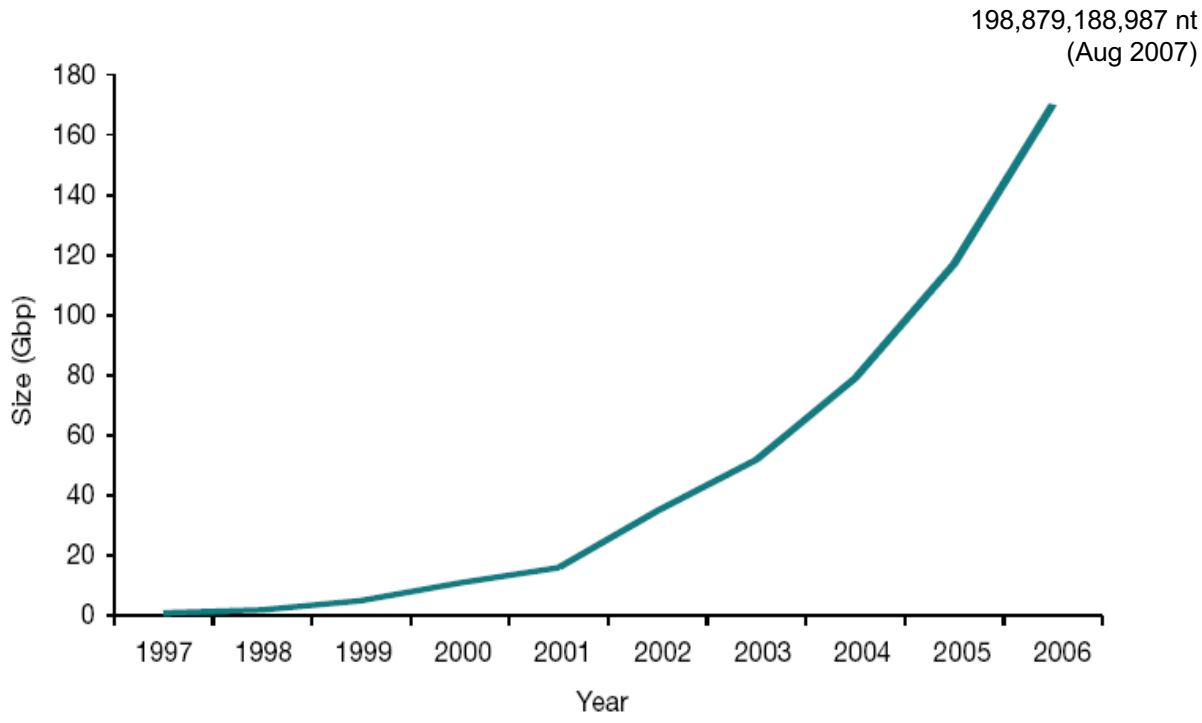
4. BASI DI BIOINFORMATICA PER LO STUDIO DELLA REGOLAZIONE GENICA

La sfida dei genome browsers

aumentare sempre di più le informazioni disponibili sulle sequenze



Genome Browsers Oggi

- Ensembl Genome browser
  http://www.ensembl.org
- NCBI Map Viewer
  http://www.ncbi.nlm.nih.gov/mapview/
- UCSC Genome Browser
  http://genome.ucsc.edu

Differenze tra Ensembl, UCSC ed NCBI?

	UCSC	Ensembl	NCBI
Presentation	Genome in horizontal orientation Main page contains a single graphic displaying annotation ('tracks') Clicking on annotation element presents web page of detailed information and links to other resources	Genome in horizontal orientation Main ContigView page contains three graphics displaying annotations at different resolutions Clicking on annotation element presents box with links to other resources or Views with more detailed information	Genome in vertical orientation Annotations graphically presented in columns ('maps') Clicking on annotation elements or links in columns provides quick access to other, primarily NCBI, resources
Content	13 vertebrate, 15 invertebrate Many cross-species annotations including conservation across eight species ENCODE Project annotations	13 vertebrate, six invertebrate Heavy focus on gene annotations such as Ensembl genes and VEGA HapMap project-related Views	11 vertebrate, five invertebrate, one protozoan, 12 plant, eight fungi Annotations primarily from NCBI resources
Functionality	Text search, BLAT sequence search, isPCR primer search Advanced annotation extraction using Table Browser Ability to upload and view own	Text search, BLAST and SSAHA sequence search, e-PCR primer search Advanced annotation extraction using BioMart Ability to upload and view own annotations	Text search, BLAST sequence search, e-PCR primer search Basic annotation extraction

Human genome assemblies

- GRCh38 (aka hg38)
  - No gaps
  - www.ensembl.org
  - Most up-to-date and supported

- GRCh37 (aka hg19)
  - 250 gaps
  - grch37.ensembl.org
  - Limited data and software updates

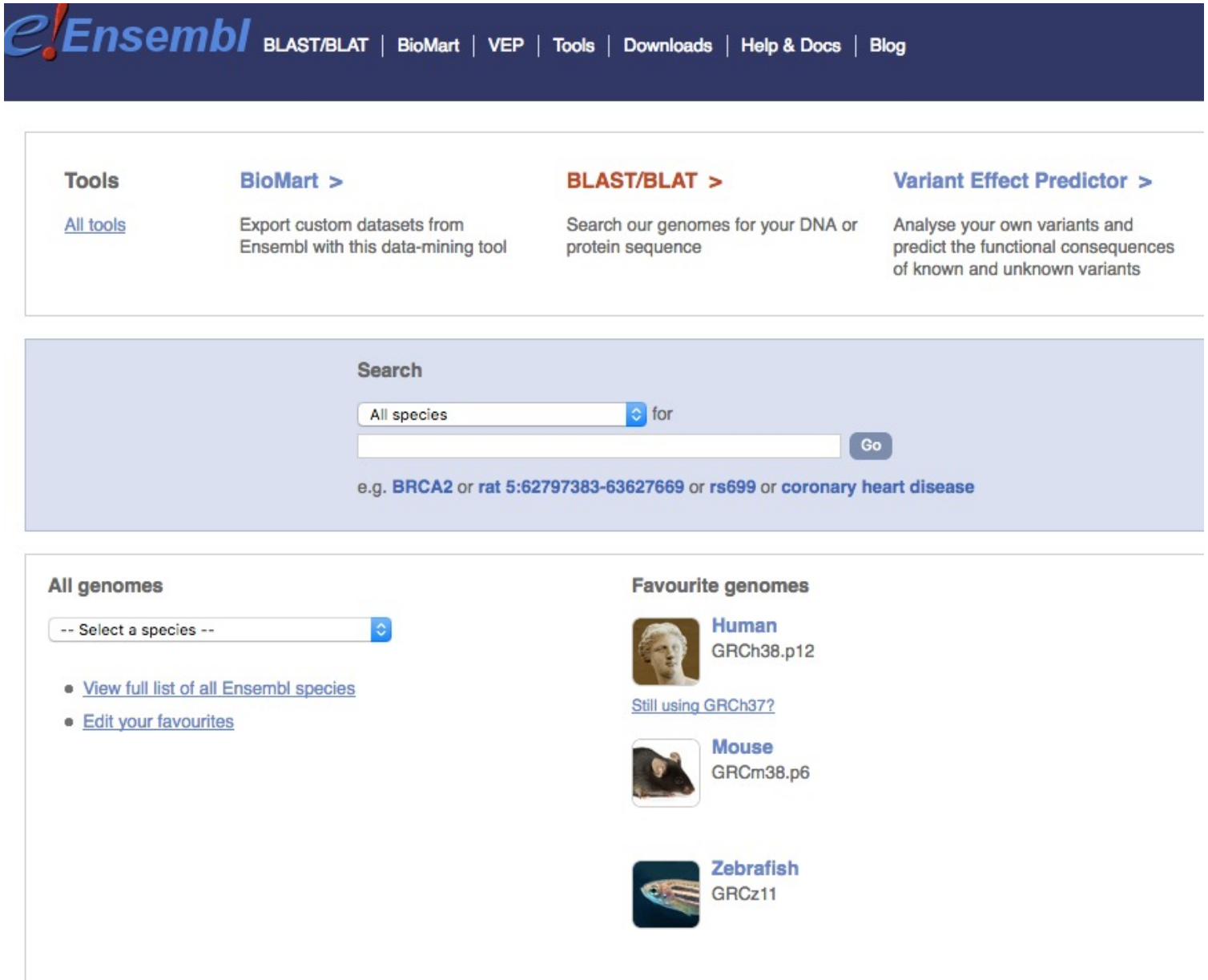
- NCBI36 (aka hg18)
  - 150,000 gaps
  - ncbi36.ensembl.org
  - No longer updated

Cosa impariamo da queste annotazioni?

- **All'interno di un genoma: elementi regolatori, ordine dei geni, struttura della cromatina…….**

- **Facendo studi comparativi: evoluzione, regioni conservate, riarrangiamenti…predizione di geni.**

Studiare i genomi con Ensembl



Nomenclatura in Ensembl

- **ENSG###** — **Ensembl Gene ID**
- **ENST###** — **Ensembl Transcript ID**
- **ENSP###** — **Ensembl Peptide ID**
- **ENSE###** — **Ensembl Exon ID**

- **Per specie diverse dall'uomo è aggiunto un suffisso**

MUS (*Mus musculus*) for mouse: ENSMUSG###
DAR (*Danio rerio*) for zebrafish: ENSDARG###, etc.

Quali annotazioni sono disponibili?

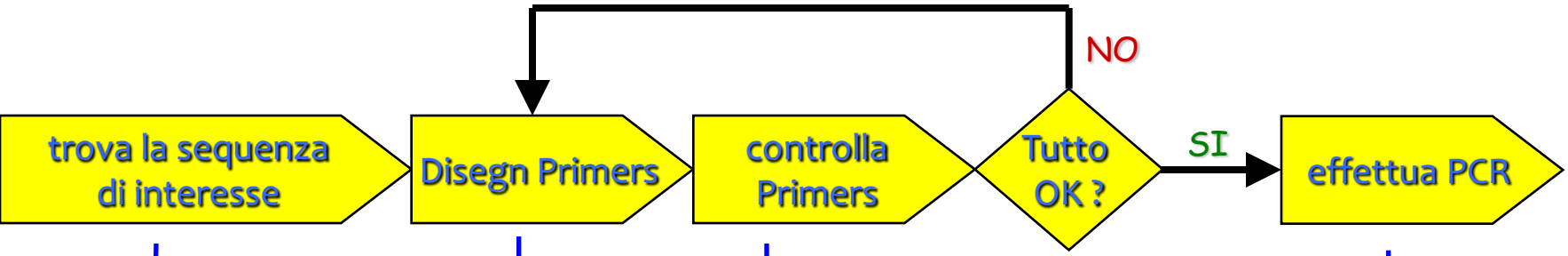
- Gene/transcript/peptide models (coding and noncoding (ncRNAs))
- IDs in other database
- Mapped cDNAs, peptides, micro array probes, BAC clones etc.
- Cytogenetic bands, markers, repeats etc.
- Comparative data:

  orthologues and paralogues, protein families, whole genome alignments, syntenic regions
- Variation data:

  Single Nucleotide Polymorphisms (SNPs)
- Regulatory data:

  “best guess” set of regulatory elements from ENCODE
- Data from external sources (DAS)

Primer design-come procedere



NCBI o Ensembl

Primer3
…o software simile

Per i parametri desiderati
- Tm
- Grandezza amplificato
- strutture secondarie
- complementarietà
- specificità

Netprimer, BLAST o software simili

…e gel di elettroforesi

disegno dei primers

Un buon disegno dei primers è la chiave di una PCR di successo !!

Proprietà importanti di una buona coppia di primers

Ogni primers dovrà avere

- lunghezza basi compresa tra 18-24
- 40-60% G/C
- Distribuzione bilanciata di basi G/C e A/T
- T_m che permette un annealing tra 55-65° C
- NO strutture secondarie interne (hair-pins)

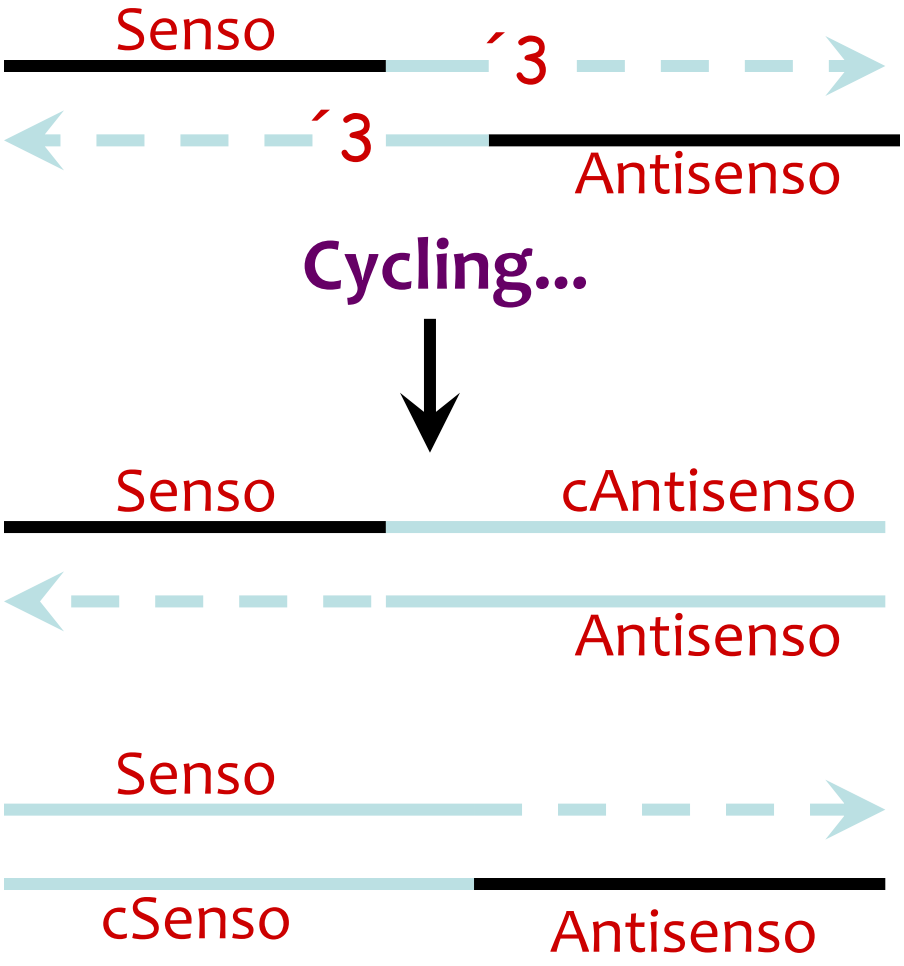
Le coppie di primers inoltre dovrebbero avere:

- T_m *simile* (*max* 2-3 ° C di differenza)
- **NO** complementarietà (> 2-3 bp) in particolare al 3'

Il problema dei dimeri di primers

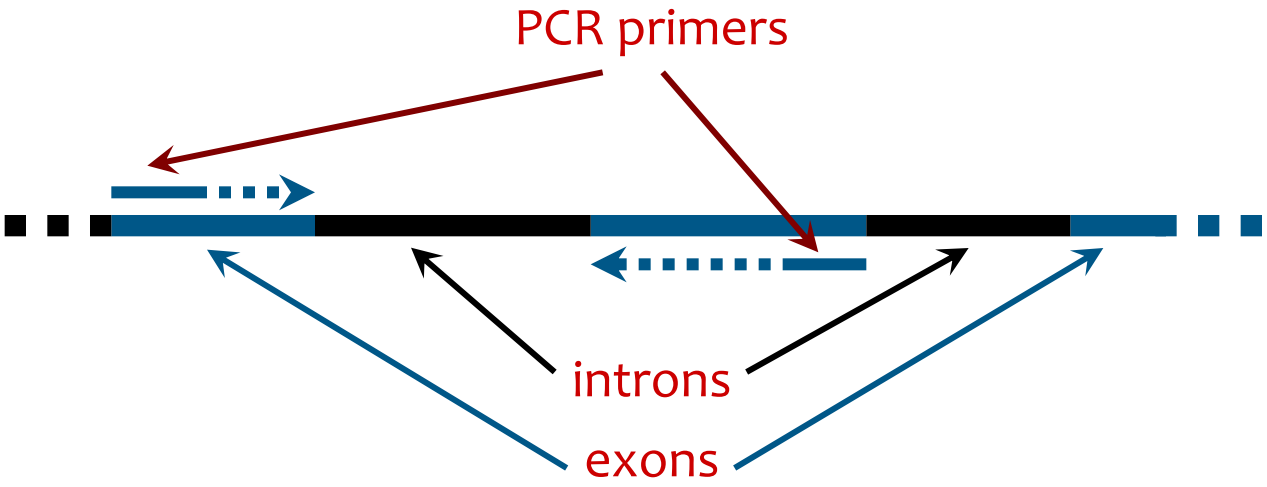
- Primers che interagiscono tra di loro sono AMPLIFICATI dalla PCR

- La formazione dei dimeri di primers compete con la PCR e può quindi compromettere l'efficienza della reazione.



Considerazioni importanti

- Evitare di avere come bersaglio della qPCR delle **strutture secondarie**
- Evitare contaminazione genomica disegnando primers che **includono esoni diversi** e tagliano introni



Links per disegno primers

- http://www.tataa.com/
- **http://www.ncbi.nlm.nih.gov/BLAST/**
- **www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html**
- ***www.ensembl.org***
- **http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi**
- **http://www.bioinfo.rpi.edu/applications/mfold/dna/form1.cgi**
- **http://primer3.ut.ee/ Primer3**

BLAST

Basic Local Alignment Search Tool

BLAST Programs

Program	Database (Subject)	Query
BLASTN	Nucleotide	Nucleotide
BLASTP	Protein	Protein
BLASTX	Protein	Nt. → Protein
TBLASTN	Nt. → Protein	Protein
TBLASTX	Nt. → Protein	Nt. → Protein

BLAST confronta le sequenze

- **usa** una sequenza *"query"*
- **La confronta con** milioni di sequenze nei database *GenBank* ® costruendo local alignments
- **elenca** quelle che sembrano simili alla query
- **dice perchè sono eventualmente omologhe**

NB

- ***BLAST suggerisce***
- ***Il ricercatore trae le sue conclusioni***

BLAST output

descrive in che modo le sequenze allineate sono simili

- *Quanto sono lunghi i segmenti allineati?*
- *BLAST ha dovuto introdurre degli spazi per allineare i segmenti?*
- *Quanto sono simili i segmenti allineati?*

Graphic Display

1. How good is the match?

- **Red = excellent!**
- Pink = pretty good
- **Green = OK, but look at other factors**
- **Blue = bad**
- **Black = really bad!**

2. How long are the matched segments?

Longer = better

BLAST fa una lista dei migliori accoppiamenti (hits)

Per ogni accoppiamento fornisce:

- **Accession number – links to Genbank flatfile**
- **Description**
- **"G" = genome link**
- **E-value**
- **Score**
  - **Link to an alignment**

Sequences producing significant alignments:

Select: All None Selected:0

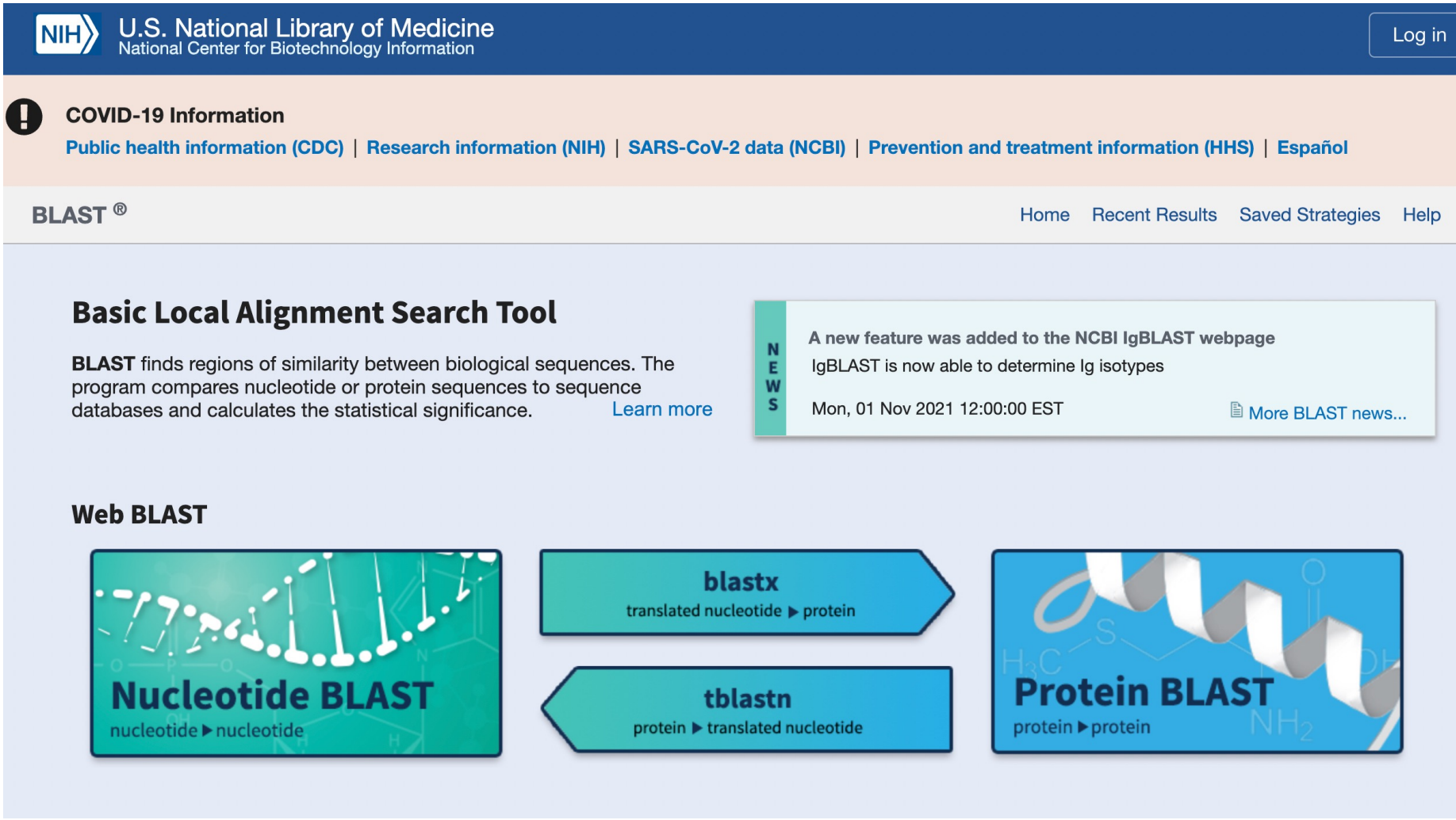
Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X8, mRNA	1585	1585	100%	0.0	100%	XM_011529250.2
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X7, mRNA	1585	1585	100%	0.0	100%	XM_011529249.2
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X6, mRNA	1585	1585	100%	0.0	100%	XM_017027878.1
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X4, mRNA	1585	1585	100%	0.0	100%	XM_011529246.2
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X5, mRNA	1585	1585	100%	0.0	100%	XM_011529248.1
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X3, mRNA	1585	1585	100%	0.0	100%	XM_011529247.1
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X2, mRNA	1585	1585	100%	0.0	100%	XM_011529245.1
PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X1, mRNA	1585	1585	100%	0.0	100%	XM_011529244.1
Homo sapiens preproenkephalin (enkB) gene, partial cds	1585	1585	100%	0.0	100%	AH002816.3
Homo sapiens prodynorphin (PDYN), transcript variant 1, mRNA	1585	1585	100%	0.0	100%	NM_024411.4
Homo sapiens prodynorphin (PDYN), transcript variant 4, mRNA	1585	1585	100%	0.0	100%	NM_001190899.2
Homo sapiens prodynorphin (PDYN), transcript variant 2, mRNA	1585	1585	100%	0.0	100%	NM_001190898.2

Che cos'è l'E-value?

- **E-value**
  - **La probabilità che quel determinato accoppiamento non è casuale**
  - **+ basso è l'E-value, + significativo è l'accoppiamento**
    - **E = 10^{-4}** è considerato il **cutoff point**
    - **E = 0** significa che le due sequenze sono statisticamente **identiche**

ESERCIZI:
http://teaching.bioinformatics.dtu.dk/teaching/index.php/Exercise:_BLAST



https://blast.ncbi.nlm.nih.gov/Blast.cgi

Software to predict CpG islands?

MethPrimer is an online platform which provides a number of tools and databases to facilitate the study of DNA methylation and epigenetics, including tools for designing prim

http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi

Software to predict miRNA binding ??

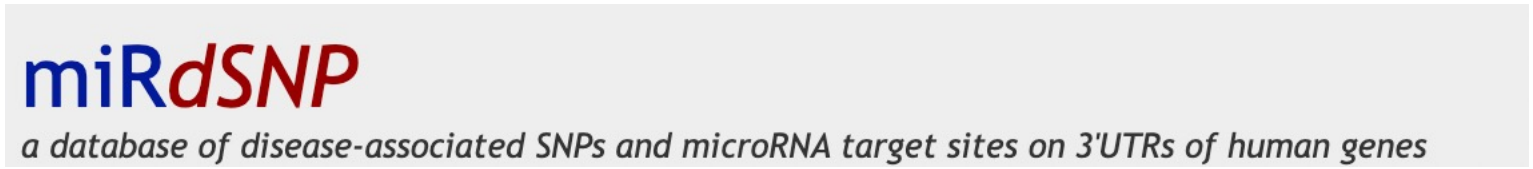
http://mirdb.org/index.html

growing evidence suggest the importance of ***both environmental and genetic factors*** *in the influence of* ***DNA methylation.***

DNA methylation can be influenced by *cis*-acting DNA sequence variation located on the same chromosome.

- *Mill et al., Am J Hum Genet 2008*
- *Zhang et al., Am J Hum Genet 2010*
- *Milani et al., Genome Res 2009*
- *Docherty et al., Behav and Brain Func 2012*
- *Ball et al., Genome Biol. 2011*

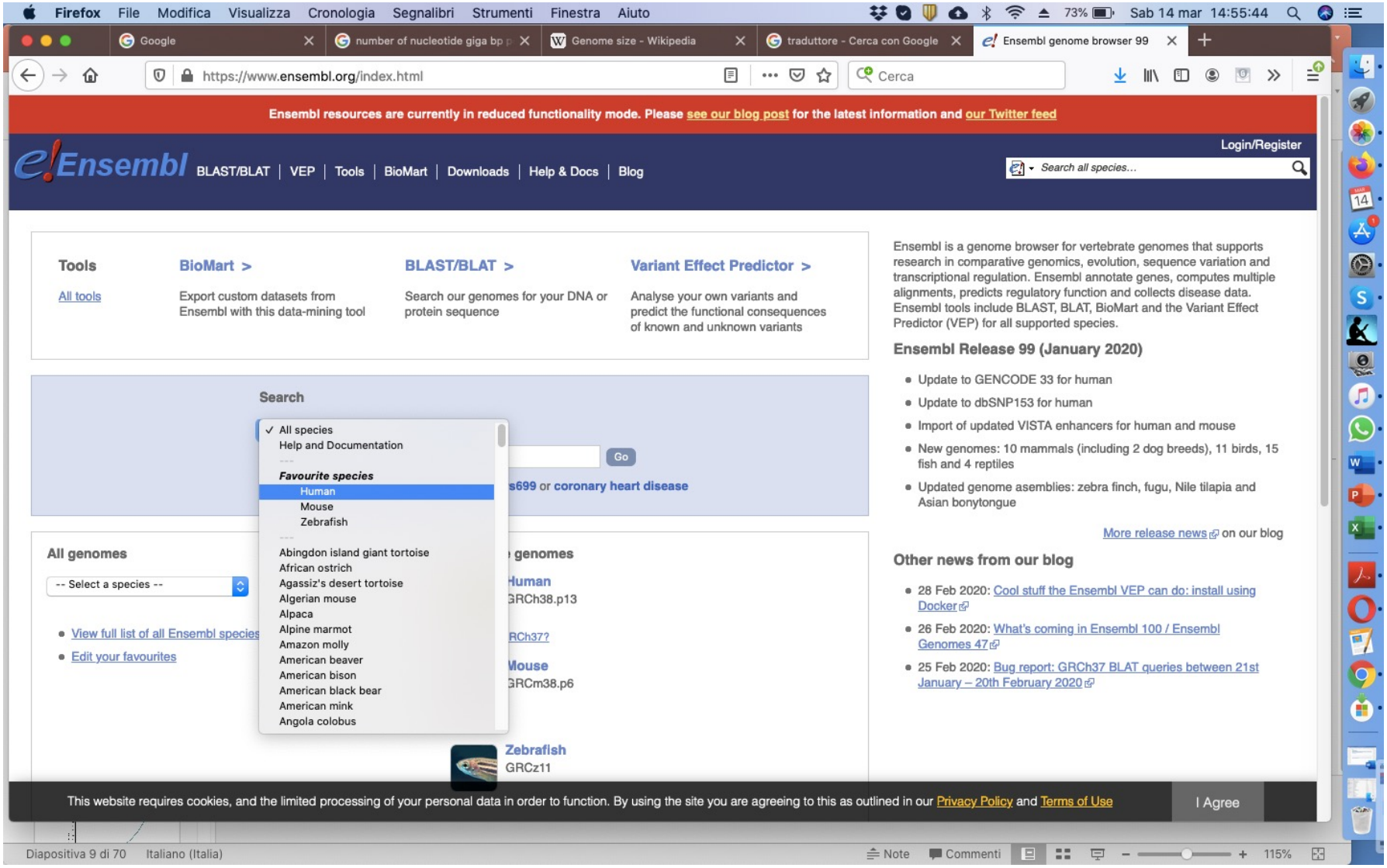
new models have to be developed to integrate genetic variants and DNA methylation.



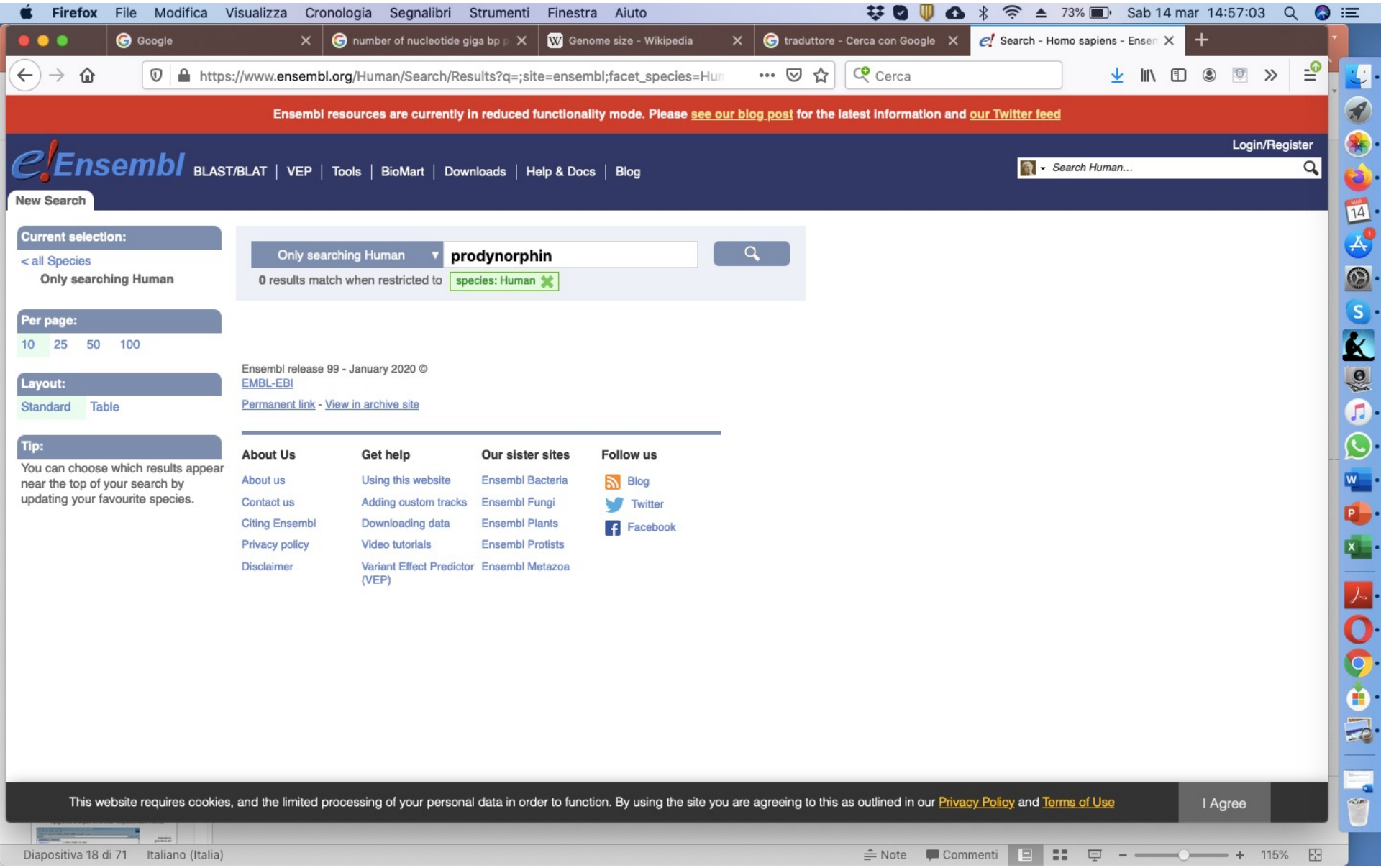
http://mirdsnp.ccr.buffalo.edu/

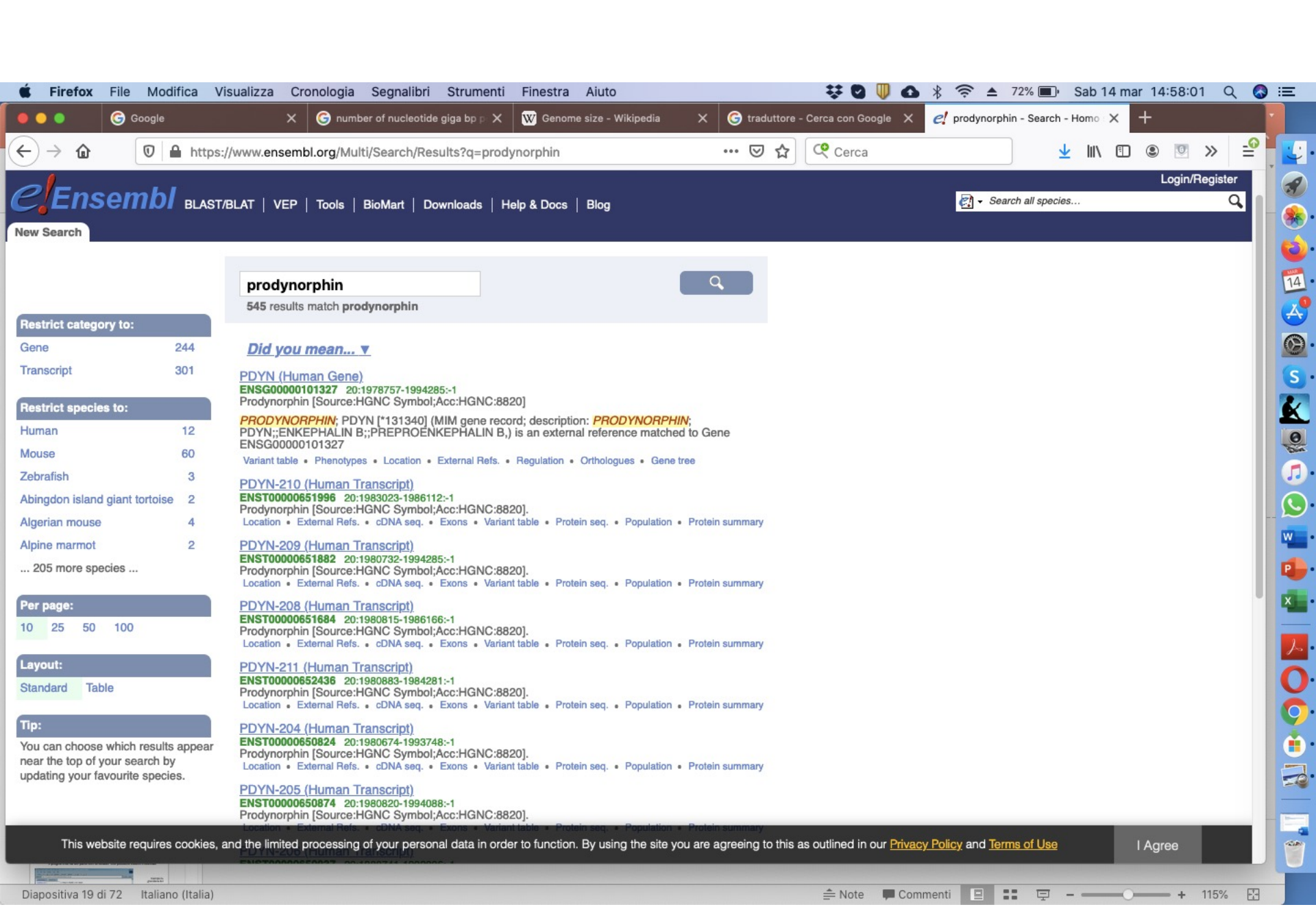
…….START

SELECT YOUR SPECIES



SELECT YOUR GENE




Firefox File Modifica Visualizza Cronologia Segnalibri Strumenti Finestra Aiuto
Sab 14 mar 14:58:01
Google
number of nucleotide giga bp p
Genome size - Wikipedia
traduttore - Cerca con Google
prodynorphin - Search - Homo
https://www.ensembl.org/Multi/Search/Results?q=prodynorphin
Cerca
Ensembl
BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog
Login/Register
Search all species...
New Search
prodynorphin
545 results match prodynorphin
Restrict category to:
Gene 244
Transcript 301
Restrict species to:
Human 12
Mouse 60
Zebrafish 3
Abingdon island giant tortoise 2
Algerian mouse 4
Alpine marmot 2
... 205 more species ...
Per page:
10 25 50 100
Layout:
Standard Table
Tip:
You can choose which results appear near the top of your search by updating your favourite species.
Did you mean... ▼
PDYN (Human Gene)
ENSG00000101327 20:1978757-1994285:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820]
PRODYNORPHIN; PDYN [*131340] (MIM gene record; description: PRODYNORPHIN; PDYN;;ENKEPHALIN B;;PREPROENKEPHALIN B,) is an external reference matched to Gene ENSG00000101327
Variant table • Phenotypes • Location • External Refs. • Regulation • Orthologues • Gene tree
PDYN-210 (Human Transcript)
ENST00000651996 20:1983023-1986112:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary
PDYN-209 (Human Transcript)
ENST00000651882 20:1980732-1994285:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary
PDYN-208 (Human Transcript)
ENST00000651684 20:1980815-1986166:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary
PDYN-211 (Human Transcript)
ENST00000652436 20:1980883-1984281:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary
PDYN-204 (Human Transcript)
ENST00000650824 20:1980674-1993748:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary
PDYN-205 (Human Transcript)
ENST00000650874 20:1980820-1994088:-1
Prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820].
This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our Privacy Policy and Terms of Use
I Agree
Diapositiva 19 di 72 Italiano (Italia)
Note Commenti
115%

Gene: PDYN ENSG00000101327

Description prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820]

Gene Synonyms ADCA, PENKB, SCA23

Location Chromosome 20: 1,978,757-1,994,285 reverse strand.
GRCh38:CM000682.2

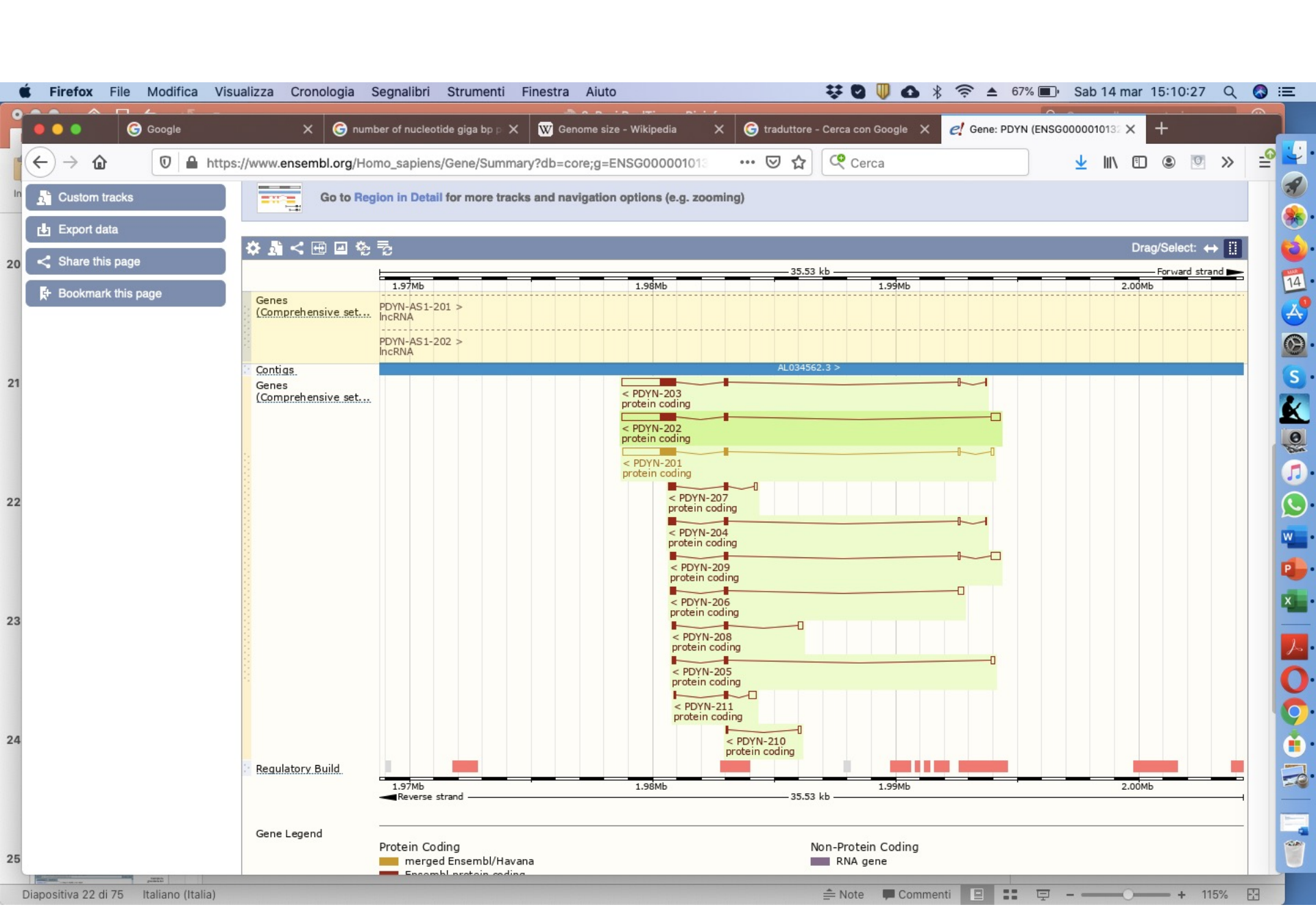
About this gene This gene has 11 transcripts (splice variants), 222 orthologues, 2 paralogues, is a member of 1 Ensembl protein family and is associated with 2 phenotypes

Transcripts Hide transcript table

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
PDYN-202	ENST00000539905.5	2725	254aa	Protein coding	CCDS13023	P01213	-	TSL:4 GENCODE basic APPRIS P1
PDYN-201	ENST00000217305.3	2556	254aa	Protein coding	CCDS13023	P01213	NM_024411.5	TSL:1 GENCODE basic APPRIS P1 MANE Select v0.7
PDYN-203	ENST00000540134.5	2472	254aa	Protein coding	CCDS13023	P01213	-	TSL:4 GENCODE basic APPRIS P1
PDYN-209	ENST00000651882.1	807	118aa	Protein coding	-	A0A494C0X3	-	CDS 3' incomplete
PDYN-206	ENST00000650937.1	589	115aa	Protein coding	-	A0A494C197	-	CDS 3' incomplete
PDYN-207	ENST00000651328.1	582	149aa	Protein coding	-	A0A494C1U2	-	CDS 3' incomplete
PDYN-204	ENST00000650824.1	555	138aa	Protein coding	-	A0A494C1J4	-	CDS 3' incomplete
PDYN-211	ENST00000652436.1	530	68aa	Protein coding	-	A0A494C1E3	-	CDS 3' incomplete
PDYN-205	ENST00000650874.1	465	89aa	Protein coding	-	A0A494C0B1	-	CDS 3' incomplete
PDYN-208	ENST00000651684.1	464	91aa	Protein coding	-	A0A494C0B3	-	CDS 3' incomplete
PDYN-210	ENST00000651996.1	196	20aa	Protein coding	-	A0A494C0Y8	-	CDS 3' incomplete

Summary


Custom tracks
Export data
Share this page
Bookmark this page
Go to Region in Detail for more tracks and navigation options (e.g. zooming)
Drag/Select:
35.53 kb
Forward strand
1.97Mb
1.98Mb
1.99Mb
2.00Mb
Genes (Comprehensive set...
PDYN-AS1-201 > lncRNA
PDYN-AS1-202 > lncRNA
Contigs
AL034562.3 >
Genes (Comprehensive set...
< PDYN-203 protein coding
< PDYN-202 protein coding
< PDYN-201 protein coding
< PDYN-207 protein coding
< PDYN-204 protein coding
< PDYN-209 protein coding
< PDYN-206 protein coding
< PDYN-208 protein coding
< PDYN-205 protein coding
< PDYN-211 protein coding
< PDYN-210 protein coding
Regulatory Build
Reverse strand
Gene Legend
Protein Coding
merged Ensembl/Havana
Non-Protein Coding
RNA gene

Ensembl resources are currently in reduced functionality mode. Please see our blog post for the latest information and our Twitter feed

Transcript: PDYN-202 ENST00000539905.5

Description prodynorphin [Source:HGNC Symbol;Acc:HGNC:8820]

Gene Synonyms ADCA, PENKB, SCA23

Location Chromosome 20: 1,978,757-1,994,285 reverse strand.

About this transcript This transcript has 3 exons, is annotated with 17 domains and features, is associated with 4279 variant alleles and maps to 290 oligo probes.

Gene This transcript is a product of gene ENSG00000101327.9 Hide transcript table

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
PDYN-202	ENST00000539905.5	2725	254aa	Protein coding	CCDS13023	P01213	-	TSL:4 GENCODE basic APPRIS P1
PDYN-201	ENST00000217305.3	2556	254aa	Protein coding	CCDS13023	P01213	NM_024411.5	TSL:1 GENCODE basic APPRIS P1 MANE Select v0.7
PDYN-203	ENST00000540134.5	2472	254aa	Protein coding	CCDS13023	P01213	-	TSL:4 GENCODE basic APPRIS P1
PDYN-209	ENST00000651882.1	807	118aa	Protein coding	-	A0A494C0X3	-	CDS 3' incomplete
PDYN-206	ENST00000650937.1	589	115aa	Protein coding	-	A0A494C197	-	CDS 3' incomplete
PDYN-207	ENST00000651328.1	582	149aa	Protein coding	-	A0A494C1U2	-	CDS 3' incomplete
PDYN-204	ENST00000650824.1	555	138aa	Protein coding	-	A0A494C1J4	-	CDS 3' incomplete
PDYN-211	ENST00000652436.1	530	68aa	Protein coding	-	A0A494C1E3	-	CDS 3' incomplete
PDYN-205	ENST00000650874.1	465	89aa	Protein coding	-	A0A494C0B1	-	CDS 3' incomplete
PDYN-208	ENST00000651684.1	464	91aa	Protein coding	-	A0A494C0B3	-	CDS 3' incomplete
PDYN-210	ENST00000651996.1	196	20aa	Protein coding	-	A0A494C0Y8	-	CDS 3' incomplete

Summary

Google | number of nucleotide giga bp | Genome size - Wikipedia | traduttore - Cerca con Google | Transcript: PDYN-202 (ENST00

https://www.ensembl.org/Homo_sapiens/Transcript/Exons?db=core;g=ENSG0000010

Cerca

- Population comparison
- Comparison image
- External References
- Supporting evidence
- ID History
  - Transcript history
  - Protein history

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

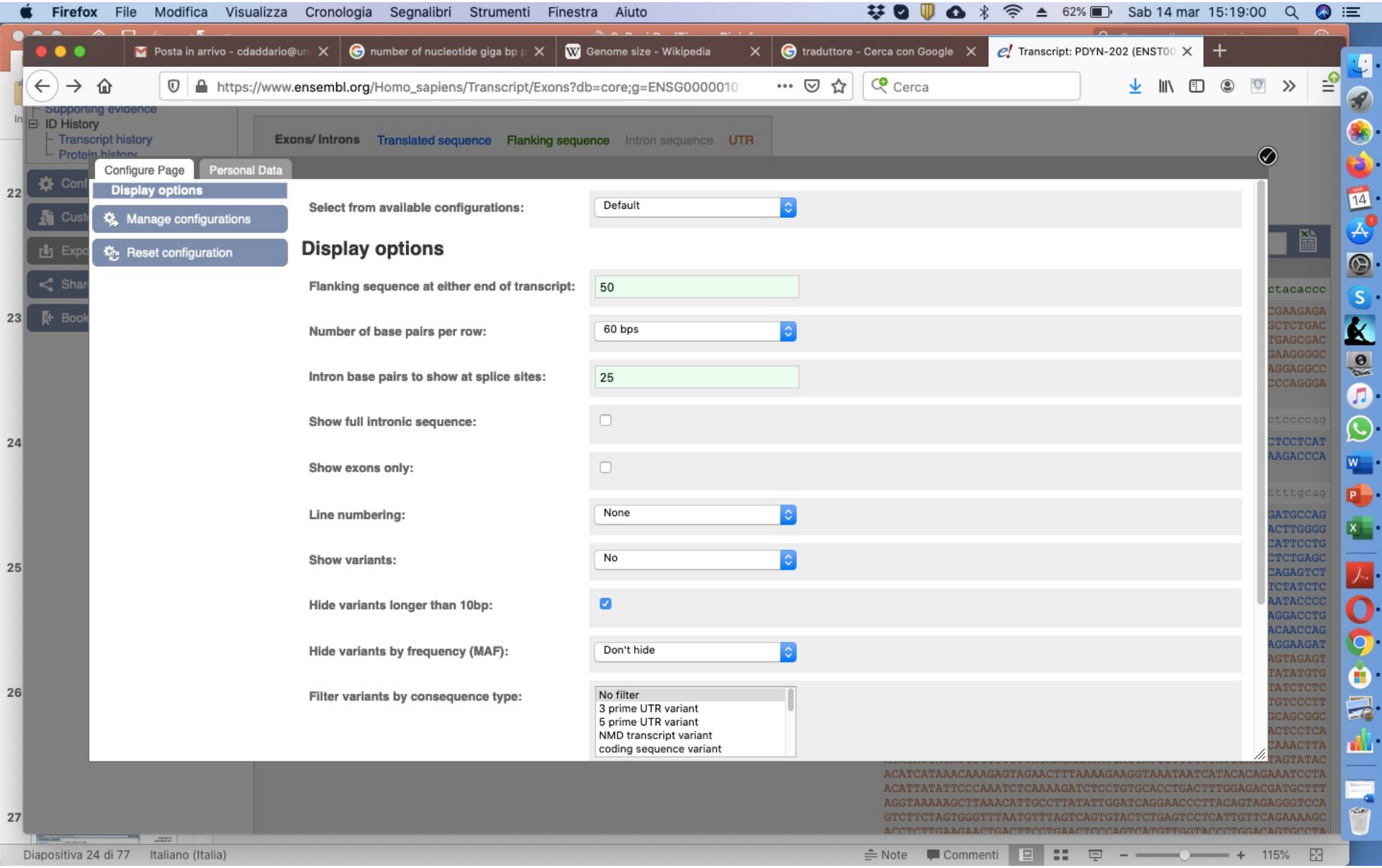
Download sequence

Exons/ Introns | Translated sequence | Flanking sequence | Intron sequence | UTR

Markup loaded

Show/hide columns | Filter

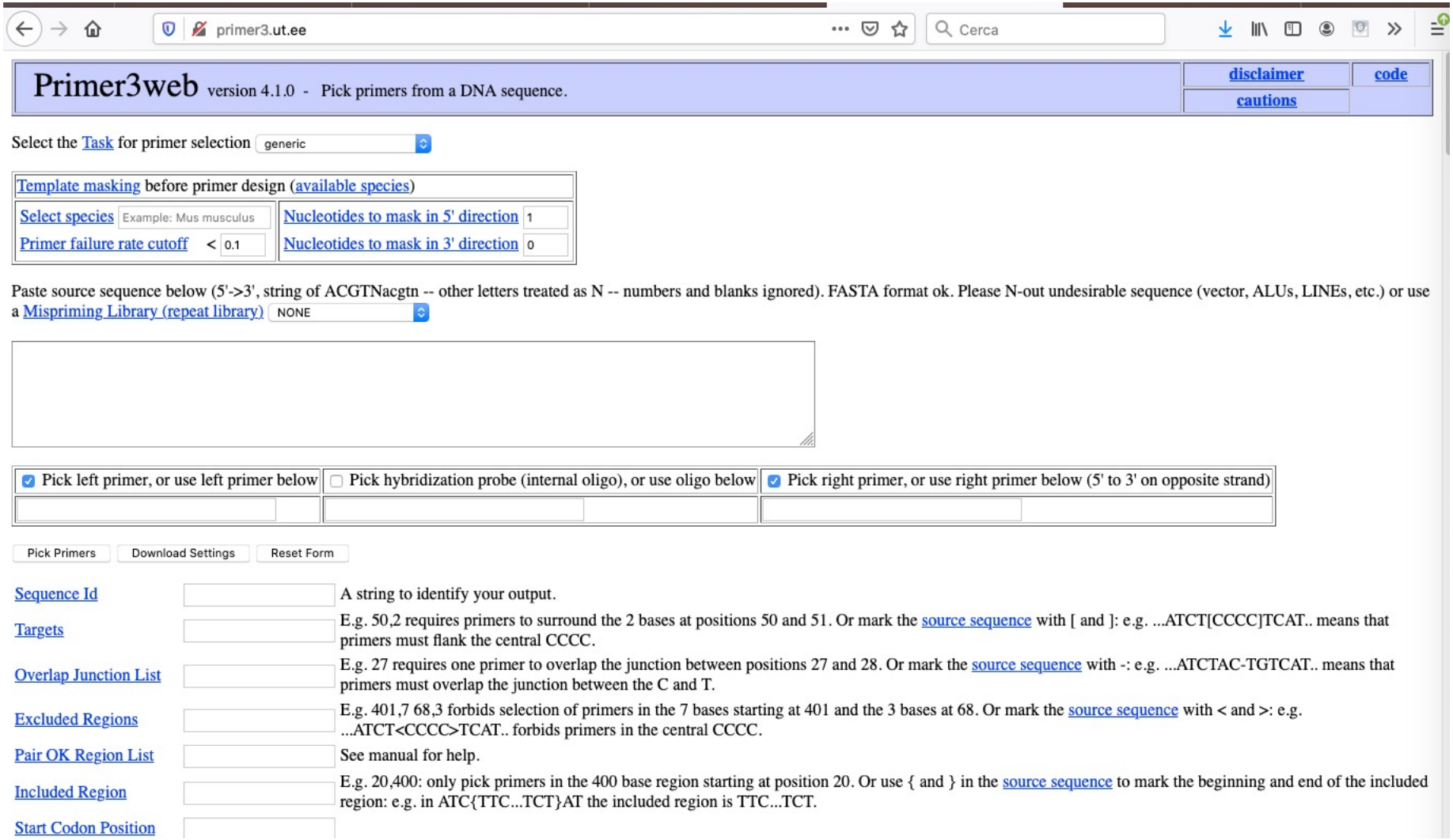
No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence						..........catcaaaaactgaagtggccggatttcaagtgacaaacagcactacaccc
1	ENSE00002202717	1,994,285	1,993,911	-	-	375	AGAATCCGCCCCCTCCCCAACACGGCTGCCTTCCTCCTCCACATCCCTGTCACGAAGAGA AGCCTATTGTGTCAGGCCCAGGGAGTTCGAGTTGAAGGGCCTGGGGGTCTGTGCTCTGAC TGCTCTCAGCCACTTCCCCATTGGCTCCCAGCAGCCTGTGCTCAGCAAGGGCTGAGCGAC AGGGGAGGCTCTCGTCCATAAAAGGGGGGAAGAGGCACCAGAACTGCCATTTGAAGGGGC TTTGGTGGTGTTCACAGCTGCCTCTTTGGCACCTCCTCCCAAGCCGGAGTCAAGGAGGCC CCTGAGCCTTGGACCAGCCACTGCCACCTCCGACCTGCTCGGCCAGAAGCTGCCCAGGGA CAAAGCAGAGTGCAG
	Intron 1-2	1,993,910	1,983,104			10,807	gtaacatccagaggggcactggaa.....ttatcttctctttcttcctccccag
2	ENSE00000655739	1,983,103	1,982,956	-	0	148	CAGGAATTGCTGAGACAGGATGGCCTGGCAGGGGCTGGTCCTGGCTGCCTGCCTCCTCAT GTTCCCCTCCACCACAGCGGACTGCCTGTCGCGGTGCTCCTTGTGTGCTGTAAAGACCCA GGATGGTCCCAAACCTATCAATCCCCTG
	Intron 2-3	1,982,955	1,980,959			1,997	gtaggtttcaggcaaggttcttcaa.....tgtggtctttttgggtctttgcag
3	ENSE00002253018	1,980,958	1,978,757	0	-	2,202	ATTTGCTCCCTGCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGGAGAGATGCCAG AGCTTTCTGTCTTTTTTCACCCCCTCCACCCTTGGGCTCAATGACAAGGAGGACTTGGGG AGCAAGTCGGTTGGGGAAGGGCCCTACAGTGAGCTGGCCAAGCTCTCTGGGTCATTCCTG AAGGAGCTGGAGAAAAGCAAGTTTCTCCCAAGTATCTCAACAAAGGAGAACACTCTGAGC AAGAGCCTGGAGGAGAAGCTCAGGGGTCTCTCTGACGGGTTTAGGGAGGGAGCAGAGTCT GAGCTGATGAGGGATGCCCAGCTGAACGATGGTGCCATGGAGACTGGCACACTCTATCTC GCTGAGGAGGACCCCAAGGAGCAGGTCAAACGCTATGGGGGCTTTTTGCGCAAATACCCC AAGAGGAGCTCAGAGGTGGCTGGGGAGGGGGACGGGGATAGCATGGGCCATGAGGACCTG TACAAACGCTATGGGGGCTTCTTGCGGCGCATTCGTCCCAAGCTCAAGTGGGACAACCAG AAGCGCTATGGCGGTTTTCTCCGGCGCCAGTTCAAGGTGGTGACTCGGTCTCAGGAAGAT CCGAATGCTTACTCTGGAGAGCTTTTTGATGCATAAGCACCTCTTTTCATGGAGTAGAGT CAGGAGAAACCCCTGACACCTTTTCAGGTTGGAGTGCATTCATTCATCCTCTTATATGTG CCCCTTCCCCATGCTCAGCTCAGCATTGTGTACAAAATATCCAAGCCCAGCCTATCTCTC TTCTGCGTGGGAGTATGTTATTTCTCTGGGGTCTGTGATGGGGAAGGGTGGATGTCCCTT CCCCACAATAGGCTTAGTGCTTGGCTCAGACACCTAGACTCTAAAACTATCAGCAGCGGC AGCAGCAGCAGCAGCAGCAGTTTGTGATCTGTCCTTCCAACCTGTTCACGTGACTCCTCA ATTCCAGGGAACCAGAGCGATGTGTTCTTTGTACCTGTAGGTCTATGATGTCCAAACTTA ACAGATCACATGCCCCTCTTAGAAGAAATATGAGCATGCTCCCTCATGCAGATAGTATAC ACATCATAAACAAAGAGTAGAACTTTAAAAGAAGGTAAATAATCATACACAGAAATCCTA ACATTATATTCCCAAATCTCAAAAGATCTCCTGTGCACCTGACTTTGGAGACGATGCTTT


Firefox File Modifica Visualizza Cronologia Segnalibri Strumenti Finestra Aiuto
Sab 14 mar 15:19:00
Posta in arrivo - cdaddario@un
number of nucleotide giga bp p
Genome size - Wikipedia
traduttore - Cerca con Google
Transcript: PDYN-202 (ENST00
https://www.ensembl.org/Homo_sapiens/Transcript/Exons?db=core;g=ENSG0000010
Cerca
Supporting evidence
ID History
Transcript history
Protein history
Exons/ Introns
Translated sequence
Flanking sequence
Intron sequence
UTR
Configure Page
Personal Data
Display options
Manage configurations
Reset configuration
Select from available configurations:
Default
Display options
Flanking sequence at either end of transcript:
50
Number of base pairs per row:
60 bps
Intron base pairs to show at splice sites:
25
Show full intronic sequence:
Show exons only:
Line numbering:
None
Show variants:
No
Hide variants longer than 10bp:
Hide variants by frequency (MAF):
Don't hide
Filter variants by consequence type:
No filter
3 prime UTR variant
5 prime UTR variant
NMD transcript variant
coding sequence variant
Diapositiva 24 di 77
Italiano (Italia)
Note
Commenti
115%

Primer 3

http://primer3.ut.ee/

Step 1 – copia la sequenza di interesse nella finestra



Primer3web version 4.1.0 - Pick primers from a DNA sequence.

disclaimer | code | cautions

Select the Task for primer selection generic

Template masking before primer design (available species)

Select species Example: Mus musculus

Primer failure rate cutoff < 0.1

Nucleotides to mask in 5' direction 1

Nucleotides to mask in 3' direction 0

Paste source sequence below (5'->3', string of ACGTNacgtn -- other letters treated as N -- numbers and blanks ignored). FASTA format ok. Please N-out undesirable sequence (vector, ALUs, LINEs, etc.) or use a Mispriming Library (repeat library) NONE

```
ATGGCCTGGCAGGGGCTGGTCCTGGCTGCCTGCCTCCTCATGTTCCCCTCCACCACAGCGGACTGCCTGTCGCGGTGCTCCTTGTGTGCTGTAAAGA
CCCAGGATGGTCCCAAACCTATCAATCCCCTGATTTGCTCCCTGCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGGAGAGATGCCAG
AGCTTTCTGTCTTTTTTCACCCCCTCCACCCTTGGGCTCAATGACAAGGAGGACTTGGGGAGCAAGTCGGTT
```

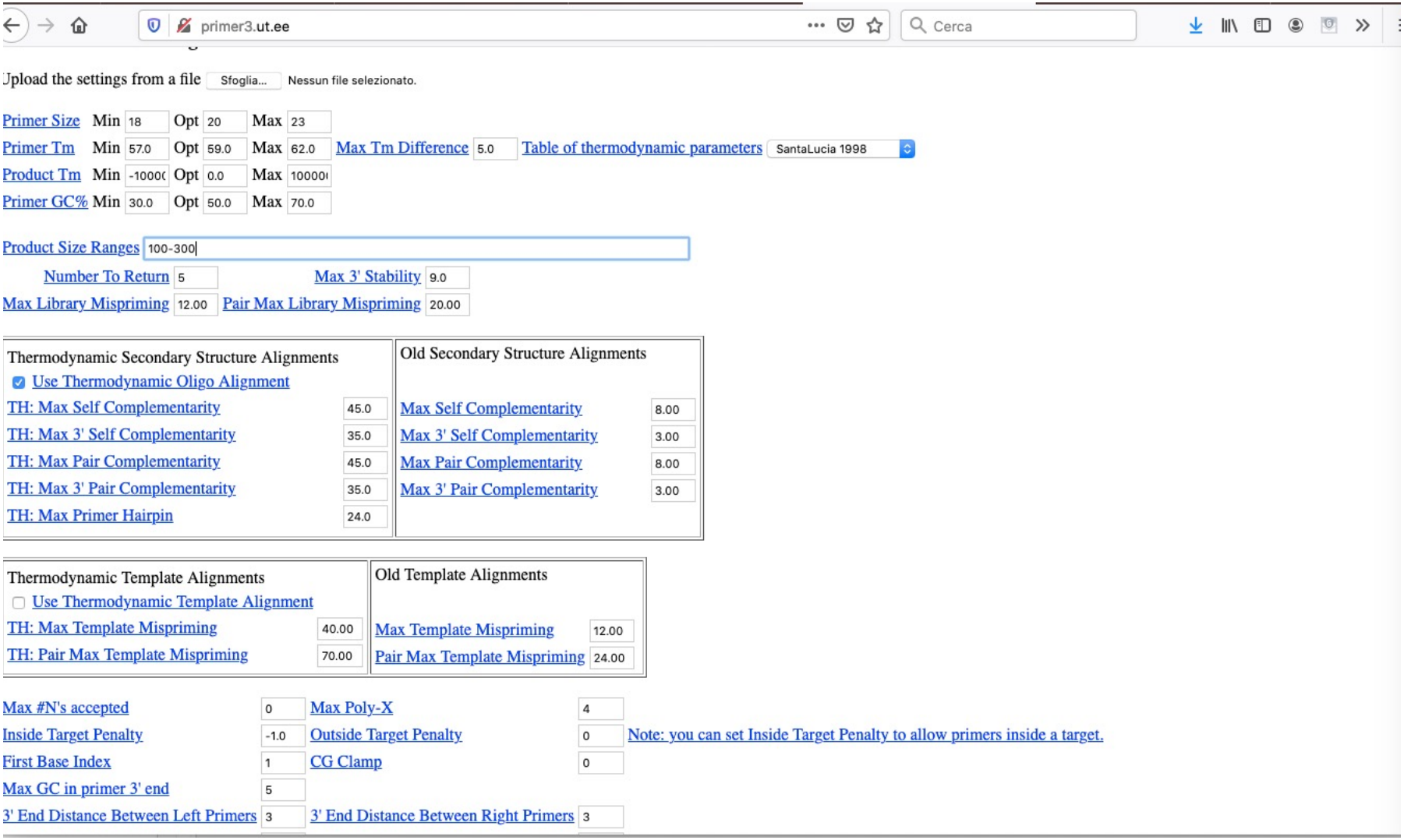
Pick left primer, or use left primer below | Pick hybridization probe (internal oligo), or use oligo below | Pick right primer, or use right primer below (5' to 3' on opposite strand)

Pick Primers | Download Settings | Reset Form

Sequence Id		A string to identify your output.
Targets		E.g. 50,2 requires primers to surround the 2 bases at positions 50 and 51. Or mark the source sequence with [and]: e.g. ...ATCT[CCCC]TCAT.. means that primers must flank the central CCCC.
Overlap Junction List		E.g. 27 requires one primer to overlap the junction between positions 27 and 28. Or mark the source sequence with -: e.g. ...ATCTAC-TGTCAT.. means that primers must overlap the junction between the C and T.
Excluded Regions		E.g. 401,7 68,3 forbids selection of primers in the 7 bases starting at 401 and the 3 bases at 68. Or mark the source sequence with < and >: e.g. ...ATCT<CCCC>TCAT.. forbids primers in the central CCCC.
Pair OK Region List		See manual for help.
Included Region		E.g. 20,400: only pick primers in the 400 base region starting at position 20. Or use { and } in the source sequence to mark the beginning and end of the included region: e.g. in ATC{TTC...TCT}AT the included region is TTC...TCT.
Start Codon Position		

Step 3 – setting parametri

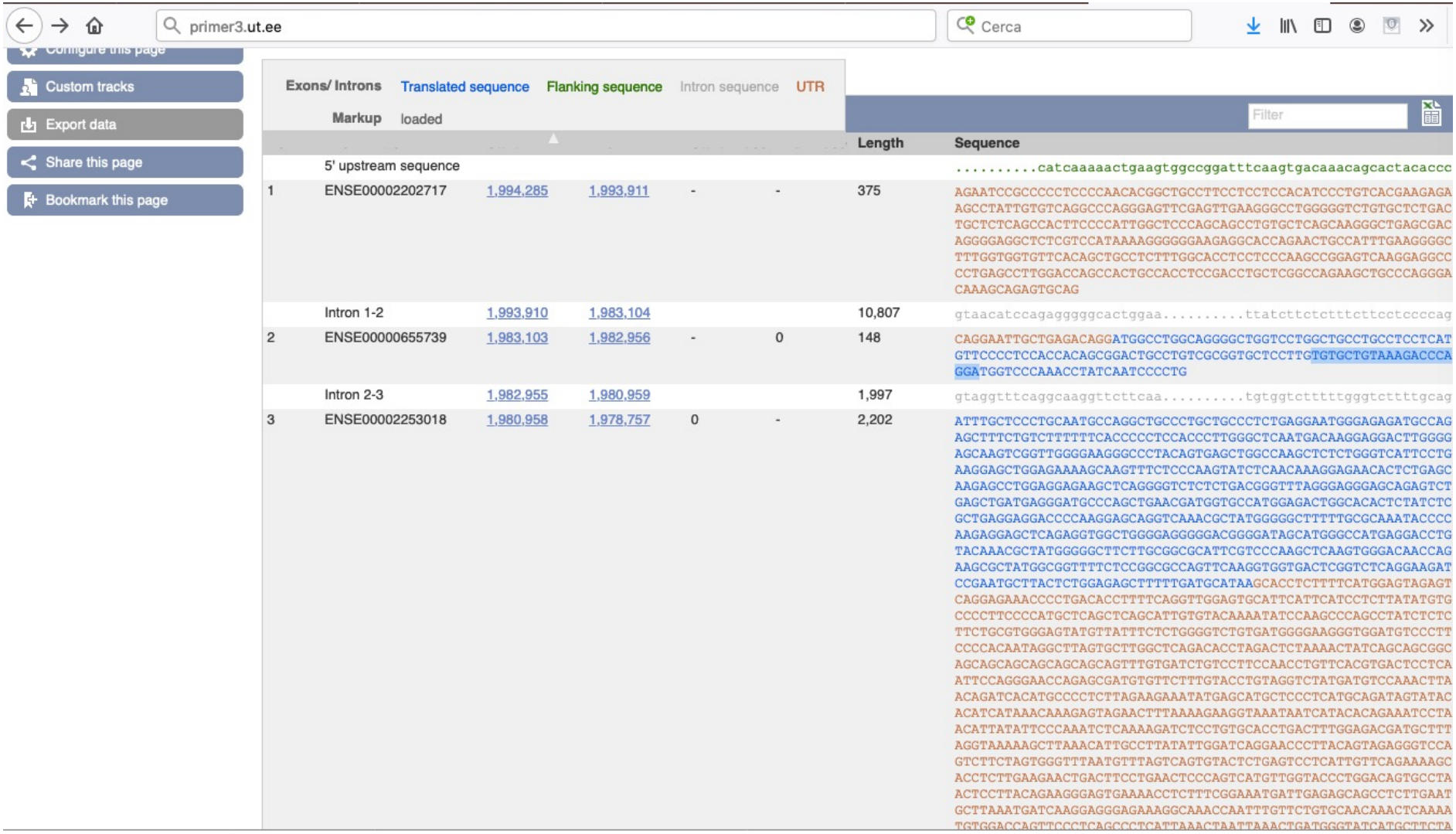
Il programma ha dei parametri di default che possono essere modificati



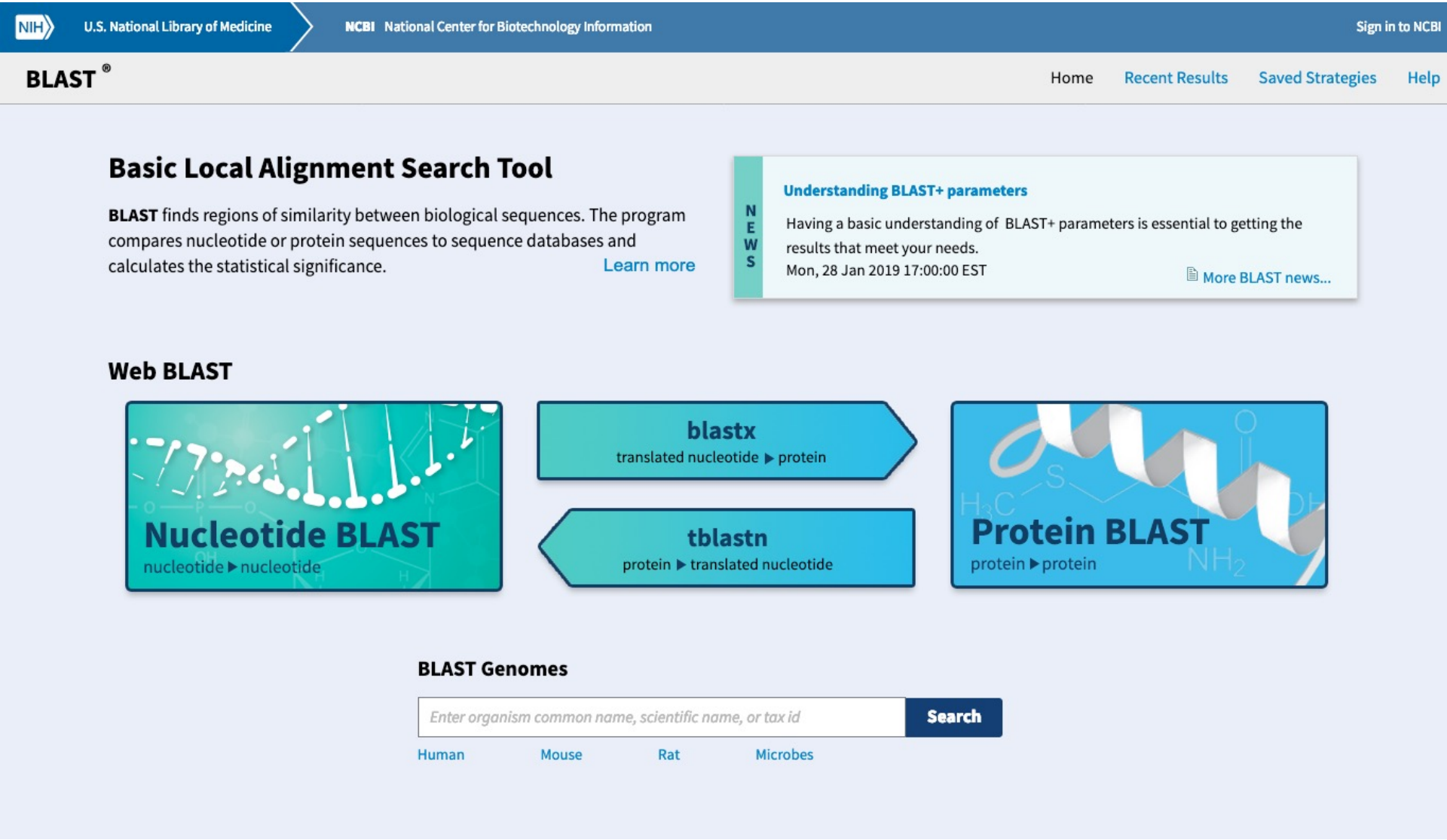
Step 4 –
Scegli i primers

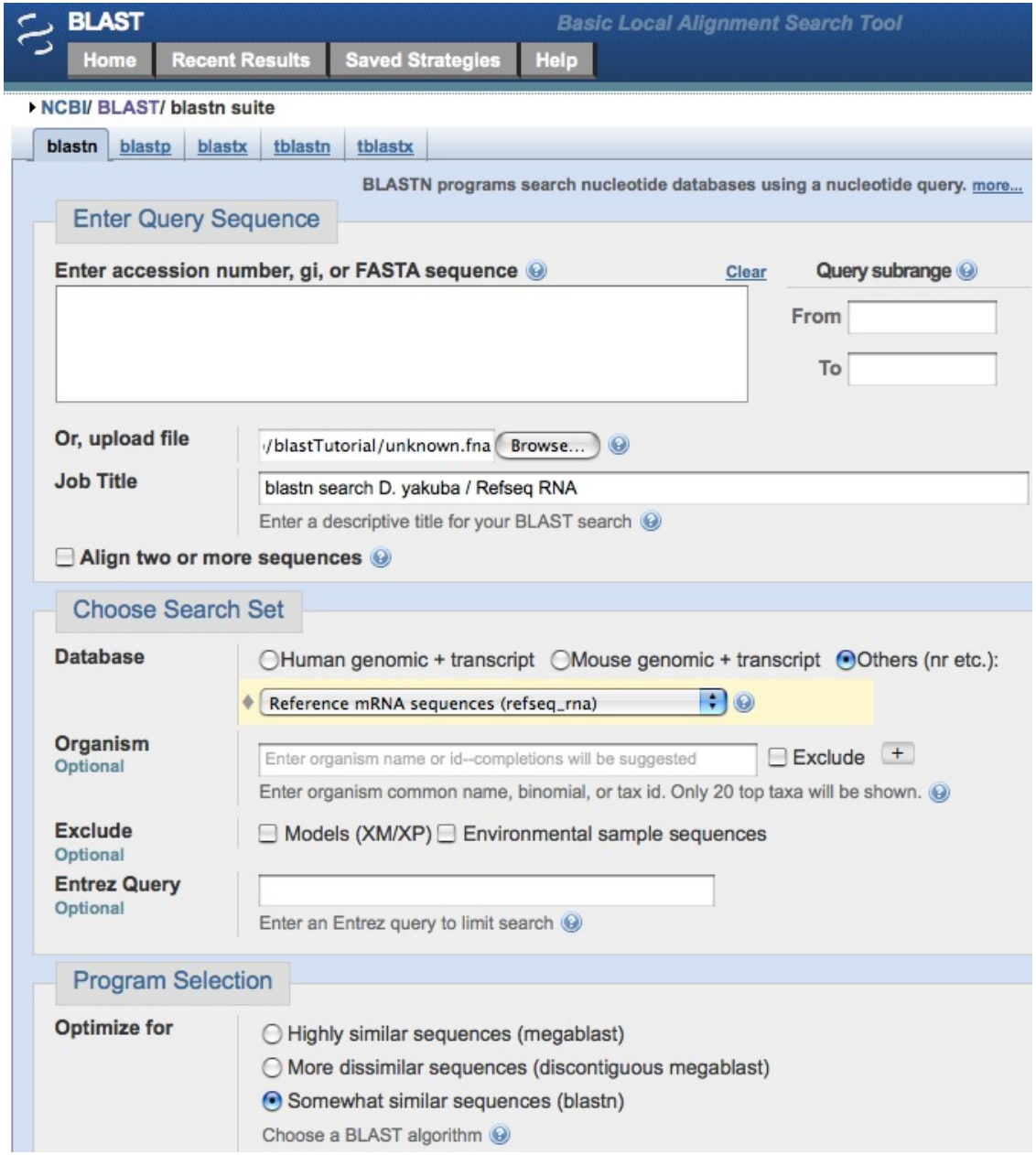
Primer3 Output

```
PRIMER PICKING RESULTS FOR

Template masking not selected
No mispriming library specified
Using 1-based sequence positions
OLIGO            start  len       tm      gc%  any_th  3'_th  hairpin  seq
LEFT PRIMER         85   20    58.57    50.00    0.00   0.00     0.00 TGTGCTGTAAAGACCCAGGA
RIGHT PRIMER       187   20    58.86    55.00    0.00   0.00     0.00 GGCATCTCTCCCATTCCTCA
SEQUENCE SIZE: 261
INCLUDED REGION SIZE: 261

PRODUCT SIZE: 103, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

    1 ATGGCCTGGCAGGGGCTGGTCCTGGCTGCCTGCCTCCTCATGTTCCCCTCCACCACAGCG


   61 GACTGCCTGTCGCGGTGCTCCTTGTGTGCTGTAAAGACCCAGGATGGTCCCAAACCTATC
                           >>>>>>>>>>>>>>>>>>>>

  121 AATCCCCTGATTTGCTCCCTGCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGGAG
                                                 <<<<<<<<<<<<<

  181 AGATGCCAGAGCTTTCTGTCTTTTTTCACCCCCTCCACCCTTGGGCTCAATGACAAGGAG
      <<<<<<<

  241 GACTTGGGGAGCAAGTCGGTT


KEYS (in order of precedence):
>>>>>> left primer
<<<<<< right primer

ADDITIONAL OLIGOS
                    start  len       tm      gc%  any_th  3'_th  hairpin  seq

 1 LEFT PRIMER         34   20    59.46    60.00    0.00   0.00     0.00 CTCCTCATGTTCCCCTCCAC
   RIGHT PRIMER       190   20    58.86    55.00    0.00   0.00     0.00 TCTGGCATCTCTCCCATTCC
   PRODUCT SIZE: 157, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

 2 LEFT PRIMER        122   20    59.45    55.00    0.00   0.00     0.00 ATCCCCTGATTTGCTCCCTG
   RIGHT PRIMER       242   20    59.38    55.00    0.00   0.00     0.00 TCCTCCTTGTCATTGAGCCC
   PRODUCT SIZE: 121, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

 3 LEFT PRIMER         37   20    60.04    60.00    0.00   0.00     0.00 CTCATGTTCCCCTCCACCAC
   RIGHT PRIMER       141   19    59.00    57.89    0.00   0.00     0.00 CAGGGAGCAAATCAGGGGA
   PRODUCT SIZE: 105, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

 4 LEFT PRIMER         80   21    59.12    52.38    0.00   0.00     0.00 CCTTGTGTGCTGTAAAGACCC
```


primer3.ut.ee
Cerca
Configure this page
Custom tracks
Export data
Share this page
Bookmark this page
Exons/ Introns
Translated sequence
Flanking sequence
Intron sequence
UTR
Markup
loaded
Filter
Length
Sequence
5' upstream sequence
.....catcaaaaactgaagtggccggatttcaagtgacaaacagcactacaccc
1
ENSE00002202717
1,994,285
1,993,911
-
-
375
AGAATCCGCCCCCTCCCCAACACGGCTGCCTTCCTCCTCCACATCCCTGTCACGAAGAGA
AGCCTATTGTGTCAGGCCCAGGGAGTTCGAGTTGAAGGGCCTGGGGGTCTGTGCTCTGAC
TGCTCTCAGCCACTTCCCCATTGGCTCCCAGCAGCCTGTGCTCAGCAAGGGCTGAGCGAC
AGGGGAGGCTCTCGTCCATAAAAGGGGGGAAGAGGCACCAGAACTGCCATTTGAAGGGGC
TTTGGTGGTGTTCACAGCTGCCTCTTTGGCACCTCCTCCCAAGCCGGAGTCAAGGAGGCC
CCTGAGCCTTGGACCAGCCACTGCCACCTCCGACCTGCTCGGCCAGAAGCTGCCCAGGGA
CAAAGCAGAGTGCAG
Intron 1-2
1,993,910
1,983,104
10,807
gtaacatccagaggggcactggaa.....ttatcttctctttcttcctccccag
2
ENSE00000655739
1,983,103
1,982,956
-
0
148
CAGGAATTGCTGAGACAGGATGGCCTGGCAGGGGCTGGTCCTGGCTGCCTGCCTCCTCAT
GTTCCCCTCCACCACAGCGGACTGCCTGTCGCGGTGCTCCTTGTGTGCTGTAAAGACCCA
GGATGGTCCCAAACCTATCAATCCCCTG
Intron 2-3
1,982,955
1,980,959
1,997
gtaggtttcaggcaaggttcttcaa.....tgtggtctttttgggtcttttgcag
3
ENSE00002253018
1,980,958
1,978,757
0
-
2,202
ATTTGCTCCCTGCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGGAGAGATGCCAG
AGCTTTCTGTCTTTTTTCACCCCCTCCACCCTTGGGCTCAATGACAAGGAGGACTTGGGG
AGCAAGTCGGTTGGGGAAGGGCCCTACAGTGAGCTGGCCAAGCTCTCTGGGTCATTCCTG
AAGGAGCTGGAGAAAAGCAAGTTTCTCCCAAGTATCTCAACAAAGGAGAACACTCTGAGC
AAGAGCCTGGAGGAGAAGCTCAGGGGTCTCTCTGACGGGTTTAGGGAGGGAGCAGAGTCT
GAGCTGATGAGGGATGCCCAGCTGAACGATGGTGCCATGGAGACTGGCACACTCTATCTC
GCTGAGGAGGACCCCAAGGAGCAGGTCAAACGCTATGGGGGCTTTTTGCGCAAATACCCC
AAGAGGAGCTCAGAGGTGGCTGGGGAGGGGGACGGGGATAGCATGGGCCATGAGGACCTG
TACAAACGCTATGGGGGCTTCTTGCGGCGCATTCGTCCCAAGCTCAAGTGGGACAACCAG
AAGCGCTATGGCGGTTTTCTCCGGCGCCAGTTCAAGGTGGTGACTCGGTCTCAGGAAGAT
CCGAATGCTTACTCTGGAGAGCTTTTTGATGCATAAGCACCTCTTTTCATGGAGTAGAGT
CAGGAGAAACCCCTGACACCTTTTCAGGTTGGAGTGCATTCATTCATCCTCTTATATGTG
CCCCTTCCCCATGCTCAGCTCAGCATTGTGTACAAAATATCCAAGCCCAGCCTATCTCTC
TTCTGCGTGGGAGTATGTTATTTCTCTGGGGTCTGTGATGGGGAAGGGTGGATGTCCCTT
CCCCACAATAGGCTTAGTGCTTGGCTCAGACACCTAGACTCTAAAACTATCAGCAGCGGC
AGCAGCAGCAGCAGCAGTTTGTGATCTGTCCTTCCAACCTGTTCACGTGACTCCTCA
ATTCCAGGGAACCAGAGCGATGTGTTCTTTGTACCTGTAGGTCTATGATGTCCAAACTTA
ACAGATCACATGCCCCTCTTAGAAGAAATATGAGCATGCTCCCTCATGCAGATAGTATAC
ACATCATAAACAAAGAGTAGAACTTTAAAAGAAGGTAAATAATCATACACAGAAATCCTA
ACATTATATTCCCAAATCTCAAAAGATCTCCTGTGCACCTGACTTTGGAGACGATGCTTT
AGGTAAAAAGCTTAAACATTGCCTTATATTGGATCAGGAACCCTTACAGTAGAGGGTCCA
GTCTTCTAGTGGGTTTAATGTTTAGTCAGTGTACTCTGAGTCCTCATTGTTCAGAAAAGC
ACCTCTTGAAGAACTGACTTCCTGAACTCCCAGTCATGTTGGTACCCTGGACAGTGCCTA
ACTCCTTACAGAAGGGAGTGAAAACCTCTTTCGGAAATGATTGAGAGCAGCCTCTTGAAT
GCTTAAATGATCAAGGAGGGAGAAAGGCAAACCAATTTGTTCTGTGCAACAAACTCAAAA
TGTGGACCAGTTCCCTCAGCCCTCATTAAACTAATTAAACTGATGGGTATCATGCTTCTA

https://blast.ncbi.nlm.nih.gov/Blast.cgi



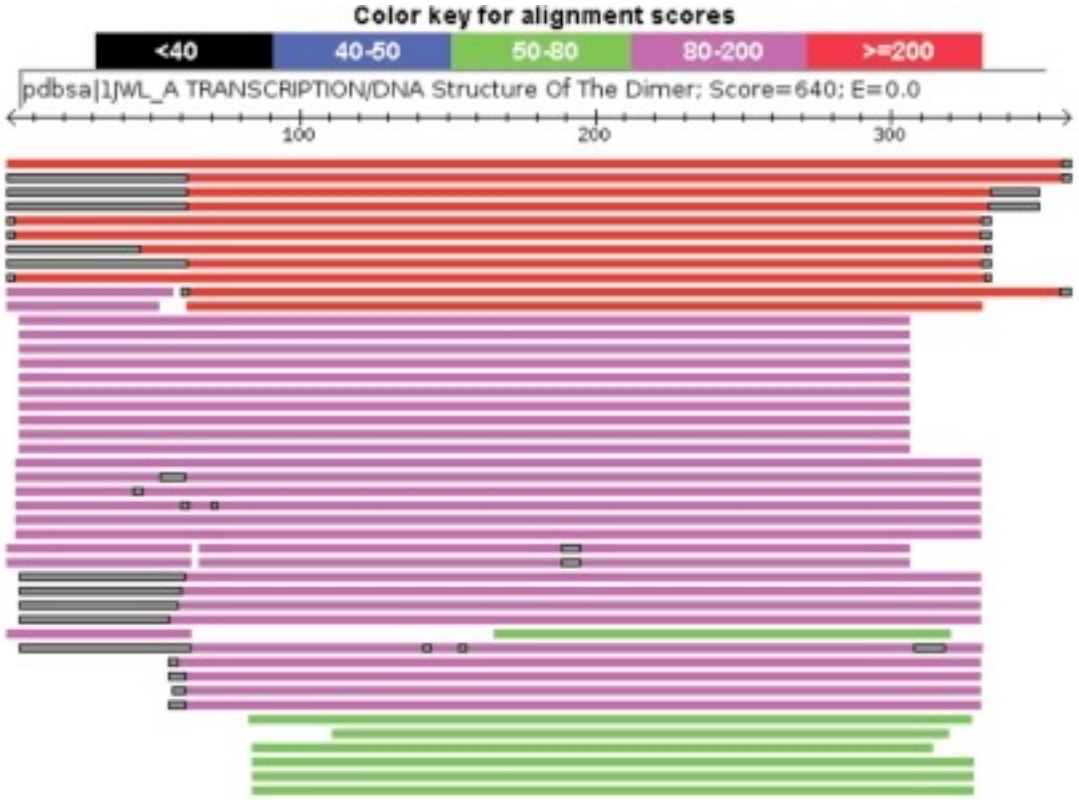


STEP 1

Setting up our blastn search of our unknown sequence against the NCBI Refseq RNA database

When the NCBI web server is busy, the search may take 5 minutes or more

BLAST report



Sequences producing significant alignments:		Score (bits)	E value	Source DB	NCBI Entrez	Cath Prot/Chain
pdbsa|1LBG_A	TRANSCRIPTION/DNA Lactose Operon Repressor Bo...	688	0	SDB	NCBI	CATH Prot
pdbsa|1LBH_A	TRANSCRIPTION REGULATION Intact Lactose Opero...	688	0	SDB	NCBI	CATH Prot
pdbsa|1JYF_A	TRANSCRIPTION Structure Of The Dimeric Lac Re...	669	0	SDB	NCBI	CATH Prot
pdbsa|1JYE_A	TRANSCRIPTION Structure Of A Dimeric Lac Repr...	666	0	SDB	NCBI	CATH Prot
pdbsa|1JWL_A	TRANSCRIPTION/DNA Structure Of The Dimeric La...	640	0	SDB	NCBI	CATH Prot
pdbsa|1EFA_A	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa|1EFA_C	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa|1JWL_C	TRANSCRIPTION/DNA Structure Of The Dimeric La...	640	0	SDB	NCBI	CATH Prot
pdbsa|1EFA_B	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa|1TLF_A	TRANSCRIPTION REGULATION Unprecedented Quater...	575	1e-164	SDB	NCBI	CATH Prot