

## ПРИЛОЖЕНИЕ

### Количественная оценка неслучайных контактов плазмид с хроматином.

В результате эксперимента мы получили количество контактов плазмид № 1 и № 2 с ДНК *H. sapiens* и *M. musculus*, выраженное в виде числа выравниваний на соответствующие референсные геномы. Эти значения были представлены в виде долей от общего числа контактов плазмид и обозначены буквами (табл. П1).

**Таблица П1.** Сокращенные обозначения контактов плазмиды № 1 и № 2 с хроматинами двух типов

Символ	Значение
доля контактов плазмиды № 1, трансфицированной в клетки человека	
$X$	с хроматином <i>H. sapiens</i>
$Y$	с хроматином <i>M. musculus</i>
доля контактов плазмиды № 2, добавленной к клеткам после фиксации	
$A$	с хроматином <i>H. sapiens</i>
$B$	с хроматином <i>M. musculus</i>

Доля контактов плазмиды № 1 с ДНК человека ( $X$ ) состоит из доли контактов, образовавшихся внутри ядра и отражающих биологические закономерности распределения плазмиды ( $X_{cis}$ ), и контактов, образовавшихся в результате столкновений молекул плазмиды с фрагментами ДНК в растворе в результате случайной диффузии ( $X_{trans}$ ). Значение  $Y$  – доля контактов плазмиды № 1 с геномом *M. musculus* отражает «случайные» взаимодействия плазмиды с ДНК в растворе. Величина  $X_{trans}$  пропорциональна  $Y$  в соответствии с количественным соотношением ДНК *H. sapiens* и добавленной к ней ДНК *M. musculus*, поскольку столкновение плазмид в растворе в результате диффузии равновероятно как с ДНК *H. sapiens*, так и с ДНК *M. musculus* ( $X_{trans} = Y$  в случае, если соотношение молекул хроматина *H. sapiens* и *M. musculus* равно 1 : 1).

Чтобы учесть неравное количество молекул хроматина двух видов, мы оцениваем долю каждого по доле контактов плазмиды № 2 с этим хроматином. Соотношение хроматинов *H. sapiens* и *M. musculus* пропорционально  $\frac{A}{B}$ . Плазида № 2 была добавлена в предварительно фиксированные клетки и образовывала контакты в растворе в результате случайной диффузии при отсутствии биологических предпочтений, поэтому лигирования плазмиды № 2 к ДНК *H. sapiens* и *M. musculus* являются равновероятными событиями, частота которых отражает количественное соотношение хроматинов. Таким образом,  $X_{trans}$  может быть выражено как  $X_{trans} = Y \cdot \frac{A}{B}$ . Соответственно,  $X_{cis} = X - X_{trans}$ .

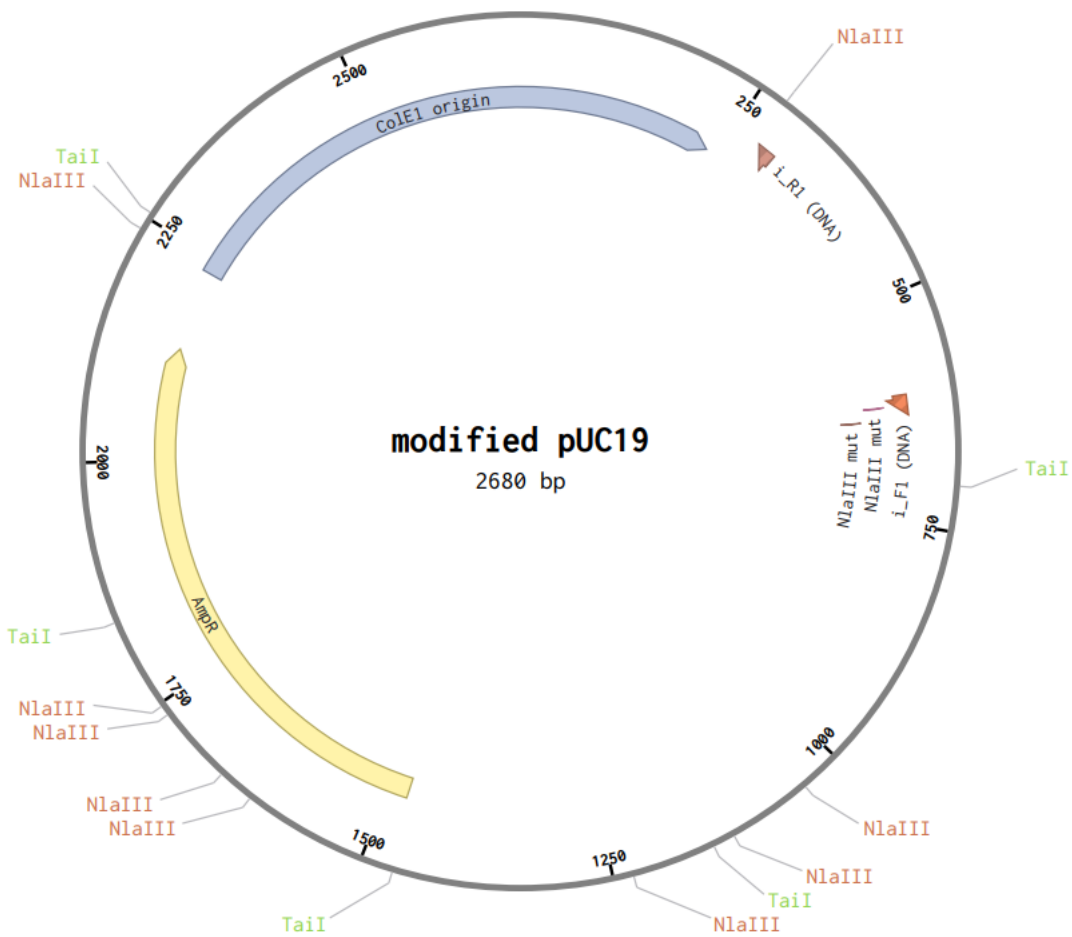
Согласно результатам, полученным в трех независимых экспериментах, среднее число контактов плазмиды, отражающее биологические закономерности распределения плазмид в ядре ( $X_{cis}$ ) составляет 44,2% (см. табл. П2). Несмотря на высокий уровень «шума» (~50% *транс*-контактов,  $X_{trans}$ ), это значение сопоставимо со статистикой, полученной для экспериментов Hi-C. Например, в Hi-C-библиотеках можно наблюдать, что доля *транс*-контактов варьирует от 50 до 15%, в зависимости от того, насколько «шумной» является Hi-C-библиотека [20]. Хотя *транс*-контакты в Hi-C-библиотеках оцениваются с использованием другого подхода, и значения нельзя напрямую сравнивать с приведенными выше оценками  $X_{trans}$ , они находятся в одном диапазоне, что говорит о приемлемом качестве 4C-данных.

**Таблица П2.** Количественная оценка контактов плазмиды с хроматином *H. sapiens* и *M. musculus*

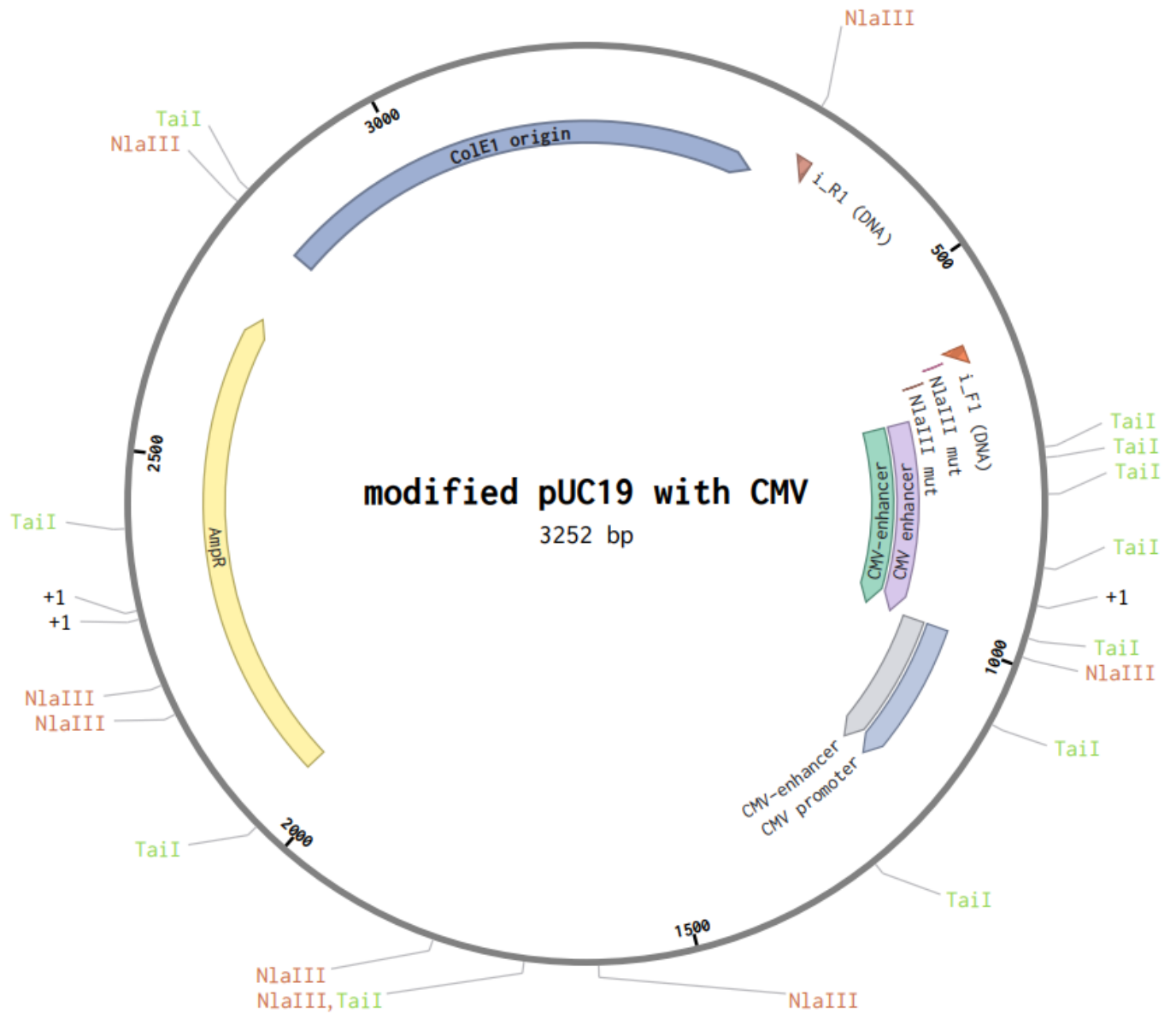
Символ	1	2	3	Среднее значение
$X$	75,21	83,97	80,05	79,74
$Y$	24,79	16,03	19,95	20,26
$A$	55,22	69,42	66,53	63,73
$B$	44,78	30,58	33,47	36,27
$X_{cis}$	<b>44,63</b>	<b>47,56</b>	<b>40,38</b>	<b>44,19</b>

Примечание. Значения представлены в виде доли числа выравниваний на соответствующий референсный геном, выраженной в процентах, от общего числа выравненных прочтений библиотеки.

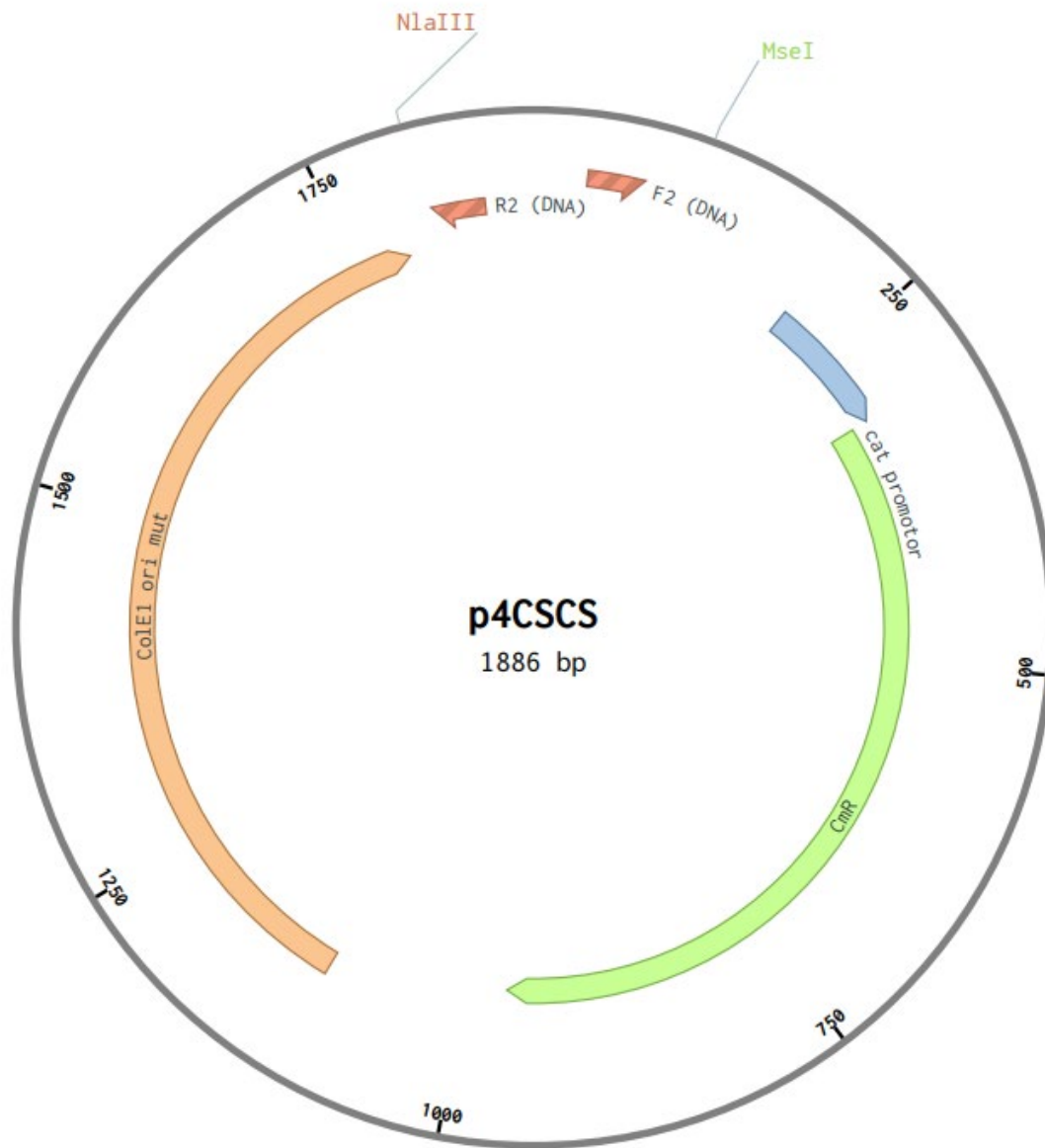
Кроме того, мы сравнили  $X$  (доля контактов плазмиды № 1 с хроматином человека) и  $A$  (доля контактов плазмиды № 2 с хроматином человека) с помощью точного теста Фишера. Полученное значение  $p$ -value пренебрежимо мало для всех трех реплик.



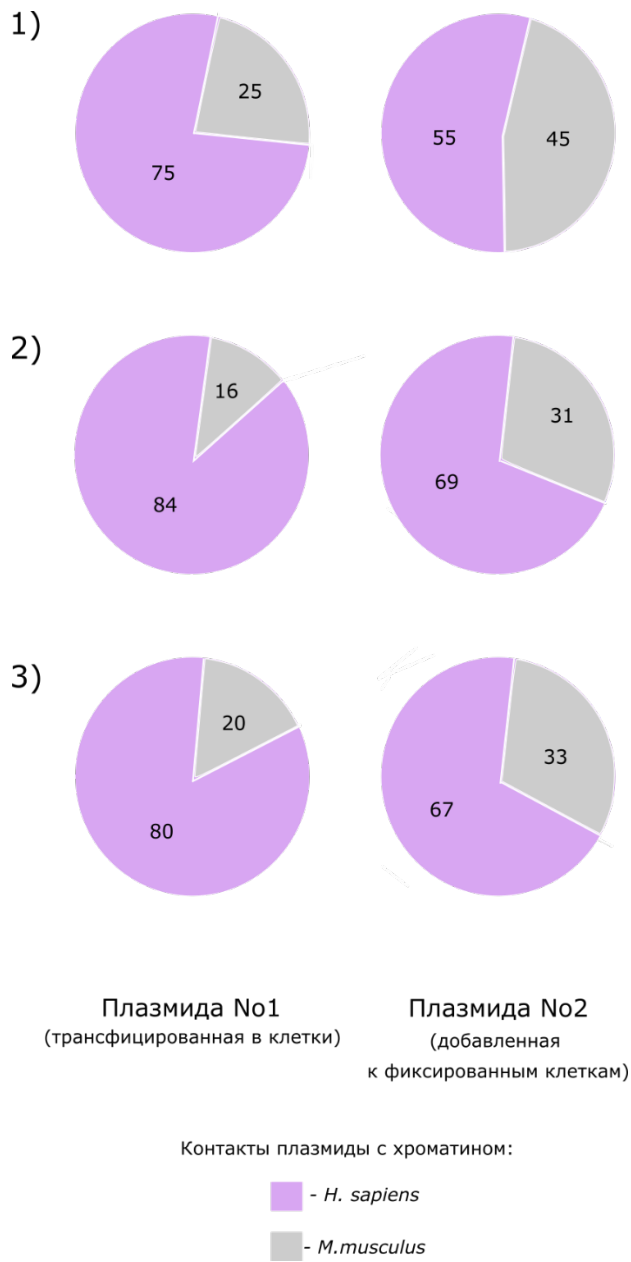
**Рис. П1.** Карта плазмидного вектора: модифицированный pUC19. Праймеры i\_F1 и i\_R1 содержат последовательности для отжига секвенирующих праймеров illumina F2, R2 (см. текст 2).



**Рис. П2.** Карта плазмидного вектора: модифицированный pUC19 со встройкой CMV-промотора. Праймеры i\_F1 и i\_R1 содержат последовательности для отжига секвенирующих праймеров illumina F2, R2 (см. текст 2)



**Рис. ПЗ.** Карта плазмидного вектора: p4CSCS. Праймеры F2 и R2 содержат служебные последовательности, необходимые для секвенирования illumina (см. текст 2)



**Рис. П4.** Оценка уровня фонового шума в 4С-протоколе с плазмидой для отдельных повторностей. Для всех трех реплик различие частот контактов было значимым. Согласно точному тесту Фишера,  $p$ -value пренебрежимо мало; подробности см. в табл. П2

**Текст 1: последовательности плазмидных векторов.**

> modified\_pUC19

gagatacctacagcgtgagctatgagaaagcggccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggctgg  
aacaggagagcgcacgagggagcttccaggggaaacgcctggatctttatagtcctgtcgggttcgccacctctgactgagcgtcgtat  
ttttgtgatgctcgtcagggggggcggagcctatgaaaaacgccagcaacgcggcctttttacggctcctggccttttctggccttttctca  
catgttcttctcgtctatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcggccagccgaacgaccgag  
cgcagcagtcagtgagcaggaagcgggaagagcggccaatacgcgaaccgctctccccgcggttggccgattcattaatgcagctg  
gcacgacaggttcccgactggaaagcgggagtgagcgcgaacgcaattaatgtgagttagctcactcattagccaccccagcctttacac  
tttatgcttccggctcgtatgttggattgtgagcggataacaatttcacacaggaacagctatgaccattacgccaagcttggcctgca  
ggctgactctagaggatccccgggtaccgagctcgaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttaccca  
cttaatgccttgcagcacatcccccttccgagctggcgtaatagcgaagaggcccgaccgatgcccttccaacagttgcgcagcct  
gaatggcgaatggcgcctgatgggtatttctccttacgactctgtcgggtatttcacaccgcatatggtgactctcagtacaatctgctctg  
atgccgcatagttaaaccagccccgacaccgccaacaccgctgacgcgacctgacgggctgtctgctcccggcatccgcttacagac  
aagctgtgaccgtctccgggagctgcatgtgtcagaggtttcaccgtcatcaccgaaacgcgcgagacgaaaggcctcgtgatacgcct  
attttataggttaatgtcatgataataatggttcttagacgtcaggtggcacttttcgggaaatgtgcgcggaaccctatttgttttttctaa  
atacattcaaatatgatccgctcatgagacaataaccctgataaatgctcaataatattgaaaaggaagagatgagtattcaacatttccgt  
gtcgccttattcccttttttgcggcattttgccttctgttttctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg  
cacgagtggtttacatcgaactggatctcaacagcggtaagatccttgagagtttcccccgaagaacgtttccaatgatgagcacttttaa  
agttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggctgccgcatacactattctcagaatgacttggttga  
gtactcaccagtcacagaaaagcatcttaccggatggcatgacagtaagagaattatgcagtctgccataaccatgagtataactgcggg  
ccaactacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactgccttgatcgttggg  
aaccggagctgaatgaagccataccaacgacgagcgtgacaccagatgcctgtagcaatggcaacaacgttgcgcaactattaactg  
gcgaactacttacttagcttcccggcaacaattaatagactggatggaggcggataaagtgcaggaccacttctgcgctcggccctccg  
gctggctggtttattgctgataaatctggagccggtagcgtgggtctcgcggtatcattgcagcactggggccagatggttaagccctccc  
tatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgattaagcatt  
ggtaactgtcagaccaagttactcatatatacttttagattgatttaaacttcatttttaattaaaaggatctaggtgaagatccttttgataatctc  
atgacaaaatccctaacgtgagtttcttccactgagcgtcagaccctgagaaaagatcaaaggatcttctgagatccttttttctgcgc  
gtaatctgctgcttgcacaacaaaaaccaccgctaccagcgggtggttgtttgccgatcaagagctaccaactcttttccgaaggtaactg  
gctcagcagagcgcagataccaataactgttcttctagttagccgtatgtagccaccacttcaagaactctgtagcaccgcctacatact  
cgctctgtaaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggc  
gcagcggctcgggctgaacgggggggtcgtgcacacagcccagcttggagcgaacgacctacaccgaact

> modified\_pUC19\_with\_CMV

gagatacctacagcgtgagctatgagaaagcggcagctcccgaagggagaaaggcggacaggtatccggtaaagcggcagggctgg  
aacaggagagcgcacgaggagctccaggggaaacgcctggatctttatagtcctgtcgggttcgccacctctgacttgagcgtcga  
ttttgtgatgctcgtcagggggcggagcctatgaaaaacgccagcaacgcggcctttttacggctctggccttttctggccttttctca  
catgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcggccagccgaacgaccgag  
cgcagcagtcagtgagcaggaagcgggaagagcggccaatacgcgaaccgcctctccccgcggttggccgattcattaatgcagctg  
gcacgacaggtttcccactggaaagcgggcagtgagcgcgaacgaattaatgtgagttagctcactcattagccaccccagcctttacac  
tttatgcttccggctcgtatgttggtaattgtgagcggataacaatttcacacaggaaacagctatgaccattacgccaagcttggcctgca  
ggtcACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGT  
AAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGAC  
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATT  
TACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC  
TATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTAT  
GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG  
CGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC  
GTCTCCACCCCAATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTT  
TCCAAAATGTCGTAACAACCTCCGCCCAATTGACGCAAATGGGCGGTAGGCGTGTACG  
GTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAgactctagagg  
atccccgggtaccgagctcgaattcactggcctcgttttacaacgtcgtgactgggaaaaccctggcgttaccacacttaatgccttgcag  
cacatccccctttgccagctggcgtaatagcgaagaggcccgcaccgatcgccttccaacagttgdcgagcctgaatggcgaatggc  
gcctgatgcggtattttctcttacgcatctgtgcggtatttcacaccgcatatggtgactctcagtacaatctgctctgatgccgcatagtaa  
gccagccccgacaccgccaacaccgctgacgcgccctgacgggcttctctcctccggcatccgcttacagacaagctgtgaccgtct  
ccgggagctgcatgtgtcagaggtttaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctattttataggttaatg  
tcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaaccctatttgttttttctaaatacattcaaatatgt  
atccgctcatgagacaataaccctgataaatgctcaataatattgaaaaggaagagtatgagtattcaacattccgtgtcgccttattccct  
ttttgcggcattttgccttctgtttttgtcaccagaaacgctggtaagtaaaagatgctgaagatcagttgggtgcacgagtggttac  
atcgaactggatcacaacagcggtaagatccttgagagttttcggccgaagaacgctttccaatgatgagcacttttaaagttctgctatgtg  
cgcggtattatcccgtattgacggcgggaagagcaactcggtcgccgatacactattctcagaatgacttgggtgagtactaccagtcac  
agaaaagcatcttacggatggcatgacagtaagagaattatgcagtgtgccataaccatgagtataaacactcggccaacttactctgac  
aacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactgccttgatcgttgggaaccggagctgaat

gaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacggtgcgcaaacattataactggcgaactacttact  
tagctcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcctccggctggctggttattg  
ctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagcctcccgtatcgtagtatctac  
acgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgattaagcattggtaactgtcagacc  
aagttactcatatatactttagattgattaaaacttcatttttaatttaaaggatctaggtgaagatccttttgataatctcatgacaaaatccct  
taacgtgagtttctccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttagatcctttttctgcgctaatctgctgcttgc  
aaacaaaaaaccaccgctaccagcgggtggttggttgccggtcaagagctaccaactctttccgaaggtaactggcttcagcagagcg  
cagataccaaactgttcttctagttagccgtagttaggccaccactcaagaactctgtagcaccgcctacatacctcgtctgtaactct  
gttaccagtggctgctgccagtgccgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggtcgggc  
tgaacggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaact

> 4CSCS

cggaacaggagagcgcacgaggagctccagggggaacgcctggtatctttatagctctgctcgggttcgccacctctggcttgagcgt  
cgattttctgatgctcgtcagggggcggagcctatggaaaaacgccagcaacgcggccTCTAGACATCTGACCATGA  
GGAGAACATAGATCGGAAGAGCACACGTCTGAACTCCAGTCACTGTCTCCACTACT  
CAACCGCTGCTTTAGGCTGACAGACTAGGTATCGCTGACCGTATCAGCAAGAGACTA  
CACTCTTCCCTACACGACGCTCTTCCGATCTTGACTATGGTTGCAAGAGACGAAGCT  
TCGTCTCAGTCTTAAGTACTGATCTTGGTCATAGCTGTTTCCTGGGAGGATCCTCAAC  
AGGTTGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGACACGTTGATCGGC  
ACGTAAGAGGTTCCAACCTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTT  
GAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGA  
TATAACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTC  
AGTTGCTCAATGcACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTgAAGA  
CCGTAAAGAAAAATAAGCACAAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCT  
GATGAATGCTCATCCGGAATTTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGG  
GATAGTGTTACCCCTTGTTACACCGTTTTCCAcGAGCAAACCTGAAACGTTTTTCATCGC  
TCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTGCAAGATGT  
GGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTT  
TCGTtTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTgAACGTGGCGAATATG  
GACAACTTCTTCGCCCCGTTTTCACTATGGGCAAATATTATACGCAAGGCGACAAGG  
TGCTGATGCCGCTGGCGATTCAGGTTTCATCAcGCCGTCTGTGATGGCTTCCAcGTCGGC

AGAATGCTCAATGAATTACAACAGTATTGCGATGAGTGGCAGGGCGGGGCGTAAtttgat  
atcgagctcgcttgactcctgttgatagatccagtaatgacctcagaactccatctggattgttcagaacgctcggttgcgcccggcggtttt  
ttattggtgagaatccaattTCCAACCAGAATTCCgagatctatgacaaaatccctataacgagagttttcgttccactgagcgt  
cagaccccgtagaaaagatcaaggatcttcttgatcctttttctgcgcgtaatctgctgcttgcaaacaaaaaaccaccgctaccagc  
ggtggtttgttgcggatcaagagctaccaactctttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgt  
agccgtagtaggcccacttcaagaactctgtagcaccgctacatacctcgtctgcttaatcctgttaccagtggctgctccagtggcg  
ataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctgggctgaacggggggtcgtgcacacagc  
ccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccacgcttcccgaagggagaaaggc  
ggacaggtatccgtaagcggcagggt

## **Текст 2: ПЦР праймеры для приготовления NGS библиотек.**

Последовательности праймеров для секвенирования NGS (5'→3'):

i\_F1: TCTTCCCTACACGACGCTCTTCCGATCTAACAGCTATGACCATTACGCCAA

i\_R1: CTGGAGTTCAGACGTGTGCTCTTCCGATCTGAATCAGGGGATAACGCAGGAA

F2: AATGATACGGCGACCACCGAGATCTACAC-NNNNNNNN-  
ACACTCTTCCCTACACGAC

R2: CAAGCAGAAGACGGCATAACGAGAT-NNNNNNNN-  
GTGACTGGAGTTCAGACGTGT

Используемые праймеры состоят из индексных последовательностей (восемь букв «N»), по которым секвенируемые образцы отличаются друг от друга, и служебных последовательностей, необходимых для секвенирования illumina.

Для 4С-экспериментов с векторами на основе pUC19 (модифицированный pUC19 и модифицированный pUC19 с CMV-промотором) проводили вложенную ПЦР с использованием праймеров i\_F1, i\_R1 на первом этапе и с F2, R2 – на втором. Вектор p4CSCS содержит в своей последовательности сайты отжига праймеров F2 и R2, поэтому

присоединение к концам целевых фрагментов олигонуклеотидных адаптеров, необходимых для секвенирования illumina, осуществлялось в один этап.