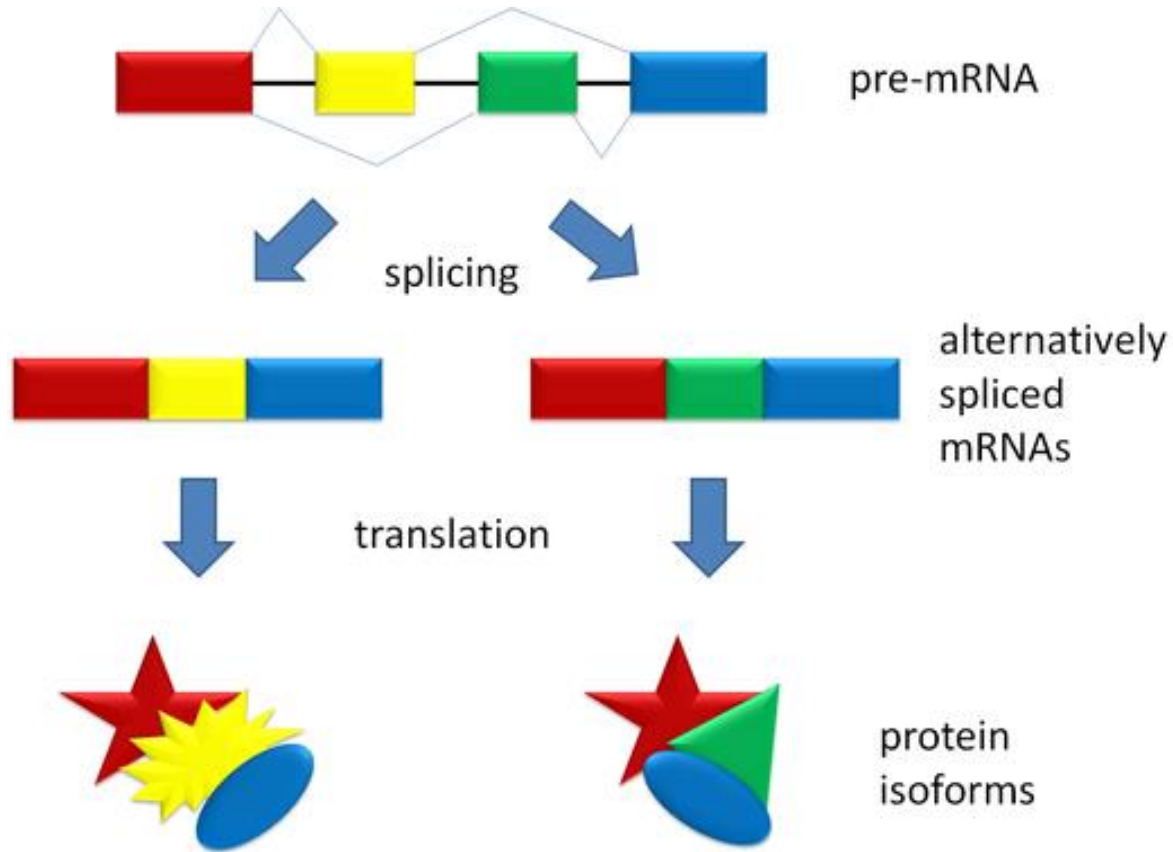


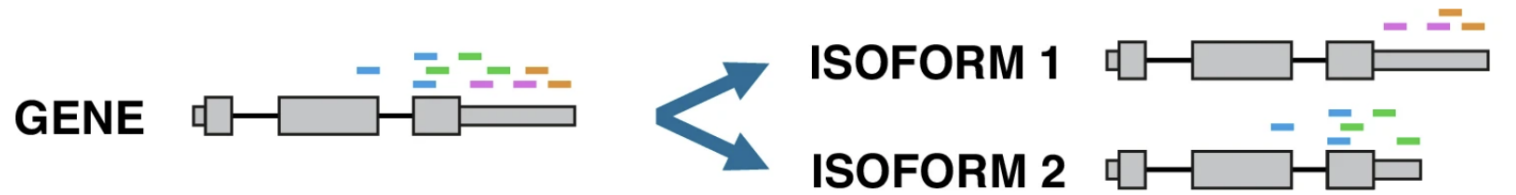
# What is an isoform?



Each person has about **20,000** coding genes

x5

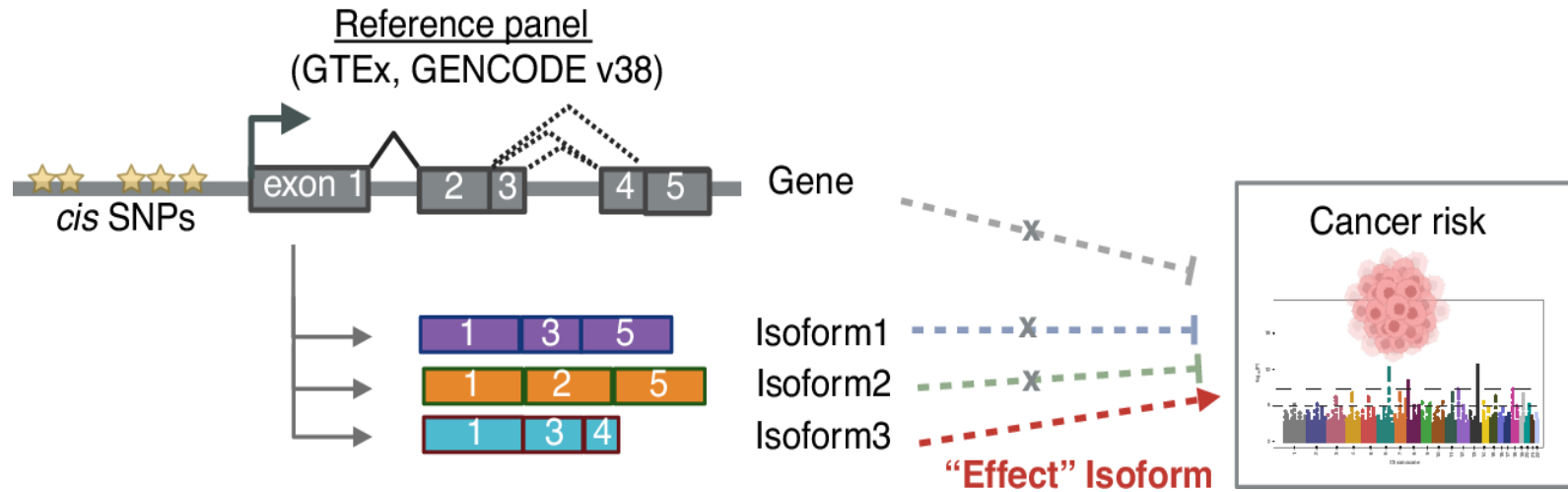
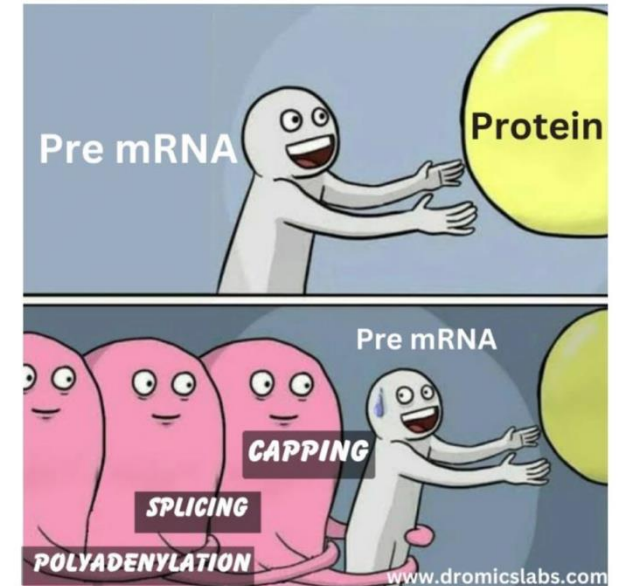
Proteins are essential for life, and the human body has over **100,000** different proteins performing many different functions.



	Cell 1	Cell 2	...	Cell N
Gene A	8	4	...	6
Gene B	2	0	...	3
Gene C	0	2	...	1
...			...	
Gene X	5	7	...	6

		Cell 1	Cell 2	...	Cell N
Gene A	<b>Iso 1</b>	7	1	...	0
	<b>Iso 2</b>	1	3	...	6
Gene B	<b>Iso 1</b>	2	0	...	3
...					
Gene X	<b>Iso 1</b>	3	3	...	4
	<b>Iso 2</b>	2	4	...	2

Are isoforms good for us?



## ALGORITHM ;)

```
assign random values to all  $f(i)$ 
while not converged do
  E-step:
  initialize all  $n(j)$  to 0
  for each read  $r$  do
     $\text{sum} = \sum_{j:w_{r,j}>0} w_{r,j} f(j)$ 
    for each isoform  $j$  with  $w_{r,j} > 0$  do
       $n(j)+ = w_{r,j} f(j) / \text{sum}$ 
    end for
  end for
  M-step:
   $s = \sum_j n(j) / (l(j) - \mu + 1)$ 
  for each isoform  $j$  do
     $f(j) = \frac{n(j) / (l(j) - \mu + 1)}{s}$ 
  end for
end while
```

# DPI calculation (Delta Percent Isoform)

$$DPI = \frac{\sum |\Delta PI_i|}{2}$$

2 DPI= all of the isoforms pi  
Delts PI = absolute value of pi's

**pos** → **acceptor (3' splice site)**

**neg** → **donor (5' splice site)**

2 most found isoforms are shown, where are the others??

gid	p_val	dpi	pos_iso_1	pos_iso_2	pos_iso_1_dpi	pos_iso_2_dpi	neg_iso_1	neg_iso_2	neg_iso_1_dpi	neg_iso_2_dpi	adj_p_val	gname
<b>ENSG0000079150</b>	0.0014534999464932115	<b>28.125</b>	ENST00000419184.1	ENST00000470945.1	14.543270111083984	6.911057949066162	ENST00000424785.6	ENST00000464248.1	-23.4375	-4.6875	0.011643364380367663	FKBP7

$$\sum |\Delta PI| = 14.54327 + 6.91106 + 23.4375 + 4.6875 = 49.57933$$

Now divide by 2:

$$DPI = \frac{49.57933}{2} = 24.79$$

$$\sum |\Delta PI_{all}| = 2 \times DPI = 2 \times 28.125 = 56.25$$

$$\sum |\Delta PI| = 49.57933 + 6.67067 = 56.25$$

$$DPI = \frac{56.25}{2} = 28.125$$

# Motivation

## Standard DPI / isoform analyses:

- Typically report only the **top 2 positive and top 2 negative isoforms**

## What if I am interested in other isoforms of genes??

- Remaining isoforms are often treated as **noise or biologically negligible**

This study introduces **Dpi<sub>3</sub>**, a quantitative metric defined as:

$$\mathbf{Dpi}_3 = (2 \times \mathbf{dpi}) - \sum |\text{dominant isoform DPI values}|$$

<u>Gen</u>	<u>Dpi_3</u>	<u>Significance</u>
<b>CFLAR</b>	~30.9	Very powerful alternative splicing/apoptotic switch
<b>CALD1</b>	~16.5	Cytoskeletal reorganization
<b>SEC31A</b>	~16.0	ER–Golgi transport regulation
<b>TPM1</b>	~10.2	Actin filament dynamics
<b>RTKN</b>	~11.2	Rho-GTPase signal

# Output file called “third\_iso.csv”

gid	p_val	Dpi_3	gname
ENSG000000079150	0.0014534999464932115	21.35	FKBP7

$\geq 10$

- Write this gene to the visualization code
- Go to this web page:  
<https://freese.gitbook.io/swan/tutorials/visualization>



**civciv**   
@ciivciivv

**Yazılımı** bırakıyorum.....  
çok sıkıldım çok fazla sorun var  
yetişemiyorum

19:01 · 16 Haz 15

mRNA and Protein(s)

1. [NM\\_001135212.2](#) → [NP\\_001128684.1](#) **peptidyl-prolyl cis-trans isomerase FKBP7 isoform b precursor**

[See identical proteins and their annotated locations for NP\\_001128684.1](#)

Status: REVIEWED

Description	Transcript Variant: This variant (2) uses an alternate in-frame splice site in the coding region, compared to variant 1. The resulting isoform (b) is 1 amino acid shorter than isoform a.
Source sequence(s)	<a href="#">AA447628</a> , <a href="#">AC009948</a> , <a href="#">AY353086</a> , <a href="#">BC009711</a> , <a href="#">DA020387</a>
Consensus CDS	<a href="#">CCDS46462.1</a>
UniProtKB/Swiss-Prot	<a href="#">Q9Y680</a>
Related	<a href="#">ENSP00000415486.2</a> , <a href="#">ENST00000434643.6</a>

Conserved Domains (3) [summary](#)

<a href="#">pfam00254</a> Location:46 → 140	FKBP_C; FKBP-type peptidyl-prolyl cis-trans isomerase
<a href="#">pfam13499</a> Location:151 → 213	EF-hand_7; EF-hand domain pair
<a href="#">c108302</a> Location:151 → 212	EFh; EF-hand, calcium binding motif; A diverse superfamily of calcium sensors and calcium signal modulators; most examples in this alignment model have 2 active canonical EF hands. Ca2+ binding induces a conformational change in the EF-hand motif, leading to ...

2. [NM\\_001410972.1](#) → [NP\\_001397901.1](#) **peptidyl-prolyl cis-trans isomerase FKBP7 isoform c precursor**

[See identical proteins and their annotated locations for NP\\_001397901.1](#)

Status: REVIEWED

Source sequence(s)	<a href="#">AC009948</a> , <a href="#">AK299215</a>
Consensus CDS	<a href="#">CCDS92911.1</a>
UniProtKB/TrEMBL	<a href="#">B4DRE2</a>
Related	<a href="#">ENSP00000509396.1</a> , <a href="#">ENST00000464248.1</a>

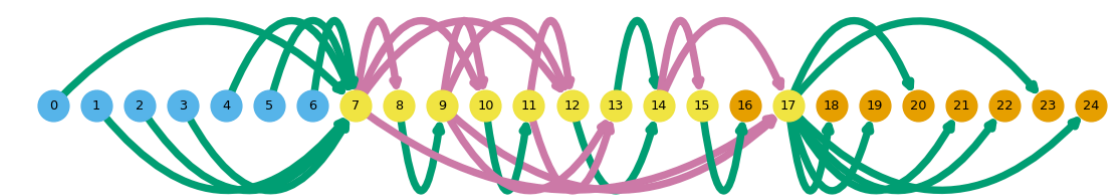
3. [NM\\_181342.3](#) → [NP\\_851939.1](#) **peptidyl-prolyl cis-trans isomerase FKBP7 isoform a precursor**

[See identical proteins and their annotated locations for NP\\_851939.1](#)

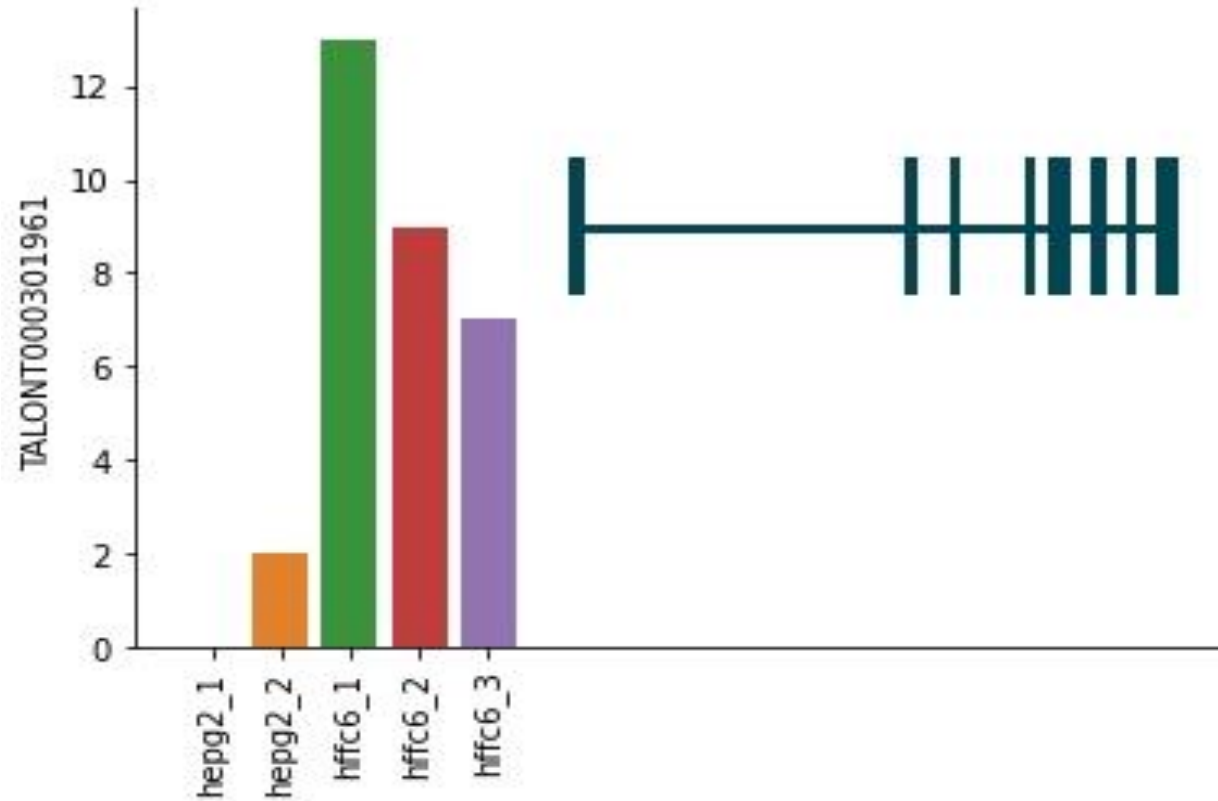
Status: REVIEWED

Description	Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (a).
Source sequence(s)	<a href="#">AA447628</a> , <a href="#">AC009948</a> , <a href="#">BC009711</a> , <a href="#">DA020387</a>
Consensus CDS	<a href="#">CCDS2280.1</a>
UniProtKB/Swiss-Prot	<a href="#">Q4ZG70</a> , <a href="#">Q6V3B2</a> , <a href="#">Q86U65</a> , <a href="#">Q96DA4</a> , <a href="#">Q9Y680</a> , <a href="#">Q9Y6B0</a>
Related	<a href="#">ENSP00000413152.2</a> , <a href="#">ENST00000424785.7</a>

Conserved Domains (2) [summary](#)



Swan assigns a node to each unique splice site.



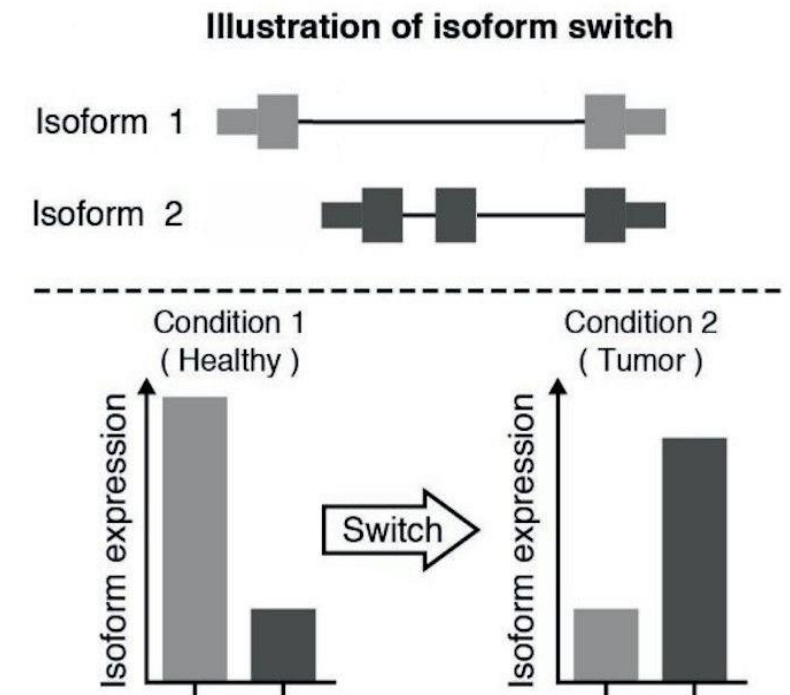
TALONT00311961L isoform shows:

**Very low** expression in hepg2\_1 and hepg2\_2 cell lines

**High expression** in hffc6\_2 and hffc6\_3 cell lines

This shows the **change in usage** across different cell lines → **isoform switch potential**

Gene is the same  
Isoform is the same  
Isoform function is different! Why?





- Bioinformatics Tutorials -> <https://ngs101.com/tutorials/>

For advanced isoform analysis ->

<https://bioconductor.org/packages/release/bioc/vignettes/IsoformSwitchAnalyzeR/inst/doc/IsoformSwitchAnalyzeR.html#importing-data-from-salmon-via-tximeta>

Github101 -> <https://github.com/nuryagli/git-github-101-networking>

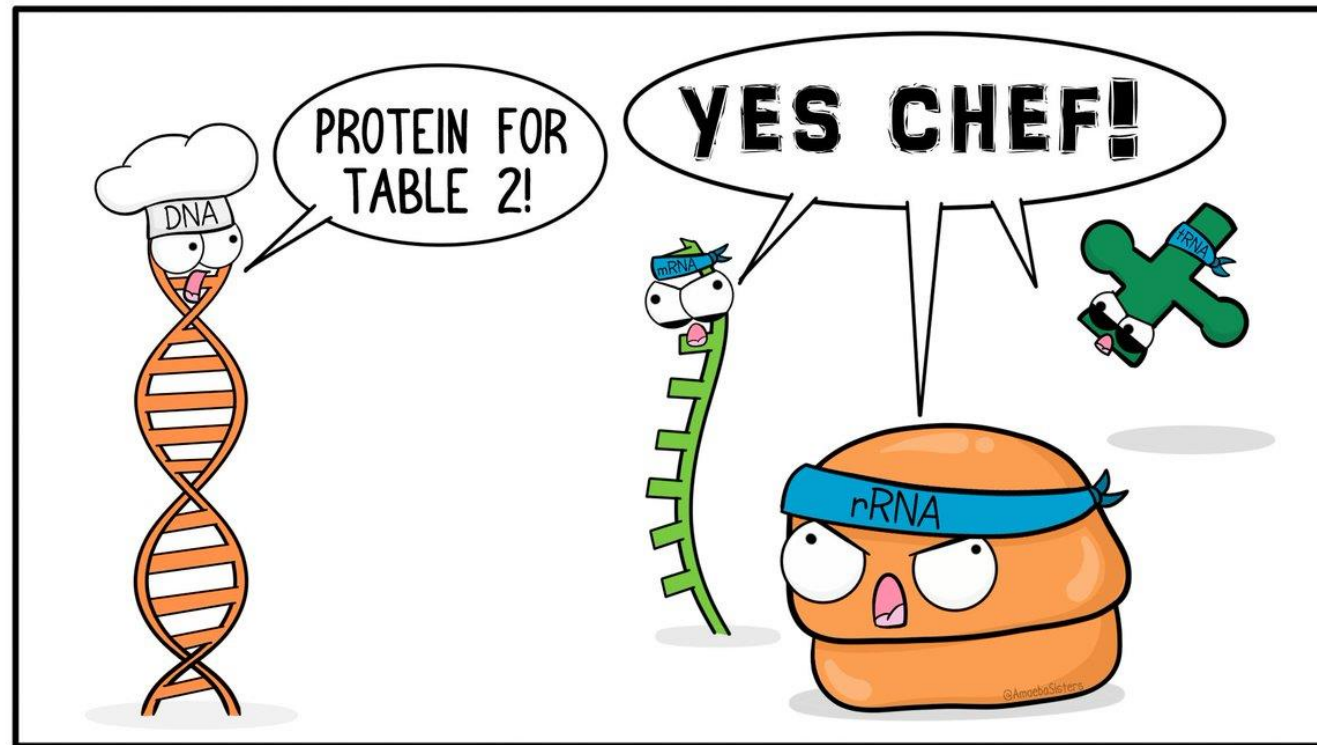


# Thanks for attending!!

## Any questions???

Paramecium Parlor

@AmoebaSisters



RNA: If you can't help make proteins,  
get out of the kitchen.