# Variaciones genómicas y expresión génica

Valencia, Septiembre 2011
CEFIRE

Francisco García
fgarcia@cipf.es
http://bioinfo.cipf.es/fgarcia
Departamento de Bioinformática y Genómica
Centro de Investigacion Principe Felipe (CIPF)
Valencia (España)

INB

PRINCIPE FELIPE
CENTRO DE INVESTIGACION

er ciberer
CENTRO DE INVESTIGACION BIOMEDICA EN RED
DE ENFERMEDADES RARAS

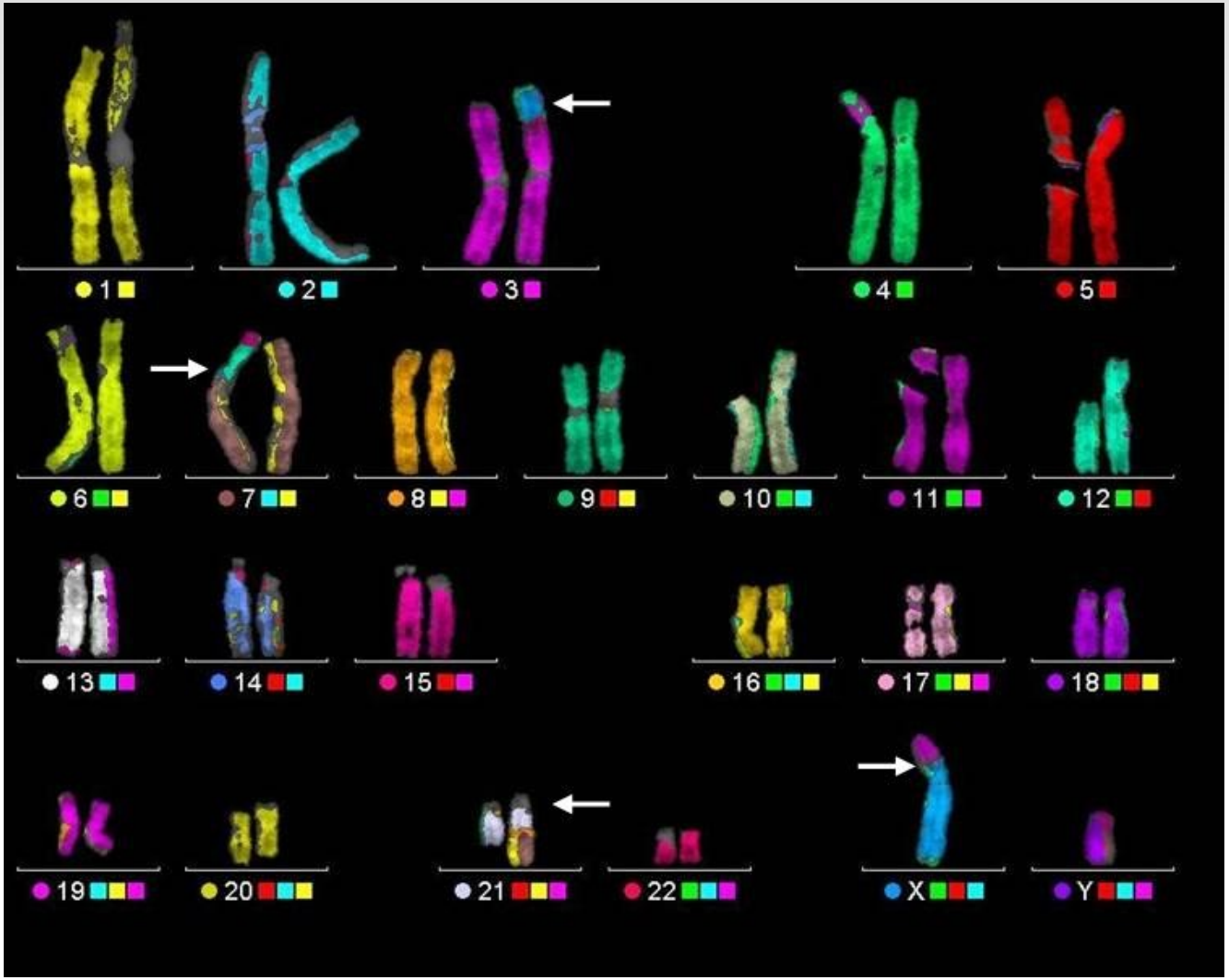
# Información biológica

**ADN** → **ARN** → **proteína**

**Genoma** → **Transcriptoma** → **Proteoma**

# Variaciones genómicas

# Cariotipo

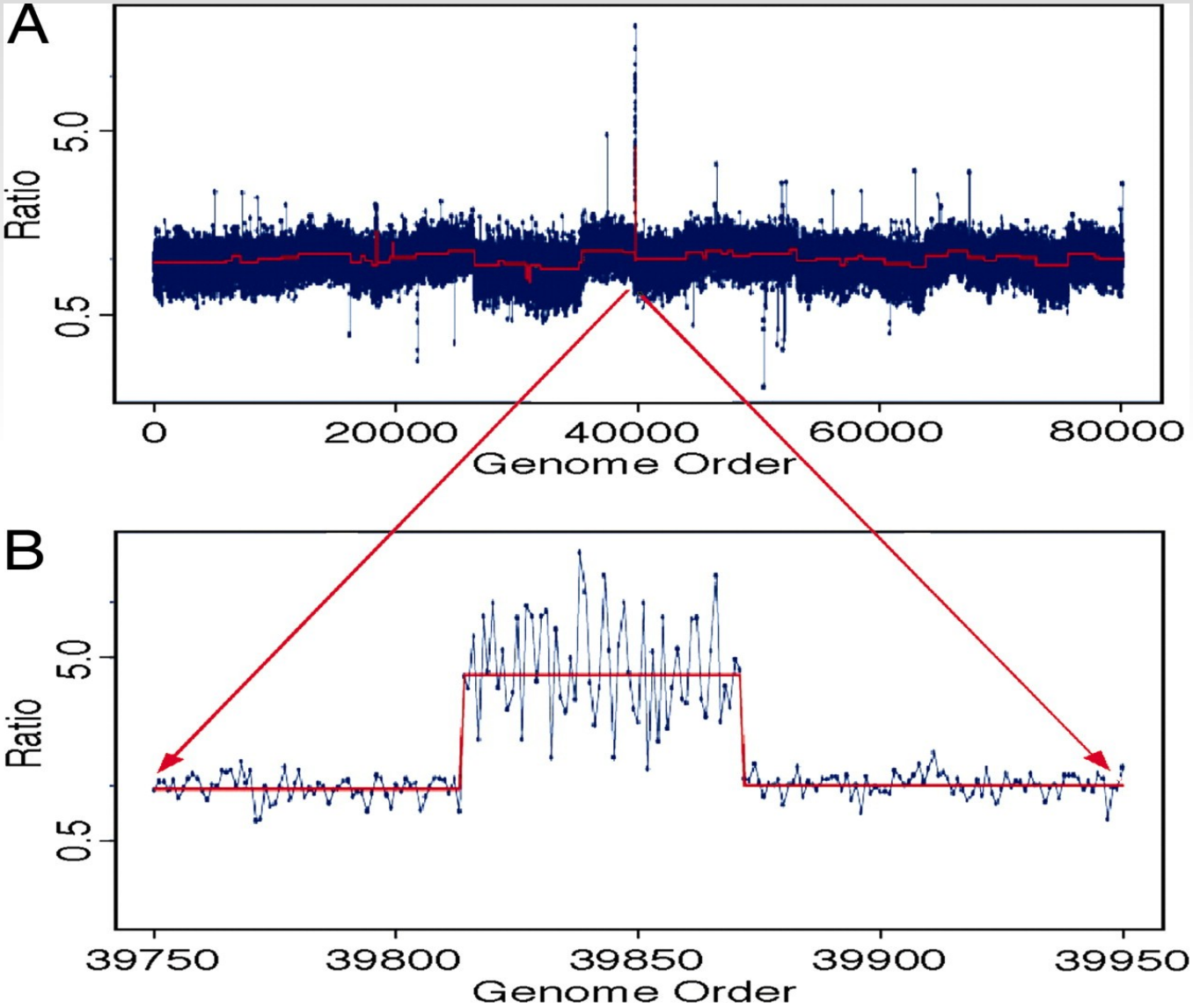


# Tipos de variaciones genómicas

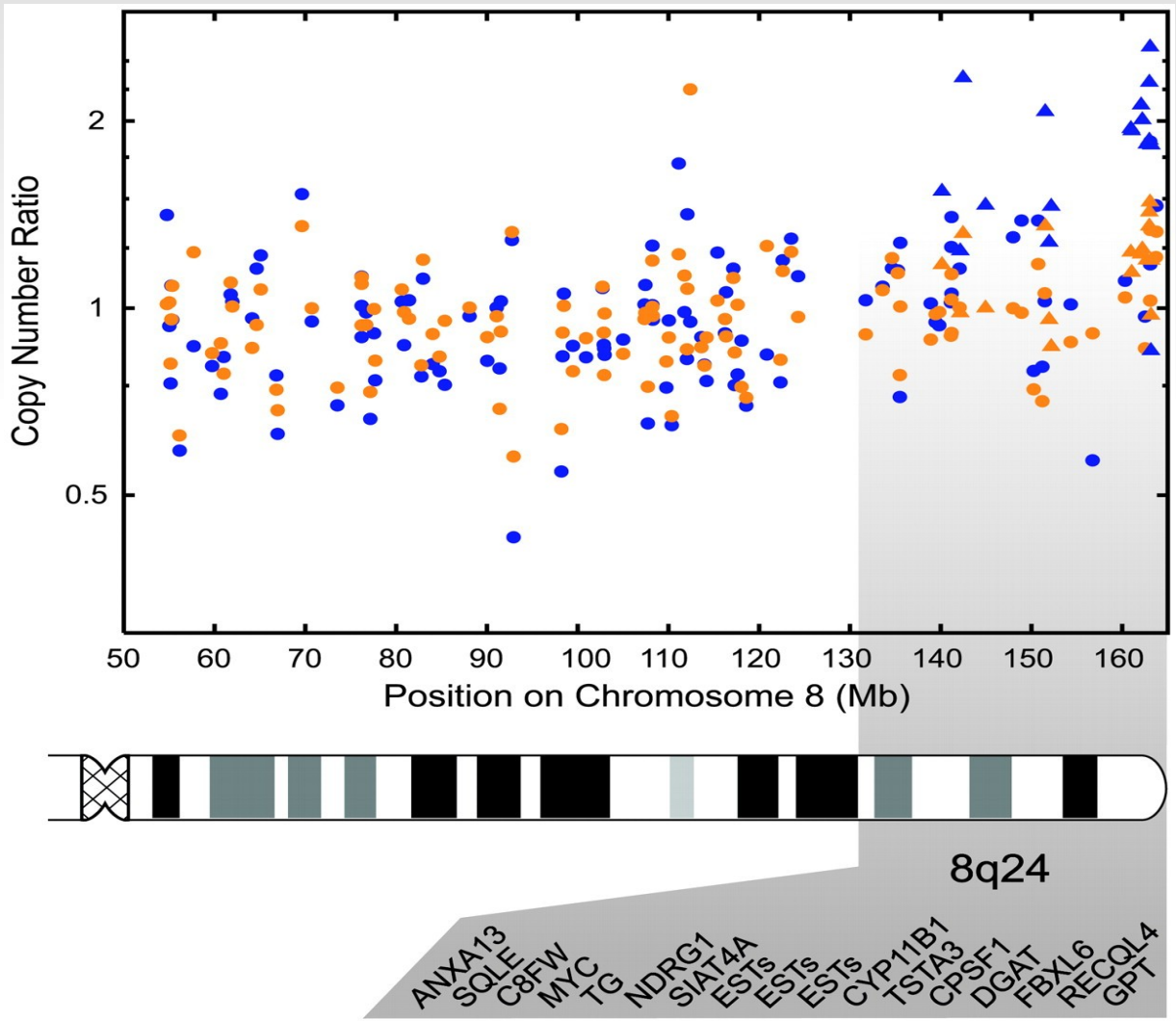
- Variación del número de copias (CNV)
  - Duplicidad – Amplificación
  - Deleción

- Polimorfimos
  - Polimorfismo de un solo nucleótido (SNPs)

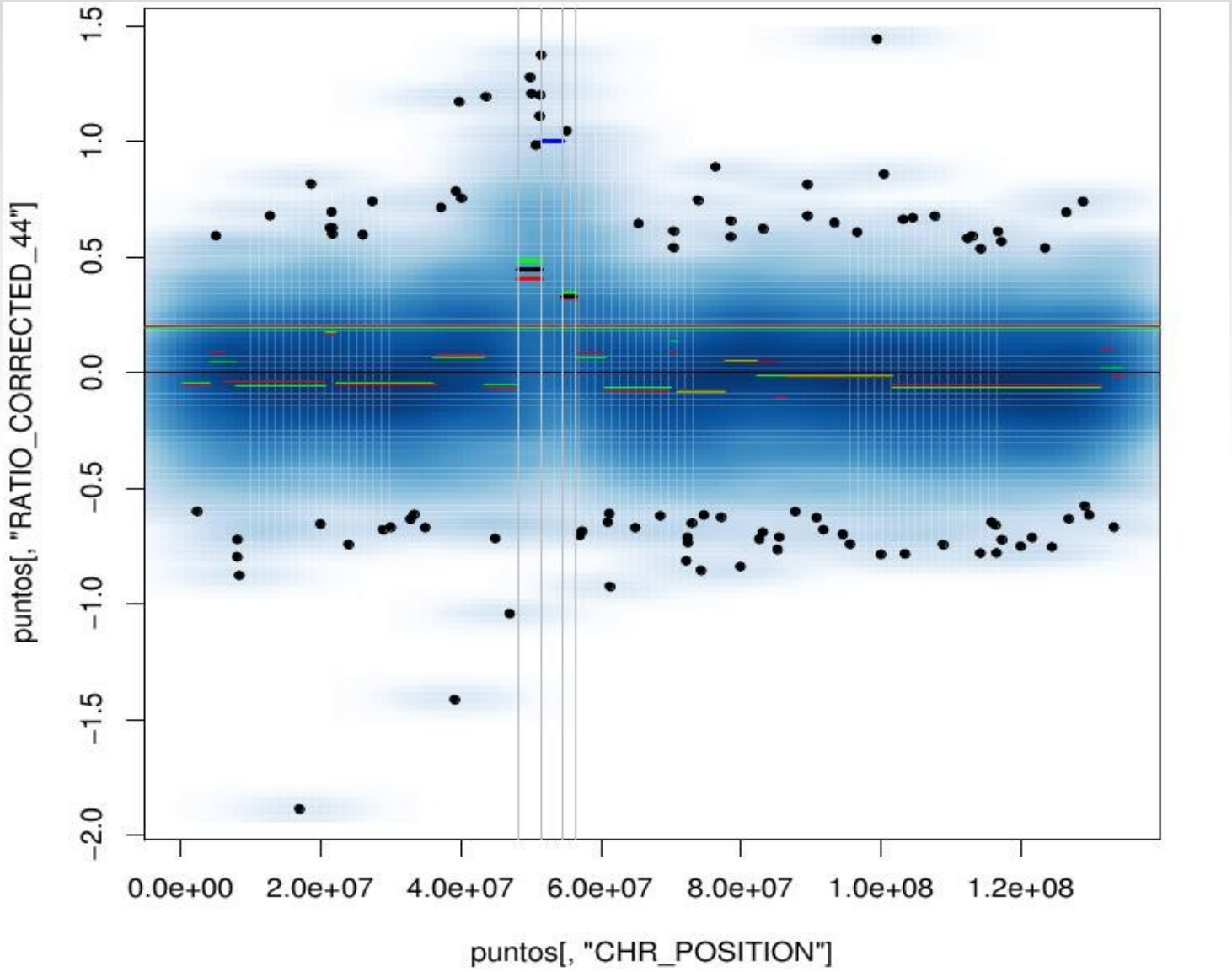
# Análisis de segmentación



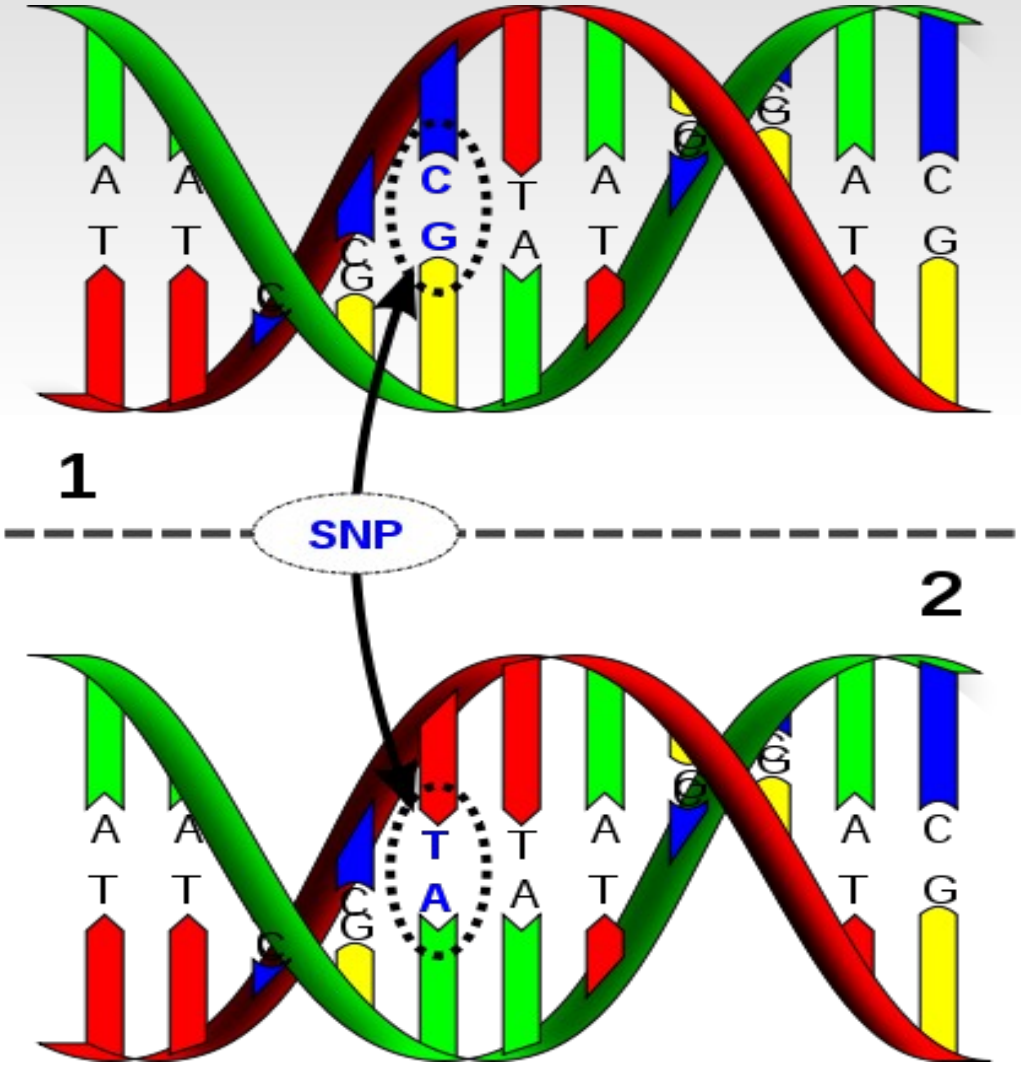
# Combinando datos de varias muestras



# Búsqueda de regiones comunes



# Single nucleotide polymorphisms (SNP)



# Estudios de asociación de SNPs

| individuo | SNP1 | SNP2 | Enfermedad |
|---|---|---|---|
| array1 | A/A | A/B | caso |
| array2 | A/B | A/A | caso |
| array3 | A/B | A/B | caso |
| array4 | A/A | A/B | caso |
| array5 | A/B | A/B | control |
| array6 | A/B | A/A | control |
| array7 | B/B | A/B | control |
| array8 | A/A | A/B | control |
| array9 | B/B | A/B | control |
| array10 | B/B | A/A | control |
| array11 | B/B | A/B | control |
| array12 | B/B | A/B | control |

| | A | B |
|---|---|---|
| caso | **6** | **2** |
| control | **4** | **8** |

$$\frac{6 \cdot 8}{2 \cdot 4} > 1 \longrightarrow$$ Alelo A asociado a caso

# Estratificación

**Población 1**

| | A | B |
|---|---|---|
| caso | **2** | **1** |
| control | **2** | **1** |

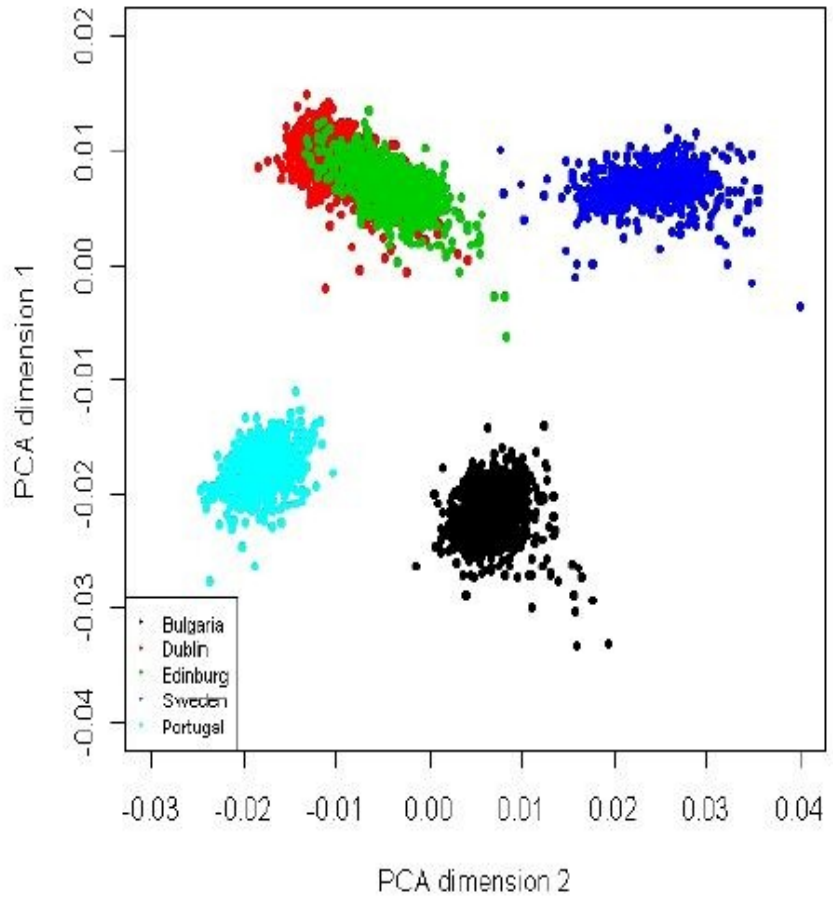
$$\frac{2 \cdot 1}{2 \cdot 1} = 1$$

**Población 2**

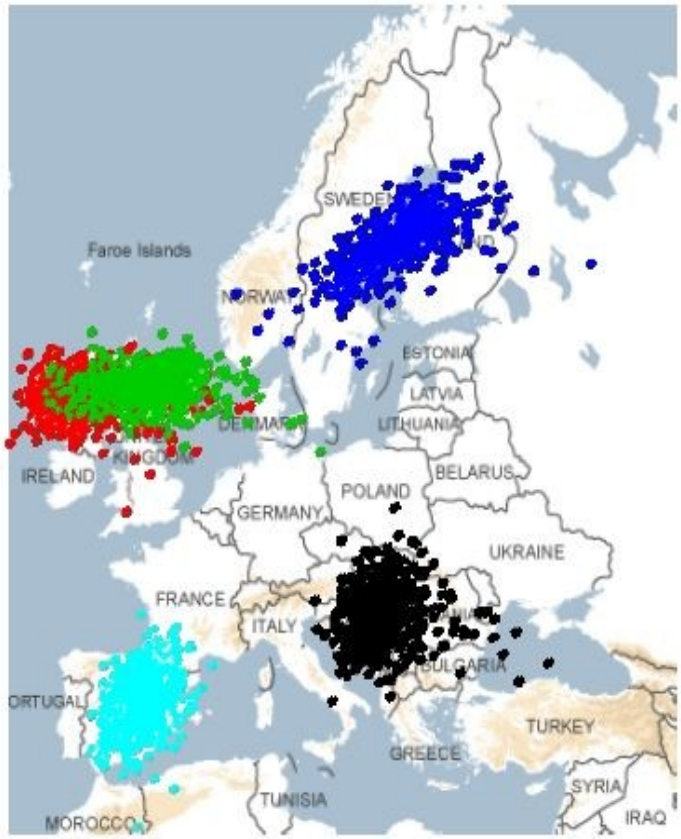
| | A | B |
|---|---|---|
| caso | **4** | **1** |
| control | **2** | **7** |

$$\frac{4 \cdot 7}{2 \cdot 1} >> 1$$

# Estratificación de la población




a.

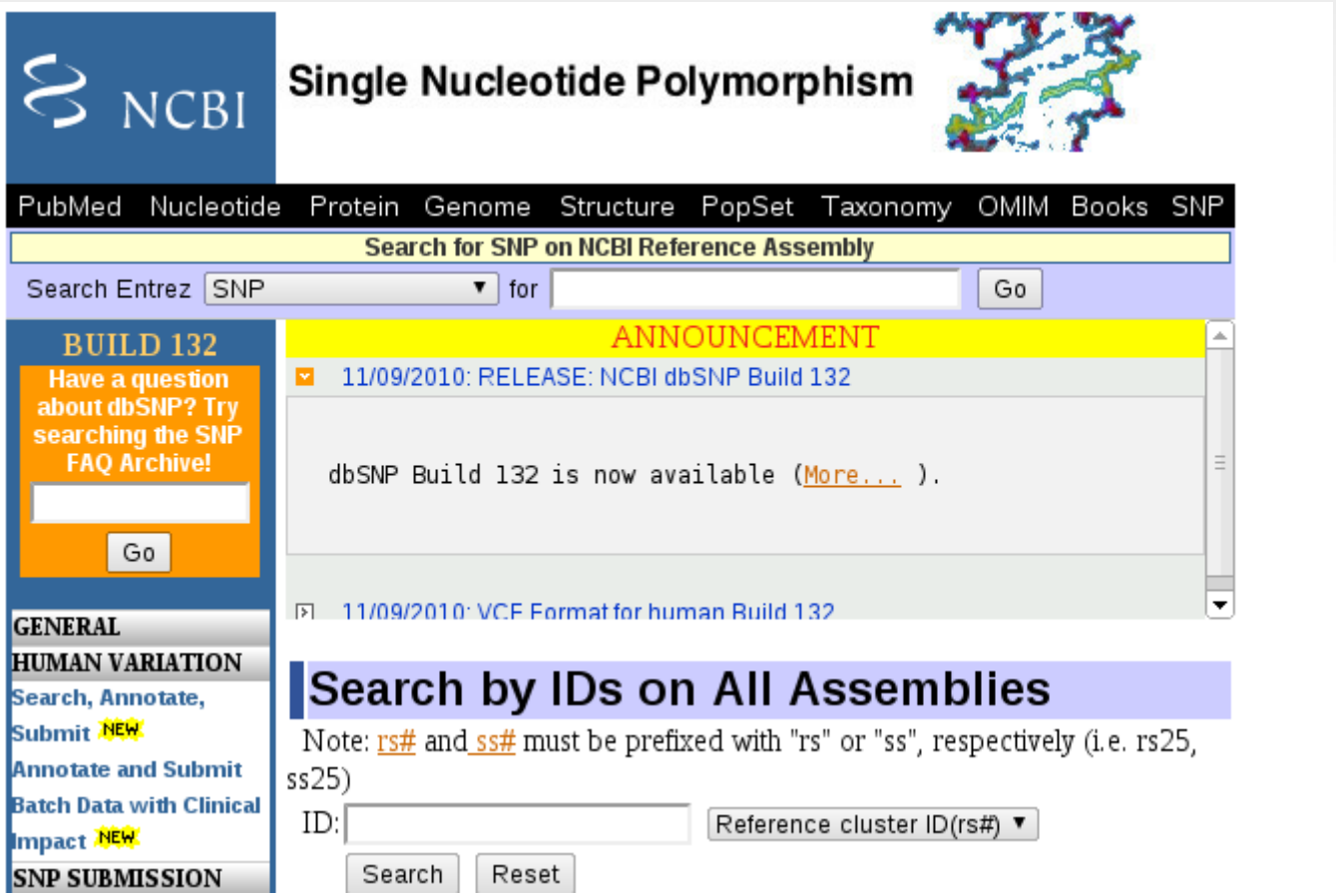



b.

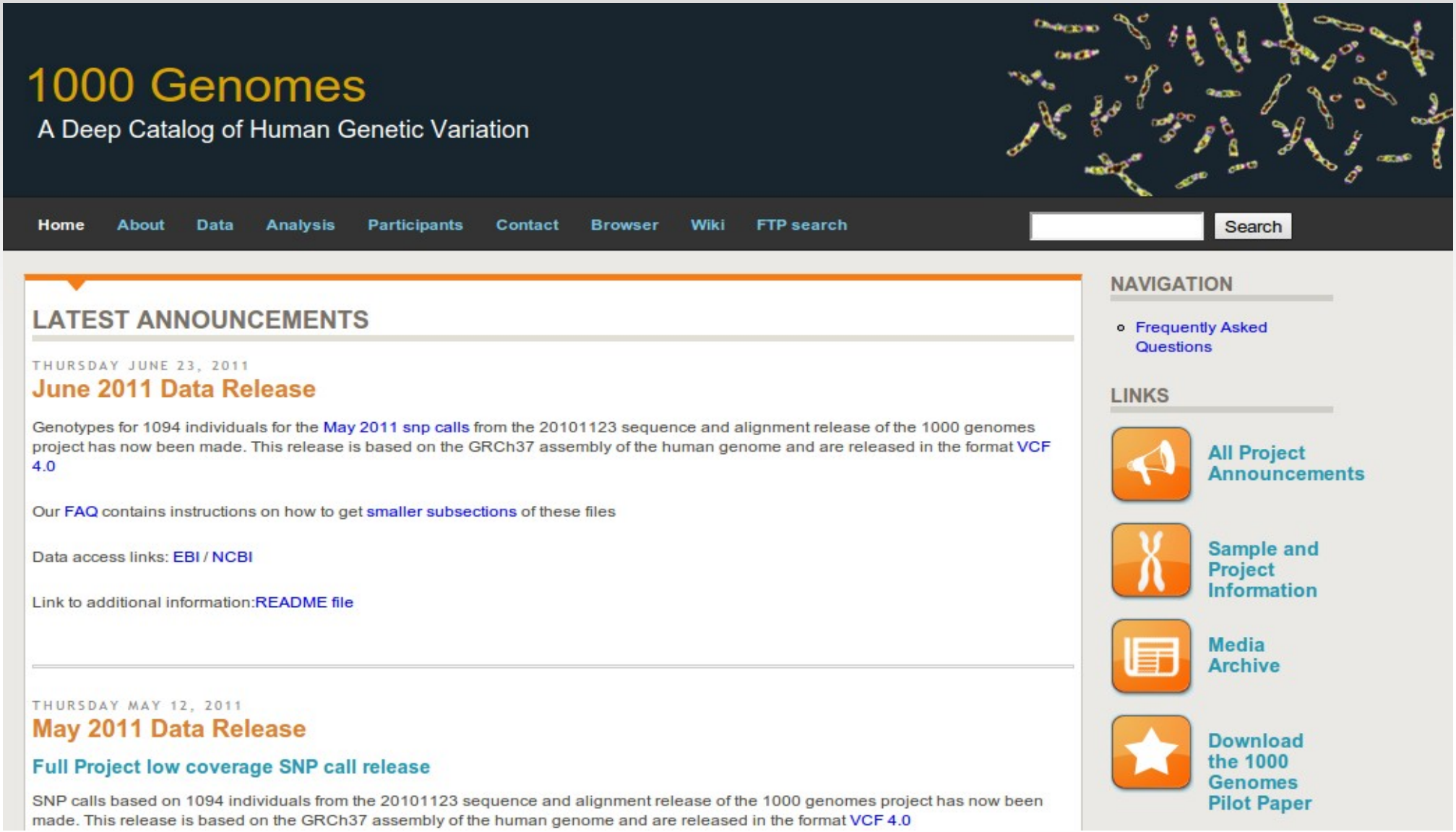
Valentina Moskvina et al.
Human Heredity Vol. 70, No. 2, 2010

# Recursos – Bases de datos

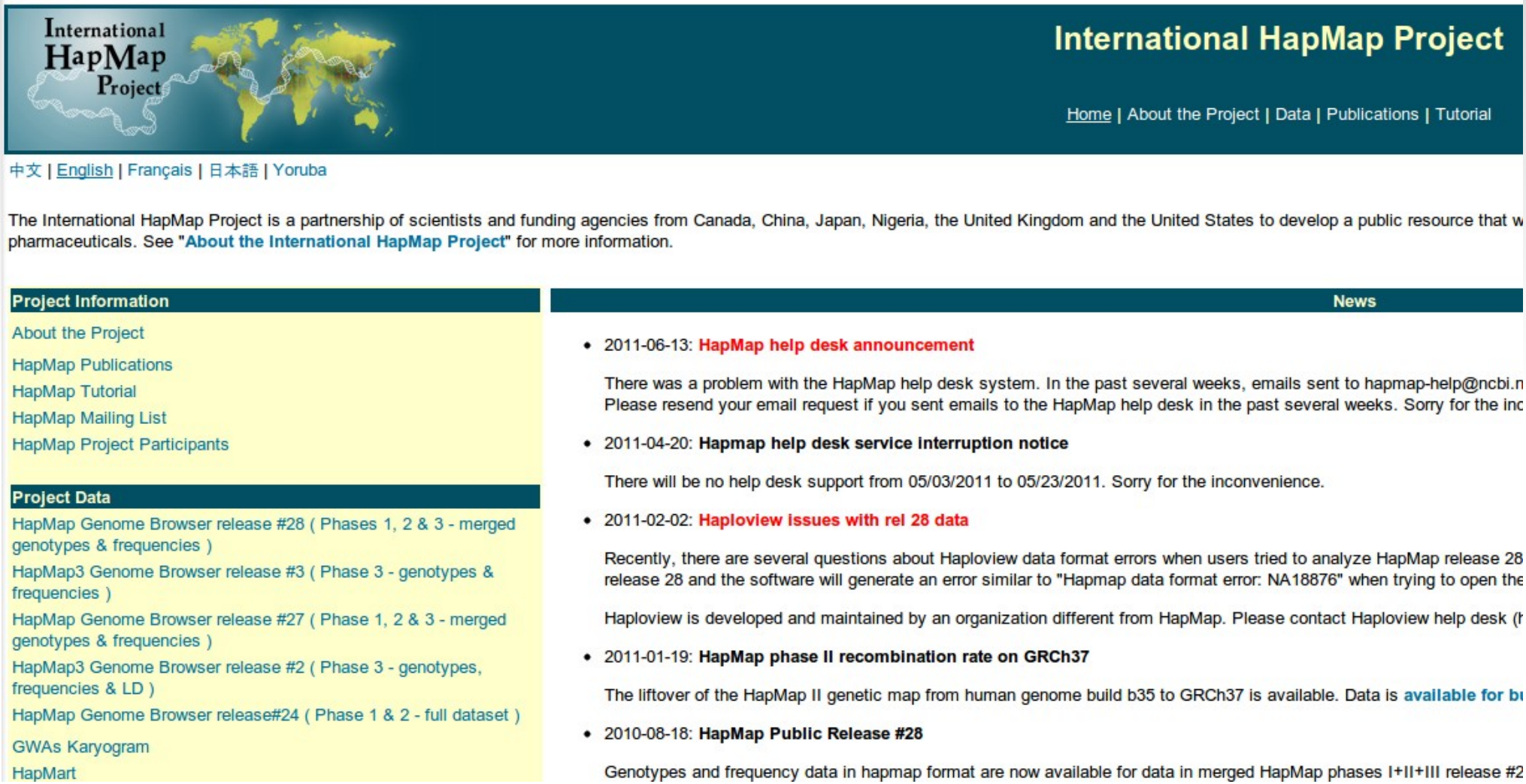
- Búsqueda, anotación e inclusión de SNPs



# Recursos- Variaciones individuales

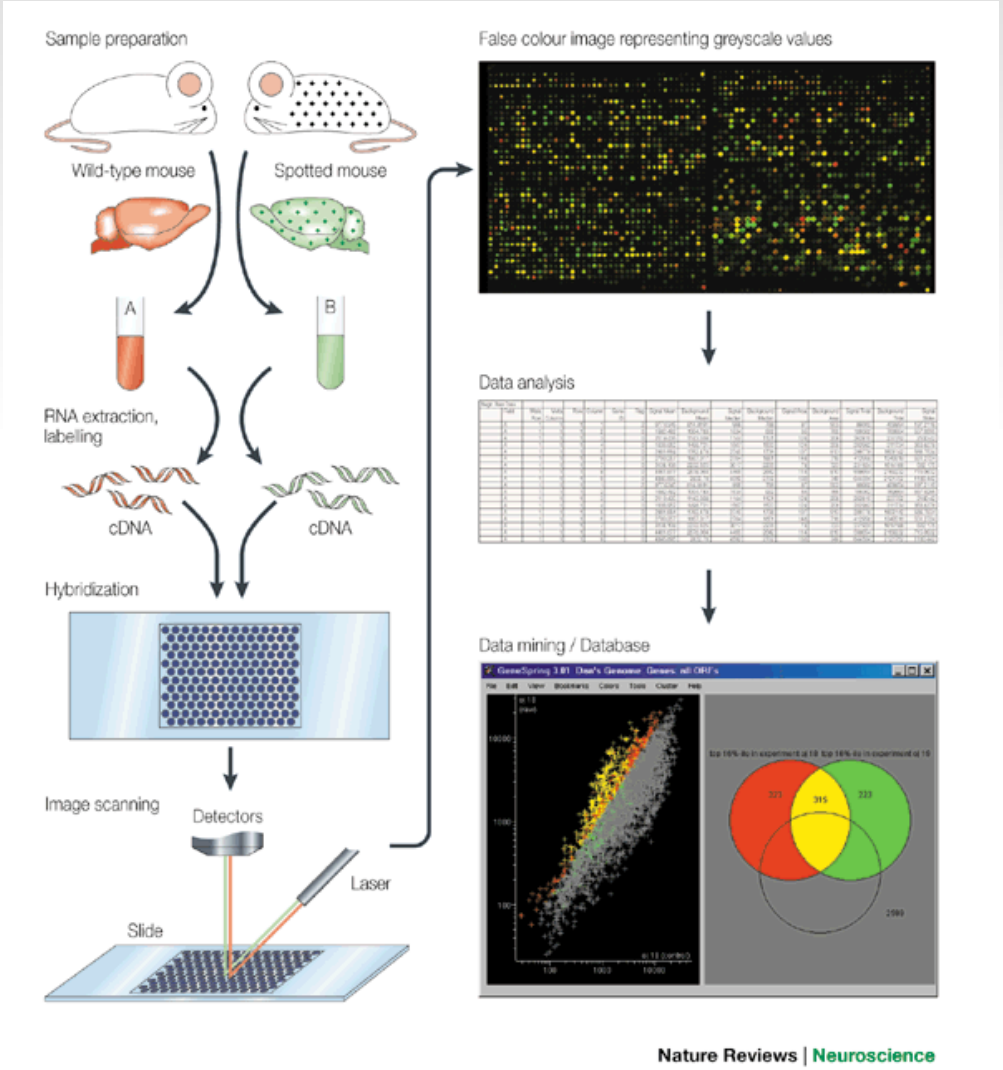


# Recursos- Mapas de haplotipos

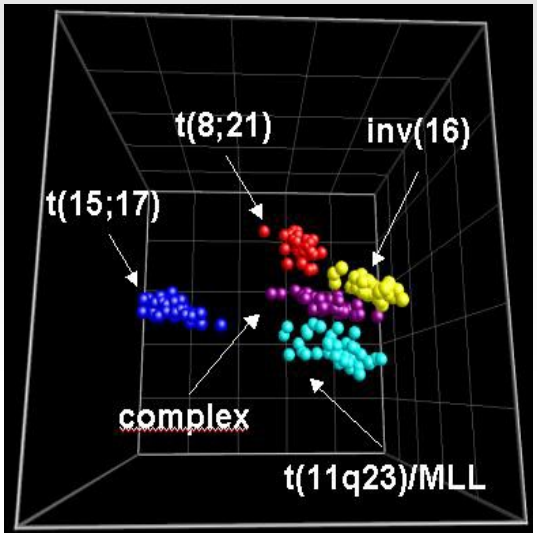


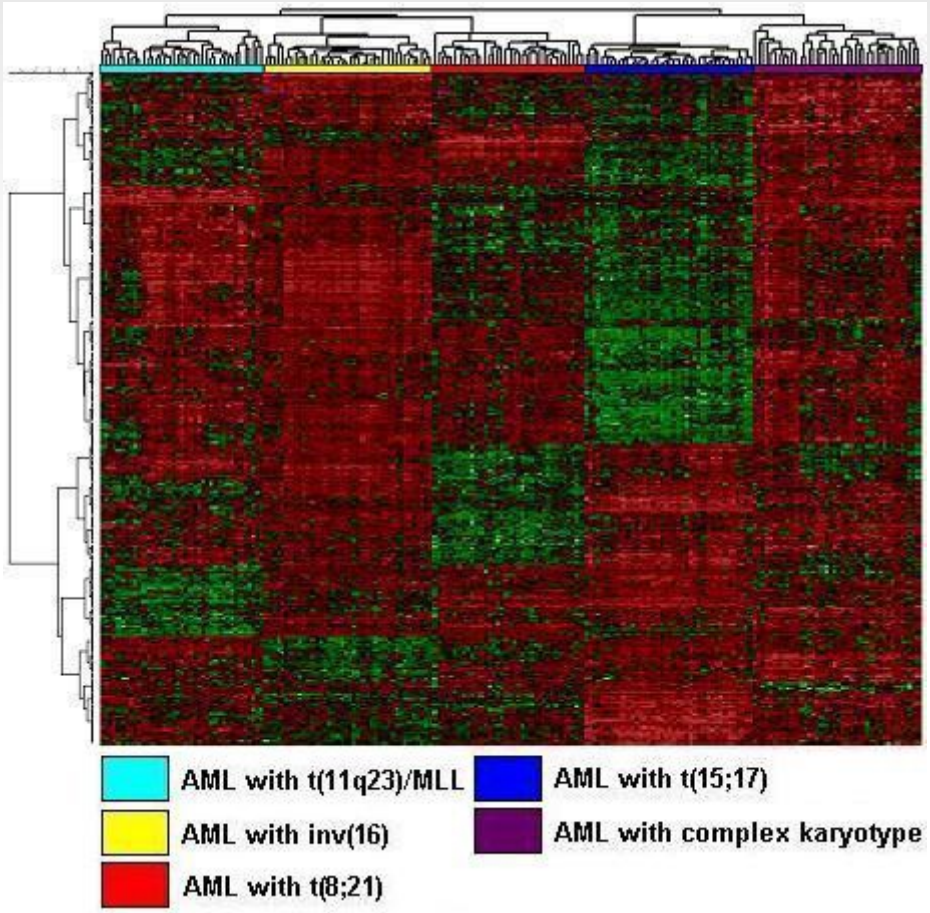
# Estudios de expresión génica

# Comparación de clases

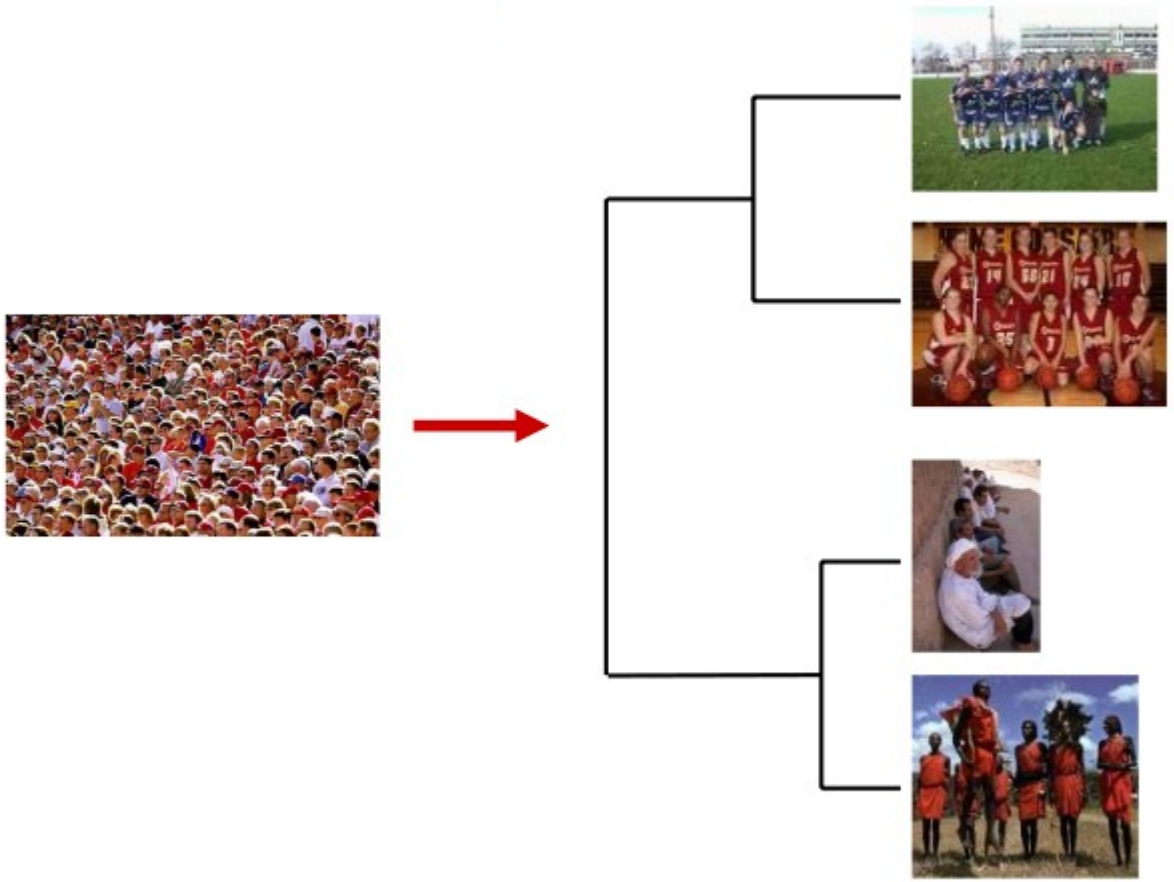


# Descubrimiento de clases

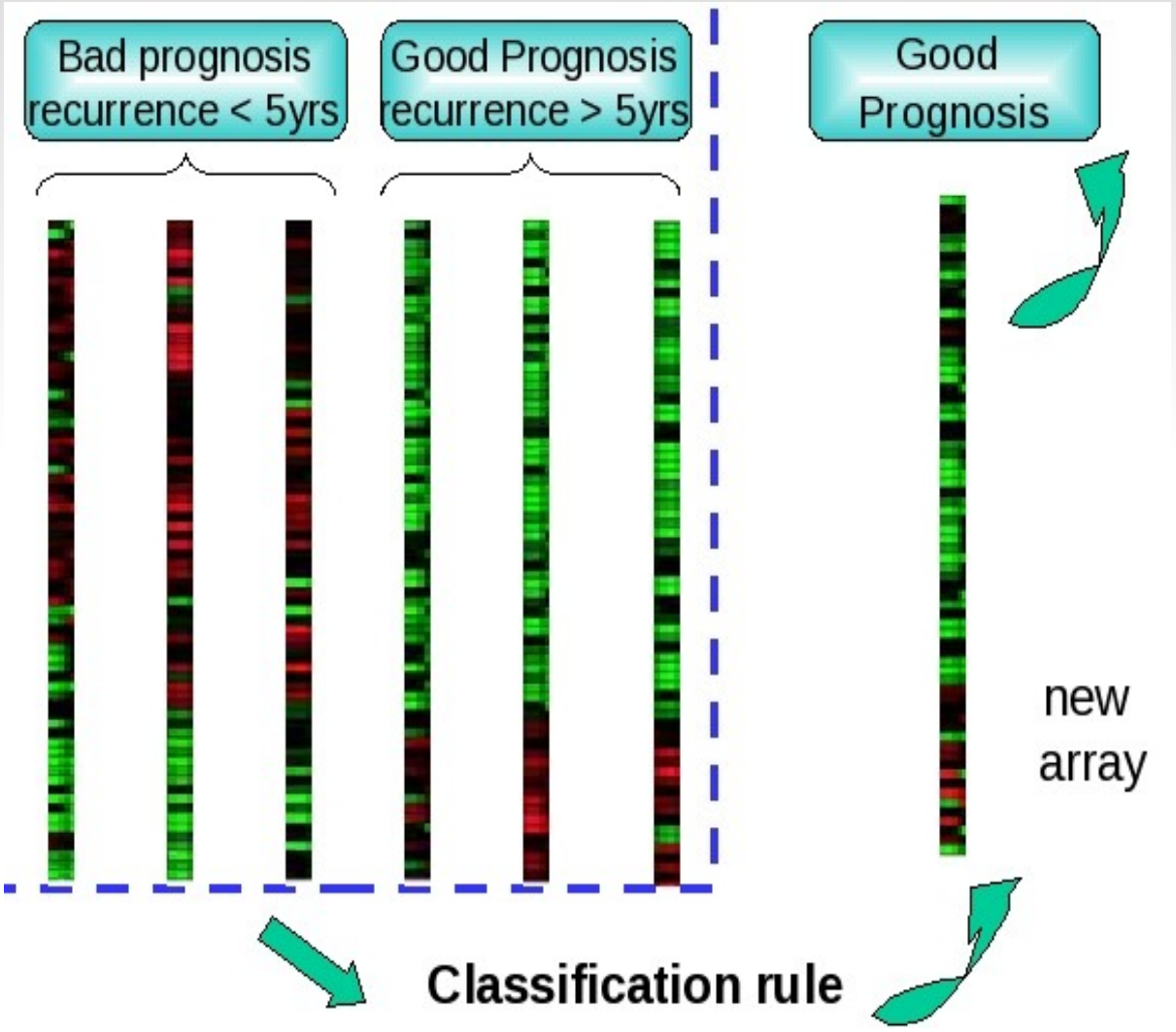


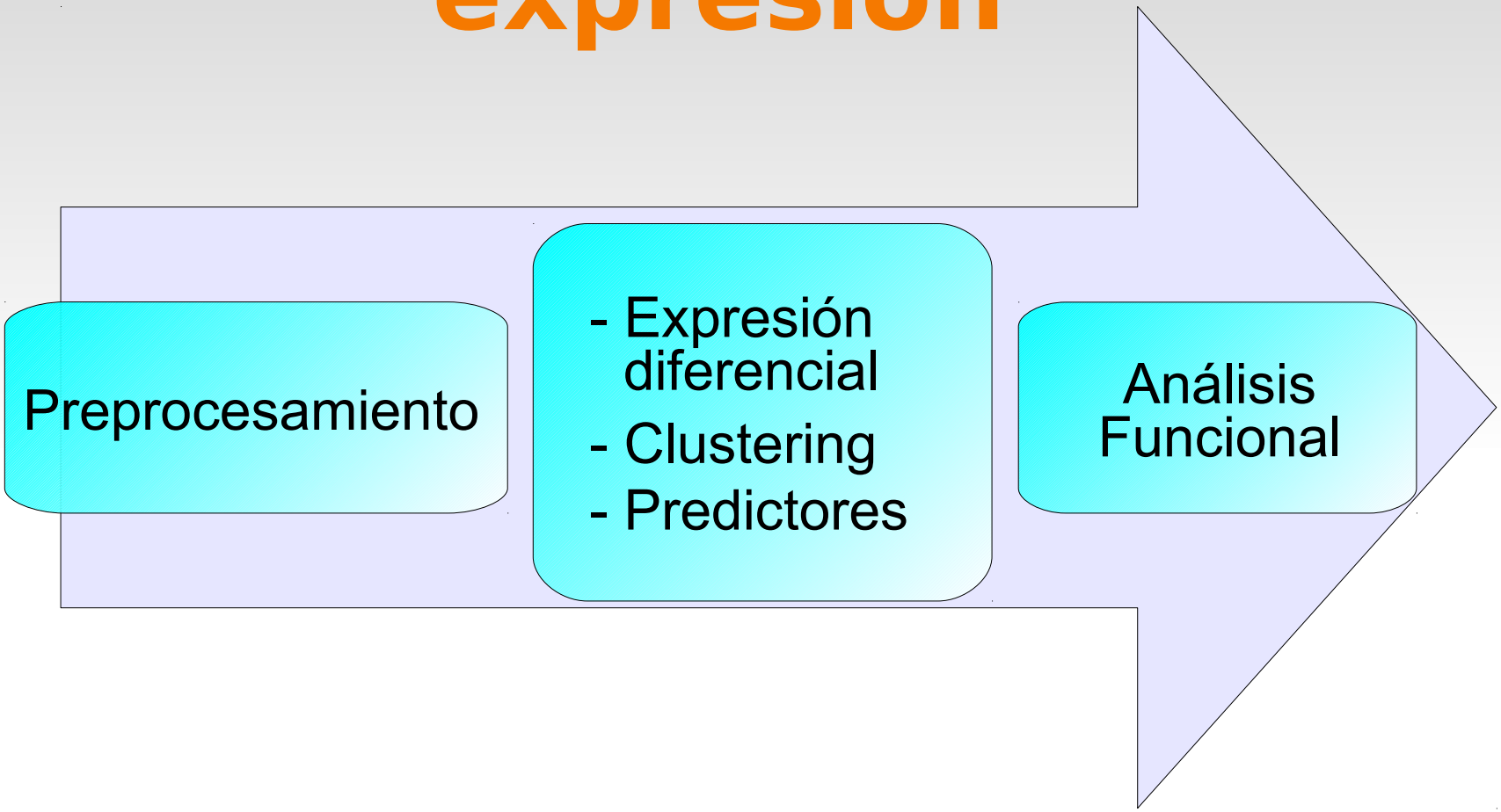
# Clustering



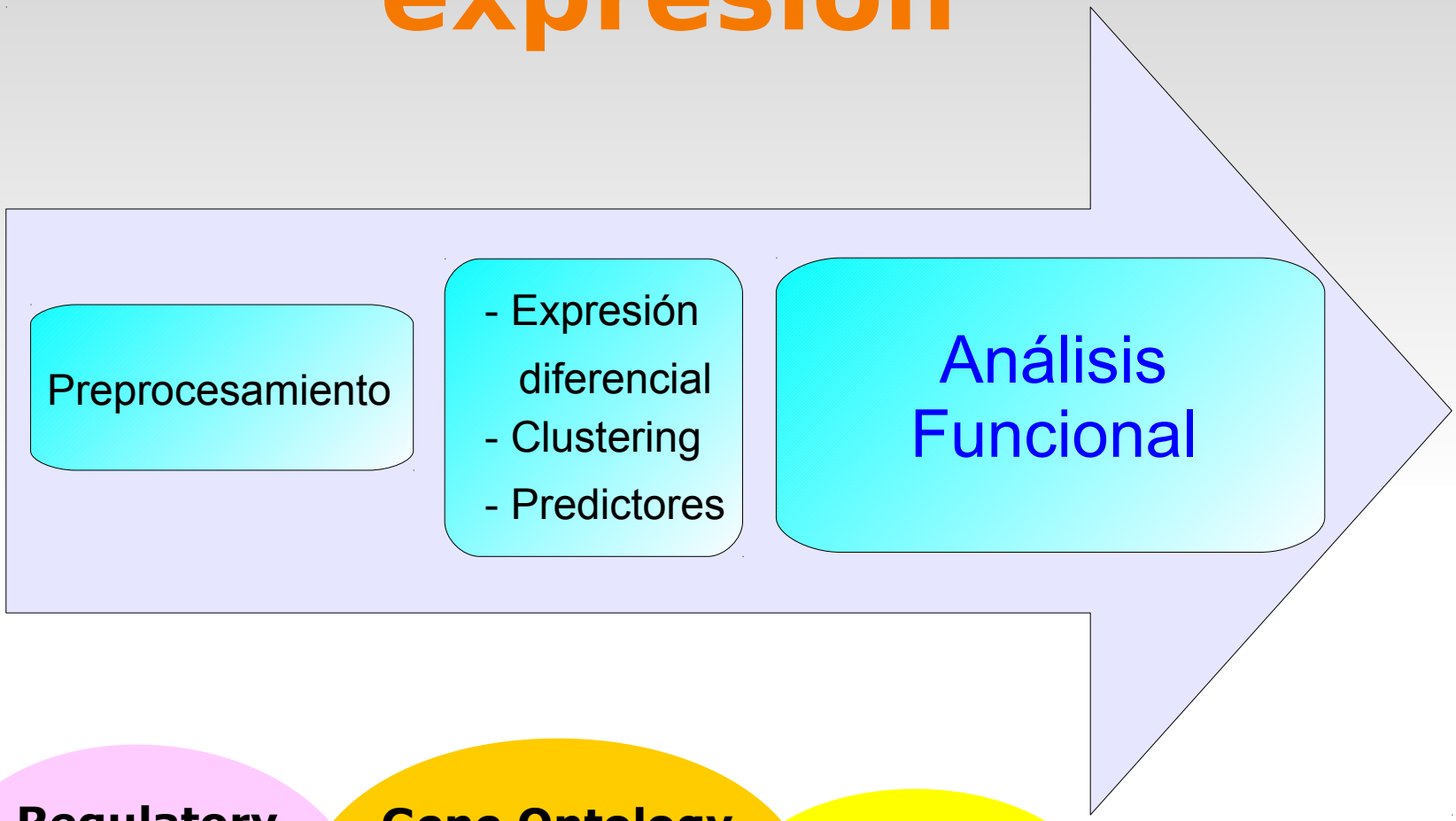
# Predicción de clases



# Análisis de datos de expresión



# Análisis de datos de expresión




**Regulatory elements**

**miRNA, CisRed**

**Transcription Factor Binding Sites**

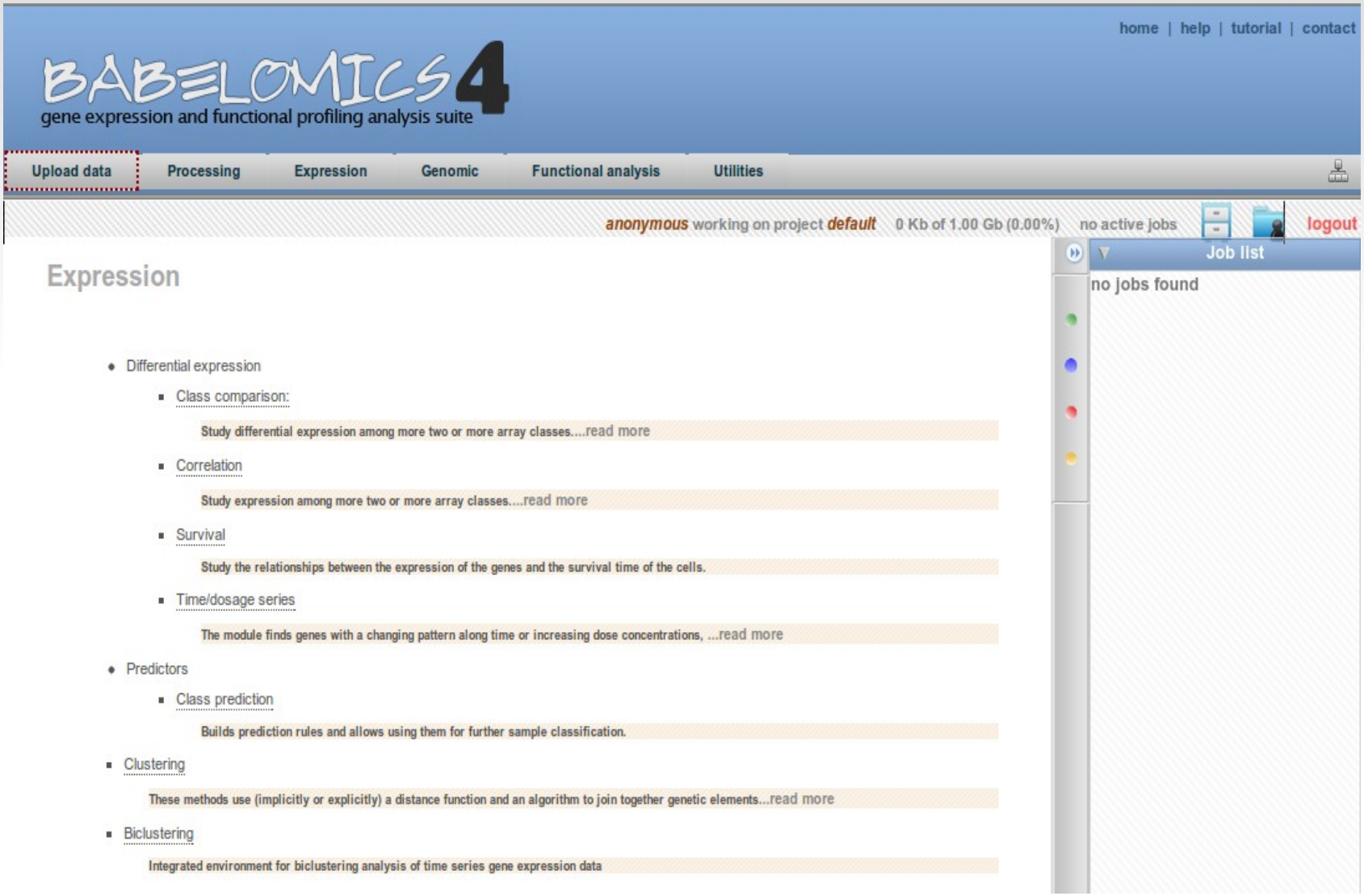
**Gene Ontology**

Biological Process
Molecular Function
Cellular Component

**KEGG pathways**

**Bases de datos biológicos**

# Babelomics


home | help | tutorial | contact
BABELOMICS 4
gene expression and functional profiling analysis suite
Upload data
Processing
Expression
Genomic
Functional analysis
Utilities
anonymous working on project default
0 Kb of 1.00 Gb (0.00%)
no active jobs
logout
Expression
Differential expression
Class comparison:
Study differential expression among more two or more array classes....read more
Correlation
Study expression among more two or more array classes....read more
Survival
Study the relationships between the expression of the genes and the survival time of the cells.
Time/dosage series
The module finds genes with a changing pattern along time or increasing dose concentrations, ...read more
Predictors
Class prediction
Builds prediction rules and allows using them for further sample classification.
Clustering
These methods use (implicitly or explicitly) a distance function and an algorithm to join together genetic elements...read more
Biclustering
Integrated environment for biclustering analysis of time series gene expression data
Job list
no jobs found

# Ejemplo de análisis con Babelomics

**Descripción de los datos**

- 24 muestras correspondientes a 12 individuos entrenados y 12 sedentarios
- Arrays Affymetrix, 22.000 genes

**¿Qué queremos hacer?**

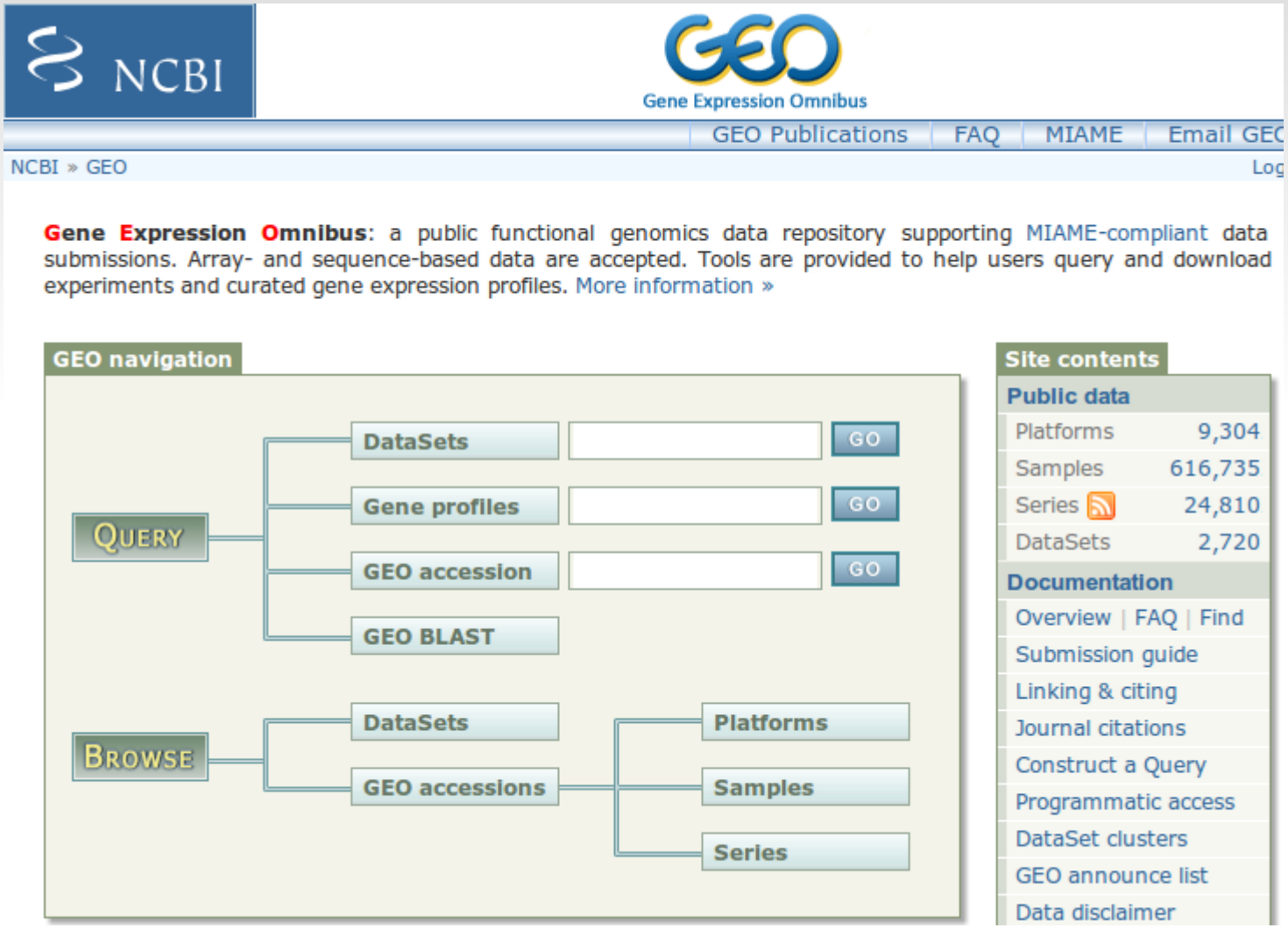
- Análisis de expresión diferencial

# Ejemplo de análisis con Babelomics

## Formato de los datos

| #VARIABLE | entrenado | entrenado | sedentario | sedentario | sedentario | entrenado |
|---|---|---|---|---|---|---|
| #NOMBRES | ARRAY1 | ARRAY2 | ARRAY3 | ARRAY4 | ARRAY5 | ARRAY6 |
| AFFX-BioB-5_at | 7.521 | 6.933 | 6.659 | 7.589 | 6.347 | 7.623 |
| AFFX-BioB-M_at | 8.394 | 7.695 | 7.722 | 8.404 | 7.055 | 8.487 |
| AFFX-BioB-3_at | 8.163 | 7.610 | 7.545 | 8.113 | 7.207 | 8.176 |
| AFFX-BioC-5_at | 8.959 | 8.367 | 8.269 | 9.160 | 7.940 | 9.190 |
| AFFX-BioC-3_at | 8.381 | 7.679 | 7.713 | 8.551 | 7.568 | 8.682 |
| AFFX-BioDn-5_at | 9.173 | 8.245 | 8.314 | 9.009 | 8.766 | 8.975 |
| AFFX-BioDn-3_at | 11.918 | 11.398 | 11.353 | 11.726 | 11.512 | 11.797 |
| AFFX-CreX-5_at | 12.275 | 11.640 | 11.781 | 12.007 | 11.685 | 12.003 |
| AFFX-CreX-3_at | 12.919 | 12.274 | 12.306 | 12.482 | 12.343 | 12.567 |
| AFFX-DapX-5_at | 3.649 | 3.625 | 3.639 | 3.500 | 3.746 | 3.714 |
| AFFX-DapX-M_at | 3.879 | 3.907 | 3.864 | 3.926 | 3.879 | 3.854 |
| AFFX-DapX-3_at | 3.558 | 3.602 | 3.550 | 3.562 | 3.627 | 3.599 |
| AFFX-LysX-5_at | 3.546 | 3.662 | 3.614 | 3.584 | 3.546 | 3.575 |
| AFFX-LysX-M_at | 4.272 | 4.383 | 4.305 | 4.469 | 4.297 | 4.260 |
| AFFX-LysX-3_at | 3.521 | 3.580 | 3.578 | 3.466 | 3.504 | 3.489 |
| AFFX-PheX-5_at | 3.832 | 3.904 | 3.763 | 3.717 | 3.716 | 3.667 |
| AFFX-PheX-M_at | 3.724 | 3.655 | 3.567 | 3.704 | 3.705 | 3.595 |

# Recursos – Bases de datos



# Algunas ideas clave

**Diferentes niveles de información biológica, distintas situaciones de interés:**

**1. Variaciones genómicas: CNV, SNP's**

**2. Expresión génica**

**3. Interacciones entre proteínas**