Распознавание сайтов

Задание №1

**Поиск регуляторных мотивов транскрипции в бактериальных**

п**оследовательностях**

Необходимо найти регуляторный мотив (набор сайтов) в полученных

последовательностях с помощью программы MEME.

Приведены последовательности перед генами, экспрессия которых регулируется пуриновым репрессором

**PurR**.

>codB

catgctcagatgcctgatgcgacgctgacgcgtcttatcaggcctacccactgtttttacaccgataatttttcccccacctttttgcactcattcatataaaaaatatatttcccca***cgaaaacgattgcttt***ttatcttcagatgaatagaatgcggcggattttttgggtttcaaacagcaaaaagggggaatttcg

>purE

tggtactgggtgcctggcatacggaaggttcaatggtgaaagtcacggcggatgacgttgagctgattcattttccgttttaaaaaacccgcaactttgctgatttcacagcca***cgcaaccgttttcctt***gctctctttccgtgctattctctgtgccctctaaagccgagagttgtgcaccacaggagttttaagacgc

>pyrC

gcgaagatggcagccagggcgaagcaatggcgaaaaaactggcgaaaggcattgaagtgaagccaggcgaaattgtcattccatttactgattaatcacgagggcgcattcgcgccctttatttttcgtgca***aaggaaaacgtttccg***cttatcctttgtgtccggcaaaaacatcccttcagccggagcatagagatta

>purR

ttaccacttccccttttcgtcaagatcggccaaaattccacgcttacactatttgcgtactggccattgaccccttcctgacgctccgtgtcgtttttccggcgtaccgcaacacttttgttgtgcgtaaggtgtgta***aaggcaaacgtttacc***ttgcgattttgcaggagctgaagttagggtctggagtgaaatggaa

>cvpA

cgtaggtcggataaggcgtttacgccgcatccgacacgcattgcccgatgccgcaaaggcataaaaagtcgatggcgttgaatattttttcagcgccatttttattgatgcgcgggaaggaaatcccta***cgcaaacgttttcttt***ttctgttagaatgcgccccgaacaggatgacagggcgtaaaatcgtgggacacat

>purM

atattcagatggttatgaaagaagattattccatccgaaaactaacctttaccctggcacaagtcttctttcgccgcgcgcctggggaaaagacgtgcaaaaaggttgtgtaaagcagtct***cgcaaacgtttgcttt***ccctgttagaattgcgccgaattttatttttctaccgcaagtaacgcgtggggacccaagcag

>guaB

actaacggtaaaaattgcaggggattgagaaggtaacatgtgagcgagatcaaattctaaatcagcaggttattcagtcgatagtaacccgcccttcggggatagcaagcattttttgcaaaaaggggta***gatgcaatcggttacg***ctctgtataatgccgcggcaatatttattaaccactctggtcgagatattgccc

>glnB

aactcaactatttgcgtaagctgctgcaaatcaccaaaggcaacgtcacccacgcggcgagaatggcggggcgcaaccggacagaattttataaactgctttcccgacacgagctg***gatgcaaacgatttca***aggaatgaattggcgttatgtgttacgtttagcagatcaaaagacaggcgaccttttcaaggaatagc

>purL

aggtttactcataaatgagcggcattttgcgtaaacctgcgccagatggcaacttattacagccattggcggcacgcgttgctaattcacgatggtgattttatttc***cacgcaaacggtttcg***tcagcgcatcagattctttataatgacgcccgtttcccccccttgggtacaccgaaagcttagaagacgagagactt

>purA

accaatttgcccgataatattttacgtcgttttggcggtggacttgtggttgcgggcgttgtggtctactacatgttga***ggaaaacgattggctg***aacaaaaaacagactgatcgaggtcatttttgagtgcaaaaagtgctgtaactctgaaaaagcgatggtagaatccatttttaagcaaacggtgattttgaaaaa

>folD

caaggaaataaatcagttccagatttacaacgccatcatggacgaaaaatgaagctttcagtctcagcgacggtgcgcctcacctt***cgcaagaggtcgcttc***acgcgataaatctgaaacgaaacctgacagcgcgccccgcttctgacaaaataggcgcatccccttcgatctacgtaacagatggaatcctctctctg

>rpiA

gaacggtgaactggtgcgcctgtttgcaccgtggagcggcatacccagaccgttgtatgctttatttgcgggg***cgaaaggggatgcctg***ccattgcgcgatattttatggatgagttaaccacgcggcttgccaacggggtctgaatcgctttttttgtatataatgcgtgtgaaattt

В данных последовательностях

* синим выделены экспериментально установленные сайты;
* сайты, найденные с помощью программы MEME с параметром **«One per sequence»** выделены *курсивом* ;
* сайты, найденные с помощью программы MEME c параметром **«Zero or one per sequence»**  выделены **жирным шрифтом;**
* все сайты на сером фоне

Таким образом, программа МЕМЕ распознала с большой точностью сайты в 9 из 10 последовательностей( в одной нашел сайт, не совпадающий с экспериментальным).В двух нашел сайты, хотя экспериментально они не были там обнаружены.

Программа MEME обладает высокой специфичностью и чувствительностью.

У MEME достаточно удобный интерфейс, подробное описание опций и много возможностей. Эта программа хорошо подходит для применения предсказания ,поиска регуляторных сайтов в бактериальных геномах. Однако, доверять только одной программе всё-таки не стоит.

Задание №2

**Поиск сайтов в эукариотических последовательностях**

Необходимо найти сайты для известных сигналов в полученных последовательностях с помощью программы rVISTA. Для этого требуется сделать попарные выравнивания последовательности из генома человека с последовательностями из геномов других млекопитающих(корова и крыса).

Все последовательности представляют собой промоторные области генов, экспрессирующихся в мышечной ткани. Поэтому нужно пытаться найти в них сайты связывания мышечно-специфичных факторов транскрипции.

 AP2 GATA1 MEF2 MEF3

MYOD SRF TEF1 TEF

**>Human Human\_Desmin:1-3350 (+)**

**>Rat Rat\_Desmin:1-2485 (+)**

000000001 CCTCTTTCACCTGCCTTCCAAGCCTCGGGGTGTTGATTTGCAGCCCTCCATGGTAGCAAA 000000060

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000061 TCCTCAGGGGTGGAACTGAATTTTGAAAACAGTCCCAAGTGGTTCGGGGCCAAGTCTGGT 000000120

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000121 GGTTGTTGATTTTGACTAGAGAATAGATGGCATGATTTATTTGTCTTCTGCCTCTTTTCA 000000180

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000181 CCAAGGAACTGAGGACTTGGGCTGGTCCAAAGGCCAATTTCCTTATGTGGTCTGGAGATC 000000240

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000241 ATTCACAGAGAAGAGCTCCAGAAATGTTGTCCCAGTAGCATCATAGAAATAGGTGTAGAG 000000300

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000301 TCCCCTTCAGGAAGATTAATCTGGTTCTTTGGGCTGTAAGTGATCAAAGCAATACCAAGT 000000360

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000361 GCCTTAAATTAAAAAAAAAAAAAAAAAAAAAAACGGAAAGTTTGTTGGCCCAAGGCCAGG 000000420

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000421 AAGGACAGTGTGGGAGCTCAACTCACAAAGTTGAGGGAAGCACTGCAGGAACCAAGGGGC 000000480

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000481 TGGCCTGCTCCTCCTCTCCAGCCTCCTCTGCTTCTTACATATTGACCTCTCTTTCTTCCT 000000540

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000541 ACTCCCCCAGGGGGCAGGAAACATGGCTTCCACAGGTTCCAGTTGAAGAATCCCAGTTCC 000000600

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000601 GTCTATAAATTCCAGGGAAGGTCTCTGATTGGCCCTGCTCATTCCCAGGCCCATTCCTTG 000000660

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000661 ACCCAGTCACTGAAGTCAGGGAGATGCAGTAATAAGACTGGCTGGAATCAGGGTCTTTAG 000000720

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000721 GGGTGGAGGGATGGGGAGGAGGCACAGCATGTCATCAAAATAAGGAAATTGCAAAAGAAA 000000780

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000781 GCTTGCAGGCTACTTTGAATGACAATGAGAAAGACG-GTGCTGCCTGAGTG-TGTTAAGG 000000838

>>>>>>>>> ||||||||||| || ||||||||| | |||||||| <<<<<<<<<

000000001 -----------------AATGACAATGACAACTGAATGTGCTGCCTAGGATATGTTAAGG 000000043

TEF1

GATA1

TEF

SRF

AP2

MYOD

MEF2

000000839 ATCCACATGGTCTCCAAAAT**CCTCCAGGAGCATACAG**TCTAGTCTGGGAGATGAGACACA 000000898

>>>>>>>>> ||||| | ||||||| |||| | |||||||| ||||| ||| || |||| | | <<<<<<<<<

000000044 ATCCATGCAGCCTCCAAACC**CCTCTACTAGCATACAA**TCTAGATTGGAAGCTGAGGCCCC 000000103

TEF1

GATA1

TEF

SRF

AP2 +++++++++++++++++

MYOD

MEF2

000000899 AAAATAACCAGAACAC**AACAGCTTGCACTGACTCGAGGGCTGGATAAGAATATCTGGAAC** 000000958

>>>>>>>>> | || | | |||||| | ||||||||||||||| || | ||||||||||||||||||| | <<<<<<<<<

000000104 AGAAGAGCTAGAACAT**ACCAGCTTGCACTGACTTGAAGCCTGGATAAGAATATCTGGAGC**  000000163

TEF1

GATA1 ++++++++++++++++++++++

TEF

SRF

AP2 +++++++++++++++++++++++++ ++

MYOD

MEF2 ++++++++++++

000000959 **TCCCCCATCTATTTCAGA**AGCT**TGTCTCTTGGATGAAAATTAGA**C**ACTTAATGGGAAAGG** 000001018

>>>>>>>>> ||||||||| |||||| || | |||||| |||| |||||||| ||||||| |||| | <<<<<<<<<

000000164 **TCCCCCATCCATTTCA**-AAACG**TGTCTCGGGGATAAAAATTAGC**T**GCTTAATGCGAAA-G** 000000221

TEF1

GATA1 ++++++++++

TEF

SRF

AP2 +++++++++++++ +++++++++++++ +++++++++++++++

MYOD

MEF2 ++++++++++++++ ++++++++++++

000001019 GCT**TTGAAAAGAGTG**CAGTAACAAAGCCCCCTTTACAATTTAC---CCGGCACATTCACA 000001075

>>>>>>>>> |||||||||||| ||||||| ||| |||| ||| || || | | |||| <<<<<<<<<

000000222 GCT**TTGAAAAGACTG**CAGTAGCAA---GTCCTTCACAGTTCACTTGCTGTAGA---CACA 000000275

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2 ++++++++++++

000001076 CCCATCCTGAGGCCAAAGCCACAGGCTGTGAGGTCTCACTGTCTCAGCT-----T----- 000001125

>>>>>>>>> ||| ||| ||||||||| ||| ||| ||| | | ||| <<<<<<<<<

000000276 TCCAGTCTGGAGCCAAAGCC---TAGTGTTTGGTTTCA--GACGCAGGGACAAGAGACTG 000000330

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000000331 AGGTGCATATTTTAGACACCAAAGCAGACCTGACCTCCAGGTTCTCCCAGCATCCCTCAG 000000390

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000000391 TCACTACCTCGAACACCCTGCCTCCAACCCTGAACTTTCCAGCCAGGGGCTGGGGTGCCC 000000450

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001126 ---------------------------------------------------------CCT 000001128

>>>>>>>>> || <<<<<<<<<

000000451 TTTCTGTAGAGGTTCTTCCCTGTATAATCCAGACATTTTGGTCTTTCTTCGCCTCTCTCT 000000510

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001129 GAGCTATAAAATGGGAATGATGCTAGTGTCTACCT----------C-------------- 000001164

>>>>>>>>> |||||| | | | | ||| || | <<<<<<<<<

000000511 GAGCTAGA----------GCCCTTTCTCTCTTTCTCTTCCCTCTCCCCCAAGCCCTCCCC 000000560

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000000561 ATGGCAACTCCCCTGGCCTCAATCCTTGGGGCAGTGAACTCCCCTCAGAGCAGCTCCCCA 000000620

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000000621 ATAAACCTACCTTTCATAGAATCTAATCTGGCTTGAATTGGTTCATTTCACCAGCAGAGA 000000680

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000000681 AATAACCTATCACCCAGACTCAGGATTCCTCCCTGTGCTATGCTACAAAAATGGCATCTA 000000740

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001165 -----------------CTAGGGTTG**GAGAATTGGGG-GTCATGGGTGT**GAAGTGCTCAG 000001206

>>>>>>>>> | ||| ||||| |||| |||||||||| | ||||| <<<<<<<<<

000000741 CACAGCCTCTGCTCATGGT-CTGTTA**GAGAACAGGGGAGTCATGGGTGC**AGTGCTCTCAG 000000799

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++

MYOD

MEF2

000001207 CAGCTTGGCCCACACTAGGTGGTCAGTACATGTAAGGTA------TTATTGTTGCTACAT 000001260

>>>>>>>>> ||| ||||||||| ||||| ||||||||| | | <<<<<<<<<

000000800 CAGGGA----------------TCAGTACATATAAGGCCTTGTTGTTATTGTTGTTGTAA 000000843

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001261 ACATTAGTAGGGCCTGGGCCTCTTTAAA----CCTTTATAGGGTAGCATGGCAAGGCTAA 000001316

>>>>>>>>> | | | | || ||||| || ||| | | || | || |||| | | <<<<<<<<<

000000844 GCTTCATTGTCTCCCTGGCCTGTTCAAAGGCTCGATCATTGCACAGTCATGCAAAGATGC 000000903

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001317 CCATCCTCACTTTATATCTGACAAGCTGGGGCTCAG------------AGAGGACGTGC- 000001363

>>>>>>>>> | || ||| ||| || | || || || | ||||| ||| <<<<<<<<<

000000904 TCGGCCCAACTCTATGACTAA-----TGCCCCTGAGCATTAGGGAGGTAAAGGACCTGCT 000000958

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001364 -CTGAGC-----TGGGGCTCAGACAAGGACACACCTACTAGTAACCCCTCCAGCTGGTGA 000001417

>>>>>>>>> || ||| | || ||| || | | | | ||||| | | <<<<<<<<<

000000959 TCTTAGCACCTCTTTCTTTCTGACCAGCCCTGGGTACCGA-------CAACAGCTCCTAA 000001011

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001418 TGGCAGG-------TCTAGGGTAGGACCAGTGACTGG**CTCCTAATCGAGCA**CTCTATTTT 000001470

>>>>>>>>> | | |||||||| | | | || |||||||| || |||| ||| <<<<<<<<<

000001012 TCCAGTGCTCTGTTTCTAGGGTTGAGCTATGGAT----**TCCTAATCTAGTA**CTCCATTCC 000001067

TEF1

GATA1 +++++++++++++

TEF

SRF

AP2

MYOD

MEF2

000001471 CAGGGTTTGC**ATTCCAAAAGGGTC**AGGTCCAAG---AGGGACCTGGAGTGCCAAGTGGA- 000001526

>>>>>>>>> |||| |||| ||| | |||||||||| | ||| || ||| | || | ||| | <<<<<<<<<

000001068 CAGGCTTTGT**GTTCTAGAAGGGTC**AGGCCTAAGGGAAGAGAC----ACAGCAATGTGCAC 000001123

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++

MYOD

MEF2

000001527 GGTGTAGAGGCACGGCCAGTACCCATGGAGAATG**GTGGATGTCCTTAGGGGTTAGC**AAGT 000001586

>>>>>>>>> ||||| || |||| |||||||||| || ||| <<<<<<<<<

000001124 A--------------------------GAGAAAT**GTAGATGCCCTTAGGGGTCAGT**GAGT 000001157

TEF1

GATA1 ++++++++++

TEF

SRF +++++++++++++++

AP2 ++++++++++++++++++

MYOD

MEF2

000001587 GCCGTGTGCTAAGGAGGGGGCTTTGG--AGGTTGGGCAGGCCCTCTGTGGGGCTCCATTT 000001644

>>>>>>>>> ||| ||||| || | ||| | | |||| | | || ||| | |||| <<<<<<<<<

000001158 GCCTCGTGCTTGGGGAGTCAGGATGGGAAAGGAAGGCAAACACACTCCAGGGGTTCATT- 000001216

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001645 TTGTGGGGGTGGGGGCTGGAGCATTATAGGGGGTGGGA------AGTGATTGGGGCTGTC 000001698

>>>>>>>>> | ||| |||||| || ||| <<<<<<<<<

000001217 -----------------------------CTGTCGGGGTCCTAGAGTGATGGGAGCT--- 000001244

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001699 ACCCTAGCCTTCCTTATCTGACGCCCACCCATGCCTCCTCAGGT--------ACCCCCTG 000001750

>>>>>>>>> | |||||||||| |||| || ||| |||||| | <<<<<<<<<

000001245 ---CCAGCCTTCCTT--------GACACCTCTGTCTCTTCAGGTGCCTGGCTC----CCA 000001289

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001751 CCCCCCACAGCTCCTCTCCTGTGCCTTGTTTCCCAGCCATGCGT-TC**TCCTCTATAAATA** 000001809

>>>>>>>>> ||||| | | ||||||||| | | | || | ||||||||||||||| <<<<<<<<<

000001290 GTCCCCAGAACGCCTCTCCTGCACTTCGCTTGGTGCTGCTTTCCGTC**TCCTCTATAAATA**  000001349

TEF1

TEF

GATA1

SRF +++++++

AP2

MYOD

MEF2 ++++++++++++

000001810 **CCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGATGGTTGGGTTGACATG** 000001869

>>>>>>>>> || ||||||||||| | |||||||||||||||||||||||||||| ||| ||||| || <<<<<<<<<

000001350 **CCTACTCTGGTATTTCGCCTTGGCAGCTGTTGCTGCCAGGGAGATGGCTGGCTTGACGTG** 000001409

TEF1

GATA1 ++++++++++

TEF

SRF ++++++++++++++

AP2 +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD +++++++++++++++++++

MEF2 +++

000001870 **CGGCTCCTGACAAAACACAAACCCC-TGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGG**  000001928

>>>>>>>>> | |||||||||||||||||||||| ||||||| ||||| ||||| |||||||||||||| <<<<<<<<<

000001410 **CACCTCCTGACAAAACACAAACCCCGTGGTGTGAGTGGGTGTGGGCGGTGTGAGTAGGGG** 000001469

TEF1 ++

GATA1 +++

TEF

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2 ++++++++++++

000001929 **G-ATGAATCAGGGAGGGGGCGGGGGAC**CCAGG--GGGCAGGAGCCACACA**AAGTCTGTGC**  000001985

>>>>>>>>> | ||||||||| ||||||| | |||| || | | ||||||| || ||||| | ||| <<<<<<<<<

000001470 **GGATGAATCAGAGAGGGGGTGAGGGAG**ACA-GGGGCGCAGGAGTCAGGCA**AAGGCGATGC** 000001528

TEF1 +++++

GATA1 ++++++++++++

TEF

SRF

AP2 +++++++++++++++++++++++++++ ++++++++++

MYOD

MEF2

000001986 **GGGGGTGGG**AGCGCACATAGCAATTGGAAACTGAAAGCTTATCAGACCCTTTCTGGAAAT 000002045

>>>>>>>>> ||||||| || ||| || |||||||| | | | | |||||||| <<<<<<<<<

000001529 **GGGGGTGTG**ACTACACGCAG---TTGGAAACAGCCATCAGAATCTC-----TCTGGAAAC 000001580

TEF1

TEF

GATA1

SRF

AP2 +++++++++

MYOD

MEF2

000002046 CAGCCCACTGTTTATAAACTTGAGGCCCCACCCTCGACAGTACCGGGGAGGAAGAG**GGCC** 000002105

>>>>>>>>> || ||| ||||||||||||| |||| <<<<<<<<<

000001581 TCTCCTGCTGGCTATAAACTTGAGGGATGCAGAA----------------------**GGCC** 000001618

TEF1

TEF

GATA1

SRF

AP2 ++++

MYOD

MEF2

000002106 **TGCACTAGTCCAGAGGGAAACTGAGGCTCAGGGCTAG**CTCGCCCATAGACATACATGGCA 000002165

>>>>>>>>> |||||| ||| |||||||||||||||||| | || | | || | ||| | <<<<<<<<<

000001619 **TGCACTCCTCCCTAGGGAAACTGAGGCTCAGAGTTAA**AACCC----AGGTGTCAGTGGTA 000001674

TEF1

TEF

GATA1

SRF

AP2 +++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002166 GGCA----------GGCTTTGGCCAGGATCCCTCCGCCTGCCAGGCGTCTCCCTGCCCTC 000002215

>>>>>>>>> ||| ||| ||| |||| | ||| ||| | | | ||| ||| <<<<<<<<<

000001675 TGCATAAGCAATCAGGCACTGGTCAGGGTTACTC--TCTGACTACCATAGGCCTATCCTA 000001732

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002216 CCTTCCTGCCTAGAGACCCCCACCCTCAAGCCTGGCTGGT------CTTTGCC------- 000002262

>>>>>>>>> | || ||||||||| || | |||| || || | <<<<<<<<<

000001733 CAGGCCCACCTAGAGACTCC------TTTGATGGGCTAGTAGACTCCTGTCTGGGCTTCG 000001786

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000000 ------------------------------------------------------------ 000000000

>>>>>>>>> <<<<<<<<<

000001787 CCCACAGTCGTTGGGAACTTCAGCATTTTCTAGGCAATTTTATTTGCAACTTTATTTATC 000001846

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002263 --------------------TGAGACCCAAACCTCTTCGACTTCAAGAGAATATTTAGGA 000002302

>>>>>>>>> | | |||||||| | | || | | <<<<<<<<<

000001847 GGGGGTCCTTGTTCGGTCCAATAACCTAAAACCTCTCCTCGGTGAAAATAGGCAG----- 000001901

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002303 ACAAGGTGGTTTAGGGCCTTTCCTGGGAACAGGCCTTGACCCTTTAAGAAATGACCCAAA 000002362

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002363 GTCTCTCCTTGACCAAAAAGGGGACCCTCAAACTAAAGGGAAGCCTCTCTTCTGCT-GTC 000002421

>>>>>>>>> |||||||| || | || |||| ||||| <<<<<<<<<

000001902 -------------------------CCTCAAACAAATG-AAATTCTCTAGCCTGCTTT-- 000001933

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002422 TCCCCTGACCCCACTCCCCCCCACCCCAGGACGAGGAGATAACCAGGGCT-------GAA 000002474

>>>>>>>>> |||||||| || |||| | ||| <<<<<<<<<

000001934 -----------------------CCCCAGGATAAG-------ACAGGCATCCAAATGGAA 000001963

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002475 AGAGGCCCGCCTGG-----GGGCTGCAGACATGCTTGCTGCCTGCCCTGGCGAAGGATTG 000002529

>>>>>>>>> | || | || || || | | || ||| ||||||||| || | <<<<<<<<<

000001964 AAGGGGGC---CGGCCAGGGGTCTCCTGTCA-----GCTCCCTGCCCTGTGAAACCCT-- 000002013

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002530 GCAGGCTTGCCCGTCACA--GGACCCCCGCTGGCTG**ACTCAGGGGCGCAGGCCTCTTGCG** 000002587

>>>>>>>>> |||||| |||| ||| || | || |||| |||||||||||||||| <<<<<<<<<

000002014 GCAGGCCTGCCTGTCTTCTGT---CCTCTTGGGGCT**GTCCAGGAGCGCAGGCCTCTTGCG** 000002070

TEF1

TEF

GATA1

SRF +++++++++++

AP2 ++++++++++++++++++++++++

MYOD

MEF2

000002588 **GGGGAGCTGGCCTCCCCGCCCCCACGGCCACGGGCCGCCCTTTCCTGGCAGGACAGCGGG** 000002647

>>>>>>>>> ||||||||||||||||||||||| || | ||| | || |||||||||||||||| ||| <<<<<<<<<

000002071 **GGGGAGCTGGCCTCCCCGCCCCCTCGCCTGCGGCCGCCCTTTTCCTGGCAGGACAGAGGG** 000002130

TEF1

GATA1 ++++

TEF

SRF +++++ +++++++++++++++++++++

AP2 +++++++++++++++++++++++++++++++ +++++++++++++++++++++++++

MYOD

MEF2

000002648 **ATCTTGCAGCTGTCAGGGGAGGGGAGGCGGGGGCTGATGTCAGGAGGGATACAAATAGTG** 000002707

>>>>>>>>> ||| |||||||||||||||||||| |||| ||| |||||||||||||| ||||||||||| <<<<<<<<<

000002131 **ATCCTGCAGCTGTCAGGGGAGGGGCGGCGAGGGGTGATGTCAGGAGGGCTACAAATAGTG** 000002190

TEF1

GATA1 ++++++ ++++++++++

TEF

SRF ++++++++++

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++ +++++++

MYOD ++++++++++++++++++

MEF2 +++++++++++++

000002708 **CCGACGGCTGGGGGCC**CT-**GTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCC** 000002766

>>>>>>>>> ||||| ||| ||| | |||||||||| | ||||||||||| |||||| ||| ||| <<<<<<<<<

000002191 **CCGACAGCTAAGGGGC**TCC**GTCTCCCCTCTTCATATCCACTCTCCAGCCGGCTGCCCGCC** 000002250

TEF1

GATA1 ++++++++++

TEF

SRF +++++

AP2 ++++++++++++++++ +++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2 ++

000002767 **CGCCGCCTCCTCCGTGCGCCCGC**-**CAGCCTCGCCCGCGCCGTCACCATGAGCCAGGCCTA**  000002825

>>>>>>>>> || |||||||| |||| |||| |||||||| || ||||| |||||||||||||||||| <<<<<<<<<

000002251 **TGCTGCCTCCTCTTTGCGTCCGC**T**CAGCCTCGTCCACGCCGCCACCATGAGCCAGGCCTA**  000002310

TEF1

TEF

GATA1

SRF

AP2 +++++++++++++++++++++++ ++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002826 **CTCGTCCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGCCCCGGGCTTCCC**  000002885

>>>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| | <<<<<<<<<

000002311 **CTCGTCCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGCGCCCCGGGCTTCTC** 000002370

TEF1

TEF

GATA1

SRF +++++++++++++++++++++

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002886 **ACTCGGCTCCCCGCTGAGTTCGCCCGTGTTCCCGCGGGCGGGTTTCGGCTCTAAGGGCTC** 000002945

>>>>>>>>> ||||| ||||||||||| || || |||||||| || || || ||||| | |||||||| <<<<<<<<<

000002371 **GCTCGGTTCCCCGCTGAGCTCTCCGGTGTTCCCACGAGCAGGCTTCGGTACCAAGGGCTC** 000002430

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002946 **CTCCAGCTCGGTGACGTCCCGCGTGTACCAGGTGTCGCGCACGTCGGGCGGGGCCGGG**GG 000003005

>>>>>>>>> ||| ||||| ||||| ||||||||||||||||||||||||||||||||||||||| <<<<<<<<<

000002431 **CTCGAGCTCAGTGACATCCCGCGTGTACCAGGTGTCGCGCACGTCGGGCGGGGCC---**-- 000002485

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD ++++++++++++++++++

MEF2

000003006 CCTGGGGTCGCTGCGGGCCAGCCGGCTGGGGACCACCCGCACGCCCTCCTCCTACGGCGC 000003065

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003066 AGGCGAGCTGCTGGACTTCTCACTGGCCGACGCGGTGAACCAGGAGTTTCTGACCACGCG 000003125

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003126 CACCAACGAGAAGGTGGAGCTGCAGGAGCTCAATGACCGCTTCGCCAACTACATCGAGAA 000003185

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003186 GGTGCGCTTCCTGGAGCAGCAGAACGCGGCGCTCGCCGCCGAAGTGAACCGGCTCAAGGG 000003245

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003246 CCGCGAGCCGACGCGAGTGGCCGAGCTCTACGAGGAGGAGCTGCGGGAGCTGCGGCGCCA 000003305

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003306 GGTGGAGGTGCTCACTAACCAGCGCGCGCGCGTCGACGTCGAGCG 000003350

>>>>>>>>> <<<<<<<<<

000000000 --------------------------------------------- 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

**>Human Human\_Desmin:1-3350 (+)**

**>Cow Cow\_Desmin:1-3058 (+)**

000000001 C**CTCTTTCACCTGCCTTCCAAGCCTCGGGGTGTTGATTTGCAGCCCTCCATGGTAGCAAA** 000000060

>>>>>>>>> |||||||||||||||||||||||||| |||||| ||||| |||||||||||||||||| <<<<<<<<<

000000001 C**CTCTTTCACCTGCCTTCCAAGCCTCCAAGTGTTGTTTTGCTGCCCTCCATGGTAGCAAA** 000000060

TEF1

TEF

GATA1 +++++

SRF ++++++++++++++++++++++++

AP2 +++++++++++++++++++++++++++++++++++ +++++++++++++++ +++

MYOD ++++++++++++++++++

MEF2

000000061 **TCCTCAGGGGTGGAACT**GAA**TTTTGAAAACAGTCCCAAGTGGTTCGGGGCCAAGTCTGGT** 000000120

>>>>>>>>> |||| ||||||||||| ||||||| |||||| ||||||||||| | ||||||||||| <<<<<<<<<

000000061 **TCCTTAGGGGTGGAACC**AAA**TTTTGGAAACAG**C**TCCAAGTGGTTCAGAGCCAAGTCTGGC**  000000120

TEF1 ++++++++++++

TEF

GATA1 ++++

SRF ++++++++++++++++++

AP2 ++++++++++++++++ +++++++++++++++++ ++++++++

MYOD ++++++++++++++++++

MEF2 ++++++++++++

000000121 **GGTTGTTGATTTTGAC**TAGAGAATAGATGGCA**TGATTTATTTGT**CTTCTGCCTCT----T 000000176

>>>>>>>>> | |||||||||||||||||||||| |||| || |||||||| | | || || | | <<<<<<<<<

000000121 **AGATGTTGATTTTGAC**TAGAGAATACATGGTA**TAATTTATTTAT**TTGCTTTCTGTCTCAT 000000180

TEF1 ++++++++++

TEF

GATA1

SRF ++++++

AP2 +++++

MYOD

MEF2 ++++++++++++ ++++++++++++

000000177 TTCACCAAG**GAACTGAGGACTTGGGCTGGTCCAAAGGCCAATTTCCTTATGTGGTCTGGA** 000000236

>>>>>>>>> ||||| |||||||||||| || |||||||||||||| |||||||||||||||||||||| <<<<<<<<<

000000181 TTCACTAAG**GAACTGAGGCCTCAGGCTGGTCCAAAGGGCAATTTCCTTATGTGGTCTGGA**  000000240

TEF1

TEF

GATA1

SRF +++++++++++++++++++++++

AP2 ++++++++++++++++++++++++++++++++++ +++++++++++++

MYOD

MEF2

000000237 **GATCA**TTCACAGAGAAGAGCTCCAGAAATGTTGTCCCAGTAGCA---TCATAGAAATAGG 000000293

>>>>>>>>> || |||||||| | ||| ||| || |||| | | | | | ||||| || <<<<<<<<<

000000241 **GACTG**TTCACAGACAGGAGTCCCAAAAGTGTTCTGGCCTCTTTAGTGTTGTGGAAATGGG 000000300

TEF1

TEF

GATA1

SRF +

AP2 +++++

MYOD

MEF2

000000294 TGTAGAGTCCCCTTCAGGAAGATTAA-TCTGGTTCTTTGGGCTGTAAGTGATCAAAG-CA 000000351

>>>>>>>>> | | |||| |||| | | |||| |||| ||| | | || || || || <<<<<<<<<

000000301 CGCACAGTCTCCTTACAGCATATTAGCCAGGGTTGTTTTGATTACAAATGGCAGTAGCCA 000000360

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000352 ATACCAAGTGCCTTAAA-TTAAAAAAAAAAAAAAAAAAAAAAACGGAAAGTTTG**TTGGCC** 000000410

>>>>>>>>> ||||||||| || | |||||||| | ||||||| |||||| <<<<<<<<<

000000361 ATACCAAGTCTCTGCAGTTTAAAAAAGA------------------AAAGTTTA**TTGGCC** 000000402

TEF1

TEF

GATA1

SRF

AP2 ++++++

MYOD

MEF2

000000411 **CAAGG-CCAGGAAG**G**ACAGTGTGGGAG**CTCAACTCACAAAGTTGAGGGAAGCACTGCAGG 000000469

>>>>>>>>> | | |||||||||| | || ||||||||| |||| | | || |||| || <<<<<<<<<

000000403 **CCAAACCCAGGAAG**G**ATACTGGGGGAG**CTCA-----CAAAAT-----GGAGAACTGAAGA 000000452

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++ ++++++++++++

MYOD

MEF2

000000470 A----A--CCAAGGGGCTGGCCTGCTCCTCCTCTCCAGCCTCCTCTGCTTCTTACATATT 000000523

>>>>>>>>> | || || | ||| | | ||||||||| | || ||| | | <<<<<<<<<

000000453 ACCAGGGTCCCGGGAGATGGAAT--TTAGGACTTCCAGCCTCACATTCTCCTTTTCTCTG 000000510

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000000524 GACCTCTCTTTC-----------------------------------TTCCT**ACTCCCCC**  000000548

>>>>>>>>> | ||||| || |||| <<<<<<<<<

000000511 GGTGTCTCTGCTGCTCCCGGCATGCTGACCTTTTTTTCCCTATGCCAGAAAG**ACCTCCCC** 000000570

TEF1

TEF

GATA1

SRF +++++

AP2 ++++++++

MYOD

MEF2

000000549 **AGGGGGCAGGAAA**CATGGCTTCCACAGGTTCCAGTT--GAAGA**ATCCCAGTTCCGTCTAT** 000000606

>>>>>>>>> || ||| ||||||||||| | | |||| ||||||| | |||||||||||| | ||| <<<<<<<<<

000000571 **AGTGGGTAGGAAA**CATGGTTGC--CAGGCTCCAGTTTAGT--A**ATCCCAGTTCCATGTAT** 000000626

TEF1

TEF

GATA1

SRF +++++++++++++ +++++++++++++

AP2 +++++++++++++ +++++++++++++

MYOD

MEF2 ++++

000000607 **AAATTCCAGGGAAGGTCTCTGATTGGCCCTGCTCATTCCCAGGCC**CATTCCTTGACCCAG 000000666

>>>>>>>>> |||||||| || |||||||||||||||||| ||| ||||||| |||||| ||| <<<<<<<<<

000000627 **AAATTCCAAGGTGGGTCTCTGATTGGCCCTGGTCAGGTCCAGGCC**TGCCCCTTGAGCCAA 000000686

TEF1

TEF

GATA1 ++++++++++

SRF +++++++++++++++++++++++

AP2 +++++++++++++ ++++++++++++++++++++

MYOD

MEF2 ++++++++

000000667 TCACT**GAAGTCAGGGAGATGCAGTAATAA**GACTGGCTGGAATCAGG-------------- 000000712

>>>>>>>>> ||| |||||| | ||| |||| |||||| ||||||| | | | <<<<<<<<<

000000687 CTGCT**GCAGTCAGAGGGATTCAGTTATAA**GAGTGGCTGGGAGCTGACTCACCTGTCCACT 000000746

TEF1 ++++++++++++

TEF

GATA1 ++++++++++

SRF +++++++++++++++

AP2

MYOD

MEF2

000000713 -------------GTCTTTAGGGGTGGAGGGATGGGGAGGAGGCACAGC**ATGTCATCAAA** 000000759

>>>>>>>>> |||| || | | | | | | | | ||| | |||| | |||| <<<<<<<<<

000000747 CATGGACAGGTGTGTCTCATGGACATGTGTGTTTGTGTGCACGCAGCGT**ATGTTAACAAA**  000000806

TEF1 +++++++++++

TEF

GATA1

SRF +++++++++

AP2

MYOD

MEF2 ++++++

000000760 **ATAAGGAA**ATTGCAAAAGA**AAGCTTGCAGGCTACTTTGAA**TGACAATGAGAAAGACGGTG 000000819

>>>>>>>>> |||||| || | ||| || ||| |||||||||||||| | | ||| <<<<<<<<<

000000807 **ATAAGGGG**ATGGGAAAGAA**AT-CTTTCAGGCTACTTTGAA**CAAGGACA---------GTG 000000856

TEF1 +

TEF

GATA1

SRF ++++++++

AP2 +++++++++++++

MYOD

MEF2 ++++++ +++++++++++

000000820 CTGCCTGAGTGTGTTAAGGAT----C--------CACATGGTCTCCAAAATC--**CTCCAG** 000000865

>>>>>>>>> ||||||| | |||| | |||||||||||| || |||||| <<<<<<<<<

000000857 CTGCCTG----TCCAAAGGGACATGCAAAGGAGTCACATGGTCTCCCCAAAATT**CTCCAG** 000000912

TEF1

TEF

GATA1

SRF

AP2 ++++++

MYOD

MEF2

000000866 **GAGCATACAGTCTAGTCTGGGAGATG**AGAC**ACAAAAATAACCAGAACA**C**AACAGCTTGCA** 000000925

>>>>>>>>> |||| |||| |||||| ||||||| ||||||||||||||||||||||||||||||||||| <<<<<<<<<

000000913 **GAGCCTACAATCTAGTTTGGGAGACG**AGAC**ACAAAAATAACCAGAACA**C**AACAGCTTGCA** 000000972

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++++ +++++++++++

MYOD

MEF2 ++++++++++++++++++

000000926 **CTGACTCGAGGGCTGGATAAGAATATCTGGAACTCCCCCATCT**ATTTCAGAAGCT**TGTCT** 000000985

>>>>>>>>> ||||||||| |||||||||||||||||||||||||||| | ||| |||||||||| <<<<<<<<<

000000973 **CTGACTCGAAGGCTGGATAAGAATATCTGGAACTCCCCACCCA**CC-TCATAAGCT**TGTCT**  000001031

TEF1

TEF

GATA1 ++++++++++++++++++++++

SRF

AP2 ++++++++++++++ ++++++++++++ +++++

MYOD

MEF2 +++++++++++

000000986 **CTTGGATGAAAATTAGA**C**ACTTAATGGGAAAGGGCTTTGAAAAGAGTG**C**AGTAACAAAGC** 000001045

>>>>>>>>> ||||||| ||||||||| ||| |||| |||||||||||| ||||| |||| |||||||| <<<<<<<<<

000001032 **CTTGGATAAAAATTAGA**A**ACTGAATGTGAAAGGGCTTTGCAAAGACAG**C**AGCAACAAAGC**  000001091

TEF1 ++++++++++++

TEF

GATA1 +++++++++++

SRF

AP2 ++++++++ ++++++++++++++++++

MYOD

MEF2 ++++++++++++ ++++++++++++ +++++++++++

000001046 **C**CCCTTTACAATT**TACCCGGCACAT**TCA**CACCCATCCTGA**GGCCAAAGCCACAGGCTGTG 000001105

>>>>>>>>> | | | ||| ||||||||||||| | |||| || ||||| || |||| || <<<<<<<<<

000001092 **T**CTTTGCAGTTA-**-ACCTGGCACAT**TCA**CACATACCCTGG**GGACAAAG-CATGGGCTCTG 000001148

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++ ++++++++++++

MYOD

MEF2 +

000001106 -AGGTCTCACT**GTCTCAGCTTCCTGAG**CT**ATAAAATGGGAA**TGATGCTAGTGTCTACCTC 000001164

>>>>>>>>> | ||| |||| |||||||||||||||| || ||||||| | | |||| |||||| <<<<<<<<<

000001149 GA-GTCAGACT**GCCTCAGCTTCCTGAG**CT**GCAACATGGGAA**GGGTACTAGCTCTTACCTC 000001207

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++ ++++++++++++

MYOD

MEF2

000001165 CTAGGGTTG----**GAGAATTGGGGGTCATGGGTG**TGAAGTGCTCAGCAGCTTGGCCCACA 000001220

>>>>>>>>> |||| || ||||||||||| ||| ||||| |||||||| ||||| || | ||| <<<<<<<<<

000001208 ACAGGGGTGTCAA**GAGAATTGGGGATCACGGGTG**GGAAGTGCTTAGCAGAGTGACACACG 000001267

TEF1

TEF

GATA1

SRF

AP2 +++++++++++++++++++++

MYOD

MEF2

000001221 CTAGGTGGTCAGTACATGTA---------------------------------------- 000001240

>>>>>>>>> | |||| |||| | | | <<<<<<<<<

000001268 GTGGGTGCTCAGAATACACAAGGTAGTAATTGTTACTTGAGCCTCTTCAAACCCATAGTC 000001327

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001241 --------AGGTATTATTGTTGCTACATACATTA--------------------GTAGGG 000001272

>>>>>>>>> | || | ||||||||| | ||| <<<<<<<<<

000001328 GGTAGACACCGCCAGGGTGACCATTCATACATTATAGTGACGACACAGAGAAGGGAAGGA 000001387

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001273 CCTGGGCCTCTTTAAACCTTTATAGGGTAGCATGGCAAGGCTAACCATCCTCACTTTATA 000001332

>>>>>>>>> | || <<<<<<<<<

000001388 CTTG-------------------------------------------------------- 000001391

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001333 TCTGACAAGCTGGGGCTCAGAGAGGACGTGCCTGAGCTGGGGCTCAGACAAGGACACACC 000001392

>>>>>>>>> || <<<<<<<<<

000001392 ----------------------------------------------------------CC 000001393

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000001393 TACTAGTAACCCCTCCAGCT**GGTGATGGCAGGTCTAGGGTA**GGACCAGTGACTGGCTCCT 000001452

>>>>>>>>> | || |||| | | | ||||||||||| ||| | || |||||| <<<<<<<<<

000001394 TGCTTGTAATTACATCTTCA**GGTGATGGCAGAGCTAAGATA**AGACCAGAC---------- 000001443

TEF1

TEF

GATA1 ++++++++++

SRF

AP2 ++++++++++++

MYOD

MEF2

000001453 AATCGAGCA**CTCTATTTTCAG**GGTTT**GCATTCCAAAAGGGTCAGGTCCAAGAGGGACCTG** 000001512

>>>>>>>>> |||||||| || |||||||||| ||||||||||| | ||||| |||||| <<<<<<<<<

000001444 ---------**CTCTATTTCCAA**ACTTT**GCATTCCTAAAGGGTCAGGCCTAAGAGAGACCTG** 000001494

TEF1 +++++

TEF

GATA1

SRF ++++++++++++++++++++++++++++++++++

AP2 ++++++++++++++ +++++++++++++++

MYOD

MEF2 +++++++++++++

000001513 **GA--GTGCCAAGTGGAGG**TGTAGAGGCACGGCCAGTACCCATGGAGAATGGTG**GATGTCC** 000001570

>>>>>>>>> | |||||| |||||| | | ||| || |||| || ||||||||| | |||||| <<<<<<<<<

000001495 **AATAGTGCCAGGTGGAGC**TTTGGAGAAACAGCCAATATTCATGGAGAAGCGGA**AATGTCC** 000001554

TEF1

TEF

GATA1

SRF ++ +++++++

AP2 ++++++++++++++++++ ++++++

MYOD

MEF2

000001571 **TTAGGGGTTAGCAAGTGCCGTGTGCTAA**GGAGGGGGCTTTGGA-----GGTTGGGCAGGC 000001625

>>>>>>>>> |||||||| ||| |||||| ||| ||||||||| | | || | || ||| || || <<<<<<<<<

000001555 **TTAGGGGTCAGCCAGTGCCTTGTACTAA**GGAGGAGACCTTAGTACAAGGGATGGACAAGC 000001614

TEF1

TEF

GATA1

SRF ++++++++

AP2 ++++++++++++++++++++++++++++

MYOD

MEF2

000001626 CCTCTGTGGGGCTCCATTTTTGTGGGGGTGG**GGGCTG-GAGCATTATAGGGGGTGGGAAG** 000001684

>>>>>>>>> ||| | || | | |||||||| |||||| | || |||||||||| <<<<<<<<<

000001615 CCT--------ATTCAGGGCTTCTCTGCTGGG**GGCTTAGAGCATGA-AGAGGGTGGGAAG** 000001665

TEF1

TEF

GATA1 ++++

SRF

AP2 +++++++++++++++++++++++++++

MYOD

MEF2

000001685 **TGATTGGGGCTG**TCACCCT--AGCCTTCCTT----ATCTGACG**CCCACCCATGCCTCCTC** 000001738

>>>>>>>>> ||||||||||| ||| ||||||||| |||||| ||||||| |||||| <<<<<<<<<

000001666 **TGATTGGGGCTC**ATCGCCTCAAGCCTTCCTCCACCATCTGA--**--TACCCATGTCTCCTC** 000001721

TEF1

TEF

GATA1 +++++++++++

SRF

AP2 ++++++++++++ +++++++++++++++++

MYOD

MEF2

000001739 **AGGTACC**CCCTG-----**CCCCCCACAGCTCCTCTCCTGTGC-CTTG**TTTCCCAGCCATGC 000001792

>>>>>>>>> |||||||| ||||||| ||||||||||||||| |||||||||| | | <<<<<<<<<

000001722 **AGGTACC**CGACCCCCAC**CCCCCCAGAGCTCCTCTCCTGTGTGCTTG**TTTCCCGGACTGAT 000001781

TEF1

TEF

GATA1

SRF

AP2 +++++++ ++++++++++++++++++++++++++++++

MYOD

MEF2

000001793 G--TTC**TCCTCTATAAATACCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGG** 000001850

>>>>>>>>> | |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| <<<<<<<<<

000001782 GCATTC**TCCTCTATAAATACCAGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGG** 000001841

TEF1

TEF

GATA1 ++

SRF +++++++++++++++++++++

AP2 ++++++++++++++++++++++++++++++++

MYOD +++++++++++++++++

MEF2 ++++++++++++++++ ++++++++++++

000001851 **AGATGGTTGGGTTGACATGCGGCTCCTGACAAAACACAAACCCCTGGTGTGTG**TGGGCGT 000001910

>>>>>>>>> |||||| |||||||||||| ||||||||||||||||||||||||||| | | <<<<<<<<<

000001842 **AGATGGCTGGGTTGACATGAGGCTCCTGACAAAACACAAACCCCTGGGGGTGG**------- 000001894

TEF1

TEF

GATA1 ++++++++

SRF

AP2 +++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2 ++++++++++++

000001911 **GGGTGGTGTGAGT-AGGGGGATGAATCAGGGAGGGGGCGG**GGGACCCAGGGGGCAGGAGC 000001969

>>>>>>>>> |||||||||||| ||||||||||||||| ||||||| <<<<<<<<<

000001895 **GGGTGGTGTGAGGAAGGGGGATGAATCAGAGAGGGGGT--**-------------------- 000001932

TEF1 ++++++

TEF

GATA1 ++++++++++++++

SRF

AP2 +++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000001970 CACACAAAGTCTGTGCGGGGGTGGGAGCGCACATAGC----------------------- 000002006

>>>>>>>>> ||||||||| | || | <<<<<<<<<

000001933 ----------------GGGGGTGGGCTCACAGGGACAGAAGCCCAGAGAAGTTGGCAGTG 000001976

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002007 --------------------------------AATTGGA**AACTGAAAGCTTATCAGACCC** 000002034

>>>>>>>>> ||||||||||||| ||| || ||| ||| <<<<<<<<<

000001977 TGGGGGTGAGAGGACACAGTTATAATTGTTATAATTGGA**AACTGAGAGCCTACCAGGCCC** 000002036

TEF1

TEF

GATA1

SRF ++++++++++

AP2 ++++++++++++++++++++

MYOD

MEF2

000002035 **TTTCTGGAAATCAGCCCACTGTTT**A**TAAACTTGAGGCCCCACCCTCG**ACAGTACCGGGGA 000002094

>>>>>>>>> ||||||||||||| || ||||| |||||||||||| ||||||| ||| || | <<<<<<<<<

000002037 **TTTCTGGAAATCACCCTGTTGTTT**G**TAAACTTGAGGCTGCACCCTCA**CTGGTAGGGGTGG 000002096

TEF1

TEF

GATA1 ++++++++++

SRF ++++++++++++++

AP2 ++++++++++++++++++++++++ ++++++++++++++++++++++

MYOD

MEF2 ++++++++++++++

000002095 GGAAGAG---GGCC**TGCACTAGTCCAGAGGGAAACTGAGGCTCAGGGCTAG**CTCGCCCAT 000002151

>>>>>>>>> || || | ||| || | |||||||||||||||||||| | ||| | | | || <<<<<<<<<

000002097 GGTGGAAAAGGACC**TTCATGAATCCAGAGGGAAACTGAGGCTGGGAGCTGG**AACTCAGAT 000002156

TEF1

TEF

GATA1

SRF

AP2 +++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002152 A------GACATACATG-GCAGGCAGGCTTTGGCC----AGGATCCCTCCGCCTGCCAGG 000002200

>>>>>>>>> ||||||| || | ||| |||| ||| || | ||| || | ||| <<<<<<<<<

000002157 GCTAGTGGACATACCTGAGGAGGTGGGCTCTGGTGGGGATGGCTGCCTATCCCAGGGAGG 000002216

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002201 CGT-------CTCCCTGCCCTCCCTTCCTGCCTAGAGACCCCCACCCT----------C- 000002242

>>>>>>>>> | | ||||||||| | |||||| | | || | <<<<<<<<<

000002217 CCTTCCCATCCTCCCTGCCTGTC----------AGAGACGCTCCTCCACTAGGCCTGGCT 000002266

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002243 -AAGCCTGGCTG**GTCTTTGCCTGAGACC**C---AAACCTCTTCGACTTCAAGAGAATATT- 000002297

>>>>>>>>> || | |||||||||| ||||| || ||| | | | |||| |||||| || <<<<<<<<<

000002267 GGGGCATCCCTG**GTCTTTGTCTGAGCCC**TCCTAAATGTTTGCCACTTTCGGAGAATCTTT 000002326

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++

MYOD

MEF2

000002298 -------TAGGAACAAGGTGGTTTAGGGCCTTTCCTGGGAACAGGCCTT**GACCCTTTAAG** 000002350

>>>>>>>>> | || |||||| ||| | || || ||| || |||||||| | ||| | | <<<<<<<<<

000002327 AGGGAAAAAAAAAAAAGGTGATTTGCAGACTGTCTTGGAAATAGGCCTT**GTCTCTTCAGG** 000002386

TEF1 +

TEF

GATA1

SRF

AP2 +++++++++++

MYOD

MEF2

000002351 **AAATGACCCAAAGTCT**CTCCTTGACCAAAAAGGGG**A-CCCTCAAACTAAAGG**GA**AGCCTC** 000002409

>>>>>>>>> |||||||||||||| |||| ||| ||| | ||||||||||||| |||||||| <<<<<<<<<

000002387 **AAATGACCCAAAGTAT**CTCTTTGGAGCCAAAAGAA**GCCCCTCAAACTAAATA**GA**AGCCTC** 000002446

TEF1 +++++

TEF

GATA1

SRF

AP2 + +++++++++++++ +++++++++++++++++++++++++

MYOD

MEF2

000002410 **TCTTCTGCTGTCT**CCCCTGACCCCACTCCCCCCCACCCCAGGACGAGGAGATAACCAGGG 000002469

>>>>>>>>> || | ||| ||||| |||| ||||||||||| ||| | ||| <<<<<<<<<

000002447 **TCCCCAGCTCTCT**CCGCTGAT------------CACCCCAGGACCAGG--GAGGCAAGGA 000002492

TEF1

TEF

GATA1

SRF

AP2 +++++++++++++

MYOD

MEF2

000002470 CTGAAAGAGGCCCG--CCTG**GGGGCTGCAGACATGCTTGCTGCC**TGCCCTGGCGAAGGAT 000002527

>>>>>>>>> ||||||| || | | | ||||||| |||||||| |||| || | ||| <<<<<<<<<

000002493 CTGAAAGGGGGCTGGCC--G**GGGGCTCCAGACATGTTTGCGCCC**GGGTGAAGCGGC---- 000002546

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++

MYOD

MEF2

000002528 TGGCAGGCTTGCCCGTCACAGGACCCCCGCTGGCTGACTCAGGGGCGCAGGCCTCTTGCG 000002587

>>>>>>>>> |||||||| ||||| | | || || | | | | <<<<<<<<<

000002547 TGGCAGGCCTGCCCACCTTCTGCCCTCCTAGGTCCGCCAGCCTCT--------------- 000002591

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002588 GGGGAGCTGGCCTCCCC------GCCCCCACG---------------------------- 000002613

>>>>>>>>> | | |||| |||||||| <<<<<<<<<

000002592 -CCCA------CACCCCGCAGGGGCCCCCACCCACCCGCCCGCTGAGGGGCTCAGTCCTC 000002644

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000002614 ------------------------------**GCCACGGGCCGCCCTTTCCTGGCAGGACAG** 000002643

>>>>>>>>> |||| ||||||||||||||||||||||||| <<<<<<<<<

000002645 CTGCGGGGGAGCTGGCCTCCCCGCCCCCAC**GCCAGGGGCCGCCCTTTCCTGGCAGGACAG**  000002704

TEF1

TEF

GATA1

SRF +++++++++++++++++++++

AP2 ++++++++++++++++++++++++++++++

MYOD

MEF2

000002644 **CGGGATCTTGCAGCTGTCAGGGGAGGGGAGGCGGGGGCTGATGTCAGGAGGGATACAAAT** 000002703

>>>>>>>>> ||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||| <<<<<<<<<

000002705 **CGGGATCCTGCAGCTGTCAGGGGAGGGGCGGCGGGGGCTGATGTCAGGAGGGATACAAAT** 000002764

TEF1

TEF

GATA1 ++++++++++ ++++++++++ ++++++++++

SRF ++++++

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++ +++

MYOD ++++++++++++++++++

MEF2 ++++++++

000002704 **AGTGCCGACGGCTGGGGGCCCTGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCT** 000002763

>>>>>>>>> ||||| |||||| |||||||||||||||||||||| ||||||||||||||||||||||| <<<<<<<<<

000002765 **AGTGCGGACGGCCGGGGGCCCTGTCTCCCCTCGCCACATCCACTCTCCGGCCGGCCGCCC** 000002824

TEF1

TEF

GATA1 +++++++++

SRF +++++++++

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2 ++++++

000002764 **GCCCGCCGCCTCCTCCG**TGC-**GCCCGCCAGCCTCGCCCGCGCCGTCACCATGAGCCAGGC** 000002822

>>>>>>>>> ||| ||| ||| ||||| || || |||||||||||||||||| ||| ||||||||||| <<<<<<<<<

000002825 **GCC-GCCTCCTTCTCCG**CGCC**GC---CCAGCCTCGCCCGCGCCGCCACTATGAGCCAGGC** 000002880

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++ +++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002823 **CTACTCGTCCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGCCCCGGGCTT**  000002882

>>>>>>>>> |||||||||||||||||| ||||| ||||||||||||||||||||||||||||| |||| <<<<<<<<<

000002881 CTACTCGTCCAGCCAGCGGGTGTCGTCCTAC**CGCCGCACCTTCGGCGGGGCC**CCCAGCTT 000002940

TEF1

TEF

GATA1 +++++++++++++++++++++

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002883 **CCCACTCGGCTCCCCGCTGAGTTCGCCCGTGTTCCCGCGGGCGGGTTTCGGCTCTAAGGG** 000002942

>>>>>>>>> ||| ||||||||||||||||| ||||| ||||||||||| ||||| |||||| | ||||| <<<<<<<<<

000002941 **CCCGCTCGGCTCCCCGCTGAGCTCGCCGGTGTTCCCGCGCGCGGGCTTCGGCACCAAGGG** 000003000

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD

MEF2

000002943 **CTCCTCCAGCTCGGTGACGTCCCGCGTGTACCAGGTGTCGCGCACGTCGGGCGGGGCCGG** 000003002

>>>>>>>>> |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<<<

000003001 **CTCCTCGAGCTCGGTGACGTCCCGCGTGTACCAGGTGTCGCGCACGTCGGGCGGGGCC--** 000003058

TEF1

TEF

GATA1

SRF

AP2 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

MYOD ++++++++++++++++++

MEF2

000003003 **G**GGCCTGGGGTCGCTGCGGGCCAGCCGGCTGGGGACCACCCGCACGCCCTCCTCCTACGG 000003062

>>>>>>>>> <<<<<<<<<

000000000 **-**----------------------------------------------------------- 000000000

TEF1

TEF

GATA1

SRF

AP2 +

MYOD

MEF2

000003063 CGCAGGCGAGCTGCTGGACTTCTCACTGGCCGACGCGGTGAACCAGGAGTTTCTGACCAC 000003122

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003123 GCGCACCAACGAGAAGGTGGAGCTGCAGGAGCTCAATGACCGCTTCGCCAACTACATCGA 000003182

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003183 GAAGGTGCGCTTCCTGGAGCAGCAGAACGCGGCGCTCGCCGCCGAAGTGAACCGGCTCAA 000003242

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003243 GGGCCGCGAGCCGACGCGAGTGGCCGAGCTCTACGAGGAGGAGCTGCGGGAGCTGCGGCG 000003302

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

000003303 CCAGGTGGAGGTGCTCACTAACCAGCGCGCGCGCGTCGACGTCGAGCG 000003350

>>>>>>>>> <<<<<<<<<

000000000 ------------------------------------------------ 000000000

TEF1

TEF

GATA1

SRF

AP2

MYOD

MEF2

Результаты расчетов, на сколько нуклеотидов приходится один сайт:

1) Поделите среднюю длину одной пары последовательностей (то есть той пары, которую

Вы выравнивали) на суммарное число всех найденных сайтов. Это и будет число

нуклеотидов, на которое в среднем приходится один сайт.

Human-Rat: 1900/29=66 нуклеотидов

Human-Cow:3050/110=27нуклеотидов

2) Вы искали сайты для 8 мышечно-специфичных факторов. А теперь представьте себе, что

Вам необходимо найти сайты для всех 455 факторов, имеющихся в арсенале программы

rVISTA. На какое число нуклеотидов в среднем тогда приходился бы один сайт? Чтобы

узнать это, разделите полученное значение на 56

Human-Rat : 66/56=1.12

Human-Cow:27/56=0.48