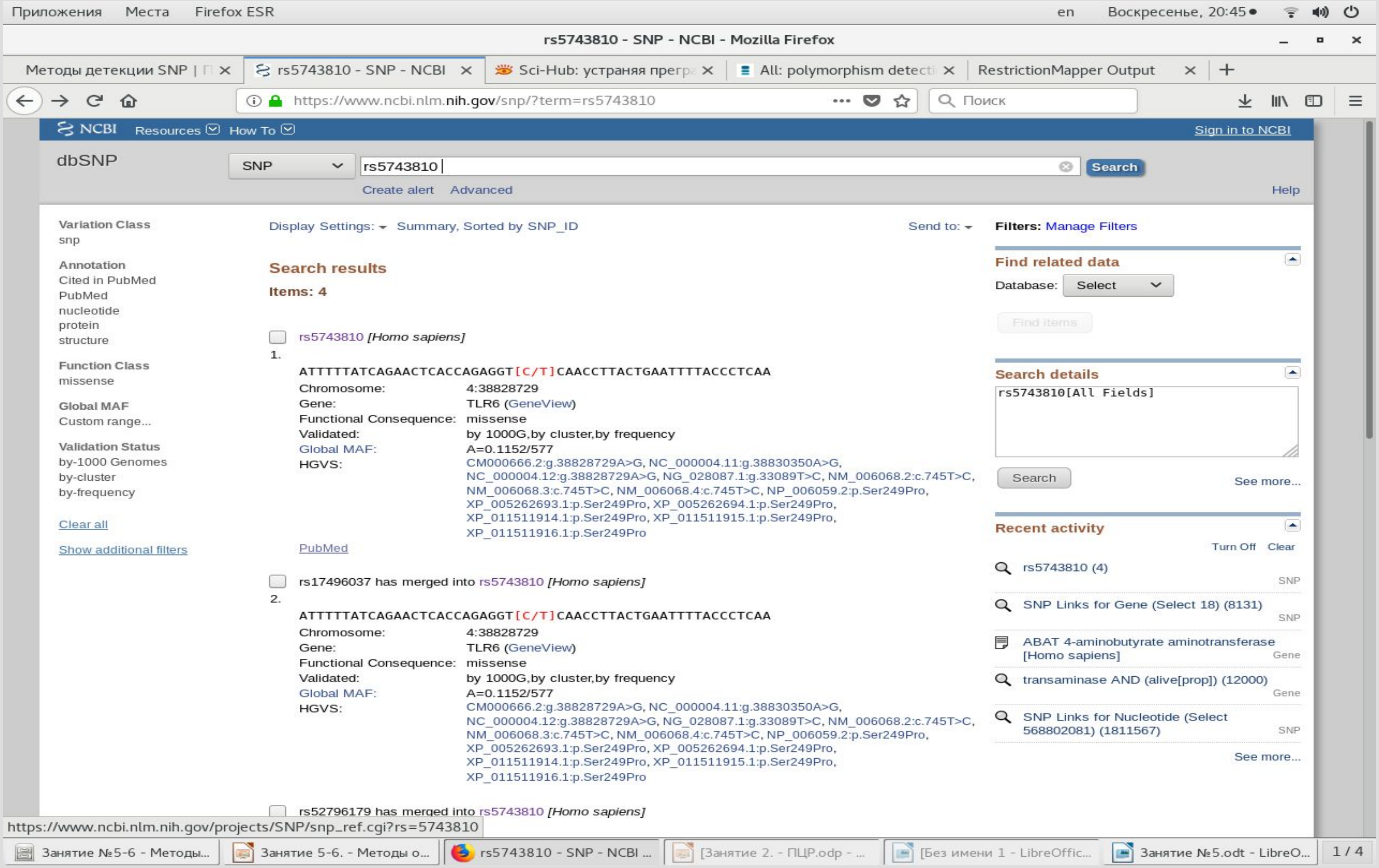
Домашнее задание №2: подбор условий для обнаружения SNP методом RFLP

1. Найти по rs номеру полиморфизм в базе данных dbSNP NCBI.



rs5743810

Variation Class
snp

Annotation
Cited in PubMed
PubMed
nucleotide
protein
structure

Function Class
missense

Global MAF
Custom range...

Validation Status
by-1000 Genomes
by-cluster
by-frequency

Clear all
Show additional filters

Search results

Items: 4

☐ 1. rs5743810 *[Homo sapiens]*

ATTTTTATCAGAACTCACCAGAGGT[C/T]CAACCTTACTGAATTTTACCCTCAA

Chromosome:	4:38828729
Gene:	TLR6 (GeneView)
Functional Consequence:	missense
Validated:	by 1000G,by cluster,by frequency
Global MAF:	A=0.1152/577
HGVS:	CM000666.2:g.38828729A>G, NC_000004.11:g.38830350A>G, NC_000004.12:g.38828729A>G, NG_028087.1:g.33089T>C, NM_006068.2:c.745T>C, NM_006068.3:c.745T>C, NM_006068.4:c.745T>C, NP_006059.2:p.Ser249Pro, XP_005262693.1:p.Ser249Pro, XP_005262694.1:p.Ser249Pro, XP_011511914.1:p.Ser249Pro, XP_011511915.1:p.Ser249Pro, XP_011511916.1:p.Ser249Pro

PubMed

☐ 2. rs17496037 has merged into rs5743810 *[Homo sapiens]*

ATTTTTATCAGAACTCACCAGAGGT[C/T]CAACCTTACTGAATTTTACCCTCAA

Chromosome:	4:38828729
Gene:	TLR6 (GeneView)
Functional Consequence:	missense
Validated:	by 1000G,by cluster,by frequency
Global MAF:	A=0.1152/577
HGVS:	CM000666.2:g.38828729A>G, NC_000004.11:g.38830350A>G, NC_000004.12:g.38828729A>G, NG_028087.1:g.33089T>C, NM_006068.2:c.745T>C, NM_006068.3:c.745T>C, NM_006068.4:c.745T>C, NP_006059.2:p.Ser249Pro, XP_005262693.1:p.Ser249Pro, XP_005262694.1:p.Ser249Pro, XP_011511914.1:p.Ser249Pro, XP_011511915.1:p.Ser249Pro, XP_011511916.1:p.Ser249Pro

☐ rs52796179 has merged into rs5743810 *[Homo sapiens]*

Search details:

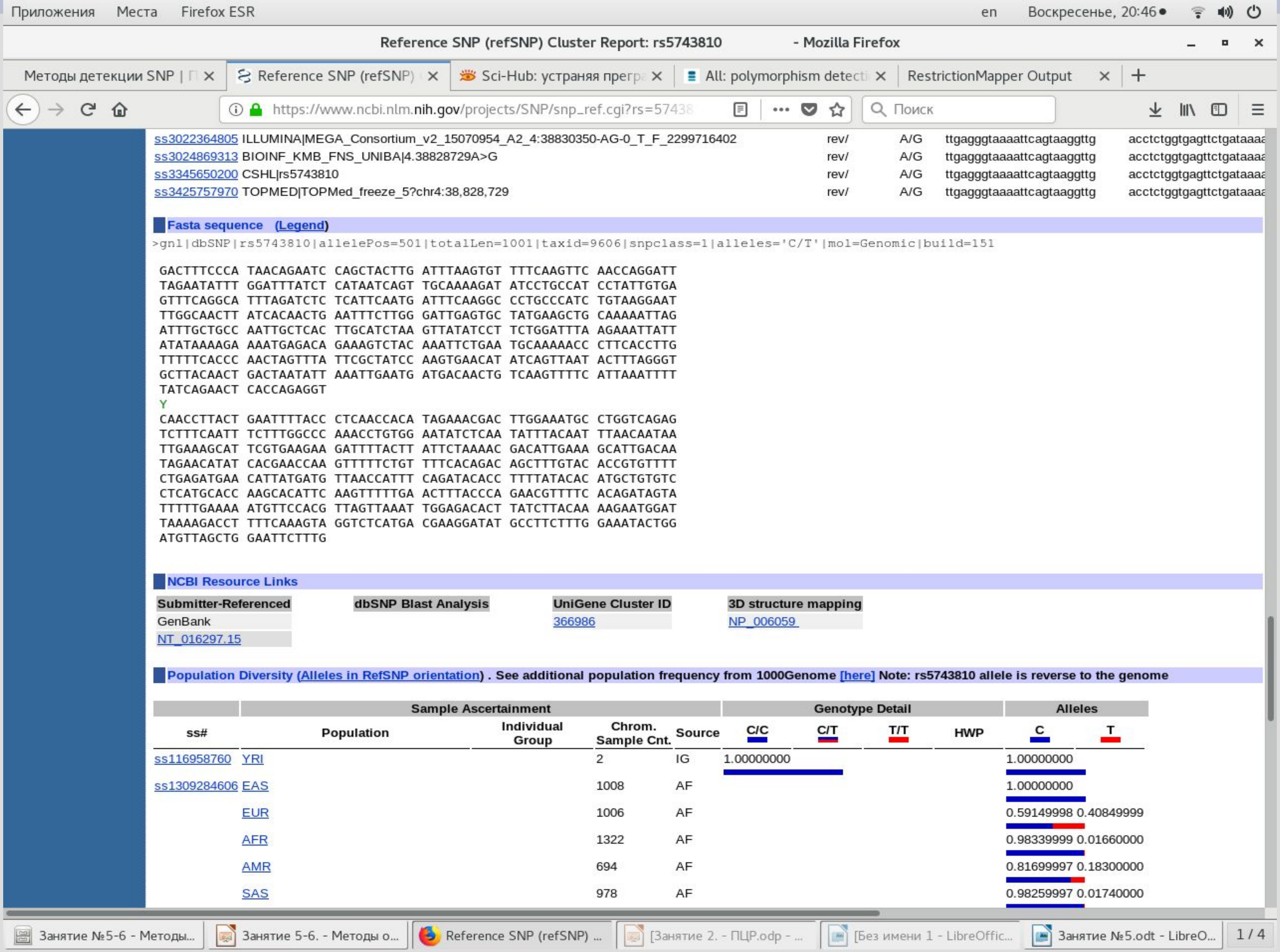
```
rs5743810[All Fields]
```

Recent activity:

- rs5743810 (4) — SNP
- SNP Links for Gene (Select 18) (8131) — SNP
- ABAT 4-aminobutyrate aminotransferase [Homo sapiens] — Gene
- transaminase AND (alive[prop]) (12000) — Gene
- SNP Links for Nucleotide (Select 568802081) (1811567) — SNP

https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=5743810

2. Скопировать последовательность около полиморфизма SNP±250 нуклеотидов (в примере SNP обозначен как Y, т.е. вариация пиримидинов C/T) в текстовый файл.



ss3022364805	ILLUMINA|MEGA_Consortium_v2_15070954_A2_4:38830350-AG-0_T_F_2299716402	rev/	A/G	ttgagggtaaaattcagtaaggttg	acctctggtgagttctgataaa
ss3024869313	BIOINF_KMB_FNS_UNIBA|4.38828729A>G	rev/	A/G	ttgagggtaaaattcagtaaggttg	acctctggtgagttctgataaa
ss3345650200	CSHL|rs5743810	rev/	A/G	ttgagggtaaaattcagtaaggttg	acctctggtgagttctgataaa
ss3425757970	TOPMED|TOPMed_freeze_5?chr4:38,828,729	rev/	A/G	ttgagggtaaaattcagtaaggttg	acctctggtgagttctgataaa

Fasta sequence (Legend)

```
>gnl|dbSNP|rs5743810|allelePos=501|totalLen=1001|taxid=9606|snpclass=1|alleles='C/T'|mol=Genomic|build=151

 GACTTTCCCA TAACAGAATC CAGCTACTTG ATTTAAGTGT TTTCAAGTTC AACCAGGATT
 TAGAATATTT GGATTTATCT CATAATCAGT TGCAAAAGAT ATCCTGCCAT CCTATTGTGA
 GTTTCAGGCA TTTAGATCTC TCATTCAATG ATTTCAAGGC CCTGCCCATC TGTAAGGAAT
 TTGGCAACTT ATCACAACTG AATTTCTTGG GATTGAGTGC TATGAAGCTG CAAAAATTAG
 ATTTGCTGCC AATTGCTCAC TTGCATCTAA GTTATATCCT TCTGGATTTA AGAAATTATT
 ATATAAAAGA AAATGAGACA GAAAGTCTAC AAATTCTGAA TGCAAAAACC CTTCACCTTG
 TTTTTCACCC AACTAGTTTA TTCGCTATCC AAGTGAACAT ATCAGTTAAT ACTTTAGGGT
 GCTTACAACT GACTAATATT AAATTGAATG ATGACAACTG TCAAGTTTTC ATTAAATTTT
 TATCAGAACT CACCAGAGGT
 Y
 CAACCTTACT GAATTTTACC CTCAACCACA TAGAAACGAC TTGGAAATGC CTGGTCAGAG
 TCTTTCAATT TCTTTGGCCC AAACCTGTGG AATATCTCAA TATTTACAAT TTAACAATAA
 TTGAAAGCAT TCGTGAAGAA GATTTTACTT ATTCTAAAAC GACATTGAAA GCATTGACAA
 TAGAACATAT CACGAACCAA GTTTTTCTGT TTTCACAGAC AGCTTTGTAC ACCGTGTTTT
 CTGAGATGAA CATTATGATG TTAACCATTT CAGATACACC TTTTATACAC ATGCTGTGTC
 CTCATGCACC AAGCACATTC AAGTTTTTGA ACTTTACCCA GAACGTTTTC ACAGATAGTA
 TTTTTGAAAA ATGTTCCACG TTAGTTAAAT TGGAGACACT TATCTTACAA AAGAATGGAT
 TAAAAGACCT TTTCAAAGTA GGTCTCATGA CGAAGGATAT GCCTTCTTTG GAAATACTGG
 ATGTTAGCTG GAATTCTTTG
```

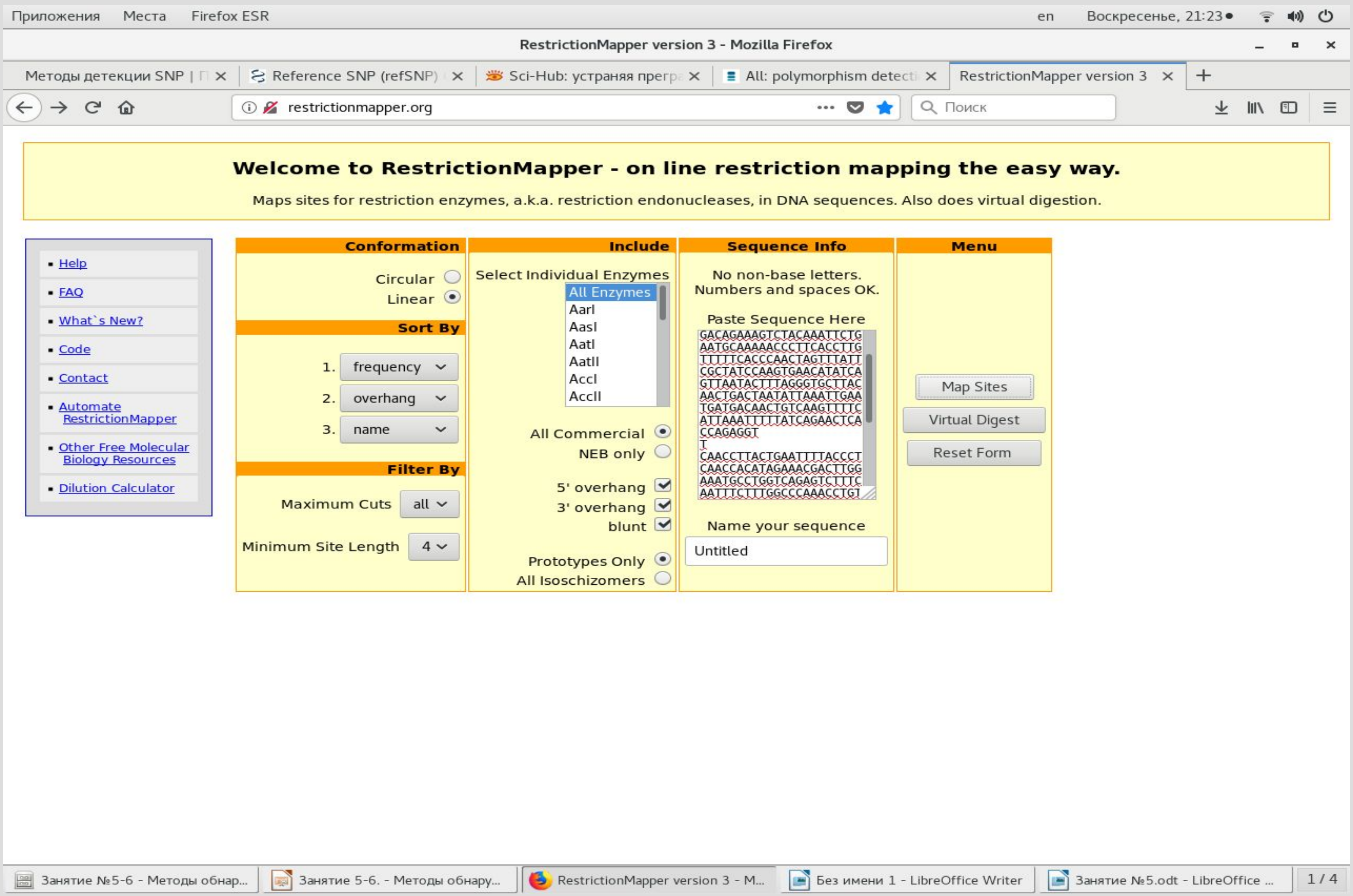
NCBI Resource Links

Submitter-Referenced	dbSNP Blast Analysis	UniGene Cluster ID	3D structure mapping
GenBank		366986	NP_006059
NT_016297.15			

Population Diversity (Alleles in RefSNP orientation) . See additional population frequency from 1000Genome [here] Note: rs5743810 allele is reverse to the genome

	Sample Ascertainment				Genotype Detail				Alleles	
ss#	Population	Individual Group	Chrom. Sample Cnt.	Source	C/C	C/T	T/T	HWP	C	T
ss116958760	YRI		2	IG	1.00000000				1.00000000	
ss1309284606	EAS		1008	AF					1.00000000	
	EUR		1006	AF					0.59149998	0.40849999
	AFR		1322	AF					0.98339999	0.01660000
	AMR		694	AF					0.81699997	0.18300000
	SAS		978	AF					0.98259997	0.01740000

3. На сайте проекта RestrictionMapper вставить в окно формы последовательность SNP±250 нуклеотидов и нажать Map Sites для поиска сайтов рестрикции (на месте SNP указать какой-то его вариант)



Welcome to RestrictionMapper - on line restriction mapping the easy way.

Maps sites for restriction enzymes, a.k.a. restriction endonucleases, in DNA sequences. Also does virtual digestion.

Conformation

Circular ○
Linear ●

Sort By

1. frequency
2. overhang
3. name

Filter By

Maximum Cuts: all
Minimum Site Length: 4

Include

Select Individual Enzymes: All Enzymes

All Commercial ●
NEB only ○

5' overhang ☑
3' overhang ☑
blunt ☑

Prototypes Only ●
All Isoschizomers ○

Sequence Info

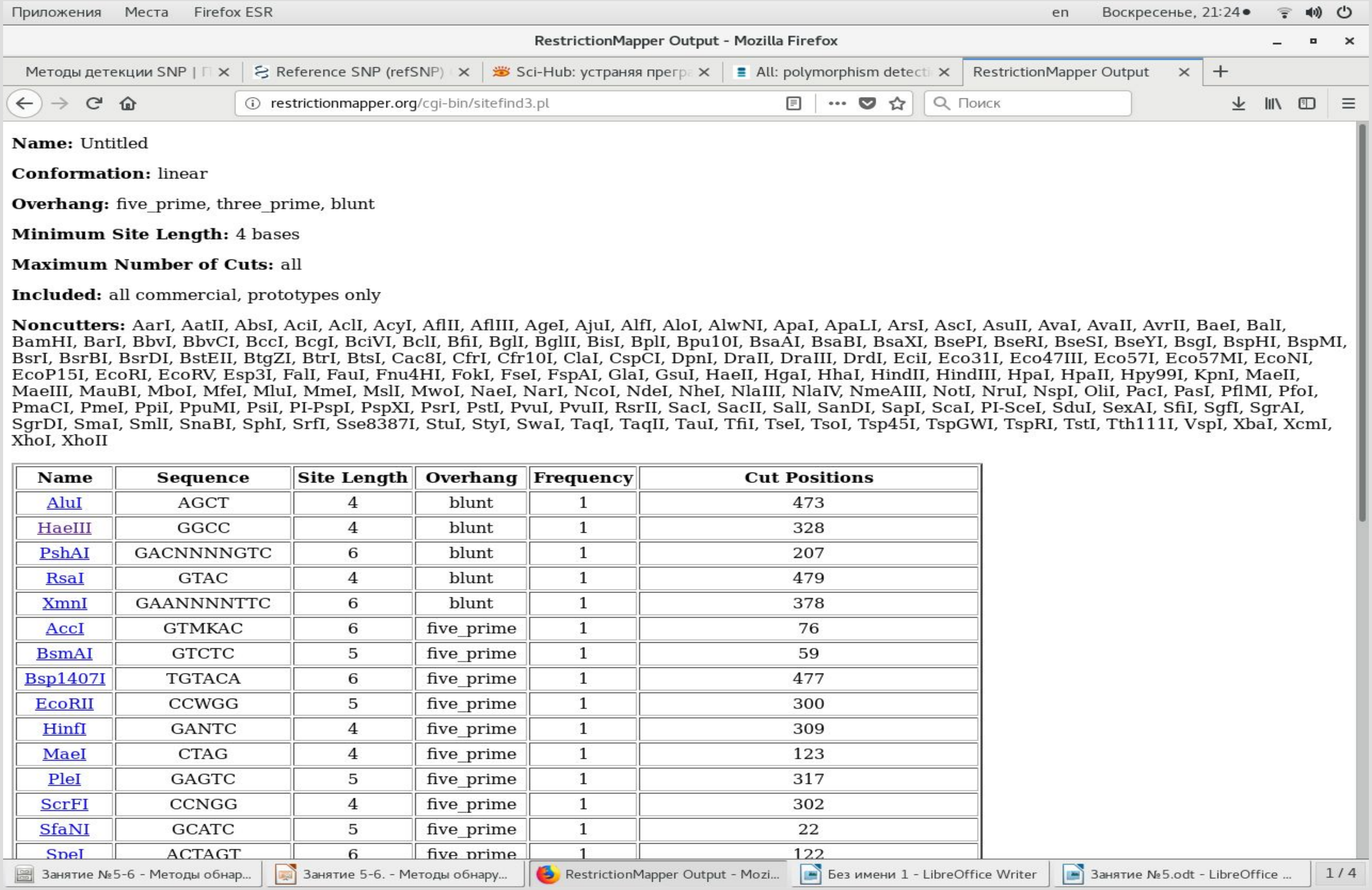
No non-base letters.
Numbers and spaces OK.

Paste Sequence Here

```
GACAGAAAGTCTACAAATTCTG
AATGCAAAAACCCTTCACCTTG
TTTTTCACCCAACTAGTTTATT
CGCTATCCAAGTGAACATATCA
GTTAATACTTTAGGGTGCTTAC
AACTGACTAATATTAAATTGAA
TGATGACAACTGTCAAGTTTTC
ATTAAATTTTTATCAGAACTCA
CCAGAGGT
T
CAACCTTACTGAATTTTACCCT
CAACCACATAGAAACGACTTGG
AAATGCCTGGTCAGAGTCTTTC
AATTTCTTTGGCCCAAACCTGT
```

Name your sequence: Untitled

4. В появившемся отчёте выбрать рестриктазы с единичным сайтом разрезания (Cut Position) недалеко от позиции SNP (251). Выполнить пп.3-4 с альтернативным вариантом SNP и выбрать рестриктазу, у которой сайт узнавания есть только для одного из вариантов.



Name: Untitled

Conformation: linear

Overhang: five_prime, three_prime, blunt

Minimum Site Length: 4 bases

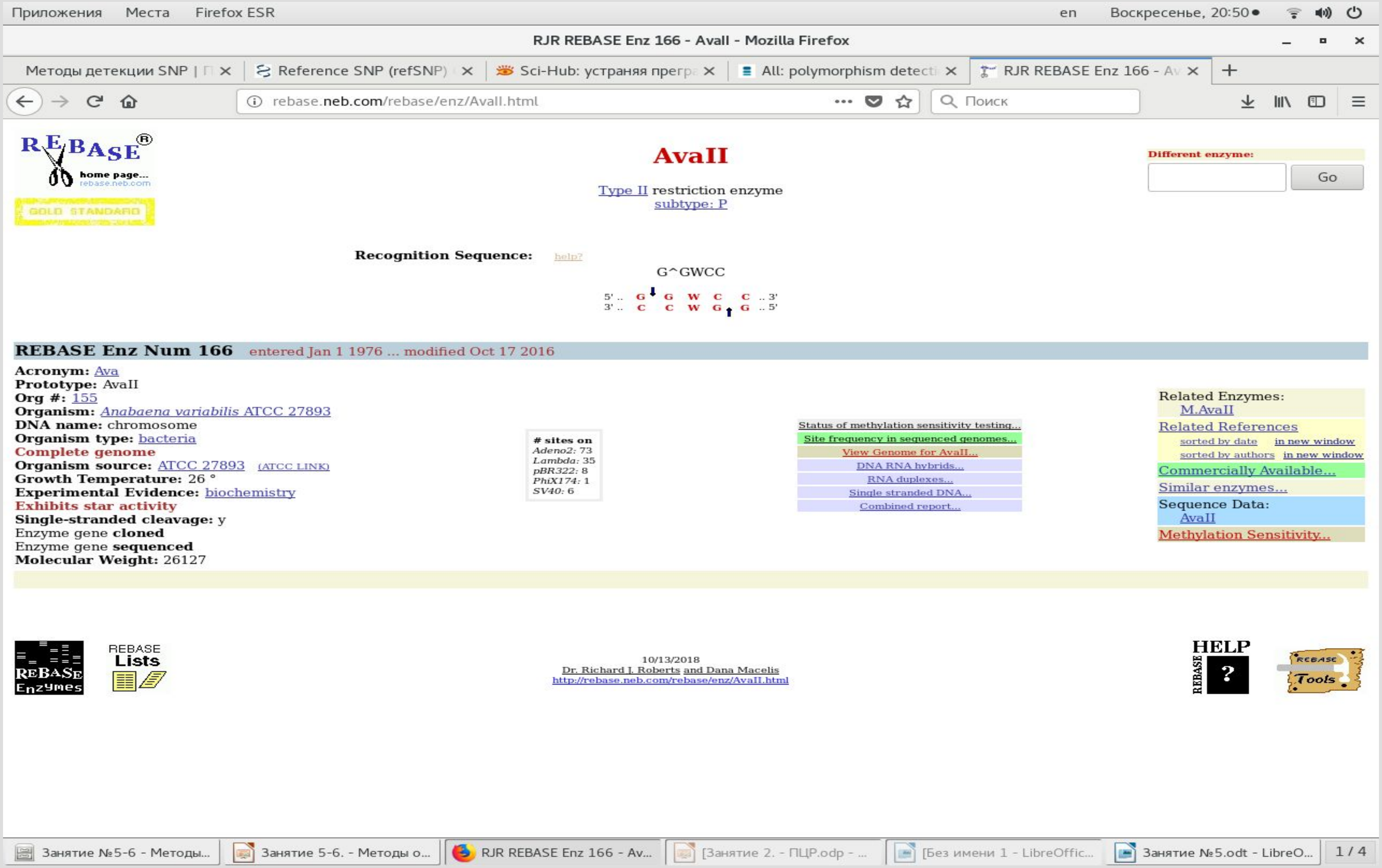
Maximum Number of Cuts: all

Included: all commercial, prototypes only

Noncutters: AarI, AatII, AbsI, AciI, AclI, AcyI, AflII, AflIII, AgeI, AjuI, AlfI, AloI, AlwNI, ApaI, ApaLI, ArsI, AscI, AsuII, AvaI, AvaII, AvrII, BaeI, BalI, BamHI, BarI, BbvI, BbvCI, BccI, BcgI, BciVI, BclI, BfiI, BglI, BglII, BisI, BplI, Bpu10I, BsaAI, BsaBI, BsaXI, BsePI, BseRI, BseSI, BseYI, BsgI, BspHI, BspMI, BsrI, BsrBI, BsrDI, BstEII, BtgZI, BtrI, BtsI, Cac8I, CfrI, Cfr10I, ClaI, CspCI, DpnI, DraII, DraIII, DrdI, EciI, Eco31I, Eco47III, Eco57I, Eco57MI, EcoNI, EcoP15I, EcoRI, EcoRV, Esp3I, FalI, FauI, Fnu4HI, FokI, FseI, FspAI, GlaI, GsuI, HaeII, HgaI, HhaI, HindII, HindIII, HpaI, HpaII, Hpy99I, KpnI, MaeII, MaeIII, MauBI, MboI, MfeI, MluI, MmeI, MslI, MwoI, NaeI, NarI, NcoI, NdeI, NheI, NlaIII, NlaIV, NmeAIII, NotI, NruI, NspI, OliI, PacI, PasI, PflMI, PfoI, PmaCI, PmeI, PpiI, PpuMI, PsiI, PI-PspI, PspXI, PsrI, PstI, PvuI, PvuII, RsrII, SacI, SacII, SalI, SanDI, SapI, ScaI, PI-SceI, SduI, SexAI, SfiI, SgfI, SgrAI, SgrDI, SmaI, SmlI, SnaBI, SphI, SrfI, Sse8387I, StuI, StyI, SwaI, TaqI, TaqII, TauI, TfiI, TseI, TsoI, Tsp45I, TspGWI, TspRI, TstI, Tth111I, VspI, XbaI, XcmI, XhoI, XhoII

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
AluI	AGCT	4	blunt	1	473
HaeIII	GGCC	4	blunt	1	328
PshAI	GACNNNNGTC	6	blunt	1	207
RsaI	GTAC	4	blunt	1	479
XmnI	GAANNNNTTC	6	blunt	1	378
AccI	GTMKAC	6	five_prime	1	76
BsmAI	GTCTC	5	five_prime	1	59
Bsp1407I	TGTACA	6	five_prime	1	477
EcoRII	CCWGG	5	five_prime	1	300
HinfI	GANTC	4	five_prime	1	309
MaeI	CTAG	4	five_prime	1	123
PleI	GAGTC	5	five_prime	1	317
ScrFI	CCNGG	4	five_prime	1	302
SfaNI	GCATC	5	five_prime	1	22
SpeI	ACTAGT	6	five_prime	1	122

5. Записать в файл название рестриктазы и место разрезания для неё (информация доступна, если нажать на название рестриктазы на предыдущем этапе).



AvaII

Type II restriction enzyme
subtype: P

Different enzyme:

Recognition Sequence: help?

G^GWCC

```
5' ..  G ↓ G  W  C    C .. 3'
3' ..  C   C  W  G ↑  G .. 5'
```

REBASE Enz Num 166 entered Jan 1 1976 ... modified Oct 17 2016

Acronym: Ava
Prototype: AvaII
Org #: 155
Organism: *Anabaena variabilis* ATCC 27893
DNA name: chromosome
Organism type: bacteria
Complete genome
Organism source: ATCC 27893 (ATCC LINK)
Growth Temperature: 26 °
Experimental Evidence: biochemistry
Exhibits star activity
Single-stranded cleavage: y
Enzyme gene **cloned**
Enzyme gene **sequenced**
Molecular Weight: 26127

 # sites on
Adeno2: 73
Lambda: 35
pBR322: 8
PhiX174: 1
SV40: 6

Status of methylation sensitivity testing...
Site frequency in sequenced genomes...
View Genome for AvaII...
DNA RNA hybrids...
RNA duplexes...
Single stranded DNA...
Combined report...

Related Enzymes:
M.AvaII
Related References
sorted by date in new window
sorted by authors in new window
Commercially Available...
Similar enzymes...
Sequence Data:
AvaII
Methylation Sensitivity...

10/13/2018
Dr. Richard J. Roberts and Dana Macelis
http://rebase.neb.com/rebase/enz/AvaII.html

6. Подобрать к данной последовательности (501 нуклеотид) праймеры (см. Предыдущую работу) с таким расчётом, чтобы ампликон составил 300-400 нуклеотидов, а рестрикция давала два куска с соотношением 2:3 (приблизительно). Записать в файл последовательность ампликона, праймеры, их температуру плавления, температуру отжига для ПЦР (см. соответствующее занятие) и фрагменты после рестрикции

- SNP для ДЗ:
- 1). rs4833095
- 2). rs5744168
- 3). rs11096955
- 4). rs11096956
- 5). rs4129009