

---

# Scienza patrimonio dell'umanità

---

- ❖ I dati sperimentali, le sequenze di acidi nucleici, le strutture delle proteine pubblicati e prodotti da centri di ricerca pubblici sono consultabili da tutti gratuitamente, basta avere accesso alla rete!



---

I dati biologici sono di tutti

---

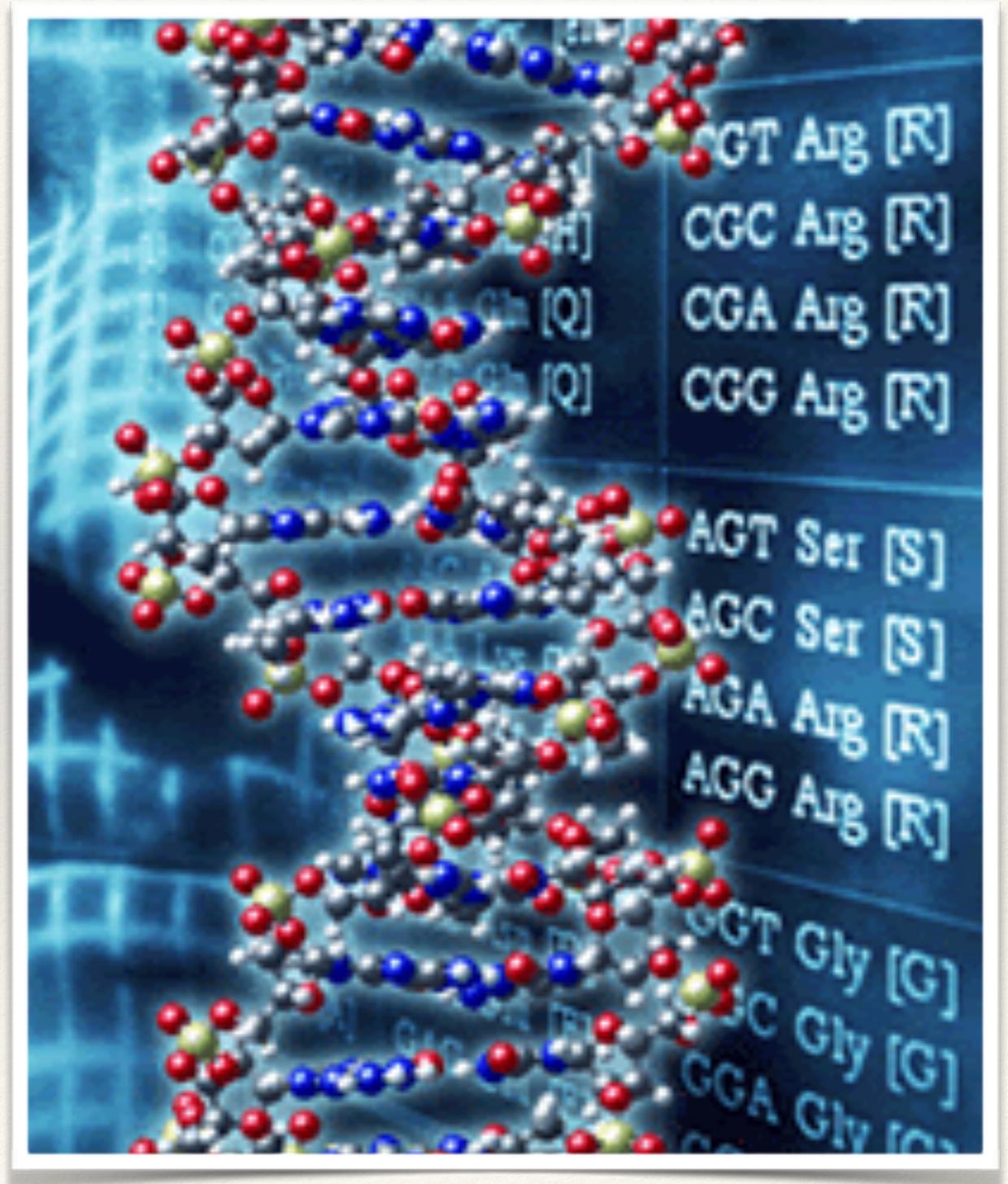


# Banche dati di sequenze

Le principali banche dati sono mantenute dagli Stati Uniti, dall'Europa e dal Giappone e sono in costante contatto

- ❖ NCBI (<http://www.ncbi.nlm.nih.gov/>)
- ❖ EBI (<http://www.ebi.ac.uk/>)
- ❖ DDBJ (<https://www.ddbj.nig.ac.jp>)

inoltre ci sono molte banche dati specializzate



---

# Il progetto ENCODE in e!Ensembl

---

- ❖ The Encyclopedia of DNA Elements (ENCODE) è un progetto pubblico lanciato dal centro di ricerca genomica americano
- ❖ L'obiettivo di ENCODE è di costruire una lista completa degli elementi funzionali del genoma umano, inclusi RNA ed elementi regolatori.
- ❖ e!Ensemble è un software prodotto dal centro europeo di bioinformatica che serve ad annotare in modo automatico, conservare e rappresentare dati genomici

---

# alla ricerca dei trascritti

---

- ❖ proviamo a cercare ora insieme i trascritti prodotti da alcuni geni usando come righello di riferimento la sequenza dei cromosomi umani sulla banca dati ENCODE utilizzando e!Ensembl
- ❖ vi ricordo che una definizione funzionale e moderna del concetto di gene potrebbe essere la seguente:

**“Un gene è una sequenza di DNA (i cui segmenti non sono necessariamente esclusivamente contigui) che produce sequenze di RNA e/o proteine ad esso correlate. Ogni gene è parte di una complessa rete di geni all’interno di un genoma e come parte di questa rete produce effetti osservabili al livello fenotipico. La sequenza di DNA di un gene e a volte anche la sua struttura vengono ereditati entrando a far parte della rete di geni del nuovo individuo nel quale contribuiscono a produrre un fenotipo.”**

*ref. Portin P & Wilkins A (2017) The Evolving Definition of the Term “Gene” GENETICS <https://doi.org/10.1534/genetics.116.196956>*

---

proviamo ad analizzare i trascritti associati a due noti geni  
espressi nel nostro cervello

---

tyrosine hydroxylase

myelin basic protein



# esempio 1: Tirosina idrossilasi

cerchiamo il gene codificante per la tirosina idrossilasi sulla mappa del genoma

- ❖ digiterete nella casella di ricerca “TH” oppure “Tyrosine hydroxylase”
- ❖ poi fra i risultati sceglierete il seguente

## [TH \(Human Gene\)](#)

**ENSG00000180176** 11:2163929-2171877:-1

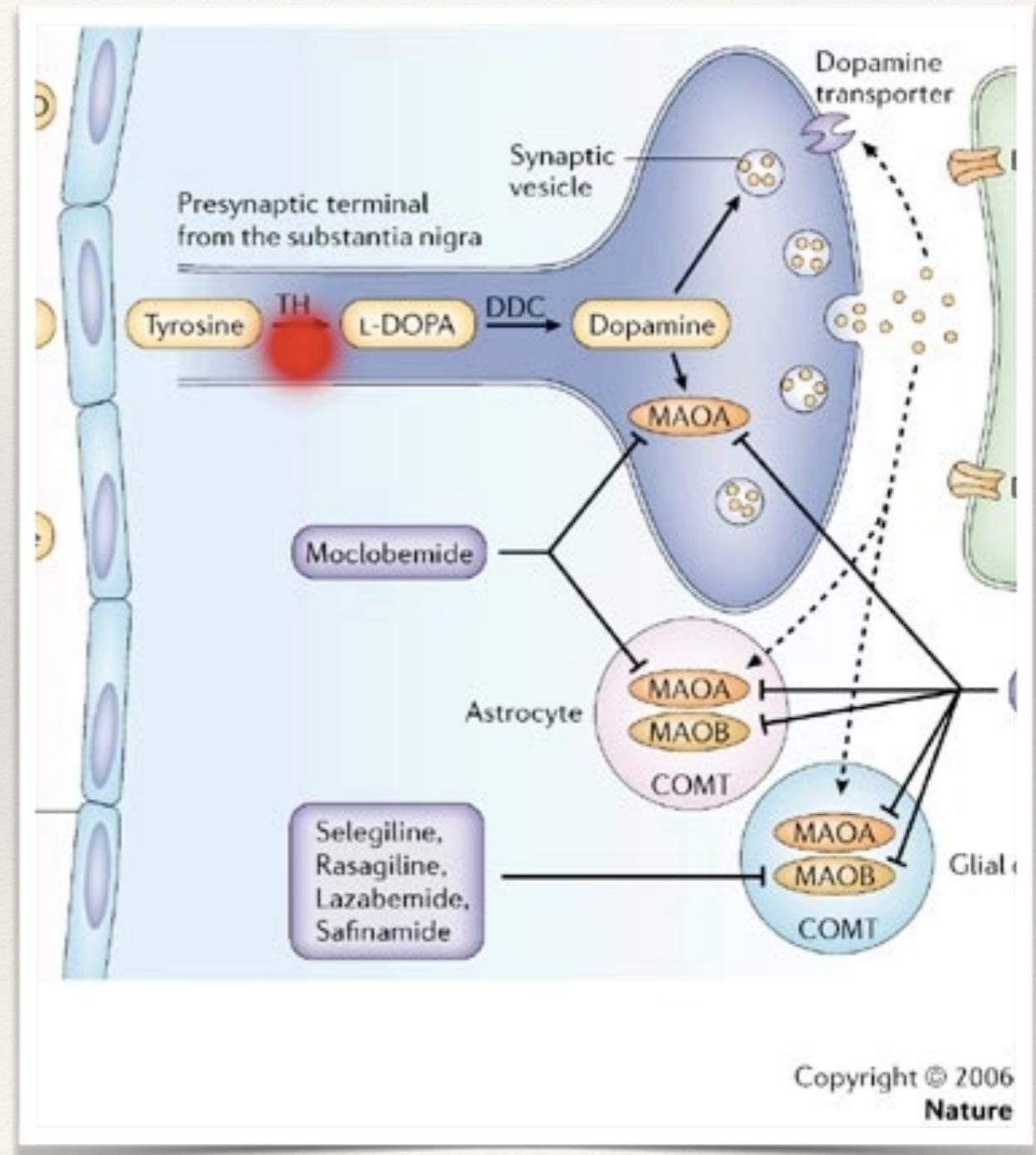
Tyrosine hydroxylase [Source:HGNC Symbol;Acc:HGNC:11782]

TYROSINE HYDROXYLASE; **TH** [\*191290] (MIM gene record; description: TYROSINE HYDROXYLASE; **TH**,) is an external reference matched to Gene ENSG00000180176

[Variant table](#) • [Phenotypes](#) • [Location](#) • [External Refs.](#) • [Regulation](#) • [Gene tree](#)

- ❖ e cliccherete su “Location” in basso al centro

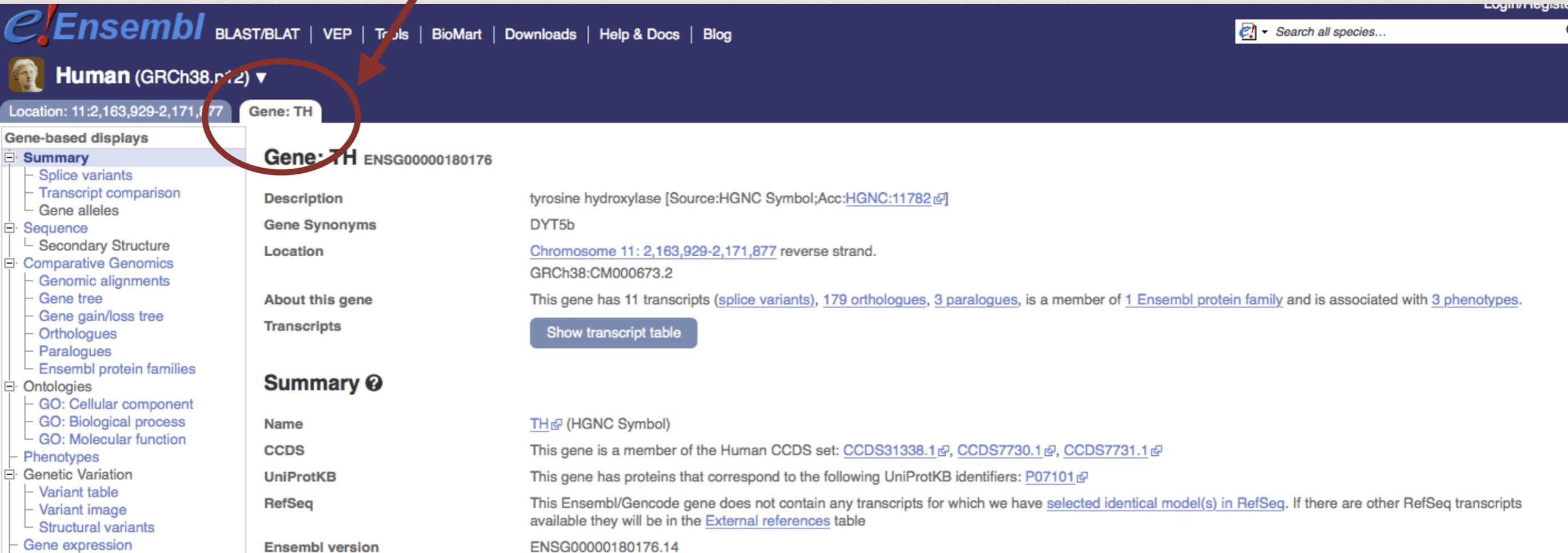
- ❖ .. e se non la trovate ecco il link



# esempio 1: Tiroosina idrossilasi

una volta trovata la pagina accertatevi di scegliere il tab denominato "Gene: TH"

la pagina apparirà come segue:



**Ensembl** BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Human (GRCh38.p12) ▼

Location: 11:2,163,929-2,171,877 Gene: TH

**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: Biological process
  - GO: Molecular function
- Phenotypes
- Genetic Variation
  - Variant table
  - Variant image
  - Structural variants
- Gene expression

**Gene: TH** ENSG00000180176

**Description** tyrosine hydroxylase [Source:HGNC Symbol;Acc:HGNC:11782]

**Gene Synonyms** DYT5b

**Location** [Chromosome 11: 2,163,929-2,171,877](#) reverse strand.  
GRCh38:CM000673.2

**About this gene** This gene has 11 transcripts ([splice variants](#)), [179 orthologues](#), [3 paralogues](#), is a member of [1 Ensembl protein family](#) and is associated with [3 phenotypes](#).

**Transcripts** [Show transcript table](#)

**Summary**

**Name** [TH](#) (HGNC Symbol)

**CCDS** This gene is a member of the Human CCDS set: [CCDS31338.1](#), [CCDS7730.1](#), [CCDS7731.1](#)

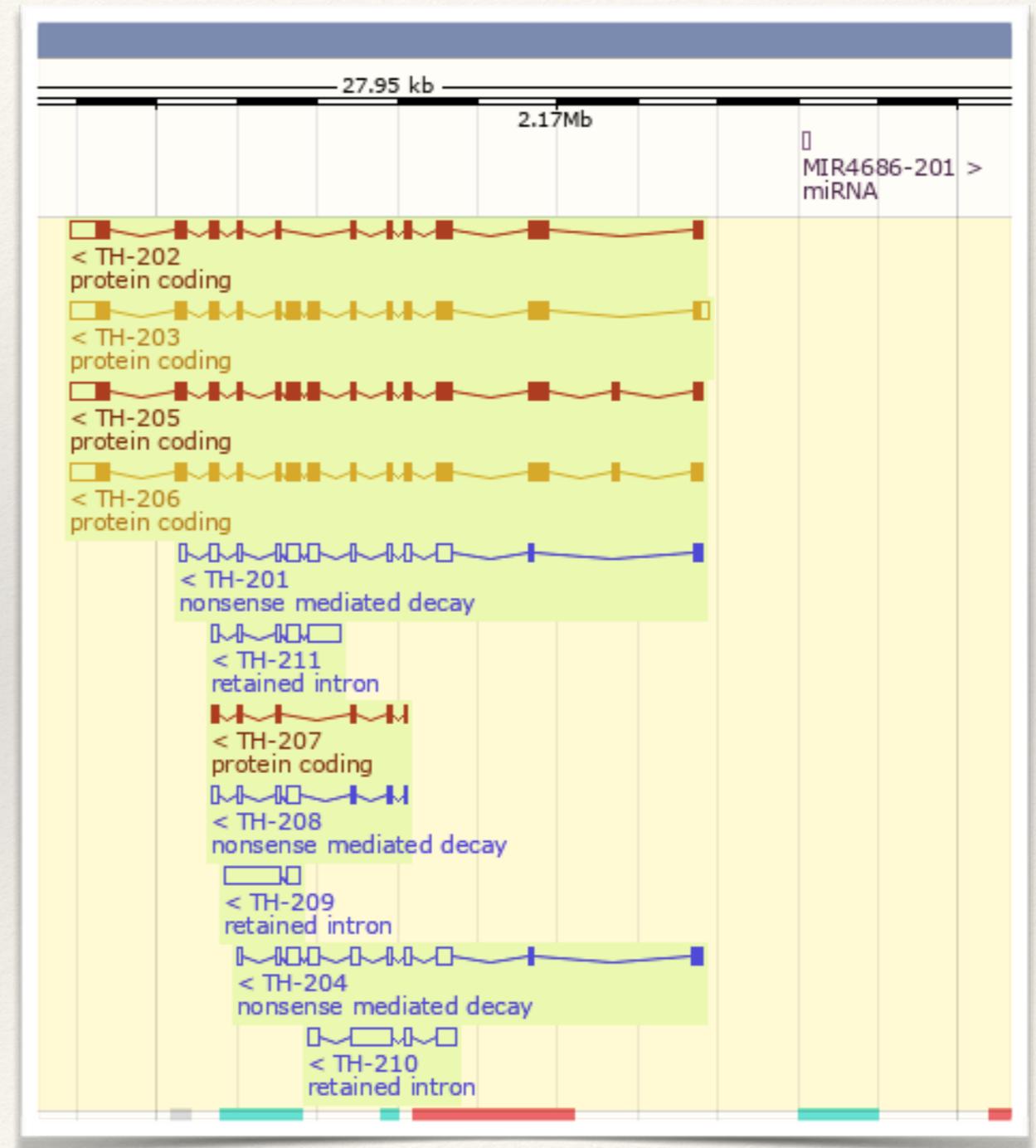
**UniProtKB** This gene has proteins that correspond to the following UniProtKB identifiers: [P07101](#)

**RefSeq** This Ensembl/Gencode gene does not contain any transcripts for which we have [selected identical model\(s\) in RefSeq](#). If there are other RefSeq transcripts available they will be in the [External references](#) table

**Ensembl version** ENSG00000180176.14

# esempio 1: Tirosina idrossilasi

- ❖ osservate con attenzione i dati che avete trovato e provate a rispondere alle seguenti domande:
  - ❖ quanti diversi trascritti trovate associati a questo gene?
  - ❖ sono trascritti tutti nella stessa direzione?
  - ❖ hanno tutti la stessa lunghezza?
  - ❖ in cosa differiscono i vari trascritti guardando l'allineamento sul genoma?
  - ❖ quanti degli RNA che trovate codificano proteine?
  - ❖ quali altre funzioni sono associate ai trascritti diverse da "protein coding"?



# esempio 1: Tiroosina idrossilasi

ora proviamo ad osservare l'espressione genica misurata per questo gene nei vari tessuti del nostro corpo, in diversi esperimenti:

- ❖ nella pagina dove siete, nel menu di sinistra selezionate “**Gene expression**”
- ❖ la pagina mostrerà una mappa del corpo umano e una rappresentazione grafica dei livelli di espressione di questo gene nei vari tessuti misurato da vari esperimenti

Location: 11:2,163,929-2,171,877 Gene: TH

**Gene: TH** ENSG00000180176

Description

Gene Synonyms

Location

About this gene

Transcripts

**Gene expression ?**

Showing 31 experiments:

♂

♀

🧠

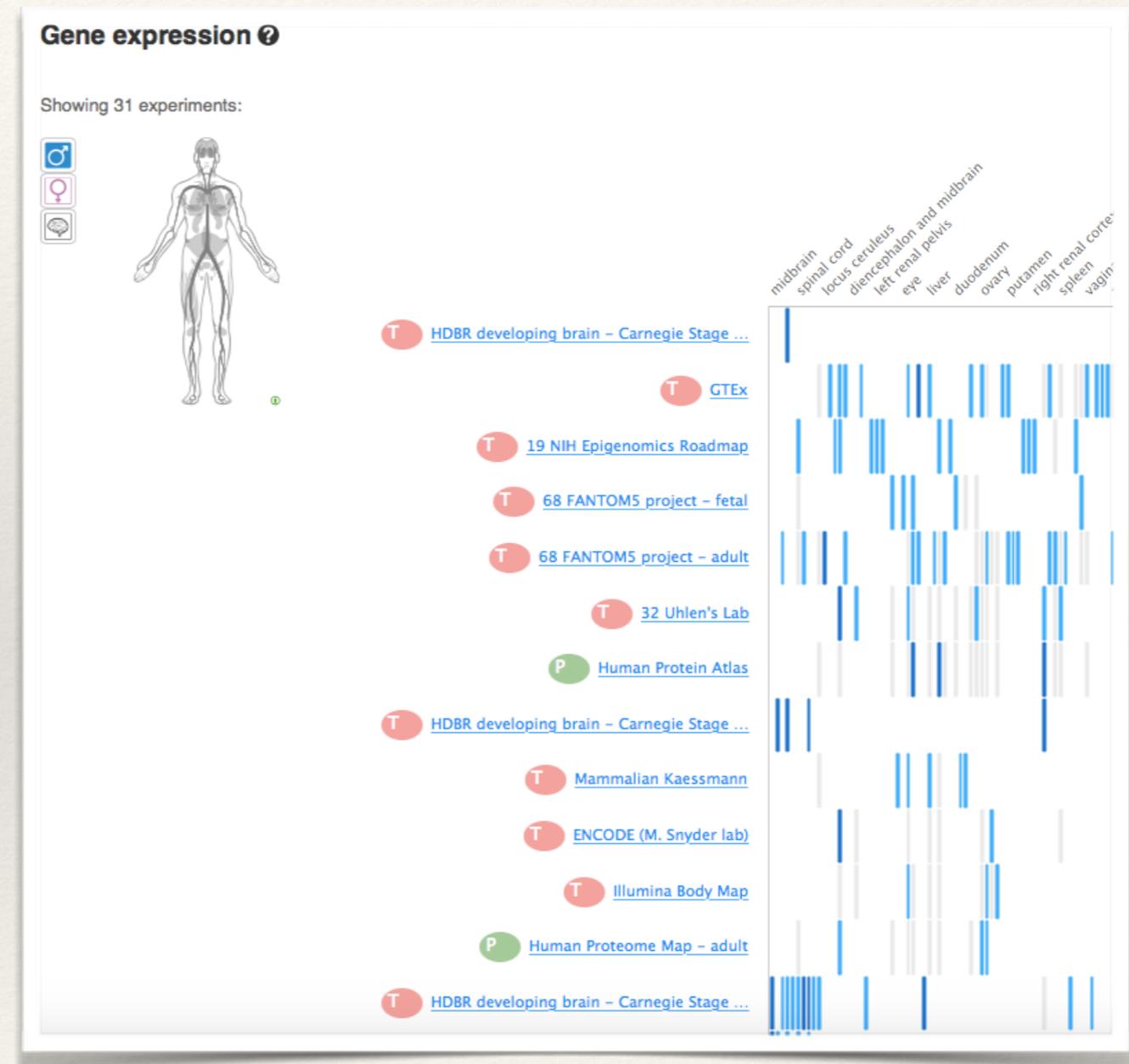
👤

⚙️ Configure this page

# esempio 1: Tirosoina idrossilasi

ora proviamo ad osservare l'espressione genica misurata per questo gene nei vari tessuti del nostro corpo, in diversi esperimenti:

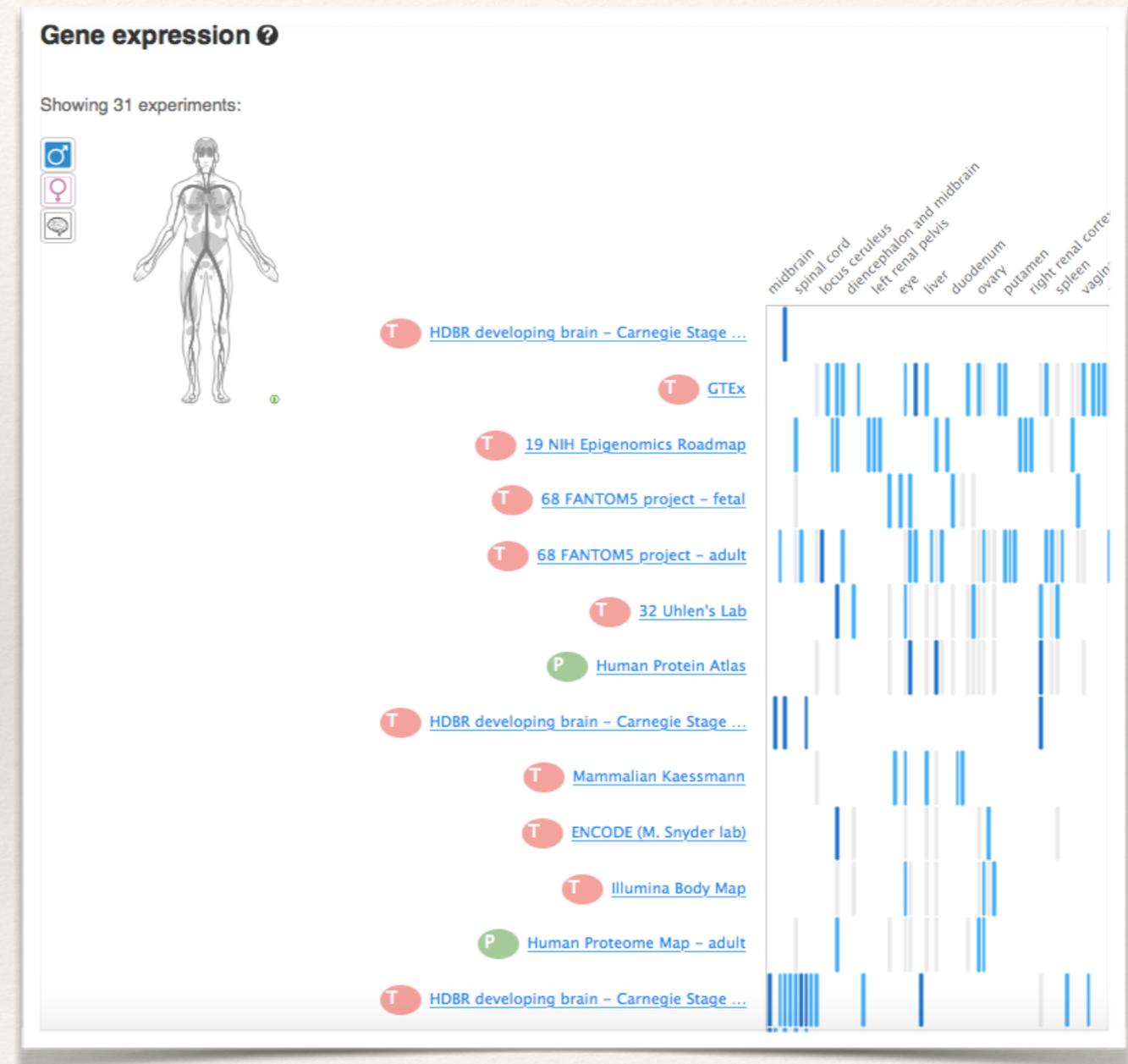
- ❖ nella pagina dove siete, nel menu di sinistra selezionate **“Gene expression”**
- ❖ la pagina mostrerà una mappa del corpo umano e una rappresentazione grafica dei livelli di espressione di questo gene nei vari tessuti misurato da vari esperimenti (potete anche ordinare i tessuti per livello di espressione, aiuta a leggere i risultati)



# esempio 1: Tirosoina idrossilasi

leggete con attenzione i risultati aiutandovi con la leggenda e provate a rispondere alle domande:

- ❖ in quali tessuti il gene risulta maggiormente espresso?
- ❖ in quali tessuti l'espressione è minima?



# esempio 1: Tirosoina idrossilasi

ora proviamo ad osservare la struttura dei trascritti nei vari tessuti del nostro corpo:

- ❖ cambiate tab e posizionatevi su “Location:..”
- ❖ dal menù di sinistra scegliete “Configure this page”

The screenshot displays the Ensembl genome browser interface for the TH gene region on Chromosome 11. The main content area shows a chromosome map with a red box indicating the region of interest. Below the map, the 'Region in detail' section displays a list of gene annotations, including protein-coding genes and lncRNAs. The left sidebar contains a navigation menu with options such as 'Chromosome summary', 'Region overview', 'Region in detail', and 'Configure this page'. Red circles and arrows highlight the 'Location: 11:2,162,275-2,174,470' field and the 'Configure this page' button.

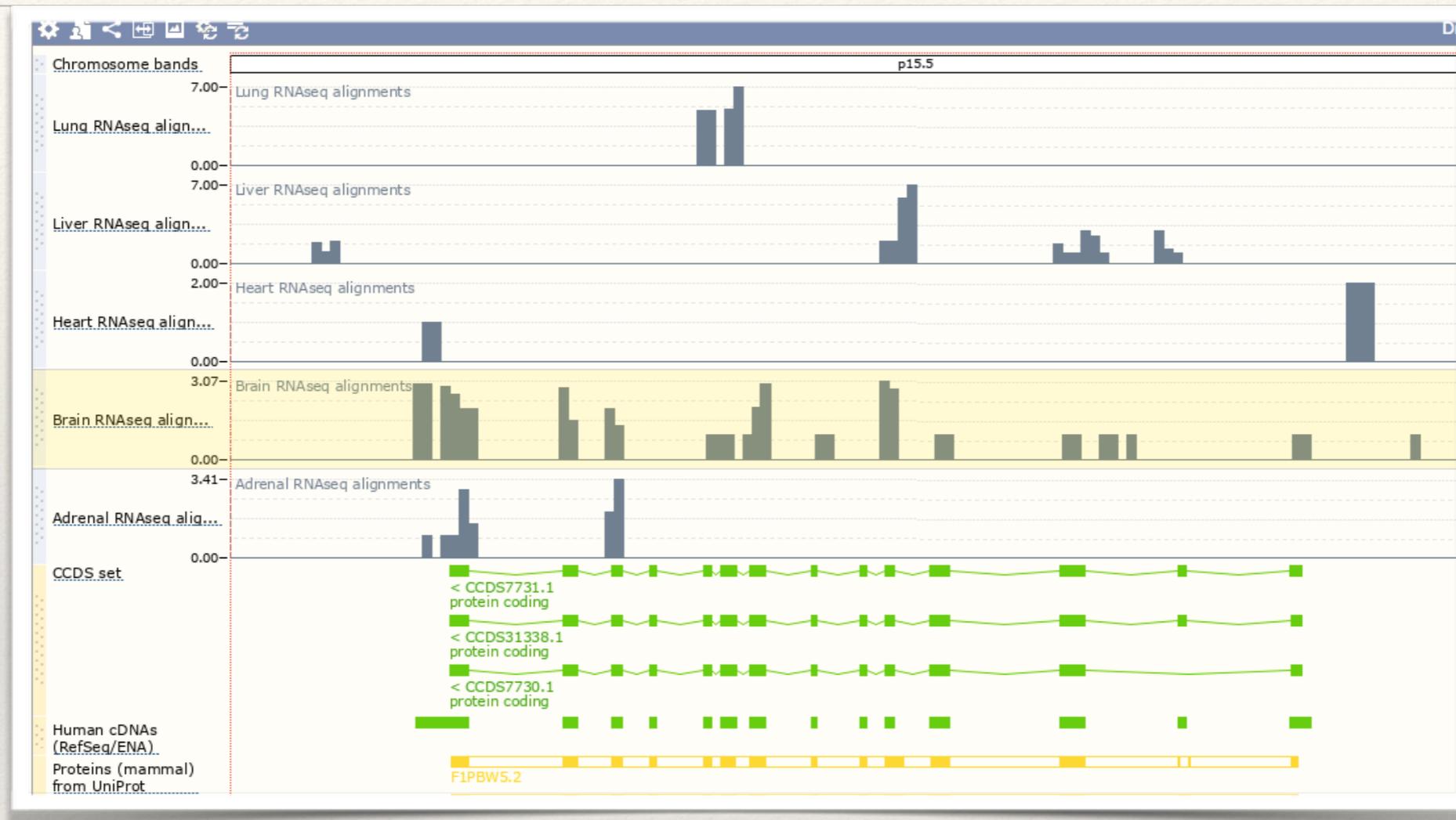
# esempio 1: Tiroosina idrossilasi

- ❖ all'interno del menù di configurazione scegliete "RNASeq models" sotto la categoria "Genes and Transcripts"
- ❖ nella pagina che si aprirà a destra attiverete i "BAM files" cliccando sul quadratino indicato dalla freccia
- ❖ e poi cliccherete sui quadratini corrispondenti ai tessuti che vi interessano, che appariranno così blu (ad esempio alcuni in cui il gene risulta molto espresso o pochissimo espresso)

The screenshot shows a web interface for configuring genomic tracks. On the left, a sidebar lists various track categories. The 'Genes and transcripts' category is expanded, and 'RNASeq models' is selected. On the right, the 'Genes and transcripts' configuration panel is visible, showing 'RNASeq models' selected under the 'Filter by' dropdown (set to 'Human BodyMap 2.0'). Below this, a 'Key' section shows 'Shown' as a dark blue square and 'Hidden' as a light blue square. At the bottom right, a 'Default style' table is shown, with a red circle highlighting the 'BAM files' column and a red arrow pointing to it from the text 'Gene models BAM files'. The table shows expression levels for various tissues.

Tissue	Gene models	Chromatin	BAM files
Adipose	0	1	1
Adrenal	0	1	1
Blood	0	1	1
Brain	0	1	1
Breast	0	1	1

# esempio 1: Tirosina idrossilasi

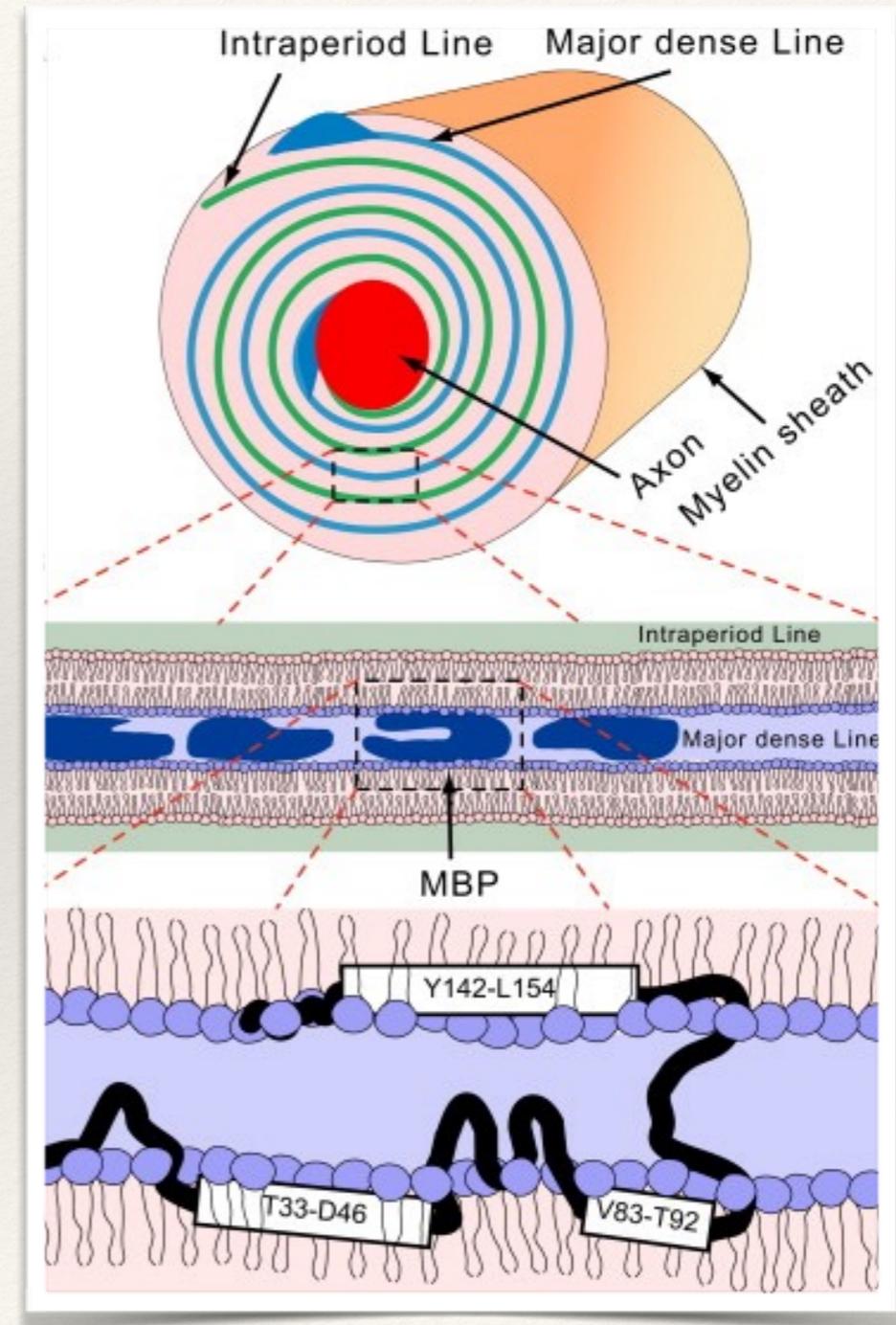


leggete con attenzione i risultati e provate a rispondere alle domande:

- ❖ la struttura dei trascritti è uguale o diversa nei vari tessuti che avete esaminato?
- ❖ in quale tessuto in particolare i trascritti espressi contengono tutti gli introni caratteristici del trascritto codificante e della proteina?

# esempio 2: myelin basic protein

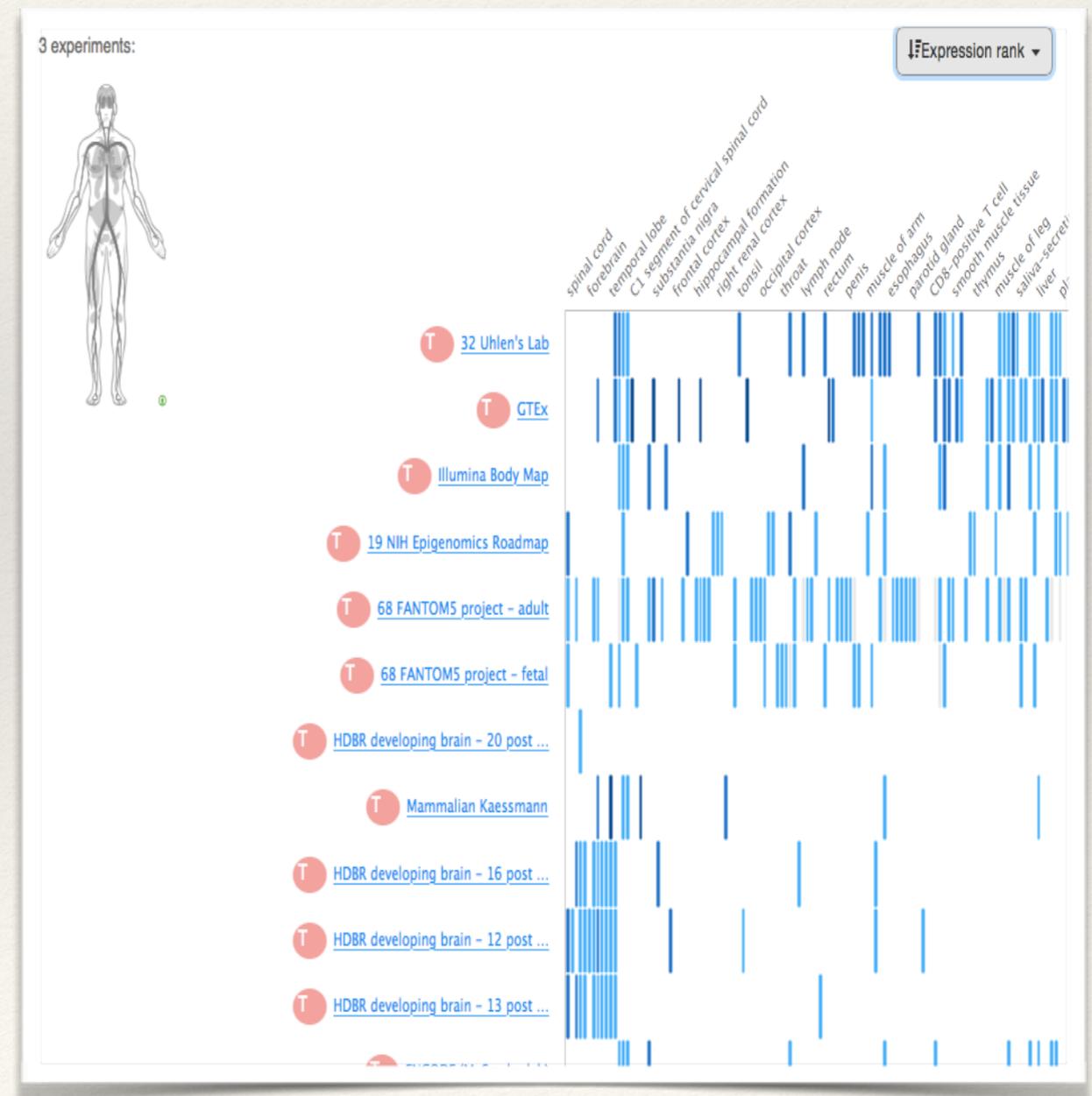
- ❖ cerchiamo il gene codificante per la “myelin basic protein” o “MBP” sulla mappa del genoma effettuando la ricerca come prima
- ❖ .. e se non la trovate ecco anche in questo caso il link
- ❖ Quante diversi trascritti trovate per questo gene?
- ❖ quali funzioni svolgono questi trascritti? (elencate le funzioni)



# esempio 2:myelin basic protein

anche in questo caso andremo ad esaminare il **livello di espressione** in vari tessuti e la **struttura dei trascritti** nei vari tessuti

- ❖ per vedere il livello di espressione come prima:
  - ❖ sceglieremo il tab “Gene:MBP”
  - ❖ cliccheremo nel menù a sinistra su “gene expression”
  - ❖ leggeremo con attenzione i risultati



# esempio 2: myelin basic protein

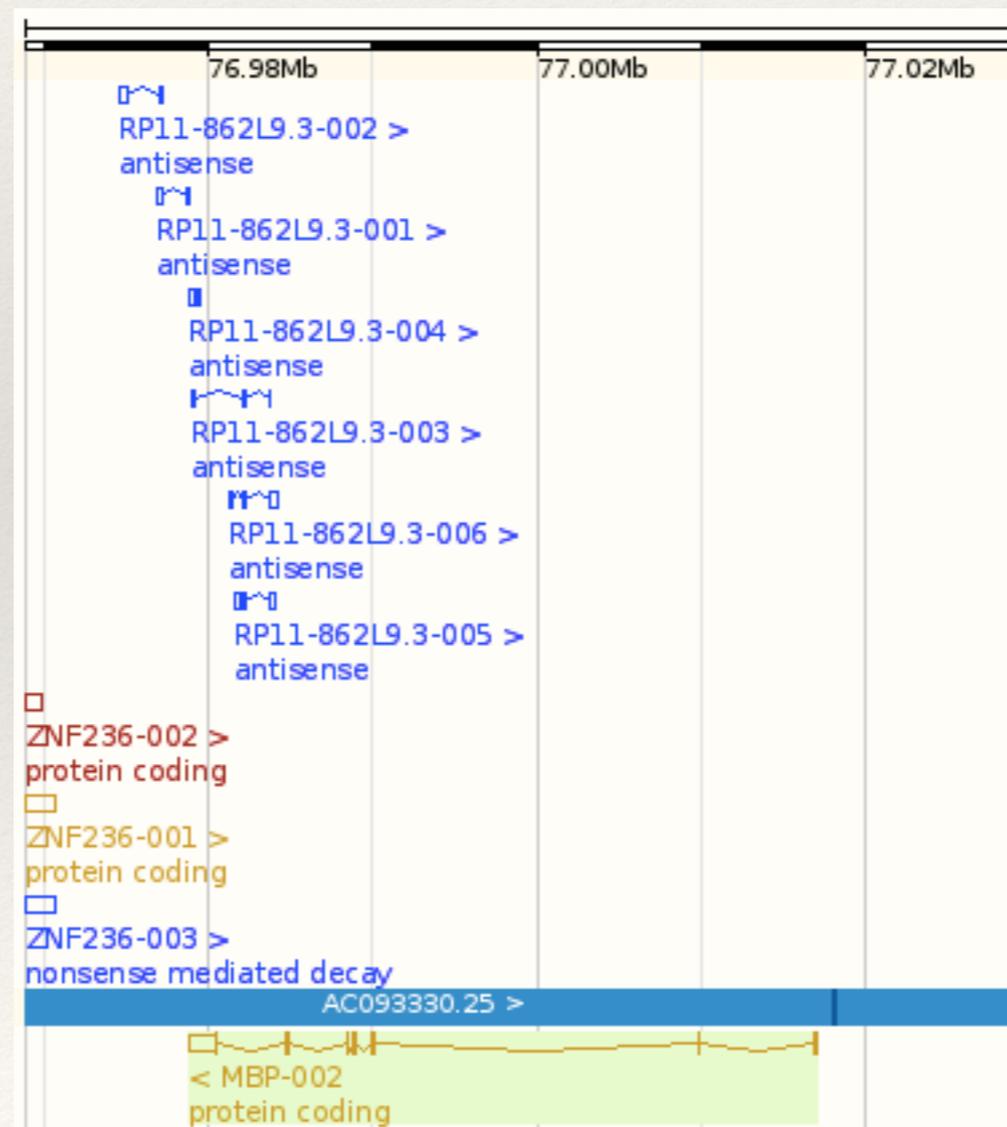
anche in questo caso andremo ad esaminare il **livello di espressione** in vari tessuti e la **struttura dei trascritti** nei vari tessuti

- ❖ per vedere la struttura dei trascritti come prima:
  - ❖ sceglieremo il tab “Location:..”
  - ❖ cliccheremo nel menù a sinistra su “Configure this page”
  - ❖ e poi su “RNASeq Models”
  - ❖ attiveremo quindi i “BAM files” cliccando sul quadratino
  - ❖ e sceglieremo i tessuti di interesse facendo diventare blu i quadratini nella colonna BAM files corrispondenti ai tessuti che vogliamo vedere



# esempio 2: myelin basic protein

- ❖ ci sono dei piccoli RNA antisenso che mappano all'inizio del gene



---

# a cosa servono gli RNA antisenso?

---

sono stati osservati sperimentalmente due effetti dei piccoli RNA noti e sequenziati:

- ❖ possono interagire con delle proteine e modificarne la funzione
- ❖ possono legarsi ad altri RNA e regolarne l'espressione

---

# i piccoli RNA che si legano ai trascritti

---

- ❖ possono legarsi ad altri RNA e regolarne l'espressione
  - ❖ ne troviamo di "cis-encoded" -> sono uguali ma in senso opposto al gene che regolano)
  - ❖ e di "trans-encoded" sRNAs -> sono codificati in una porzione del genoma diversa da quella del trascritto che regolano

---

# piccoli RNA: 20-25 basi

---

- ❖ come facciamo a scoprirne l'esistenza?
- ❖ come facciamo a trovare i "trans-endodod" RNA?

---

# small RNA: quante funzioni!

---

- ❖ una classe di piccoli RNA chiamati siRNA è principalmente coinvolta nel controllo dell'espressione genica dei trascritti con sequenza complementare che risulta nella degradazione dell'RNA complementare
- ❖ ma i siRNA sono anche coinvolti in meccanismi antivirali
- ❖ e contribuiscono a determinare la struttura del DNA genomico

# repressione dell'espressione

- ❖ meccanismo di regolazione dell'espressione
- ❖ potete immaginare qualche applicazione biotecnologica?

