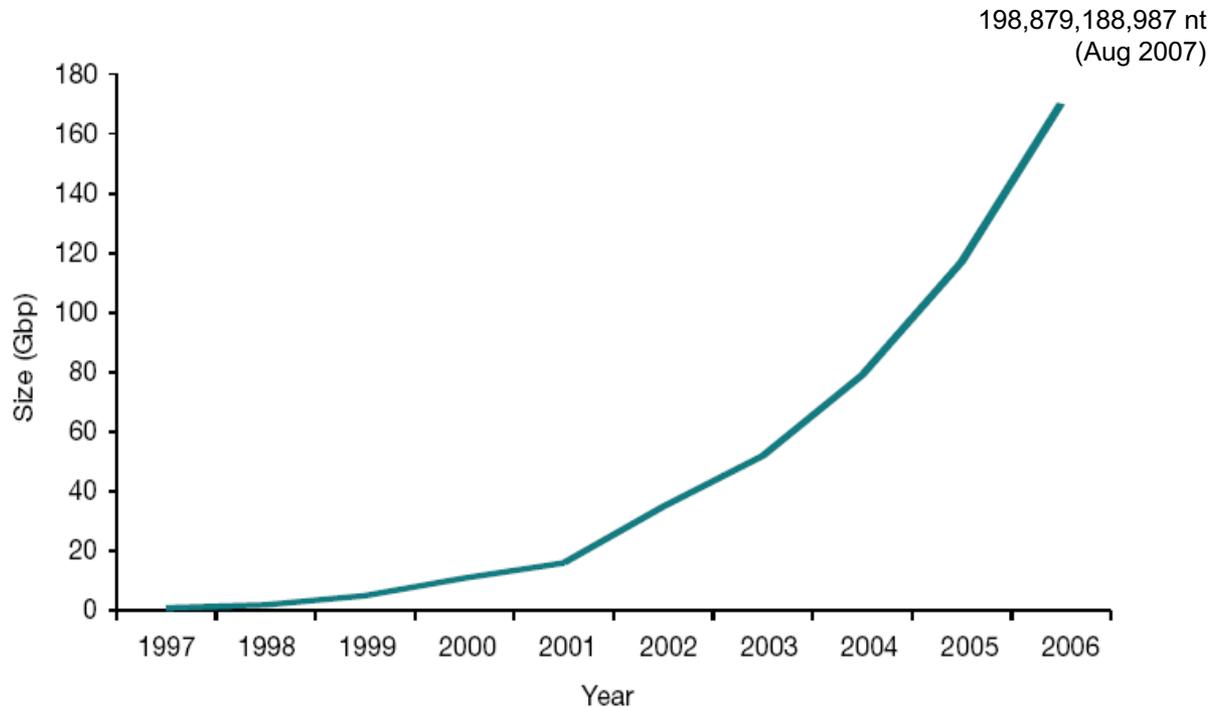


4. BASI DI BIOINFORMATICA PER LO STUDIO DELLA REGOLAZIONE GENICA

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

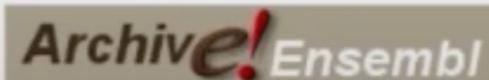
<http://genome.ucsc.edu>

Differenze tra Ensembl, UCSC ed NCBI?

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

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

[BLAST/BLAT](#)[BioMart](#)[VEP](#)[Tools](#)[Downloads](#)[Help & Docs](#)[Blog](#)

Tools

[All tools](#)

BioMart >

Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >

Search our genomes for your DNA or protein sequence

Variant Effect Predictor >

Analyse your own variants and predict the functional consequences of known and unknown variants

Search



for

e.g. [BRCA2](#) or [rat 5:62797383-63627669](#) or [rs699](#) or [coronary heart disease](#)

All genomes



- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

**Human**

GRCh38.p12

[Still using GRCh37?](#)**Mouse**

GRCm38.p6

**Zebrafish**

GRCz11

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**

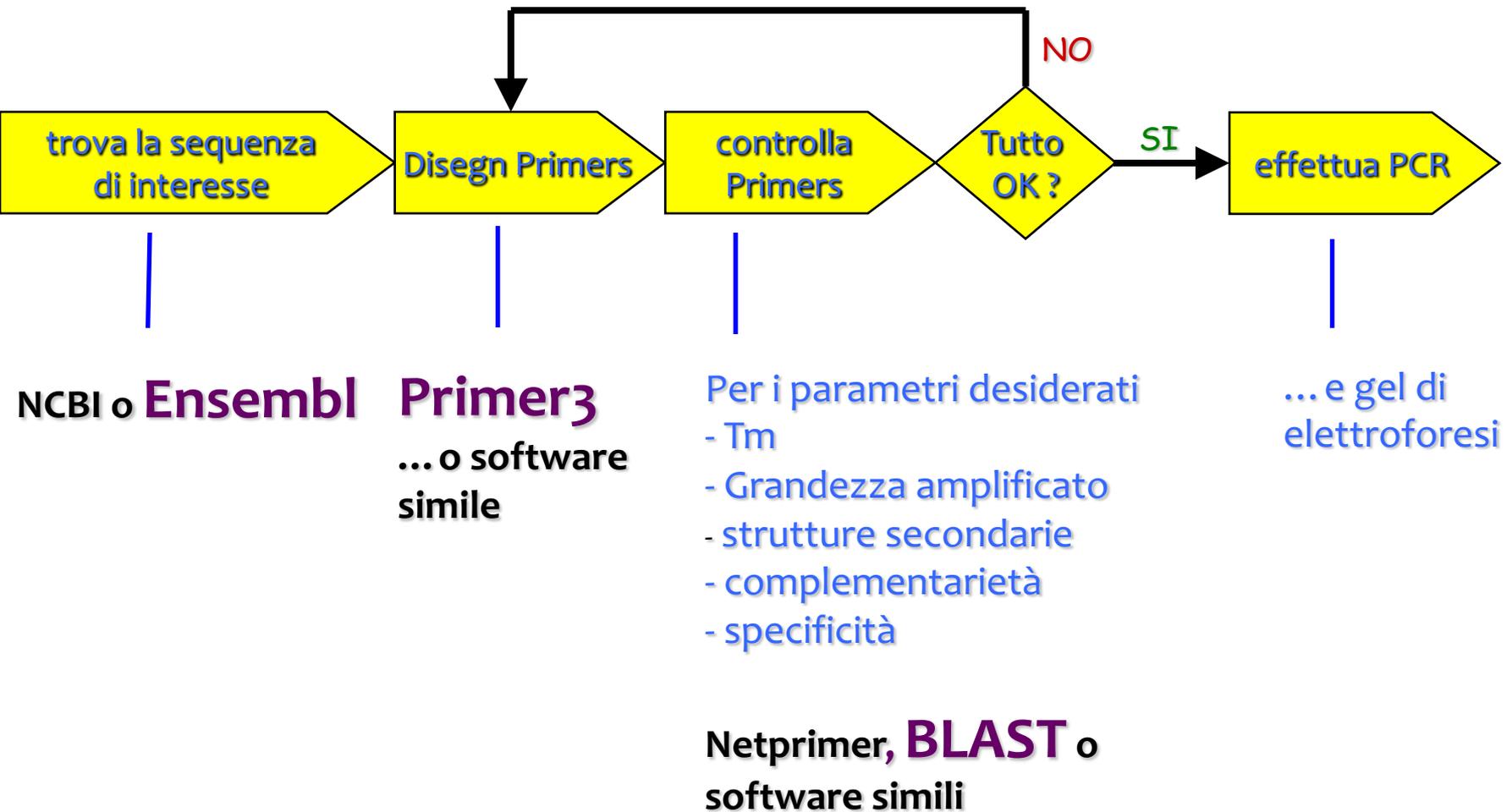
MUS (*Mus musculus*) for mouse: **ENSMUSG###**

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



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

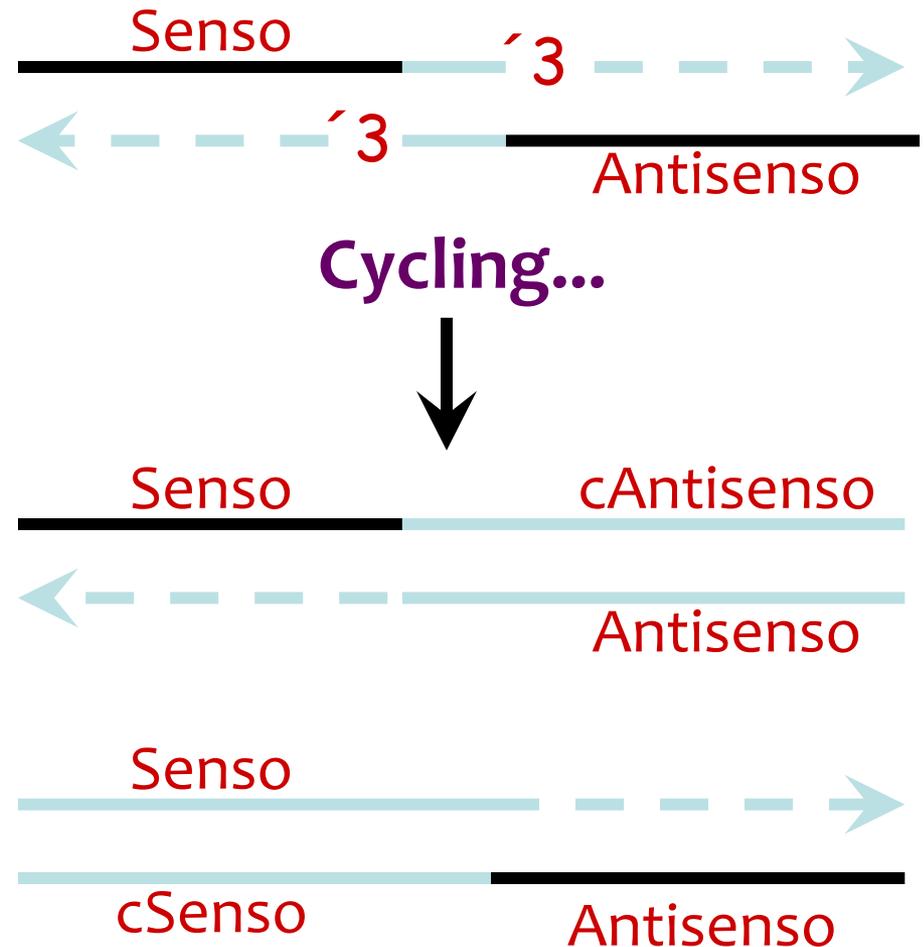
- lunghezza basi compresa tra 18-24
- 40-60% G/C
- Distribuzione bilanciata di basi G/C e A/T
- T_m che permette un annealing tra 55-65° C
- NO strutture secondarie interne (hair-pins)

Le coppie di primers inoltre dovrebbero avere:

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

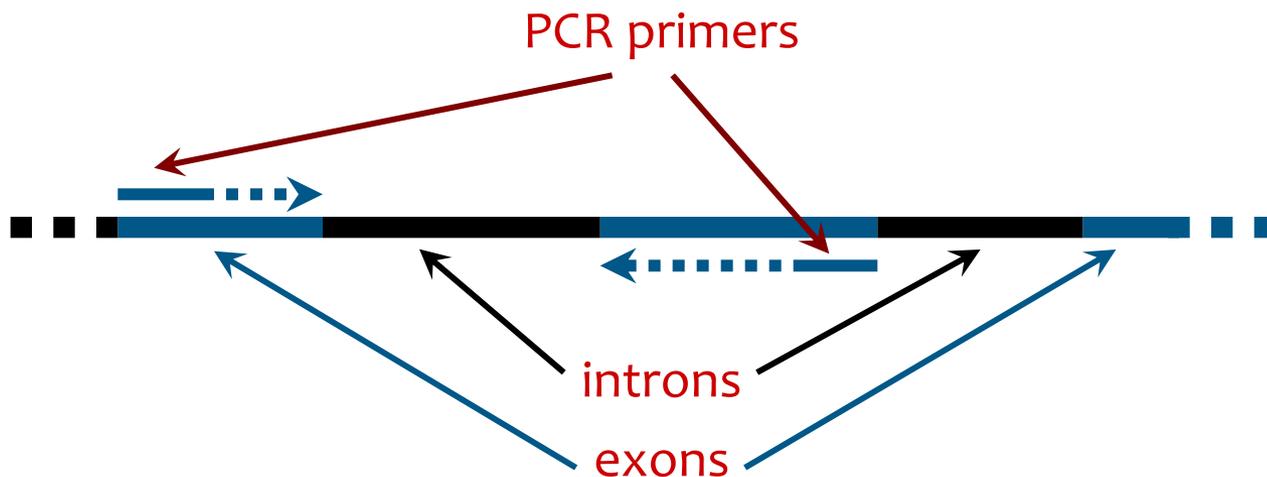
Il problema dei dimeri di primers

- 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/netplaunch/netplaunch.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
BLAST \mathbf{N}	Nucleotide	Nucleotide
BLAST \mathbf{P}	Protein	Protein
BLAST \mathbf{X}	Protein	Nt. \rightarrow Protein
TBLAST \mathbf{N}	Nt. \rightarrow Protein	Protein
TBLAST \mathbf{X}	Nt. \rightarrow Protein	Nt. \rightarrow 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
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X8, mRNA	1585	1585	100%	0.0	100%	XM_011529250.2
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X7, mRNA	1585	1585	100%	0.0	100%	XM_011529249.2
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X6, mRNA	1585	1585	100%	0.0	100%	XM_017027878.1
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X4, mRNA	1585	1585	100%	0.0	100%	XM_011529246.2
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X5, mRNA	1585	1585	100%	0.0	100%	XM_011529248.1
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X3, mRNA	1585	1585	100%	0.0	100%	XM_011529247.1
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X2, mRNA	1585	1585	100%	0.0	100%	XM_011529245.1
<input type="checkbox"/>	PREDICTED: Homo sapiens prodynorphin (PDYN), transcript variant X1, mRNA	1585	1585	100%	0.0	100%	XM_011529244.1
<input type="checkbox"/>	Homo sapiens preproenkephalin (enkB) gene, partial cds	1585	1585	100%	0.0	100%	AH002816.3
<input type="checkbox"/>	Homo sapiens prodynorphin (PDYN), transcript variant 1, mRNA	1585	1585	100%	0.0	100%	NM_024411.4
<input type="checkbox"/>	Homo sapiens prodynorphin (PDYN), transcript variant 4, mRNA	1585	1585	100%	0.0	100%	NM_001190899.2
<input type="checkbox"/>	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



COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

BLAST®

[Home](#) [Recent Results](#) [Saved Strategies](#) [Help](#)

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

A new feature was added to the NCBI IgBLAST webpage
IgBLAST is now able to determine Ig isotypes

Mon, 01 Nov 2021 12:00:00 EST

[More BLAST news...](#)

Web 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.

miRdSNP

a database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes

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

.....**START**

SELECT YOUR SPECIES

The screenshot shows the Ensembl genome browser website in a Firefox browser window. The browser's address bar displays <https://www.ensembl.org/index.html>. The website header includes the Ensembl logo, navigation links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog, and a search bar for species. A red banner at the top states: "Ensembl resources are currently in reduced functionality mode. Please see our blog post for the latest information and our Twitter feed".

The main content area features several sections:

- Tools**: Includes a link to "All tools".
- BioMart >**: Described as "Export custom datasets from Ensembl with this data-mining tool".
- BLAST/BLAT >**: Described as "Search our genomes for your DNA or protein sequence".
- Variant Effect Predictor >**: Described as "Analyse your own variants and predict the functional consequences of known and unknown variants".

The **Search** section contains a search input field and a "Go" button. A dropdown menu is open, showing "All species" (checked) and "Help and Documentation". Under "Favourite species", the following species are listed:

- Human
- Mouse
- Zebrafish

The "All genomes" section includes a dropdown menu labeled "-- Select a species --" and links to "View full list of all Ensembl species" and "Edit your favourites". Below this, a list of genomes is partially visible:

- Human**: GRCm38.p13
- Mouse**: GRCm38.p6
- Zebrafish**: GRCz11

The right sidebar contains a paragraph about Ensembl's capabilities and a section titled "Ensembl Release 99 (January 2020)" with a bulleted list of updates:

- Update to GENCODE 33 for human
- Update to dbSNP153 for human
- Import of updated VISTA enhancers for human and mouse
- New genomes: 10 mammals (including 2 dog breeds), 11 birds, 15 fish and 4 reptiles
- Updated genome assemblies: zebra finch, fugu, Nile tilapia and Asian bonytongue

Below this is a link to "More release news" and another section "Other news from our blog" with a bulleted list of recent news items.

At the bottom of the page, a cookie consent banner reads: "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](#)".

SELECT YOUR GENE

Firefox File Modifica Visualizza Cronologia Segnalibri Strumenti Finestra Aiuto

Google number of nucleotide giga bp p Genome size - Wikipedia traduttore - Cerca con Google Search - Homo sapiens - Ensembl

https://www.ensembl.org/Human/Search/Results?q=;site=ensembl;facet_species=Human Cerca

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

e!Ensembl BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog Login/Register

New Search

Current selection:
< all Species
Only searching Human

Only searching Human
0 results match when restricted to species: Human X

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.

Ensembl release 99 - January 2020 ©
[EMBL-EBI](#)
[Permanent link](#) - [View in archive site](#)

About Us About us Contact us Citing Ensembl Privacy policy Disclaimer	Get help Using this website Adding custom tracks Downloading data Video tutorials Variant Effect Predictor (VEP)	Our sister sites Ensembl Bacteria Ensembl Fungi Ensembl Plants Ensembl Protists Ensembl Metazoa	Follow us Blog Twitter Facebook
---	--	---	---

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

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].

Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary

e!Ensembl BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog Login/Register

Human (GRCh38.p13)

Location: 20:1,978,757-1,994,285 **Gene: PDYN**

- Gene-based displays**
- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
 - Sequence
 - Secondary Structure
 - Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree
 - Orthologues
 - Paralogues
 - Ensembl protein families
 - Ontologies
 - GO: Cellular component
 - GO: Molecular function
 - GO: Biological process
 - Phenotypes
 - Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
 - Gene expression
 - Pathway
 - Regulation
 - External references
 - Supporting evidence
 - ID History
 - Gene history
- Configure this page
- Custom tracks
- Export data

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](#)

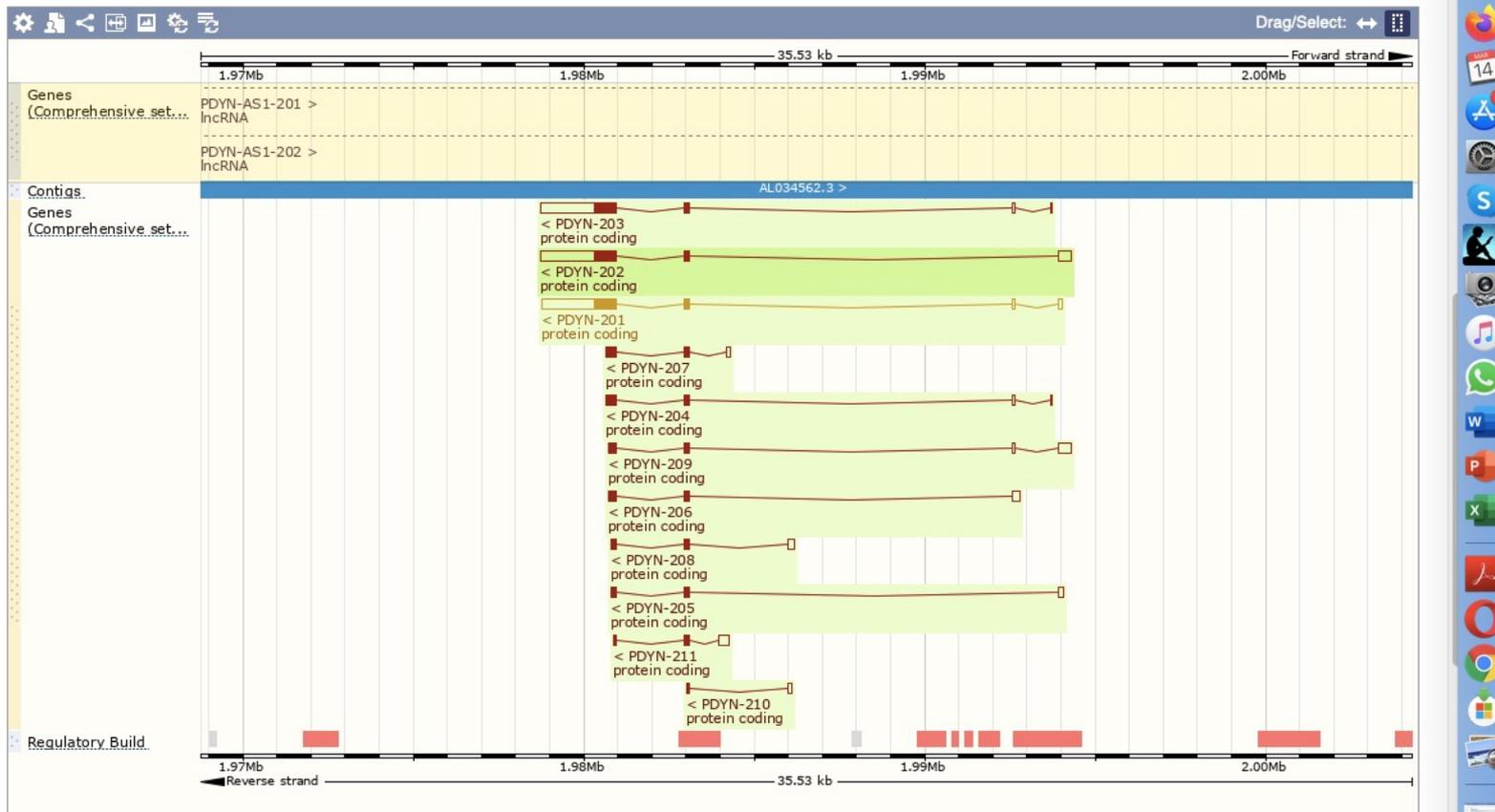
Show/hide columns (1 hidden) Filter

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)



- 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

No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence					catcaaaaaaagtgagcggattcaagtgacaacaacagcactacacc
1	ENSE00002202717	1,994,285	1,993,911	-	-	375	AGAATCCGCCCCCTCCCAACACCGCTGCCTTCCTCCATCCCTGTCACGAAGAGA AGCCTATTGTGTCAGGCCAGGAGTTCGAGTTGAAGGCCCTGGGGTCTGTCTGAC TGCTCTCAGCCACTTCCCATTTGGCTCCAGCAGCCTGTGCTCAGCAAGGGCTGAGCGAC AGGGGAGGCTCTCGTCCATAAAAGGGGGAAGAGGCCACCAGAACTGCCATTTGAAGGGCC TTTGGTGGTGTTCACAGCTGCCCTTTGGCACCTCCTCCCAAGCCGAGTCAAGGAGGCC CCTGAGCCTTGGACCAGCCACTGCCACCTCCGACCTGCTCGGCCAGAGCTGCCAGGGA CAAAGCAGAGTGCAG
	Intron 1-2	1,993,910	1,983,104			10,807	gtaacatccagagggggcactggaa.....ttatctctctctctctctctctcccaag
2	ENSE00000655739	1,983,103	1,982,956	-	0	148	CAGGAATTGTGAGACAGGATGGCTGGCAGGGGCTGGTCTGGCTGCCTGCCTCAT GTCCCTCCACCACAGCGGACTGCCTGTCGGGTGCTCCTTGTGTGCTGTAAGACCCA GGATGGTCCCAAACTTCAATCCCTG
	Intron 2-3	1,982,955	1,980,959			1,997	gtaggtttcaggcaaggtctcttcaa.....tgtggtctttttgggtcttttggag
3	ENSE00002253018	1,980,958	1,978,757	0	-	2,202	ATTTGCTCCCTGCAATGCCAGCTGCCCTGCTGCCCTCTGAGGAATGGGAGAGATGCCAG AGCTTTCTGTCTTTTTCACCCCTCCACCCTTGGGCTCAATGACAAGGAGGACTTTGGGG AGCAAGTCGGTTGGGGAAGGCCCTACAGTGAGCTGGCCAAGCTCTCTGGGTCAATCCTG AAGGAGCTGGAGAAAAGCAAGTTTCTCCCAAGTATCTCAACAAGGAGAACACTCTGAGC AAGAGCCTGGAGGAGAAGCTCAGGGGTCTCTGACGGGTTTAGGGAGGGAGCAGAGTCT GAGCTGATGAGGATGCCAGCTGAACGATGTTGCCATGGAGACTGGCACACTCTATCTC GCTGAGGAGGACCCCAAGGAGCAGGTCAACGCTATGGGGGCTTTTGGCCAAATACCCC AAGAGGAGCTCAGAGTGGCTGGGAGGGGACGGGGATAGCATGGGCCATGAGGACCTG TACAACGCTATGGGGCTTCTGCGCGCATTCGTCACAGCTCAAGTGGGACAACCAG AAGCGCTATGGCGGTTTCTCCGGCCAGTTCAGGTTGGTACTCGGCTCAGGAAGAT CCGAATGCTTACTCTGGAGAGCTTTTGTATGATAAGCACCTTTTTTATGGAGTAGAGT CAGGAGAAAACCCCTGACACCTTTTTCAGGTTGGAGTGCAATTCATCCTCTTATATGTG CCCCTTCCCATGCTCAGCTCAGCATTGTGTACAAAATATCCAAGCCAGCCTATCTCTC TTCTCGTGGGAGTATGTTATTTCTGCGGCTCTGTATGGGGAAGGGTGGATGTCCTT CCCCAATAGGCTTAGTGTGGCTCAGACACCTAGACTTAAACTATCAGCAGCGGC AGCAGCAGCAGCAGCAGCAGTTTGTGATCTGTCTTCCAACCTGTTACAGTACTCTCA ATTCCAGGGAACCAGAGCGATGTGTTCTTTGTACCTGTAGGCTATGATGTCCAACTTA ACAGATCACATGCCCTCTTAGAAGAAATATGAGCATGCTCCCTCATGCAGATAGTATAC ACATCATAAACAAAGAGTAGAACTTTAAAAGAGGTAATAATCATACACAGAAATCCTTA ACATATATTTCCAAATCTCAAAGATCTCCTGTGCACCTGACTTTGGAGACGATGCTTT

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

```

ACATCATAAACAAAGAGTAGAACTTTAAAAGAAGGTAATAATCATACACAGAAATCCTA
ACATTATATFCCCAAATCTCAAAGATCTCCTGTGCACCTGACTTTGGACGATGCTTT
AGGTAAAAAGCTTAAACATTGCCTTATATTGGATCAGGAACCCCTTACAGTAGAGGGTCCA
GTCCTCTAGTGGGTTAATGTTAGTCAGTGTACTCTGAGTCTCATTGTTTCAGAAAAGC
ACCCTTGAAGAACCTGACTTCCGTGAACCTCCAGTCATGTTGGTACCCTGGACATGGCTTA

```

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

[Template masking](#) before primer design ([available species](#))

Select species Example: Mus musculus	Nucleotides to mask in 5' direction 1
Primer failure rate cutoff < 0.1	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 \(repeat library\)](#)

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	<input type="text"/>	A string to identify your output.
Targets	<input type="text"/>	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	<input type="text"/>	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	<input type="text"/>	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	<input type="text"/>	See manual for help.
Included Region	<input type="text"/>	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	<input type="text"/>	

Select the [Task](#) for primer selection

[Template masking](#) before primer design ([available species](#))

Select species	<input type="text" value="Example: Mus musculus"/>	Nucleotides to mask in 5' direction	<input type="text" value="1"/>
Primer failure rate cutoff	<input type="text" value="< 0.1"/>	Nucleotides to mask in 3' direction	<input type="text" value="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 \(repeat library\)](#)

```
ATGGCCTGGCAGGGGCTGGTCCTGGCTGCCTCCTCATGTTCCCTCCACCACAGGGACTGCCTGTCGCGGTGCTCCTTGTGTGCTGTAAGA
CCCAGGATGGTCCAAACATCAATCCCTGATTTGCTCCCTGCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGGAGAGATGCCAG
AGCTTTCTGCTTTTTCACCCCTCCACCTTGGGCTCAATGACAAGGAGACTTGGGGAGCAAGTCGGTT
```

<input checked="" type="checkbox"/> Pick left primer, or use left primer below	<input type="checkbox"/> Pick hybridization probe (internal oligo), or use oligo below	<input checked="" type="checkbox"/> Pick right primer, or use right primer below (5' to 3' on opposite strand)
<input type="text"/>	<input type="text"/>	<input type="text"/>

- [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

Upload the settings from a file Nessun file selezionato.

[Primer Size](#) Min Opt Max

[Primer Tm](#) Min Opt Max [Max Tm Difference](#) [Table of thermodynamic parameters](#)

[Product Tm](#) Min Opt Max

[Primer GC%](#) Min Opt Max

[Product Size Ranges](#)

[Number To Return](#) [Max 3' Stability](#)

[Max Library Mispriming](#) [Pair Max Library Mispriming](#)

Thermodynamic Secondary Structure Alignments	Old Secondary Structure Alignments
<input checked="" type="checkbox"/> Use Thermodynamic Oligo Alignment	
TH: Max Self Complementarity <input type="text" value="45.0"/>	Max Self Complementarity <input type="text" value="8.00"/>
TH: Max 3' Self Complementarity <input type="text" value="35.0"/>	Max 3' Self Complementarity <input type="text" value="3.00"/>
TH: Max Pair Complementarity <input type="text" value="45.0"/>	Max Pair Complementarity <input type="text" value="8.00"/>
TH: Max 3' Pair Complementarity <input type="text" value="35.0"/>	Max 3' Pair Complementarity <input type="text" value="3.00"/>
TH: Max Primer Hairpin <input type="text" value="24.0"/>	

Thermodynamic Template Alignments	Old Template Alignments
<input type="checkbox"/> Use Thermodynamic Template Alignment	
TH: Max Template Mispriming <input type="text" value="40.00"/>	Max Template Mispriming <input type="text" value="12.00"/>
TH: Pair Max Template Mispriming <input type="text" value="70.00"/>	Pair Max Template Mispriming <input type="text" value="24.00"/>

[Max #N's accepted](#) [Max Poly-X](#)

[Inside Target Penalty](#) [Outside Target Penalty](#) [Note: you can set Inside Target Penalty to allow primers inside a target.](#)

[First Base Index](#) [CG Clamp](#)

[Max GC in primer 3' end](#)

[3' End Distance Between Left Primers](#) [3' End Distance Between Right Primers](#)

- Configure this page
- Custom tracks
- Export data
- Share this page
- Bookmark this page

Exons/ Introns		Translated sequence	Flanking sequence	Intron sequence	UTR	Markup	loaded	Length	Sequence
5' upstream sequence									
1	ENSE00002202717	1,994,285	1,993,911	-	-			375catcaaaaactgaagtggcggatttcaagtgacaacagcactacacc AGAATCCGCCCTCCCAACACGGCTGCCTTCTCTCCACATCCCTGTCCAGGAGAG AGCCTATTGTGTCAGGCCAGGGAGTTCGAGTTGAAGGGCTGGGGTCTGTGCTCTGAC TGCTCTCAGCCACTTCCCATTGGCTCCCAGCAGCTGTGCTCAGCAAGGGCTGAGCGAC AGGGGAGGCTCTGTCATAAAAGGGGGAAGAGGCACCAGAAGTCCATTTGAAGGGGC TTTGGTGGTGTTCACAGCTGCCTCTTTGGCACCTCCTCCAAGCCGGAGTCAAGGAGGCC CCTGAGCCTTGGACCAGCCACTGCCACCTCCGACCTGCTCGGCCAGAAGTCCACAGGA CAAAGCAGAGTGCAG
Intron 1-2									
2	ENSE00000655739	1,993,910	1,983,104	-	0			148	gtaacatccagagggggcactggaa.....ttatcttctcttcttctctcccccag CAGGAATTGCTGAGACAGGATGGCCCTGGCAGGGGCTGGTCTGGCTGCCTGCCTCCTCAT GTTCCCTCCACCACAGCGGACTGCCTGTGCGGGTCTCCTTGTGTGCTGTAAGAGCCCA GGTGCTGCCAAACCTATCAATCCCTG
Intron 2-3									
3	ENSE00002253018	1,982,955	1,980,959	0	-			2,202	gtaggtttcaggcaaggttcttcaa.....tgtggttcttttgggtcttttgcag ATTTGCTCCTTCCAATGCCAGGCTGCCCTGCTGCCCTCTGAGGAATGGAGAGATGCCAG AGCTTTCTGTCTTTTTACCCCTCCACCCTTGGGCTCAATGACAAGGAGGACTTGGGG AGCAAGTCGGTTGGGGAAGGGCCCTACAGTGAGCTGGCCAAGCTCTCTGGGTCAATCTCTG AAGGAGCTGGAGAAAAGCAAGTTTCTCCAAGTATCTCAACAAGGAGAACACTGTGAGC AAGAGCCTGGAGGAGAAGCTCAGGGTCTCTCTGACGGGTTTAGGGAGGAGCAGAGTCT GAGCTGATGAGGGATGCCAGCTGAACGATGGTGCATGGAGACTGGCACACTCTATCTC GCTGAGGAGGACCCCAAGGAGCAGGTCAAACGCTATGGGGGCTTTTGCGCCAAATACCCC AAGAGGAGCTCAGAGGTGGCTGGGAGGGGACGGGGATAGCATGGGCCATGAGGACCTG TACAAACGCTATGGGGGCTTCTGCGGCGCATTTCGTCCAAGCTCAAGTGGGACAACCGAG AAGCGCTATGGCGGTTTTCTCCGGCCAGTCAAGGTGGTGAATCGGCTCAGGAAGAT CCGAATGCTTACTCTGGAGAGCTTTTGTATGCATAAGCACCTCTTTTCATGGAGTAGAGT CAGGAGAAACCCCTGACACCTTTTTCAGGTTGGAGTGCAATTCATTCCTCTTATATGTG CCCCTTCCCCATGCTCAGCTCAGCATTGTGTACAAAATATCCAAGCCAGCCTATCTCTC TTCTGCGTGGGAGTATGTTATTTCTCTGGGGTCTGTGATGGGGAAGGGTGGATGTCCTT CCCCAATAGGCTTAGTGCTTGGCTCAGACACTAGACTCTAAAACATCAGCAGCGGC AGCAGCAGCAGCAGCAGCAGTTGTGATCTGTCTTCCAACCTGTTACAGTGACTCCTCA ATTCCAGGGAACAGAGCGATGTGTTCTTTGTACCTGTAGGTCTATGATGTCCAAACTTA ACAGATCACATGCCCTCTTAGAAGAAATATGAGCATGCTCCCTCATGCAGATAGTATAC ACATCATAAACAAAGAGTAGAACTTTAAAAGAAGTAAATAATCATACACAGAAATCCCTA ACATTATATCCCAAATCTCAAAGATCTCCTGTGACCTGACTTTGGAGAGCATGCTTT AGGTAAAAAGCTTAAACATTGCCTTATATTGGATCAGGAACCTTACAGTAGAGGGTCCA GTCTTCTAGTGGGTTAATGTTTAGTCAGTGTACTCTGAGTCTCATTGTTTCAGAAAAGC ACCTTTGAAGAACTGACTTCTGAACTCCAGTCAATGTTGGTACCCTGGACAGTGCCTA ACTCCTTACAGAAGGGAGTGAACCTCTTTCCGAAATGATTGAGAGCAGCCTCTTGAAT GCTTAAATGATCAAGGAGGGAGAAAGGCAACCAATTTGTTCTGTGCAACAACTCAA TGTGGACAGTTCCTCAGCCTCATTAACCTAATTAACCTGATGGGTATCATGCTCTCA

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.

[Learn more](#)

Understanding BLAST+ parameters

Having a basic understanding of BLAST+ parameters is essential to getting the results that meet your needs.

Mon, 28 Jan 2019 17:00:00 EST

[More BLAST news...](#)

Web BLAST

Nucleotide BLAST
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide

Protein BLAST
protein ► protein

BLAST Genomes

Human

Mouse

Rat

Microbes

BLAST *Basic Local Alignment Search Tool*

Home Recent Results Saved Strategies Help

► NCBI/ BLAST/ blastn suite

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide databases using a nucleotide query. [more...](#)

Enter Query Sequence

Enter accession number, gi, or FASTA sequence [Clear](#) Query subrange [Query subrange](#)

From

To

Or, upload file [Browse...](#)

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database Human genomic + transcript Mouse genomic + transcript Others (nr etc.):

Reference mRNA sequences (refseq_rna)

Organism Optional Exclude

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.

Exclude Optional Models (XM/XP) Environmental sample sequences

Entrez Query Optional

Enter an Entrez query to limit search

Program Selection

Optimize for Highly similar sequences (megablast)

More dissimilar sequences (discontiguous megablast)

Somewhat similar sequences (blastn)

Choose a BLAST algorithm

STEP 1

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



▶ [NCBI/BLAST/ Formatting Results - 56RFPSX1012](#) [\[Formatting options\]](#)

Job Title: blastn search D. yakuba / Refseq RNA search

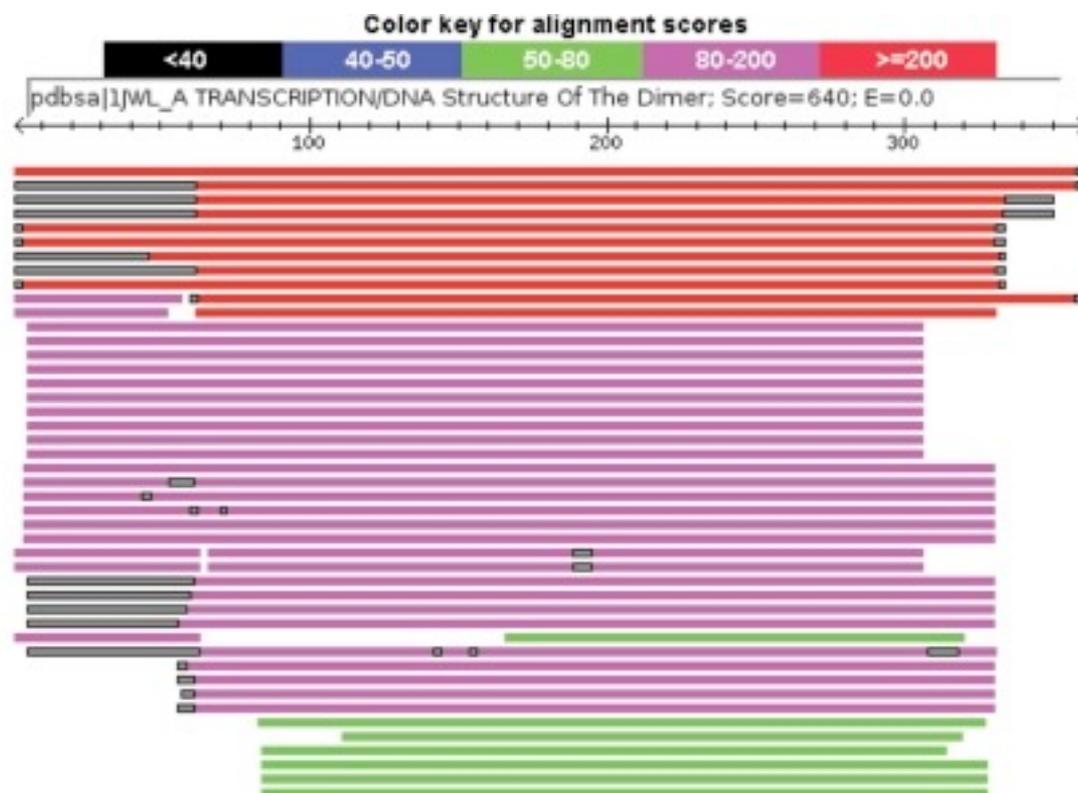
WAITING

Request ID	56RFPSX1012
Status	Searching
Submitted at	Tue May 22 17:17:42 2007
Current time	Tue May 22 17:17:45 2007
Time since submission	00:00:03

This page will be automatically updated in **13** seconds until search is done

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
pdbsa1lbg_A	TRANSCRIPTION/DNA Lactose Operon Repressor Bo...	688	0	SDB	NCBI	CATH Prot
pdbsa1lbg_A	TRANSCRIPTION REGULATION Intact Lactose Opero...	688	0	SDB	NCBI	CATH Prot
pdbsa1jyf_A	TRANSCRIPTION Structure Of The Dimeric Lac Re...	669	0	SDB	NCBI	CATH Prot
pdbsa1jye_A	TRANSCRIPTION Structure Of A Dimeric Lac Repr...	666	0	SDB	NCBI	CATH Prot
pdbsa1jwl_A	TRANSCRIPTION/DNA Structure Of The Dimeric La...	640	0	SDB	NCBI	CATH Prot
pdbsa1lfa_A	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa1lfa_C	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa1jwl_C	TRANSCRIPTION/DNA Structure Of The Dimeric La...	640	0	SDB	NCBI	CATH Prot
pdbsa1lfa_B	TRANSCRIPTION/DNA Crystal Structure Of The La...	640	0	SDB	NCBI	CATH Prot
pdbsa1tle_A	TRANSCRIPTION REGULATION Unprecedented Quater...	575	1e-164	SDB	NCBI	CATH Prot