

# Análisis de Anotaciones GO de la inversión U8

CODE ▾

Mercè Merayo Pastor

2024-12-11

- Introducción
- Preparación de Datos
- Resultados
  - Procesos Biológicos dentro de la inversión  $U_8$
  - Funciones Moleculares
  - Componentes Celulares
  - Regiones flanqueantes de los puntos de rotura
    - Tablas región flanqueante proximal
    - Tablas región flanqueante distal
    - Genes que contienen algun punto de rotura
  - Exploración de genes relacionados con la adaptación térmica en la inversión  $U_8$

## Introducción

Este documento analiza las anotaciones GO obtenidas del archivo `.gaf` del genoma de referencia  
Genome assembly UC Berk\_Dsub\_1.0 de *Drosophila subobscura*.

Se agrupan los **Procesos biológicos**, **Funciones moleculares**, y **Componentes celulares** más frecuentes para la inversión  $U_8$  tanto en las regiones flanqueantes de los puntos de rotura encontrados con Breakdancer en la cepa OF58, como en la región interna de la inversión.

## Preparación de Datos

SHOW

Agregar las descripciones de los términos GO (Gene Ontology) desde la base de datos `GO.db`.

SHOW

Carga de las anotaciones del genoma de referencia de *Drosophila subobscura* a partir de la descarga de la tabla `.tsv` desde  
[[https://www.ncbi.nlm.nih.gov/datasets/gene/GCF\\_008121235.1/](https://www.ncbi.nlm.nih.gov/datasets/gene/GCF_008121235.1/)  
([https://www.ncbi.nlm.nih.gov/datasets/gene/GCF\\_008121235.1/](https://www.ncbi.nlm.nih.gov/datasets/gene/GCF_008121235.1/))].

SHOW

SHOW

SHOW

Filtrar las anotaciones por las tres categorías principales de Gene Ontology.

SHOW

Se definen los puntos de rotura la inversión  $U_8$ .

SHOW

SHOW

# Resultados

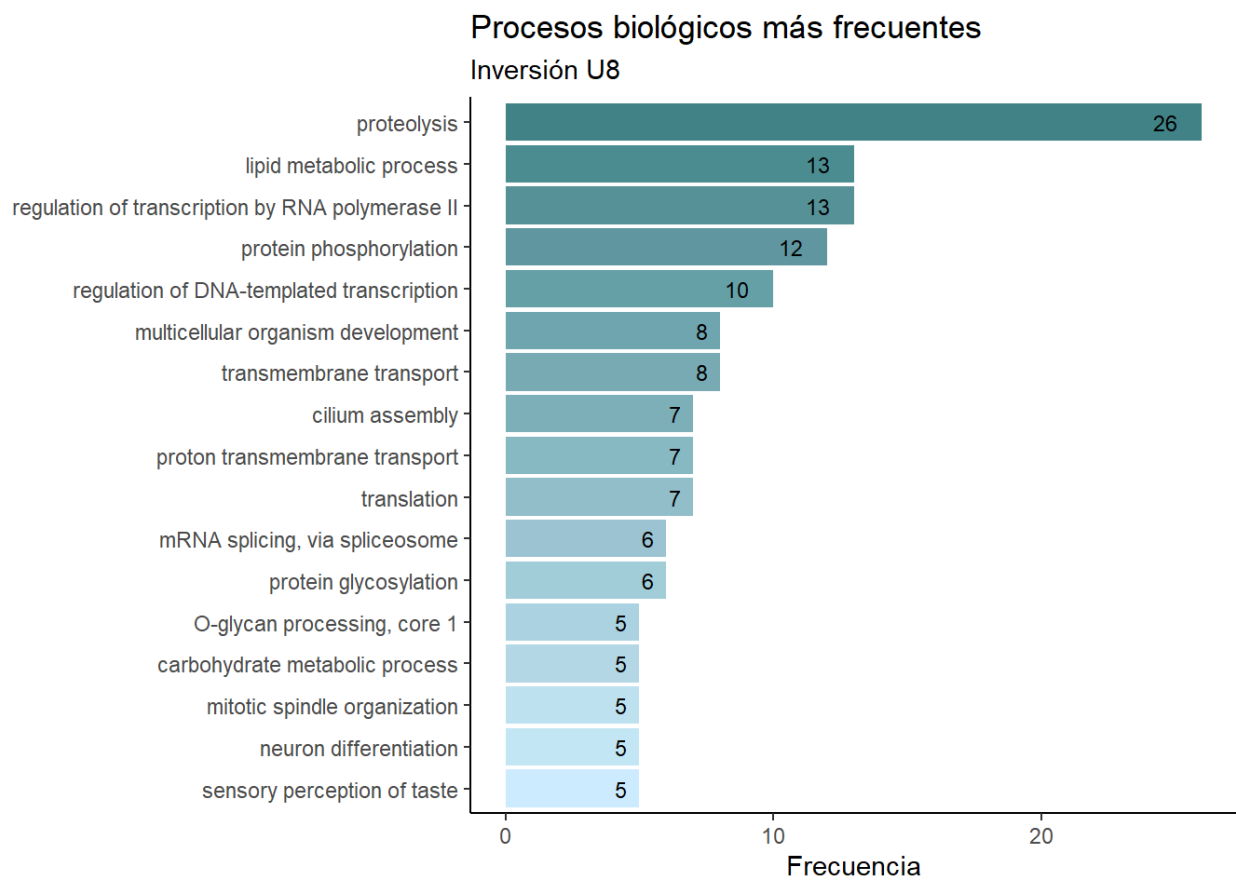
## Procesos Biológicos dentro de la inversión $U_8$

SHOW

[1] "Hay 410 procesos biológicos diferente en los genes de la región U8."

SHOW

SHOW



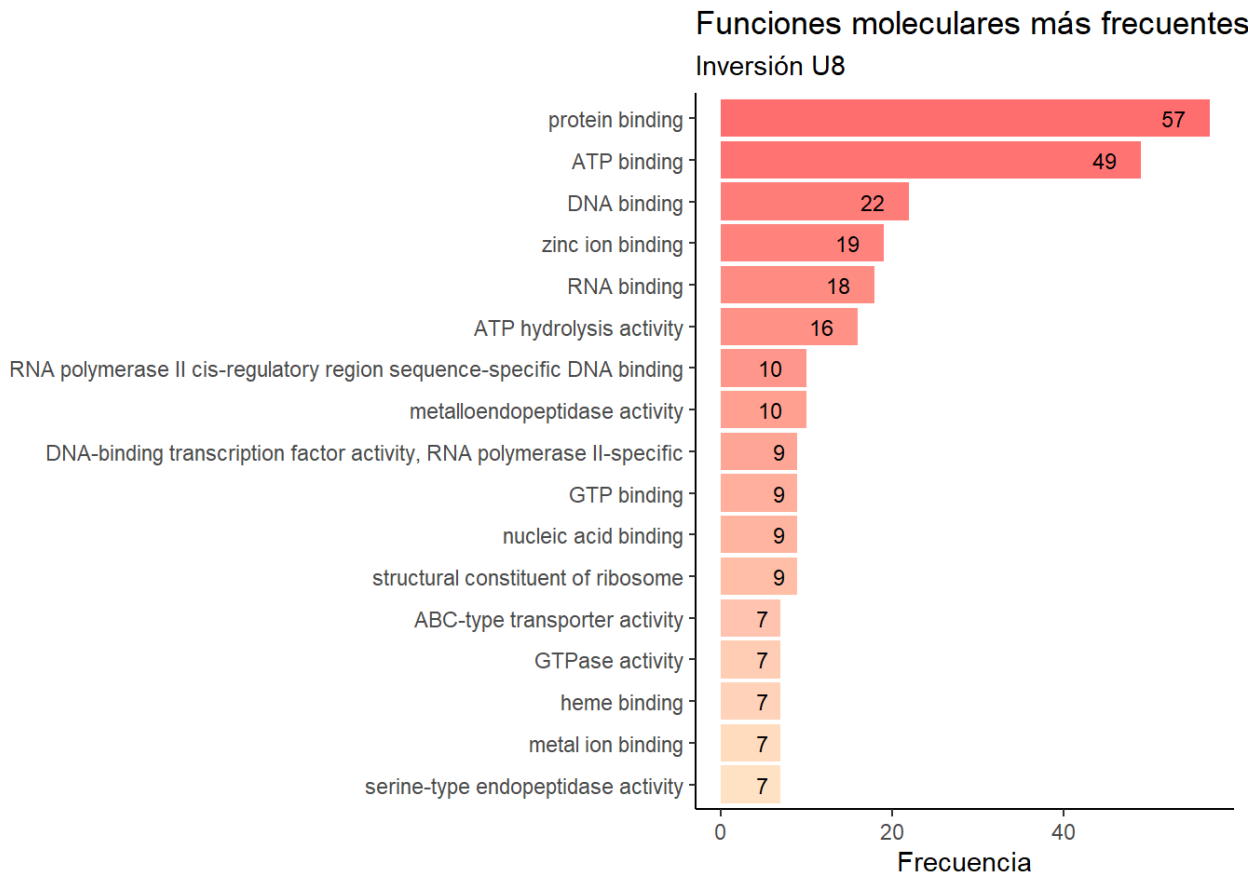
## Funciones Moleculares

SHOW

[1] "Hay 315 funciones moleculares diferentes en los genes de la región U\_8."

SHOW

SHOW



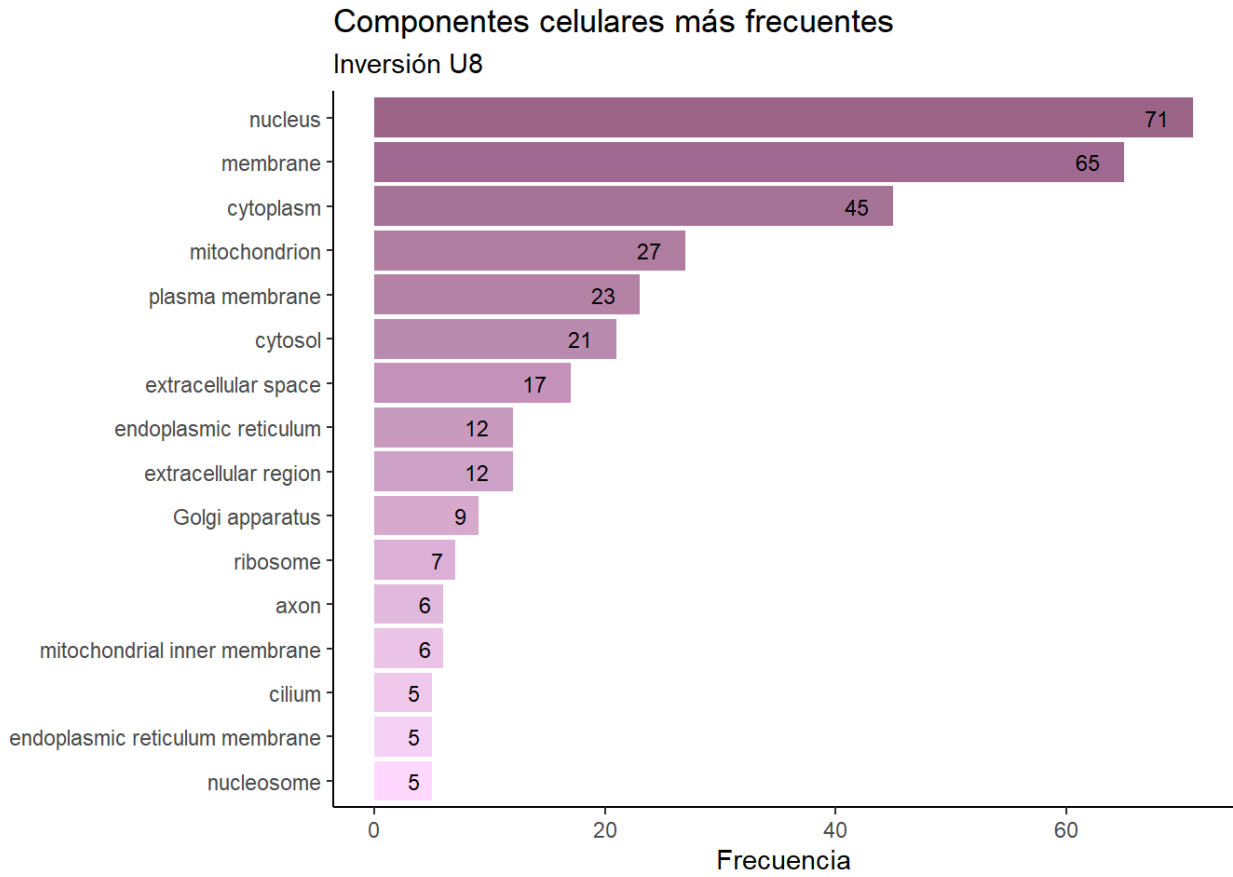
## Componentes Celulares

SHOW

[1] "Hay 227 componentes celulares diferentes en los genes de la región U\_8."

SHOW

SHOW



## Regiones flanqueantes de los puntos de rotura

### Tablas región flanqueante proximal

SHOW

SHOW

SHOW

Términos GO de los genes situados cerca del punto de rotura proximal

Gen / Nombre de la proteína	Coordenadas genómicas	Funciones Moleculares	Procesos Biológicos	Componentes Celulares
LOC117900743 - thyrostimulin beta-5 subunit	7770494-7771257	GO:0005179 (GO:0005179): hormone activity	GO:0007186 (GO:0007186): G protein-coupled receptor signaling pathway	GO:0005576 (GO:0005576): extracellular region GO:0005615 (GO:0005615): extracellular space GO:0005737 (GO:0005737): cytoplasm

Gen / Nombre de la proteína	Coordenadas genómicas	Funciones		Componentes Celulares
		Moleculares	Procesos Biológicos	
LOC117900742 - mitochondrial import inner membrane translocase subunit Tim23	7771342-7772481	GO:0008320 (GO:0008320): protein transmembrane transporter activity	GO:0030150 (GO:0030150): protein import into mitochondrial matrix	GO:0005743 (GO:0005743): mitochondrial inner membrane GO:0005744 (GO:0005744): TIM23 mitochondrial import inner membrane translocase complex
LOC117900741 - uncharacterized protein YJR142W isoform X1	7773180-7777052	GO:0044715 (GO:0044715): 8-oxo-dGDP phosphatase activity		

Tablas región flanqueante distal

SHOW

SHOW

SHOW

Términos GO de los genes situados cerca del punto de rotura proximal

Gen / Nombre de la proteína	Coordenadas genómicas	Funciones		Componentes Celulares
		Moleculares	Procesos Biológicos	
LOC117901454 - cilia- and flagella-associated protein 20 isoform X1	14719090-14720090		GO:0060271 (GO:0060271): cilium assembly GO:0060296 (GO:0060296): regulation of cilium beat frequency involved in ciliary motility GO:2000147 (GO:2000147): positive regulation of cell motility	GO:0031514 (GO:0031514): motile cilium GO:0036064 (GO:0036064): ciliary basal body
LOC117901453 - methionine aminopeptidase 1D, chloroplastic/mitochondrial	14720021-14721256	GO:0070006 (GO:0070006): metalloaminopeptidase activity	GO:0006508 (GO:0006508): proteolysis	

Gen / Nombre de la proteína	Coordenadas genómicas	Funciones Moleculares	Procesos Biológicos	Componentes Celulares
LOC117901760 - leucine-rich repeat-containing protein 15	14722075-14724843	GO:0005515 (GO:0005515): protein binding		GO:0005615 (GO:0005615): extracellular space GO:0031012 (GO:0031012): extracellular matrix
LOC117901184 - COMM domain-containing protein 5 isoform X1	14725873-14726911			GO:0005634 (GO:0005634): nucleus

Genes que contienen algun punto de rotura

SHOW

Gen	Cromosoma	Inicio	Fin	GO_ID	Tipo
1 LOC117901760	NC_048534.1	14722075	14724843	GO:0005515	F
2 LOC117901760	NC_048534.1	14722075	14724843	GO:0005615	C
3 LOC117901760	NC_048534.1	14722075	14724843	GO:0031012	C
Descripcion_GO					
1	protein binding				
2	extracellular space				
3	extracellular matrix				
Definicion_GO					
1	Binding to a protein.				
2	That part of a multicellular organism outside the cells proper, usually taken to be outside the plasma membranes, and occupied by fluid.				
3	A structure lying external to one or more cells, which provides structural support, biochemical or biomechanical cues for cells or tissues.				
Name					
1	leucine-rich repeat-containing protein 15				
2	leucine-rich repeat-containing protein 15				
3	leucine-rich repeat-containing protein 15				

Exploración de genes relacionados con la adaptación térmica en la inversión  $U_8$

Se analiza la inversión cromosómica  $U_8$  en la que se busca identificar genes asociados a la adaptación térmica dentro de esta región. Basado en el artículo de Antunes et al. (2024), que enumera genes involucrados en procesos como proteólisis, metabolismo de lípidos y plasticidad adaptativa, se quiere determinar si alguno de estos genes está presente en la región que en este TFM se ha identificado como la inversió  $U_8$ .

SHOW

Se buscan las coincidencias de genes listados en el artículo.

SHOW  
SHOW  
SHOW

Términos GO de los genes relacionados con la adaptación térmica

Gen / Nombre de la proteína	Coordenadas genómicas	Funciones Moleculares	Procesos Biológicos	Componentes Celulares
LOC117902436 - zinc finger BED domain-containing protein 4 isoform X1	8251059-8253193	GO:0003677 (GO:0003677): DNA binding		
LOC117902220 - DNA-directed RNA polymerases I and III subunit RPAC1	13174198-13175420	GO:0001054 (GO:0001054): RNA polymerase I activity GO:0001056 (GO:0001056): RNA polymerase III activity GO:0046983 (GO:0046983): protein dimerization activity	GO:0006351 (GO:0006351): DNA-templated transcription	GO:0005666 (GO:0005666): RNA polymerase III complex GO:0005736 (GO:0005736): RNA polymerase I complex