

Laboratorio 02 - Alineamiento de secuencias

En este laboratorio vamos a navegar a través de los procedimientos básicos de cómo coleccionar secuencias génicas, cómo realizar un alineamiento múltiple, y cómo diseñar partidores para PCR

Parte 1: Colectar genes homólogos

Comencemos por ir a NCBI y buscar un gen de interés.

- Ve a la base de datos de genes de [NCBI](#) y busca el gen SRY.

ncbi.nlm.nih.gov

Gene

Advanced

Search

Help

Full Report ▾

Send to: ▾

Hide sidebar >>

SRY sex determining region Y [*Homo sapiens* (human)]

Gene ID: 6736, updated on 6-Mar-2016

Summary

Official Symbol SRY provided by HGNC

Official Full Name sex determining region Y provided by HGNC

Primary source [HGNC:HGNC:11311](#)

See related [Ensembl:ENSG00000184895](#); [HPRD:08364](#); [MIM:480000](#); [Vega:OTTHUMG0000036084](#)

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as TDF; TDY; SRXX1; SRXY1

Summary This intronless gene encodes a transcription factor that is a member of the high mobility group (HMG)-box family of DNA-binding proteins. This protein is the testis-determining factor (TDF), which initiates male sex determination. Mutations in this gene give rise to XY females with gonadal dysgenesis (Swyer syndrome); translocation of part of the Y chromosome containing this gene to the X chromosome causes XX male syndrome. [provided by RefSeq, Jul 2008]

Orthologs [all](#)

Genomic context

Location: Yp11.3

Exon count: 1

See SRY in [Epigenomics](#), [MapViewer](#)

Annotation release	Status	Assembly	Chr	Location
107	current	GRCh38.p2 (GCF_000001405.28)	Y	NC_000024.10 (2786855..2787741, complement)
105	previous assembly	GRCh37.p13 (GCF_000001405.25)	Y	NC_000024.9 (2654896..2655782, complement)

Chromosome Y - NC_000024.10

CD99 → X0/2 → SRY → RNASEH2CF1 → TMR22P2

Genomic regions, transcripts, and products

Go to [reference sequence details](#)

Genomic Sequence: [NC_000024.10](#) Chromosome Y Reference GRCh38.p2 Primary Assembly ▾

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

NC_000024.10: 2.8M..2.8M (1.2Kbp) C ▾ Find:

7,900 | 2,787,800 | 2,787,700 | 2,787,600 | 2,787,500 | 2,787,400 | 2,787,300 | 2,787,200 | 2,787,100 | 2,787 K | 2,786,900 | 2,786,800

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- Phenotypes
- Variation
- Pathways from BioSystems
- Interactions
- General gene information
 - Markers, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links
 - Locus-specific Databases

Related information

- Order cDNA clone
- 3D structures
- BioAssay
- BioAssay by Target (List)
- BioAssay by Target (Summary)
- BioAssay, by Gene target
- BioAssays, RNAi Target, Active
- BioAssays, RNAi Target, Tested
- BioProjects
- BioSystems
- Books
- CCDS
- ClinVar
- Conserved Domains
- dbVar
- Full text in PMC
- Full text in PMC_nucleotide
- Gene neighbors
- Genome
- GEO Profiles

- Los resultados deberían ser familiares ya que se trata de la base de datos que se usó en el práctico anterior. Ahora, navega en esta página y busca genes ortólogos a SRY (pista: Orthologs from Annotation Pipeline)

Responde:

¿Qué función cumple el gen SRY?

¿Cuántos genes ortólogos están anotados en esa base de datos?

Dentro de esta base de datos vas a encontrar un listado de genes SRY ortólogos en diferentes especies de animales.

Gene

Gene sources: Genomic, Categories: Annotated genes, Protein-coding, Sequence content: CCDS, Ensembl, RefSeq, RefSeqGene, Status: ☒ Current, Chromosome locations, more...

Tabular 20 per page Sort by Relevance Send to:

Search results
Items: 19
Showing Current items.

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> SRY ID: 6736	sex determining region Y [<i>Homo sapiens</i> (human)]	Chromosome Y, NC_000024.10 (2786855..2787741, complement)	SRXX1, SRXY1, TDF, TDY	480000
<input type="checkbox"/> SRY ID: 449510	sex determining region Y [<i>Pan troglodytes</i> (chimpanzee)]	Chromosome Y, NC_006492.3 (26202574..26203483)		
<input type="checkbox"/> SRY ID: 103231539	sex determining region Y [<i>Chlorocebus sabaeus</i> (green monkey)]	Chromosome X, NC_023671.1 (927143..928924, complement)		
<input type="checkbox"/> SRY ID: 103680433	sex determining region Y [<i>Ursus maritimus</i> (polar bear)]	NW_007925176.1 (3318..3983)		
<input type="checkbox"/> SRY ID: 574155	sex determining region Y [<i>Macaca mulatta</i> (Rhesus monkey)]	Chromosome Y, NC_027914.1 (81625..82236, complement)		
<input type="checkbox"/> SRY ID: 102886423	sex determining region Y [<i>Pteropus alecto</i> (black flying fox)]		PAL_GLEAN10000322	
<input type="checkbox"/> SRY ID: 102505875	sex determining region Y [<i>Camelus ferus</i> (Wild Bactrian camel)]		CB1_083775001	
<input type="checkbox"/> Sry ID: 107144341	sex determining region Y [<i>Marmota marmota marmota</i> (Alpine marmot)]			
<input type="checkbox"/> SRY ID: 107131190	sex determining region Y [<i>Bos taurus</i> (cattle)]	Chromosome 13, AC_000170.1 (5424637..5425341, complement)		
<input type="checkbox"/> SRY ID: 106822352	sex determining region Y [<i>Equus asinus</i> (ass)]			
<input type="checkbox"/> SRY ID: 105003146	sex determining region Y [<i>Bison bison bison</i>]			
<input type="checkbox"/> Sry ID: 104860174	sex determining region Y [<i>Fukomys damarensis</i> (Damara mole-rat)]			
<input type="checkbox"/> SRY ID: 104667028	sex determining region Y [<i>Rhinopithecus roxellana</i> (golden snub-nosed monkey)]			
<input type="checkbox"/> SRY ID: 104264052	sex determining region Y [<i>Gavia stellata</i> (red-throated loon)]			
<input type="checkbox"/> SRY ID: 103610094	sex determining region Y [<i>Galeopterus variegatus</i> (Sunda flying lemur)]	NW_007736697.1 (12354..13031)		
<input type="checkbox"/> SRY ID: 103544114	sex determining region Y [<i>Equus przewalskii</i> (Przewalski's horse)]	NW_007681222.1 (21748..22460)		
<input type="checkbox"/> SRY ID: 103213240	sex determining region Y [<i>Orycteropus afer afer</i>]	NW_006923028.1 (29518..30178, complement)		
<input type="checkbox"/> SRY ID: 102503352	sex determining region Y [<i>Tupaia chinensis</i> (Chinese tree shrew)]		TREES_T100009460	

[Clear all](#) [Show additional filters](#)

Filters: [Manage Filters](#)

Results by taxon
Top Organisms [\[Tree\]](#)
Bos taurus (1)
Bison bison bison (1)
Camelus ferus (1)
Odobenus rosmarus divergens (1)
Ursus maritimus (1)
All other taxa (14)
[More...](#)

Find related data
Database:
Option:
Link to related Nucleotide entry

Search details
ortholog_gene_6736[group] AND alive[prop]
 [See more...](#)

Recent activity [Turn Off](#) [Clear](#)
Q ortholog_gene_6736[group] AND (alive[prop]) (19) Gene
SRY sex determining region Y [Homo sapiens] Gene
Q SRY[sym] AND (alive[prop]) (41) Gene
Q SRY AND (alive[prop]) (5715) Gene
LOC105372834 [Homo sapiens] Gene
[See more...](#)

- Descarga la secuencia en formato fasta para cada uno de los ortólogos. Para esto, haz clic en cada uno de los vínculos y navega hasta la sección "NCBI Reference Sequences (RefSeq)"
- En esa sección encontrarás un vínculo que te llevará hasta la secuencia nucleotídica del gen SRY para la especie en cuestión (pista: mRNA and Protein(s))
- Nos interesa la secuencia del mRNA del gen, no la de proteínas ni la del genoma

ncbi.nlm.nih.gov

NCBI Resources How To

Gene

Advanced

Search

Help

Full Report

Send to

Hide sidebar

SRY sex determining region Y [*Camelus ferus* (Wild Bactrian camel)]

Gene ID: 102505875, updated on 15-Nov-2015

Summary

Gene symbol SRY

Gene description sex determining region Y

Locus tag CB1_083775001

Gene type protein coding

RefSeq status MODEL

Organism [Camelus ferus](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus

Orthologs [human](#) [all](#)

Genomic context

Location: chromosome: Un

Exon count: 1

Annotation release	Status	Assembly	Chr	Location
101	current	CB1 (GCF_000311805.1)	Unplaced Scaffold	NW_006220067.1 (187..909)

NW_006220067.1

[187] [909]

SRY

Genomic regions, transcripts, and products

Genomic Sequence: NW_006220067.1 Unplaced Scaffold Reference CB1 Primary Assembly

[Go to reference sequence details](#)

[Go to nucleotide:](#) [Graphics](#) [FASTA](#) [GenBank](#)

NW_006220067.1: 78..1.0K (939bp)

Find:

Tools

Tracks

Genes, NCBI Camelus ferus Annotation Release 101, 2015-11-03

XP_006195400.1

SRY

XP_006195462.1

Genes, INSDC annotation provided by Bactrian Camels Genome Con...

CB1_083775001

EPY72333.1

wRNA-seq-determini...

RNA-seq exon coverage, aggregate (filtered), NCBI Camelus ferus Annotation Release 101 - log base 2 scaled

RNA-seq intron-spanning reads, aggregate (filtered), NCBI Camelus ferus Annotation Release 101 - log base 2 scaled

RNA-seq intron features, aggregate (filtered), NCBI Camelus fe...

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- General gene information
 - Homology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links

Related information

- BioProjects
- Conserved Domains
- Full text in PMC
- Map Viewer
- Nucleotide
- Protein
- PubMed
- PubMed(nucleotide/PMC)
- RefSeq Proteins
- RefSeq RNAs
- Taxonomy

General information

- About Gene
- FAQ
- FTP site
- Help
- My NCBI help
- NCBI Handbook
- Statistics

Related sites

- BLAST
- Genome

- Una vez que haces clic en el vínculo que te lleva al registro del mRNA del gen, haz clic en *FASTA* y simplemente copia y pega la secuencia en el bloc de notas

NCBI Nucleotide search results for **PREDICTED: Camelus ferus sex determining region Y (SRY), mRNA**. The page displays the NCBI Reference Sequence (XM_006195400.1) and the corresponding mRNA sequence. The sequence is shown in a blue box with a search bar above it. The sequence is as follows:

```
>gi|560939352|ref|XM_006195400.1| PREDICTED: Camelus ferus sex determining region Y (SRY), mRNA
ATGCAATCATATGCTTCTGCTATGTTTCGGGAACGAATGGTGATAATTACAGCTCAGCGGTACAGCAAC
AAAAATCCTTTCTTCAGGAAGCCTCCTCGCTACTTCAGAGAGACAATCGTAGCTCAAAAGGCTGTATA
TGAAGCTGGAGGAACGGTAGAGACTACATGAAGGACCGTGTCAAGCGACCCATGAACGCTTTCATTGTA
TGGGCTCGTGATCAAGGCGAAGGTGGCTCTAGAGAAATCCAAATGCAGAACGAGATCAGCAAGC
GGTGGGATACCAAGTGAATTAATTACAGAACTGAAAGCGGCGCTTCTTCGAGAGGCGCAGAGACT
CCGGGCTACATCGGGACAATACCCGGACTGTAAATACCAACCTCGTCAGAGAGCGAGGGGCTACCG
AGAACTGCAAAATCATTTCACAGACCCCTTCAACACTATGCAGCCAGGTACATGTAGACAGCCCTTTGT
ACTCTTTCACATACAGGACCATTTGTGCCAAGACCCGAGTCACGAATGGAAGGCCCGTTAAGTCCCTC
ACAACGGATGAGCATTACAGCTCACTCCCTCAGCAAGAGCACTGCAGCAACTGGACAAGCCTGCACCAC
ACTAGGGTGACACTGGCTACACAGATCTCTGGGATGGTCCCTTTTACTGTAGTTTGCAGCTGGACATT
CTCACCCTTACTTTTGGTGTGA
```

The page also includes a sidebar with various tools and information, such as "Analyze this sequence", "Articles about the SRY gene", "Reference sequence information", "More about the gene SRY", "Related information", and "LinkOut to external resources".

Perfecto! Ahora ya puedes coleccionar genes SRY homólogos (ortólogos) presentes en diferentes especies. Antes de avanzar a la siguiente sección deberías contar con un archivo de texto en el bloc de notas con todas las secuencias de mRNA homólogos del gen SRY

Parte 2: Alineamiento múltiple

Un alineamiento múltiple trata de comparar genes o regiones genéticas homólogas. La suposición es que cada columna en el alineamiento corresponde a una posición que es homóloga entre todas las secuencias que estás comparando.

- Vamos a la página del EMBL-EBI, específicamente a la de "multiple sequence alignment" ([aquí](#))

- El EMBL-EBI provee muchos servicios web entre los cuales están los de alineamiento múltiple.

Responde:

¿Qué es el EMBL-EBI?

¿Cuántos tipos de alineamiento múltiple se pueden realizar en EMBL-EBI?

¿Cuál es el programa que ellos ofrecen que funciona mejor para secuencias de proteínas?

¿Qué otros tipos de herramientas ofrece EMBL-EBI?

- Haz clic en MAFFT
- Sube tu archivo de bloc de notas con las secuencias de SRY que descargastes

Responde:

¿Cuál es el costo de abrir un gap?

¿Cuál es el costo de extender un gap?

- Haz clic en "Submit"
- Una vez que termine el programa, haz clic en "Result Summary" y luego en Jalview --> Start Jalview

This website uses cookies. By continuing to browse this site, you are agreeing to the use of our site cookies. To find out more, see our [Terms of Use](#). OK

EMBL-EBI [Services](#) [Research](#) [Training](#) [About us](#)

MAFFT

[Input form](#) [Web services](#) [Help & Documentation](#) [Share](#) [Feedback](#)

[Tools](#) > [Multiple Sequence Alignment](#) > MAFFT

Results for job mafft-I20160317-145704-0438-2777586-oy

[Alignments](#) [Result Summary](#) [Guide Tree](#) [Phylogenetic Tree](#) [Submission Details](#)

Input Sequences

[mafft-I20160317-145704-0438-2777586-oy.input](#)

Tool Output

[mafft-I20160317-145704-0438-2777586-oy.output](#)

Alignment in FASTA format

[mafft-I20160317-145704-0438-2777586-oy.fasta](#)

Guide Tree

[mafft-I20160317-145704-0438-2777586-oy.tree](#)

Phylogenetic Tree

[mafft-I20160317-145704-0438-2777586-oy.ph](#)

Percent Identity Matrix

[mafft-I20160317-145704-0438-2777586-oy.pim](#)

Jalview

[Start Jalview](#)

EMBL-EBI

[News](#)

[Brochures](#)

[Contact us](#)

[Intranet](#)

Services

[By topic](#)

[By name \(A-Z\)](#)

[Help & Support](#)

Research

[Overview](#)

[Publications](#)

[Research groups](#)

[Postdocs & PhDs](#)

Training

[Overview](#)

[Train at EBI](#)

[Train outside EBI](#)

[Train online](#)

[Contact organisers](#)

Industry

[Overview](#)

[Members Area](#)

[Workshops](#)

[SME Forum](#)

[Contact Industry programme](#)

About us

[Overview](#)

[Leadership](#)

[Funding](#)

[Background](#)

[Collaboration](#)

[Jobs](#)

[People & groups](#)

[News](#)

[Events](#)

[Visit us](#)

[Contact us](#)

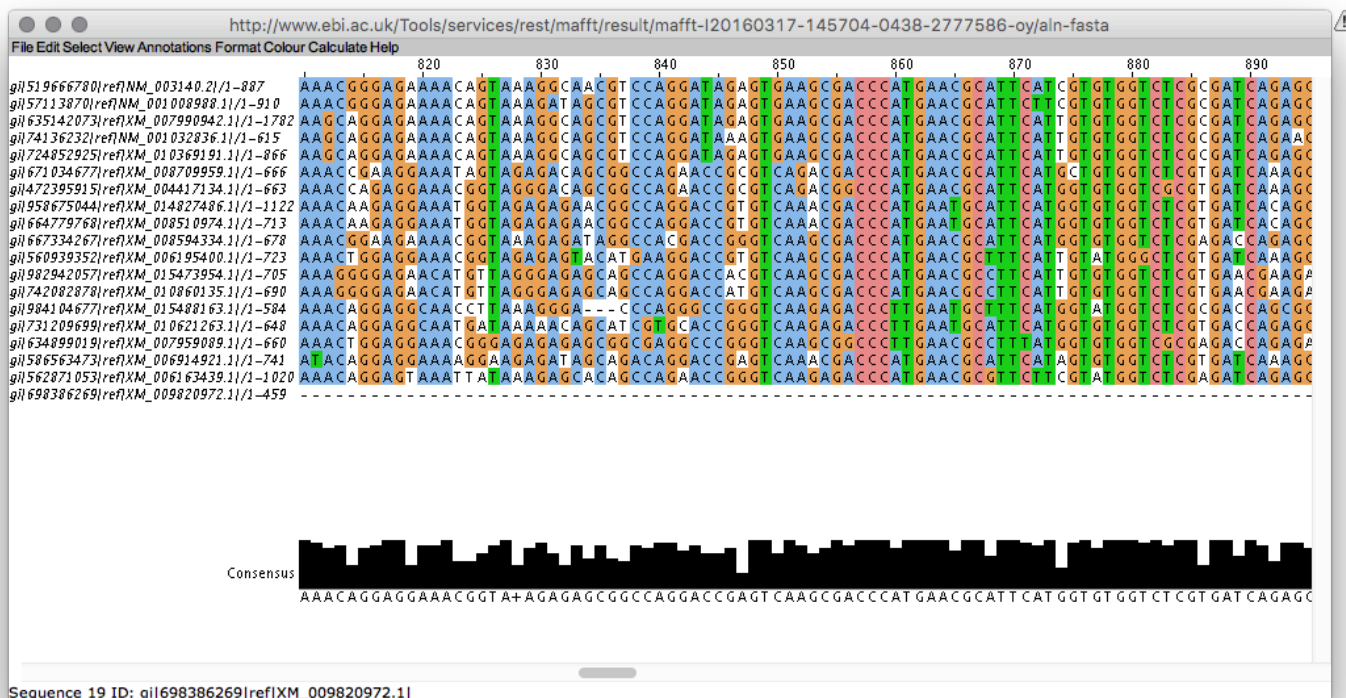
EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK +44 (0)1223 49 44 44

Copyright © EMBL-EBI 2016 | EBI is an outstation of the European Molecular Biology Laboratory | [Terms of use](#)

Jalview es un programa que depende de Java y que sirve para inspeccionar interactivamente el alineamiento

Responde:

¿Cuál es la longitud total del alineamiento?



- Ahora haz clic en la filogenia (Phylogenetic tree)
- Familiarízate con los códigos que corresponden a tus secuencias. Esos códigos naturalmente corresponden a las especies de animales donde los ortólogos de SRY fueron encontrados

Responde:

¿Cuál es la especie cuyo gen SRY está más relacionado con el gen SRY de humanos?
 ¿Cuál es el más lejano?
 ¿Cuál es la especie cuyo gen SRY es más cercana a la del burro?

Un alineamiento de secuencias es un modelo derivado de una observación directa (secuencia de ADN). Uno puede imaginar un alineamiento de secuencias como una **hipótesis de homología** donde uno afirma que cada columna del alineamiento corresponde a una posición homóloga entre las secuencias.

"All models are wrong, but some are useful" [George Box](#)

Teniendo en mente el aforismo más conocido en estadística, uno puede crear otro modelo (o alineamiento) a partir de las mismas secuencias.

- Volvamos a la página principal de MAFFT, selecciona tu archivo y haz clic en "More options"
- Vamos a modificar el costo de abrir un gap y de extender un gap.

Responde:

¿Cómo esperas que sea el alineamiento si el costo de abrir un gap aumenta? ¿Y si disminuye?

¿Cómo esperas que sea el alineamiento si el costo de extender un gap aumenta? ¿Y si disminuye?

- Prueba aumentando el costo de abrir gaps cambiando el valor de 1.53 a 2.0

Responde:

¿Cuál fue el efecto de aumentar el costo de abrir un gap en la longitud total del alineamiento?

prueba lo mismo pero esta vez disminuyendo al mínimo el costo de extender un gap. Describe cómo cambia el alineamiento

Parte 3: Diseño de partidores

Ya llegaste a la última parte de este laboratorio. Aquí vamos a seleccionar tres secuencias de SRY, las de *Chlorocebus sabaeus*, *Homo sapiens*, *Pan troglodytes* y *Macaca Mulatta*.

- Las secuencias que descargaste anteriormente corresponden a las secuencias de mRNA. Para este ejercicio necesitamos las secuencias codificantes o Coding Sequence (CDS)
- Ve nuevamente a NCBI, base de datos de genes, busca SRY y en la parte de ortólogos selecciona las especies de interés. En vez de hacer clic directamente en FASTA, baja hasta el vínculo que dice CDS
- Una pequeña ventana va a aparecer, haz clic en FASTA y copia esa secuencia en el bloc de notas.
- Repite este procedimiento para todas las especies y guarda las secuencias con el nombre primers.fasta
- Modifica los nombres de las secuencias de tal forma que los nombres de las especies estén dentro de paréntesis cuadrado

Desde

```
>gi|519666780|ref|NM_003140.2| Homo sapiens sex determining region Y (SRY), mRNA  
a  
>sry gene [Homo sapiens]
```

Siempre es más fácil copiar en vez de hacerlo tú mismo pero el costo es que no aprendes tanto. Si quieres tomar un atajo, yo ya compilé las secuencias y están [acá](#).

Finalmente, entra a la siguiente página y sigue las instrucciones. [primers4clades](#)

Responde:

Pega en tu informe los partidores que el programa encontró

Trabajo de laboratorio para la próxima semana

El trabajo de laboratorio para la próxima semana consta de dos partes. La primera parte ya la tienes parcialmente lista. Simplemente tienes que responder las preguntas que aparecen a través de esta guía y enviar un informe a tu ayudante con copia a eduardo.castro@unab.cl

Profesor	Nombre	Correo electrónico
Profesor responsable	Dr. Eduardo Castro Nallar	eduardo.castro@unab.cl
Profesor ayudante sección 1	Ingrid Araya Durán	ingrid.araya.duran@gmail.com
Profesor ayudante sección 1	Sandro Valenzuela	sandrolvalenzuelad@gmail.com
Profesor ayudante sección 2	Javier Cáceres	ja.caceresmolina@gmail.com
Profesor ayudante sección 2	Consuelo Bello	consuelobelloz@gmail.com

La segunda parte tiene que ver con leer capítulo 2 del libro *Sequence Alignment Methods, Models, Concepts, and Strategies Edited by Michael S. Rosenberg* que se llama *Insertion and Deletion Events, Their Molecular Mechanisms, and Their Impact on Sequence Alignments*. Puedes acceder al capítulo [aquí](#).