Занятие 7. Распознавание регуляторных сигналов.

Часть 1. Поиск регуляторных мотивов транскрипции в бактериальных последовательностях.

Ниже приведена разметка в исходных последовательностях (MEME26.txt) сайтов связывания PurR. Синим шрифтом выделены экспериментально установленные сайты. Сайты, найденные с помощью программы MEME с параметром «One per sequence» выделены курсивом. Сайты, найденные с помощью программы MEME с параметром «Zero or one per sequence» выделены жирным шрифтом. Было получено 17 последовательностей длины 300 bp.

codB

tacggacctgaaccgtaggtcggataaggcgctcgcgtcgcatccgacaccatgctcagatgcct
gatgcgacgctgacgcgtcttatcaggcctacccactgtttttacaccgataatttttcccccacctttt
tgcactcattcatataaaaaatatatttcc*cc**acgaaaacgattgc**tt*ttttatcttcagatgaatagaat
gcggcggatttttgggtttcaaacagcaaaaagggggaatttcgtgtcgcaagataacaactttagcca
ggggccagtcccgcagtcggcgcgg

purE

tcgcccggcggtgcatgaacttatcgccaatcagcaacctgcttttcgcgtggtactgggtgcct
ggcatacggaaggttcaatggtgaaagtcacggcggatgacgttgagctgattcattttccgttttaaaa
aacccgcaactttgctgatttcacagcc*ac**gcaaccgtttcctt**tg*ctctctttccgtgctattctctgt
gccctctaaagccgagagttgtgcaccacaggagttttaagacgcatgtcttcccgcaataatccggcgc
gtgtcgccatcgtgatggggtccaa

pyrC

gaaccaggcattacgcaattactttaaccagcaacctgcttacgtcctgcgcgaagatggcagcc
agggcgaagcaatggcgaaaaaactggcgaaaggcattgaagtgaagccaggcgaaattgtcattccatt
tactgattaatcacgagggcgcattcgcgccctttatttttcgtgc*aa**aggaaaacgtttccg*ctatcc
tttgtgtccggcaaaaacatcccttcagccggagcatagagattaatgactgcaccatcccaggtattaa
agatccgccgcccagacgactggca

purR

ttacacactgtgatgaaaaaatctcccgtcatttataatgataagtgtttttaccacttcccctt
ttcgtcaagatcggccaaaattccacgcttacactatttgcgtactggccattgaccccttcctgacgct
ccgtgtcgtttttccggcgtaccgcaacacttttgttgtgcgtaaggtgtgtaa*ag**gcaaacgtttacct*
*tg*cgattttgcaggagctgaagttagggtctggagtgaaatggaatggcaacaataaaagatgtagcgaa
acgagcaaacgtttccactacaact

cvpA

tgcctgatgcgacgctggcgcgtcttatcaggcctacgcaggggtagaaccgtaggtcggataag
gcgtttacgccgcatccgacacgcattgcccgatgccgcaaaggcataaaaagtcgatggcgttgaatat
tttttcagcgccatttttattgatgcgcgggaaggaaatccct*ac**gcaaacgtttctttt*tctgttaga
atgcgccccgaacaggatgacagggcgtaaaatcgtgggacacatatggtctggattgattacgccataa
tcgcggtgattgctttttcctctct

purM

ttttcgttgactttagtcaaaatgataacggtttgagataaagttattttatattcagatggtta
tgaaagaagattattccatccgaaaactaacctttaccctggcacaagtcttctttcgccgcgcgcctgg
ggaaaagacgtgcaaaaaggttgtgtaaagcag*tct**cgcaaacgtttgctt*tccctgttagaattgcgcc
gaattttatttttctaccgcaagtaacgcgtggggacccaagcagtgaccgataaaacctctcttagcta
caaagatgccggtgttgatattgac

guaB

acctgtcccatctcatgctcaagcagcagacgaaccgtttgattcaggcgactaacggtaaaaat
tgcaggggattgagaaggtaacatgtgagcgagatcaaattctaaatcagcaggttattcagtcgatagt
aacccgcccttcggggatagcaagcatttttgcaaaaaggggtag*at**gcaatcggttacgct**ct*gtata
atgccgcggcaatatttattaaccactctggtcgagatattgcccatgctacgtatcgctaaagaagctc
tgacgtttgacgacgttctcctcgt

glnB

gggtgaaaatacggcgctgccaacctttgttgaggcacgtaatcagtttgaactcaactatttgc
gtaagctgctgcaaatcaccaaaggcaacgtcacccacgcggcgagaatggcggggcgcaaccggacaga
attttataaactgctttcccgacacgagctggat*gcaaacgatttcaag*gaatgaattggcgttatgtgt
tacgtttagcagatcaaaagacaggcgacctttttcaaggaatagcatgaaaaagattgatgcgattataa
aacccttcaagctggacgatgtccg

purL

attctctgtgtcgtgcgcgtcccagcttgaaaaaacgtaataatagtgaaaggtttactcataaa
tgagcggcattttgcgtaaacctgcgccagatggcaacttattacagccattggcggcacgcgttgctaa
ttcacgatggtgattttatttccac*gcaaacggtttcgtca*gcgcatcagattctttataatgacgcccg
tttcccccccttgggtacaccgaaagcttagaagacgagagacttatgatggaaattctgcgtggttcgc
ctgcactgtcggcattccgaatcaa

purA

tagggccgatgctttacccgaaggcatggaagaagatgatctctgcgatgaccaatttgcccgat
aatattttacgtcgttttggcggtggacttgtggttgcgggcgttgtggtctactacatgt*tgaggaaaa*
*cgattgg*ctgaacaaaaaacagactgatcgaggtcatttttgagtgcaaaaagtgctgtaactctgaaaa
agcgatggtagaatccatttttaagcaaacggtgatttgaaaaatgggtaacaacgtcgtcgtactggg
cacccaatggggtgacgaaggtaaa

folD

aaattctttttatattgtcaggtatttcttaaattatcttaatccttagacaaggaaataaatca
gttccagatttacaacgccatcatggacgaaaaatgaagctttcagtctcagcgacggtgcgcctcacct
tcgcaagaggtcgcttcacgcgataaatctgaaacgaaacctgacagcgcgccccgcttctgacaaaata
ggcgcatccccttcgatctacgtaacagatggaatcctctctctgatggcagcaaagattattg*acggta*
*aaacgattgc*gcagcaggtgcgctc

rpiA

ttgaatggcgtggcgttattgcctcaatttgcctgtaaacaggggctt*gcgaacggtgaactgg*t
gcgcctgtttgcaccgtggagcggcatacccagaccgttgtatgctttatttgcggggcgaaaggggatg
cctgccattgcgcgatattttatggatgagttaaccacgcggcttgccaacggggtctgaatcgcttttt
ttgtatataatgcgtgtgaaatttcataccacaggcgaaacgatcatgacgcaggatgaattgaaaaaag
cagtaggatgggcggcacttcagta

carA

caatcttcttgctg*cgcaagcgttttccaga*acaggttagatgatctttttgtcgcttaatgcct
gtaaaacatgcatgagccacaaaataatataaaaaatcccgccattaagttgacttttagcgcccatatc
tccagaatgccgccgtttgccagaaattcgtcggtaagcagatttgcattgatttacgtcatcattgtga
attaatatgcaaataaagtgagtgaatattctctggagggtgttttgattaagtcagcgctattggttct
ggaagacggaacccagtttcacggt

pdhR

tgaatcggttcaattcggatttttatagtttaataatcgttaaaaaactcctttcctacgtaaag
tctacatttgtgcatagttacaactt*tgaaacgttatatatg*tcaagttgttaaaatgtgcacagtttca
tgatttcaatcaaaacctgtatggacataaggtgaatactttgttactttagcgtcacagacatgaaatt
ggtaagaccaattgacttcggcaagtggcttaagacaggaactcatggcctacagcaaaatccgccaacc
aaaactctccgatgtgattgagcag

fixA

tcgttaattttgattaataatcagtttgttatgctctgttgtgagtaaaaaataacatctgactt
tcaatattggtgatccataaaacaatattgaaaatttctttttgctacgccgtgttttcaatattggtga
ggaacttaacaatattgaaagttggatttatctgcgtgtgacattttcaatattggtgattaaagtttta
tttcaaaattaaagggcgtgatatctgtaattaacaccaccga*tatgaacgacgtttcc*ttcatgatttc
tggagatgcaatgaagattattact

rbsD

atttgcgtctgcgcggcaagctgtacttgctgctgcaacgtaatgcgctgcgtgcaccagatcaa
tttgaaatcccgccaaacagggttatcgaactgggtactcaggtcgaaatctaacgccagacgcctcctt
tcttcataagggggcgtttttgttttcatggttaatcac*catgtaaaacgtttcga*aggttgatcacattt
ccgtaacgtcacgatggttttcccaactcagtcaggattaaactgtgggtcagcgaaacgtttcgctgat
ggagaaaaaaatgaaaaaaggcacc

araB

Tcgctaac*ccaaccggtaaccccg*cttattaaaagcattctgtaacaaagcgggaccaaagccat
gacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtc
acactttgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaac

tctctactgtttctccatacccgtttttttggatg**gagtgaaacgatggcg**attgcaattggcctcgatt
ttggcagtgattctgtgcgagcttt

В большинстве случаев программа нашла правильные сайты. В последовательности purA найденные сайты практически полностью совпадают между собой, но не совпадают с экспериментально установленным сайтом.

В некоторых последовательностях (например, folD) программа, как с параметром «One per sequence», так и с параметром «Zero or one per sequence», находила сайт там, где он не был установлен экспериментально.

Часть 2. Поиск сайтов в эукариотических последовательностях.

Выравнивание: геном человека – геном коровы.

```
000000001  GCTGCTGCCCACTCTGCGACTGTGCCTGTACGTGCCAGCTCCCCGACTGCCAGAGCCTCA  000000060
>>>>>>>>>  | ||||||||| | || ||||||||||| | ||||||||||| ||||||||||||||||  <<<<<<<<<
000000001  GATGCTGCCCATTGTGTGACTGTGCCTGTGCCTGCCAGCTCCCAGACTGCCAGAGCCTCA  000000060
AP2        -----++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       --------------------------------++++++++++++--------++++++++
SRF        -++++++++++++++++++------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------+++++++++----------------------

000000061  ACTGTCTCTGCTTCGAGATCAAGCTCCGATGA-----GGACCCAGGGCCCCTGCCCTCTG  000000115
>>>>>>>>>  |||| ||||| || ||||||||||| |||||     ||| |  |||    ||  || ||  <<<<<<<<<
000000061  ACTGCCTCTGTTTTGAGATCAAGCTCAGATGAAAGATGGAACTCGGGGTGTTGATCTTTG  000000120
AP2        ++++++++++++++++++-----------------------------------------++
GATA1      -------------++++++++++-------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ++++++++++-----------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------+++++++++-----------------------------------------

000000116  GGGAGCGGCCAGCCCCCAGGGCCCATGTGCCCTCCTCCCTGAAGAGCCTTTCCCCAC---  000000172
>>>>>>>>>  ||||| || ||||| ||||         |  ||||||  | | || ||||    |||     <<<<<<<<<
000000121  GGGAGTGGACAGCCTCCAG---------GAGCTCCTCTATAATGACCCTTAGGACACCAC  000000171
AP2        ++++++++++++++-------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000173  ------GCCACTGGAACCACAGATGGCCTGCCGAGCACCCAGGCCTGGGAACTGGAAGTG  000000226
>>>>>>>>>        | ||||||| | |  ||||||  ||| |||       |||||||||    ||||  <<<<<<<<<
000000172  TACTAGGGCACTGGAGCTATGGATGGCAGGCCCAGCCTAACTACCTGGGAACCTTGAGTG  000000231
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000227  GCAGCGCAGGGCCTGGCTCCCTGCAGGGCAGGA---------CTCTTGGCCGGCTGGACG  000000277
>>>>>>>>>  |      ||||| ||| |||||| | |||||||         | ||||| | ||||      <<<<<<<<<
000000232  G------AGGGCTTGGTTCCCTGAAAGGCAGGATCCTGAGAGCCCTTGGACAGCTG----  000000281
AP2        ------------++++++++++++++++++-------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        -----------++++++++++++++++++++++---------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000278  GCAGCTCCTCTGGAGGGCCAGAAAAGAGAGGGGCTAGTGCTCGGGCAGGTGCCCTGGCTT  000000337
>>>>>>>>>                            | ||||  || ||| || |||  ||||||  || |  <<<<<<<<<
000000282  -------------------------AAAAGGGTCTGGTGTTCTGGCCAGTGCCCCAGCCT  000000316
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
```

```
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000338  CCCTTCCCCTCCACACG----TCAACGATTCTATTTGAAGTTGGGCAGGGGGGTGGCGCT  000000393
>>>>>>>>>  | | || |||||   |     | ||| |      |||| | || | | |  |  | |  |  <<<<<<<<<
000000317  CTCCTCTCCTCCCTCCCGGGATAAACTA------TTGAGGGTGAGGATGCAG--GACATT  000000368
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000394  GCTCACC-ACACACAAGTGTTATAGGAGGAGTCTGGC--CCTTGAGTACCGGGTACGCAG  000000450
>>>>>>>>>  ||||||| |  ||  ||||   | |||||  |||| |  ||| || |  | || |   ||  <<<<<<<<<
000000369  GCTCACCCAGGCAACAGTG---TGGGAGGGCTCTGCCAACCTGGACTGTCAGG-ATAAAG  000000424
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000451  GGGTGCCTCAACCACACTCCGTCCACGGACTCTCCGTTATTTTAGGAGGTGAGTGTAGTG  000000510
>>>>>>>>>  || || | || |||||| |   | |  |||||   ||||||||| ||||||||| |  ||  <<<<<<<<<
000000425  GGATGGC-CACCCACACCCTTCCTATAGACTCCTGGTTATTTTAAGAGGTGAGTATCCTG  000000483
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ----------------------------------++++++++++++---------------
MEF3       ------------------------------------------------------------
MYOD       ---------------------------------------------++++++++---------
SRF        ------------------------------------------------------------
TEF        -----------------------------------+++++++++++++-------------
TEF1       ----------------------------------+++++++++------------------

000000511  CCAGTATCTACTCTCCTTCTTAAAAAAAACCAGGGCTCCAGAGAATCAGAACAGCCACCA  000000570
>>>>>>>>>   | |  ||| |||||||  |||||||||                  || ||| |||||   <<<<<<<<<
000000484  TCTGCCTCTGCTCTCCT--TTAAAAAAAA------------------AGTACAACCACCG  000000523
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000571  TCACCGCAGGGAGTCAAGGGAGGAGGGAGATTAGAGAAGGAGCCAGGGAGGGTGGCAGGG  000000630
>>>>>>>>>   ||  |    |||||||||||||||||              ||  |  | ||    |||  <<<<<<<<<
000000524  GCA--GACACGAGTCAAGGGAGGAGGGA--------------CCCAGCGGAGTCAGTGGG  000000567
AP2        ----------+++++++++++++++++--------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000631  AGGCCACGTGATCCGAG------TCCCCTCACCCCTTTCCTTCCCACAGGTCCCTGGCCA  000000684
>>>>>>>>>  ||||  | |    | ||      |||| ||      | ||||   |||| |||| | ||  <<<<<<<<<
000000568  AGGCAGCAT----CTAGATGTTTTCCCTTCTTGTTCTGCCTT--AACAGATCCCGGACCT  000000621
AP2        ------------------------------------------------------------
```

```
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ------------------------------------------------------------

000000685   AAGATTTATTTCTCTTGACAACCAAGGGCCTCCGTCTGGATTTCCAAGGAAGAATTTCCT  000000744
>>>>>>>>>    ||||| ||||||     | || |||| |      ||| ||||| |||||||| |||| |  <<<<<<<<<
000000622   GAGATTGATTTCT-----CGACGAAGGTC-----ACTGTATTTCTAAGGAAGAGTTTCTT  000000671
AP2         ------------------------------------------------------------
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        -----------------------------------+++++++++++++-------------

000000745   CTGAAGCACCGGTGAGTGGGCAGGGGCTCC--CTCCCCATCAATAGGGCCGACCCAAGTC  000000802
>>>>>>>>>   || ||| |||||||||||  ||  ||| ||  |||| |  |||    ||  ||| ||| |  <<<<<<<<<
000000672   CTAAAG-ACCGGTGAGTGAACATTGGCACCTCCTCCTCTCCAA----GCAAACCTAAG-C  000000725
AP2         ------++++++++++++++++-------------------------------------
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ------------------------------------------------------------

000000803   TTCCTCCCCCTTCCCCCAT--GCCGGGCCCCACGATAGTGTGAATGTCAGGGGCTTCAGG  000000860
>>>>>>>>>   ||||||  | | ||||| |  || |||             |||||||||||||| |||||  <<<<<<<<<
000000726   TTCCTCTTCTTACCCCCCTAGGCAGGG------------ATGAATGTCAGGGGCCTCAGG  000000773
AP2         ---------------------------------------+++++++++++++++++++++
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------++++++++
MYOD        ---------------------------------------------+++++++++-------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ----------------------------------------+++++++++-----------

000000861   TTTCCCTAAATATAGGTCCCTGCCAGAGGATCCGTGGCGGGAAAAGGGCAGGGGTCATTA  000000920
>>>>>>>>>   |||| |||||||||||||| ||||   |||| ||||| ||| ||||||||| ||| | |  <<<<<<<<<
000000774   TTTCACTAAATATAGGTCCTTGCCGCGGGATTCGTGGTGGGGAAAGGGCAGAGGTTA-TG  000000832
AP2         ++++--------+++++++++++++++++++++++++++++++++++++++++++++--
GATA1       ------------------------------------------------------------
MEF2        -----++++++++++++-----------------------------------------
MEF3        +++++-----------------------------------------------------
MYOD        ---------------------------------------------++++++++------
SRF         -++++++++++++++++++++-------------------------------------
TEF         -++++++++++++-------------------------------------------
TEF1        -----------------+++++++++-+++++++++--+++++++++----+++++++++

000000921   GAGAAGATC-GGGGACACATGTGGGGCGG---------------------GCAGGAGCTG  000000958
>>>>>>>>>   |||||| || ||||||||  ||| || ||                     | ||||||||  <<<<<<<<<
000000833   GAGAAGGTCCGGGGACACTGGTGCGGGGGTGCGTAGGGGAGGGGTGGAGAGTAGGAGCTG  000000892
AP2         --++++++++++++++-------------------------------------------
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        -------------------------------------------------+++++++++
SRF         ----------------------------------------------------++++++
TEF         ------------------------------------------------------------
TEF1        ------------------------------------------------------------

000000959   CCTTATAACC-AGCCCGGGAACCCCTAGCTCACTCGCTGCTGACCAGGCTCTGCCGGCTC  000001017
>>>>>>>>>   |||||||||| ||||| ||  | |   |||||||||||||||||||||||||| |||||  <<<<<<<<<
```

```
000000893  CCTTATAACCCAGCCCTGGGCCGCGGGGCTCACTCGCTGCTGACCAGGCTCTG-CGGCTT  000000951
AP2        -----+++++++++++++++++----------++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       +++--------------------------------+++++++++---------------
SRF        ++++++++++++++++++++---------------++++++++++++++++------
TEF        ------------------------------------------------------------
TEF1       --++++++++++-----------------------------------------------

000000952  CTTTCACTGCACCACAGGTGGGCCGCTAGCTGGATTGG---AGGATGGGAATGGGGGTTG  000000952
AP2        ++++-----+++++++++++++++++++++++++++-------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------++++++++++++-----------------------------------
SRF        ----------++++++++++++++++++++++++++-------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000953  GGTTGGGCCAACAGGGTCCAGATGGGGTCCAGGTGAGGAGGGGAGATTTGGACGATAGGA  000000953
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000954  GCAGGGGGCTC----AGCATCTGGGAGGCAGATCAGTTCGGGGACGGATT----------  000000954
AP2        ------------------------------------------------------------
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000955  ----------TTCTTTTGGAGAAGGAAGTCAGG  000000955
AP2        ---------------------------------
GATA1      ---------------------------------
MEF2       ---------------------------------
MEF3       ---------------------------------
MYOD       ---------------------------------
SRF        -------------++++++++++++++++++--
TEF        ---------------------------------
TEF1       ---------------------++++++++++---
```

В среднем, один сайт приходится на 26 нуклеотидов.

Можно ожидать, что если бы искались все 407 факторов, то один сайт приходился бы на 0,5 нуклетоида.

Выравнивание: геном человека – геном резуса.

Для факторов MEF3 и TEF сайты не нашлись. Это странно, так как в выравнивании геном человека – геном коровы данные сайты были найдены.

```
000000001  GCTGCTGCCCACTCTGCGACTGTGCCTGTACGTGCCAGCTCCCCGACTGCCAGAGCCTCA  000000060
>>>>>>>>>  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  <<<<<<<<<
000000101  GCTGCTGCCCACTCTGCGACTGTGCCTGTACGTGCCAGCTCCCCGACTGCCAGAGCCTCA  000000160
AP2        -----+++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ---------------------------------------------+++++++++--------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000061  ACTGTCTCTGCTTCGAGATCAAGCTCCGATGAGGACCCAGGGCCCCTGCCCTCTGGGGAG  000000120
>>>>>>>>>  |||||||||||||||| |||||||||||||||| |||||||||||||||||||||||| |||  <<<<<<<<<
000000161  ACTGTCTCTGCTTCGAAATCAAGCTCCGATGAGAACCCAGGGCCCCTGCCCTCTGGAGAG  000000220
AP2        ++++++++++++++++++---+++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------+++++++++---------
SRF        ------------------------------------------++++++++++++++++++
TEF        ------------------------------------------------------------
TEF1       -------------+++++++++-----------------------------------------

000000121  CGGCCAGCCCCCAGGGCCCATGTGCCCTCCTCCCTGAAGAGCCTTTCCCCACGCCACTGG  000000180
>>>>>>>>>   |||||||||||||||||||||| |||||||||||||||||||||||| ||| ||||||||  <<<<<<<<<
000000221  TGGCCAGCCCCCAGGGCCCATGGGCCCTCCTCCCTGAAGAGCCTTTCTCCAGGCCACTGG  000000280
AP2        ++++++++++++++++++++++++++++++++++++++++++++++++++++++-----
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       -------------++++++++++++++++++-------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ---+++++++++-------------------------------------------++++

000000181  AACCACAGATGGCCTGCCGAGCACCCAGGCCTGGGAACTGGAAGTGGCAGCGCAGGGCCT  000000240
>>>>>>>>>  ||||||| |||||||||| |||||||||||||||||||||||||||||||| || ||||||  <<<<<<<<<
000000281  AACCACAAATGGCCTGCCCAGCACCCAGGCCTGGGAACTGGAAGTGGCAGTGCGGGGCCT  000000340
AP2        ----------++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ----++++++++++-----------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        -------------------++++++++++++++++++++++-------------------
TEF        ------------------------------------------------------------
TEF1       +++++-------------------------------------------------------

000000241  GGCTCCCTGCAGGGCAGGACTCTTGGCCGGCTGGACGGCAGCTCCTCTGGAGGGCCAGAA  000000300
>>>>>>>>>  |||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||||||  <<<<<<<<<
000000341  GGCTCCCTGCAGGGCGGGACTCTTGGCCGACTGGACGGCAGCTCCTCTGGAGGGCCAGAA  000000400
AP2        ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ---------------------------+++++++++--++++++++++++++--------
SRF        ----+++++++++++++++----------------++++++++++++++++++++++++++
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000000301  AAGAGAGGGGCTAGTGCTCGGGCAGGTGCCCTGGCTTCCCTTCCCCTCCACACGTCAACG  000000360
>>>>>>>>>   |||||||||||| |||| ||||||||| ||||||| || ||||||||| |||||||||  <<<<<<<<<
000000401  GAGAGAGGGGCTACTGCTGGGGCAGGTGTCCTGGCTCCCTTTCCCCTCCCCACGTCAACG  000000460
AP2        ++++++++++++++++++++++++++++++++++++++++++++++++++++++++---
```

```
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ------------------+++++++++----------------------------------

000000361   ATTCTATTTGAAGTTGGGCAGGGGGGTGGCGCTGCTCACCACACACAAGTGTTATAGGAG  000000420
>>>>>>>>>   |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||  <<<<<<<<<
000000461   ATTCTATTTGAAGGTGGGCAGGGGGGTGGCGCTGCTCACCACACACAAGTGTTATAGGAG  000000520
AP2         --------++++++++++++++++++++++++----------------------------+++
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ----------------+++++++++----------------------+++++++++------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ------------------------------------------------------------+++

000000421   GAGTCTGGCCCTTGAGTACCGGGTACGCAGGGGTGCCTCAACCACACTCCGTCCACGGAC  000000480
>>>>>>>>>   ||| |||||||||||||||||||||| || |||||||   ||||||||||||||||||||  <<<<<<<<<
000000521   GAGCCTGGCCCTTGAGTACCGGGTACACAAGGGTGCCCG-ACCACACTCCGTCCACGGAC  000000579
AP2         ++++++++++++++++++++++++++++++++++++-----------+++++++++++++++
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ++++++------------------------------------+++++++++--------

000000481   TCTCCGTTATTTTAGGAGGTGAGTGTAGTGCCAGTATCTACTCTCCTTCTTAAAAAAAAC  000000540
>>>>>>>>>   |||| |||||||||||||||||||||||||||||||||| ||||||||| ||||||||||  <<<<<<<<<
000000580   TCTCGGTTATTTTAGGAGGTGAGTGTAGTGCCAGTATCTGCTCTCCTTCCTAAAAAAAAC  000000639
AP2         ++++++-------------------------------------------------+++++
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ---+++++++++------------------+++++++++----------------------

000000541   CAGGGCTCCAGAGAATCAGAACAGCCACCATCACCGCAGGGAGTCAAGGGAGGAGGGAGA  000000600
>>>>>>>>>   |||||||||||||||||| |||||||||||||| |||||||||||  |||||||| ||||  <<<<<<<<<
000000640   CAGGGCTCCAGAGAATCACAACAGCCACCATCACGGCAGGGAGTCTGGGGAGGAGAGAGA  000000699
AP2         +++++++++++++++--------------++++++++++++++++++--------------
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        ----------------------------+++++++++------------------------

000000601   TTAGAGAAGGAGCCAGGGAGGGTGGCAGGGAGGCCACGTGATCCGAGTCCCCTCACCCCT  000000660
>>>>>>>>>   |||||| ||      |||||||||||||||||||||||||| |||||||||| ||||  <<<<<<<<<
000000700   TTAGAGGAGTCT---GGGAGGGTGGCAGGGAGGCCACGTGATCTGAGTCCCCTCATCCCT  000000756
AP2         ------------++++++++++++---------------++++++++++++--+++++++
GATA1       ------------------------------------------------------------
MEF2        ------------------------------------------------------------
MEF3        ------------------------------------------------------------
MYOD        ------------------------------------------------------------
SRF         ------------------------------------------------------------
TEF         ------------------------------------------------------------
TEF1        --------------+++++++++------------------------------++++++++

000000661   TTCCTTCCCACAGGTCCCTGGCCAAAGATTTATTTCTCTTGACAACCAAGGGCCTCCGTC  000000720
>>>>>>>>>   |||| |||||||||||||||||||||||||||||||||||||||| |||||||||||||  <<<<<<<<<
```

```
000000757  TTCCCTCCCACAGGTCCCTGGCCAAAGATTTATTTCTCTTGACAATCAAGGGCCTCCGTC  000000816
AP2        ++++++++++++++++++++++++++++++++----------+++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       ++++++++++------------------------------------------------++

000000721  TGGATTTCCAAGGAAG-AATTTCCTCTGAAGCACCGGTGAGTGGGCAGGGGCTCCCTCCC  000000779
>>>>>>>>>  |||||||||||| |||| |||||||||||||||||||||||||||||| ||||||||||||  <<<<<<<<<
000000817  TGGATTTCCAAAGAAGAAATTTCCTCTGAAGCACCGGTGAGTGGGCAAGGGCTCCCTCCC  000000876
AP2        ++++++++++++++++-----------------++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       +++++++-----++++++++++++++-----------------------------------

000000780  CATCAATAGGGCCGACCCAAGTCTTCCTCCCCCTTCCCCCATGCCGGGCCCCACGATAGT  000000839
>>>>>>>>>  | |||| |||||||||||||| ||||||||||||||| ||||||||||||||| | | ||  <<<<<<<<<
000000877  CGTCAACAGGGCCGACCCAAGGCTTCCTCCCCCTTCCTCCATGCCGGGCCCCAGGCTGGT  000000936
AP2        ++++++++++++++---------++++++++++++++++++++++++++++++++++-
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------------------
TEF        ------------------------------------------------------------
TEF1       --------------------------------------+++++++++------------+

000000840  GTGAATGTCAGGGGCTTCAGGTTTCCCTAAATATAGGTCCCTGCCAGAGGATCCGTGGCG  000000899
>>>>>>>>>  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  <<<<<<<<<
000000937  GTGAATGTCAGGGGCTTCAGGTTTCCCTAAATATAGGTCCCTGCCAGAGGATCCGTGGCG  000000996
AP2        ---++++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------------------++++++++++
TEF        ------------------------------------------------------------
TEF1       ++++++++++++------------------------------------------------

000000900  GGAAAAGGGCAGGGGTCATTAGAGAAGATCGGGGACACATGTGGGGCGGGCAGGAGCTGC  000000959
>>>>>>>>>  |||||||||||||||||||  |||||| |||||||||| ||||||||||||||||||| ||  <<<<<<<<<
000000997  GGAAAAGGGCAGGGGTCATCGGAGAAGGTCGGGGACACGTGTGGGGCGGGCAGGAGCCGC  000001056
AP2        ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------++++++++++++++++++++++++++++-
SRF        ++++++++++++++-------------++++++++++++++++++++++---------
TEF        ------------------------------------------------------------
TEF1       ----------------++++++++---------+++++++++------------------

000000960  CTTATAACCAGCCCGGGA-ACCCCTAGCTCACTCGCTGCTGACCAGGCTCTGCCGGCTCC  000001018
>>>>>>>>>  |||||||||   || ||  |||||| |||||||||||||||||||||||||||||||||  <<<<<<<<<
000001057  CTTATAACCCAGCCCGGGCACCCCTGGCTCACTCGCTGCTGACCAGGCTCTGCCGGCTCC  000001116
AP2        ----------------+++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       -++++++++++++-----------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ------------------------------------------------------------
SRF        ------------------------------------++++++++++++++++++++++++
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------
```

```
000001019  TTCGGCCTCGCCGCAGGTGGGCCCCTTGCAGGACCGGGCCGGGGTGGGGATGGGGTTGGG  000001078
>>>>>>>>>  ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||  <<<<<<<<<
000001117  TTCAGCCTTGCCGCAGGTGGGCCCCTTGCAGGACCGGGCCGGGGTGGGGATGGGGTTGGG  000001176
AP2        ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       -------+++++++++++++++++-----------------------------------
SRF        ++++++++++++++++++++++++++++++++++++++++++++++++----------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000001079  GTTGGGCCAACAGGGTCCAGATGGGGTCCAGGTGAGGAGGGGAGATTTGGACGATAGGAG  000001138
>>>>>>>>>           |||||||||||||||  |  |||||||||||||| ||||||| ||||||  <<<<<<<<<
000001177  CCG-----ACAGGGTCCAGATGGGTCCG-GGTGAGGAGGGGAGACTTGGACGGTAGGAG  000001229
AP2        ++++---------+++++++++++++++++++++++++++++++----++++++++++++
GATA1      ------------------------------------------------------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       -----------------++++++++++++++++++++++++--------------++++++
SRF        ---------------+++++++++++++++++++++++++++++----------------
TEF        ------------------------------------------------------------
TEF1       ------------------------------------------------------------

000001139  CAGGGGGCTCAGCATCTGGGAGGCAGATCAGTTCGGGGACGGATTTTCTTTTGGAGAAGG  000001198
>>>>>>>>>  ||||||||| ||||||||||| ||||||||||||||| |||||||| ||||||||| ||  <<<<<<<<<
000001230  CAGGGGGCTAAGCATCTGGGAGACAGATCAGTTCGGGGTCGGATTTTATTTTGGAGATGG  000001289
AP2        ++++++++++++++++++++++++++++++++++++++++++++++---------------
GATA1      ------------------------------------------++++++++++------------
MEF2       ------------------------------------------------------------
MEF3       ------------------------------------------------------------
MYOD       ++++++++++++-----------------------------------------------
SRF        --++++++++++++++++++++++++-----------------------++++++++++++
TEF        ------------------------------------------------------------
TEF1       -------------------+++++++++-----++++++++++++++-------------

000001199  AAGTCAGG  000001206
>>>>>>>>>  ||||||||  <<<<<<<<<
000001290  AAGTCAGG  000001297
AP2        --------
GATA1      --------
MEF2       --------
MEF3       --------
MYOD       --------
SRF        ++++++--
TEF        --------
TEF1       --------
```

В среднем, один сайт приходится на 19 нуклеотидов.

Можно ожидать, что если бы искались все 407 факторов, то один сайт приходился бы на 0,37 нуклетоида.

Таким образом, в выравнивании геном человека – геном резуса найдено почти в 1,5 раза больше сайтов, чем в выравнивании геном человека – геном коровы.