P = \frac{0.00399194}{0.006388} = 0.6249123 = 62.5\%$$

Рассчитаем вероятность иметь генотип (aa) (1 балл):

$$P = \frac{0.00237606}{0.006388} = 0.3719568 = 37.2\%$$

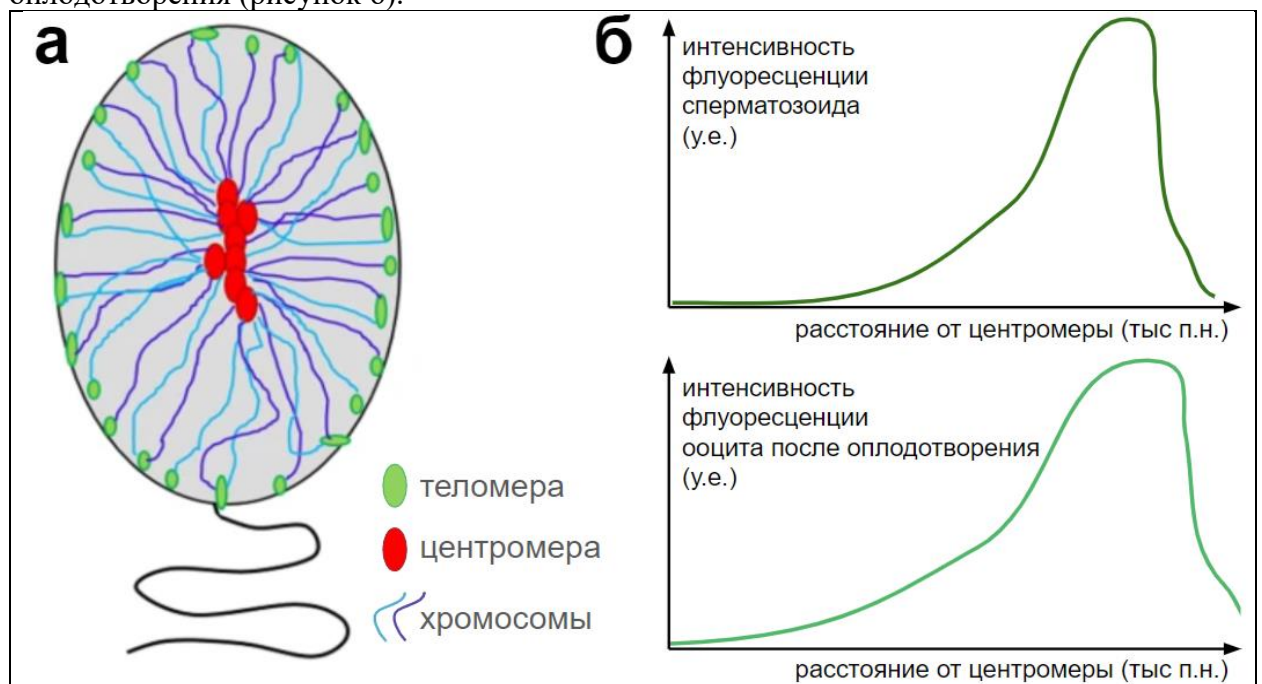
Таким образом, самый вероятный генотип в данном случае – Aa (1 балл)

Если в задании математические расчеты верны и вывод исходя из расчетов верный, но при этом использованы неверные значения вероятностей из предыдущего пункта, то задание оценивается в полный балл.

Задание 9. ДНК сперматозоидов (8 баллов).

К настоящему времени очень популярным вопросом является изучение способов организации хроматина в ядре. Среди множества различных вариантов возможных укладок хроматина особенно выделяется укладка ДНК в ядре зрелых сперматозоидов, где вместо привычных гистонов роль компактизации выполняют малые (50–100 аминокислот) положительно заряженные белки протамины. Хотя протамины и позволяют добиться крайне сильной степени компактизации ДНК и перевода ее в “молчащее” состояние, часть геномной ДНК сперматозоидов все равно остается связана с гистонами.

При изучении функции такой частичной компактизации генома в сперматозоидах проводилась оценка положения хромосом в ядре сперматозоида и экспрессии генов на разных участках хромосомы. В результате была предложена модель укладки хроматина, представленная на рисунке (а). Чтобы определить, активны ли конкретные участки хромосомы, вы вставляете в определенные участки аутосомы сперматозоида ген GFP и оцениваете флуоресценцию в сперматозоидах и в яйцеклетке на ранних этапах оплодотворения (рисунок б).



9.1. Предположите, в какой зоне ядра сперматозоида содержание протаминов будет наибольшим? Аргументируйте свой ответ (3 балла)

Решение:

Исходя из интенсивности флуоресценции мы можем сделать вывод об уровне экспрессии гена в зависимости от его положения на хромосоме. Так, наиболее активные области располагаются ближе к концу хромосом. По схеме строения сперматозоида можно утверждать, что концы хромосом расположены ближе к периферии, а середина ядра занята центромерными областями. И так как протамины, компактизуя ДНК приводят к переводу участка в молчащее состояние (указано в тексте), то скорее всего наибольшая доля протаминов содержится ближе к центру ядра

1 балл за верный ответ

2 балла за верную аргументацию

9.2. Почему интенсивность флуоресценции сперматозоида снижается к краю хромосомы? С каким типом наследования это может быть связано? (2 балла)

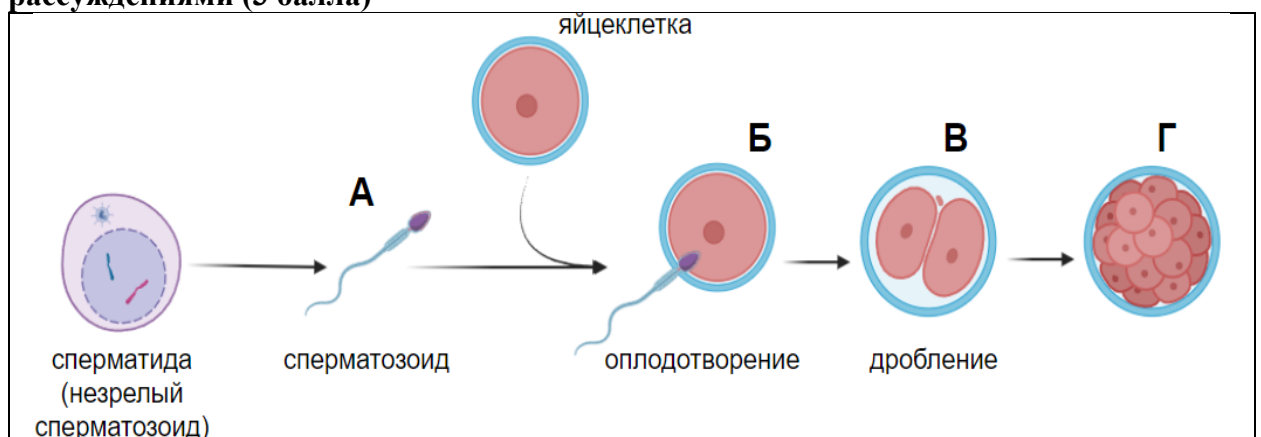
Решение:

Так как профиль интенсивности флуоресценции в ооците после оплодотворения также снижается в теломерных областях, то “замалчивание” генов на конце не связано с протаминами (они заменяются на гистоны после оплодотворения). Другой способ регуляции активности генов, не модифицируя их последовательности, заключается в модификации гистонов и метилировании ДНК. Правильно будет предположить, что в теломерных областях будет происходить эпигенетическое модифицирование генов и перевод их в неактивное состояние. Такие модификации связаны с эпигенетическим наследованием и импринтингом.

1 балл за верный ответ

1 балл за верную аргументацию

9.3. На рисунке ниже представлена схема основных стадий существования сперматозоида от момента его появления до слияния с ядром яйцеклетки и начала дробления. Расположите отмеченные на рисунке стадии (А-Г) в порядке увеличения количества активных генов в геноме сперматозоида. Подтвердите свой ответ рассуждениями (3 балла)



Решение:

Исходя из сравнения графиков интенсивности флуоресценции в сперматозоиде и ооците можно заметить, что область активного хроматина “расширяется” к центру хромосомы. Это связано с постепенной заменой протаминов на гистоны, поэтому все больше генов из сперматозоида начинают экспрессироваться после оплодотворения и в ходе

эмбрионального развития. Таким образом по количеству увеличения активных генов отмеченные стадии располагаются в порядке $A > B > B > Г$.

1 балл за верный ответ

2 балла за верную аргументацию