

iGenCo: Técnicas ómicas en el diagnóstico de enfermedades raras

Parc Científic de Barcelona, 16-17 de Novembre

La Marató





ANÁLISIS DEL EPIGENOMA

Laura Balagué, Natàlia Carreras i
Xavier Escribà

La Marató 

ISGlobal

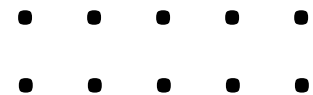


TABLA DE CONTENIDOS

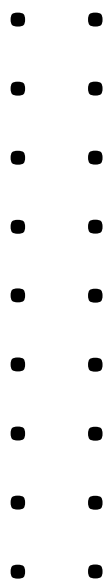
LAS EPIMUTACIONES **01**
Qué son y por qué las estudiamos

MÉTODOS DE DETECCIÓN PREVIOS **02**
Cuáles son y qué limitaciones tienen

PAQUETE *EPIMUTATIONS* **03**
Cómo funciona

MANOS A LA OBRA **04**
Caso práctico

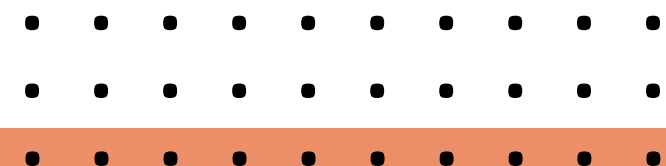
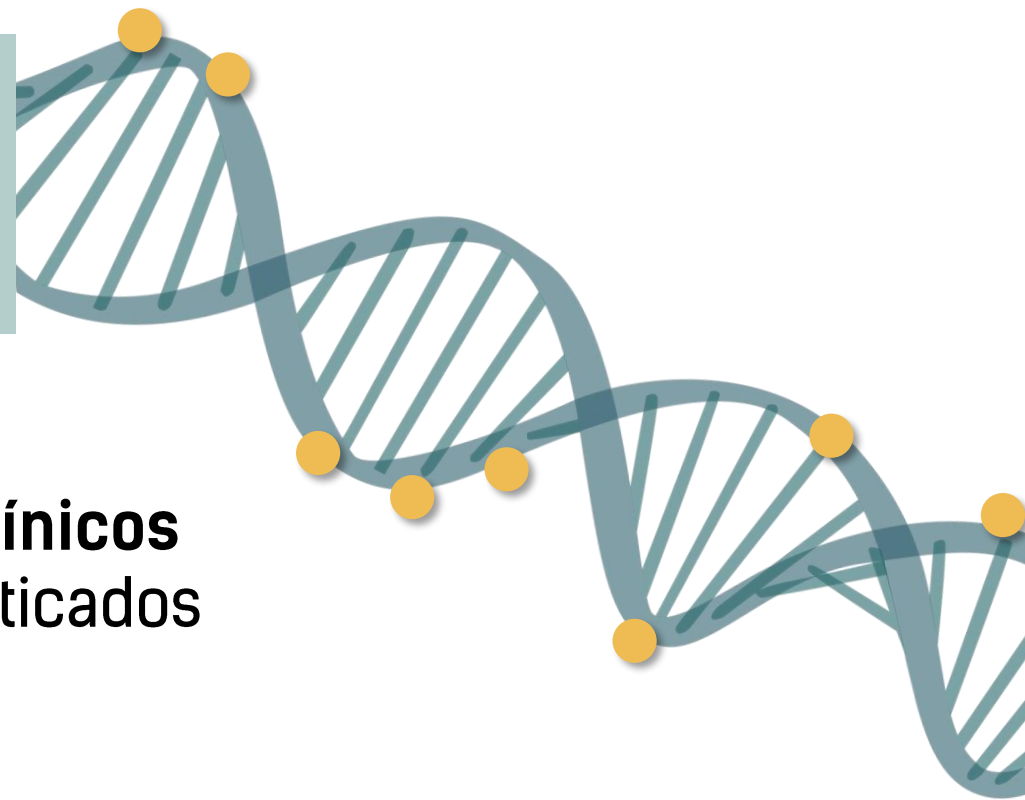
KAHOOT **05**
¿Quién se ha enterado de algo?



LAS EPIMUTACIONES

Las **epimutaciones** se definen como alteraciones raras del patrón de metilación en un loci específico.

Pueden usarse para **resolver casos clínicos ambiguos** que no pueden ser diagnosticados mediante técnicas convencionales.



MÉTODOS DE DETECCIÓN PREVIOS

01

MANOVA

Análisis multivariante
de la varianza

Aref-Eshghi et al. 2019

02

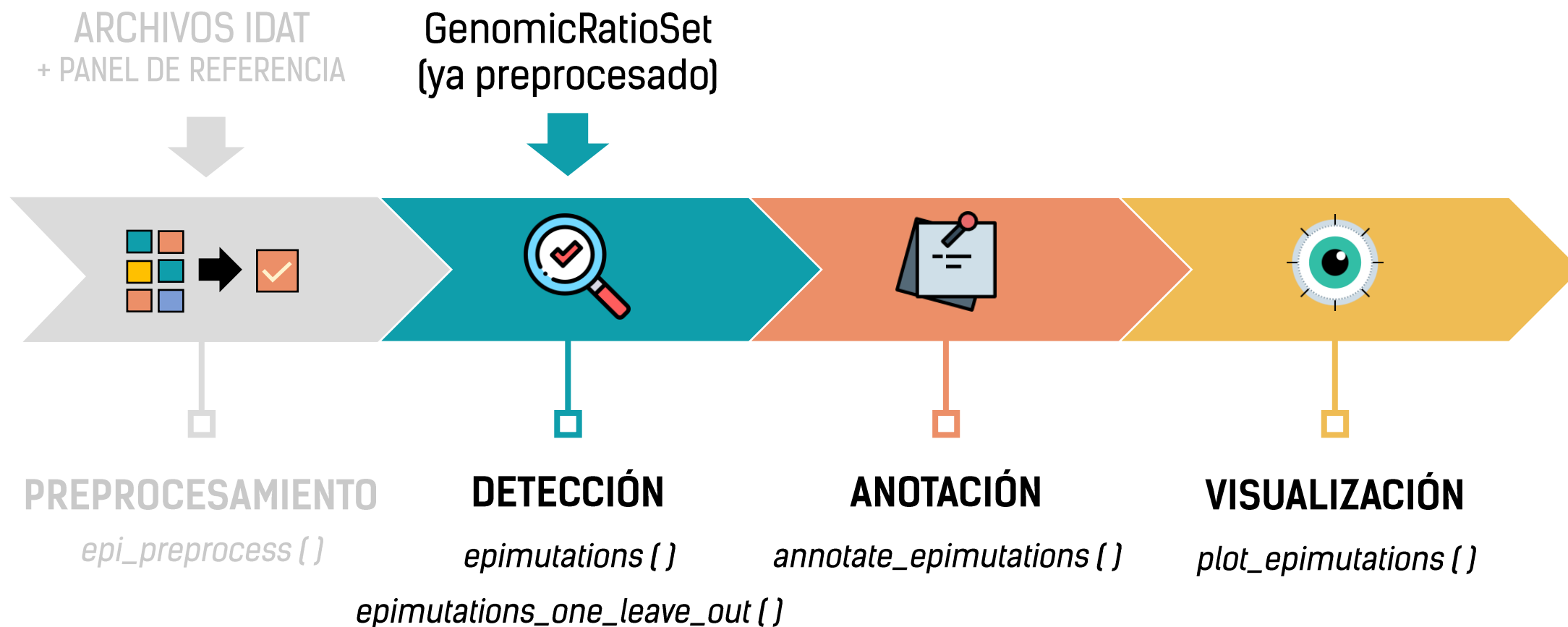
QUANTILE

Distribución de
cuantiles

Garg et al. 2020

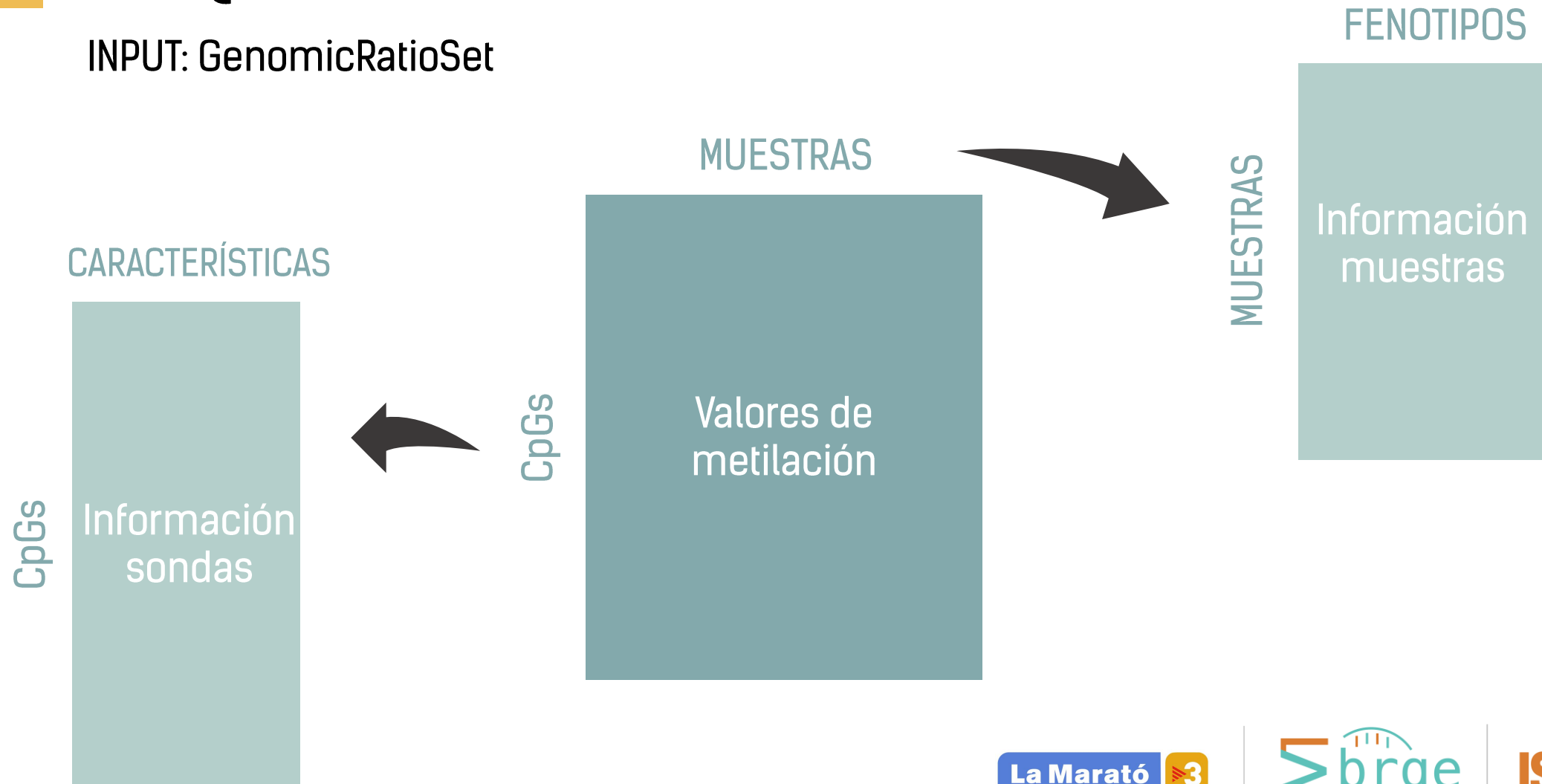
- Métodos no implementados en R
- No hay comparación de los métodos

PAQUETE *EPIMUTACIONES*



PAQUETE *EPIMUTACIONES*

INPUT: GenomicRatioSet



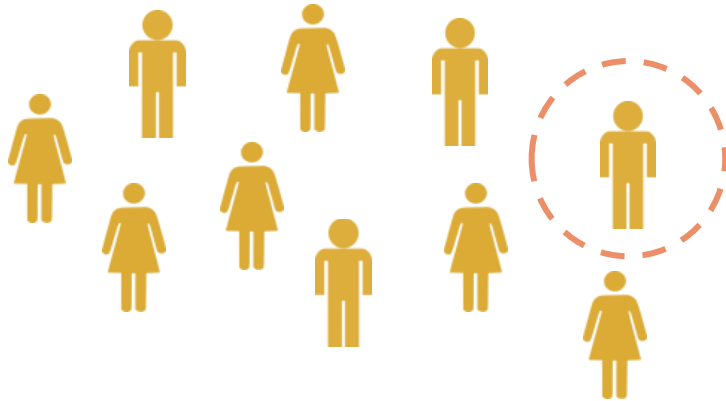


-
-
-
-

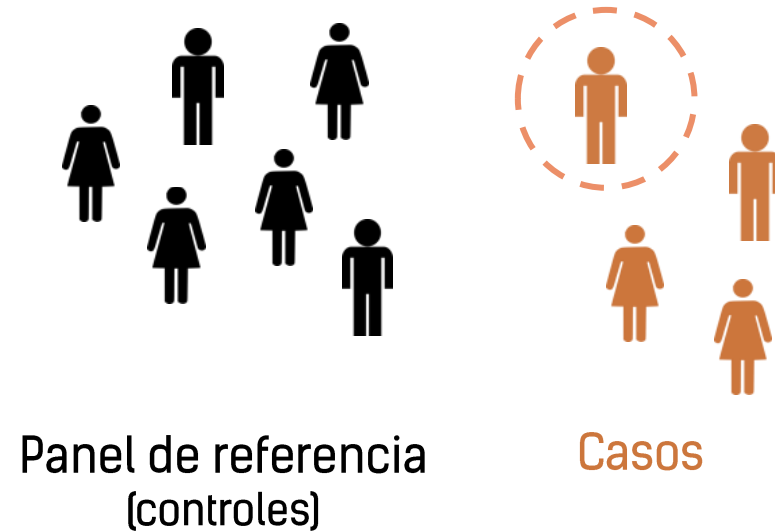
PAQUETE *EPIMUTACIONES*

DISTRIBUCIÓN DE LOS DATOS

A *Leave one out*



B Caso - Control



-
-
-
-
-

MÉTODOS DE DETECCIÓN

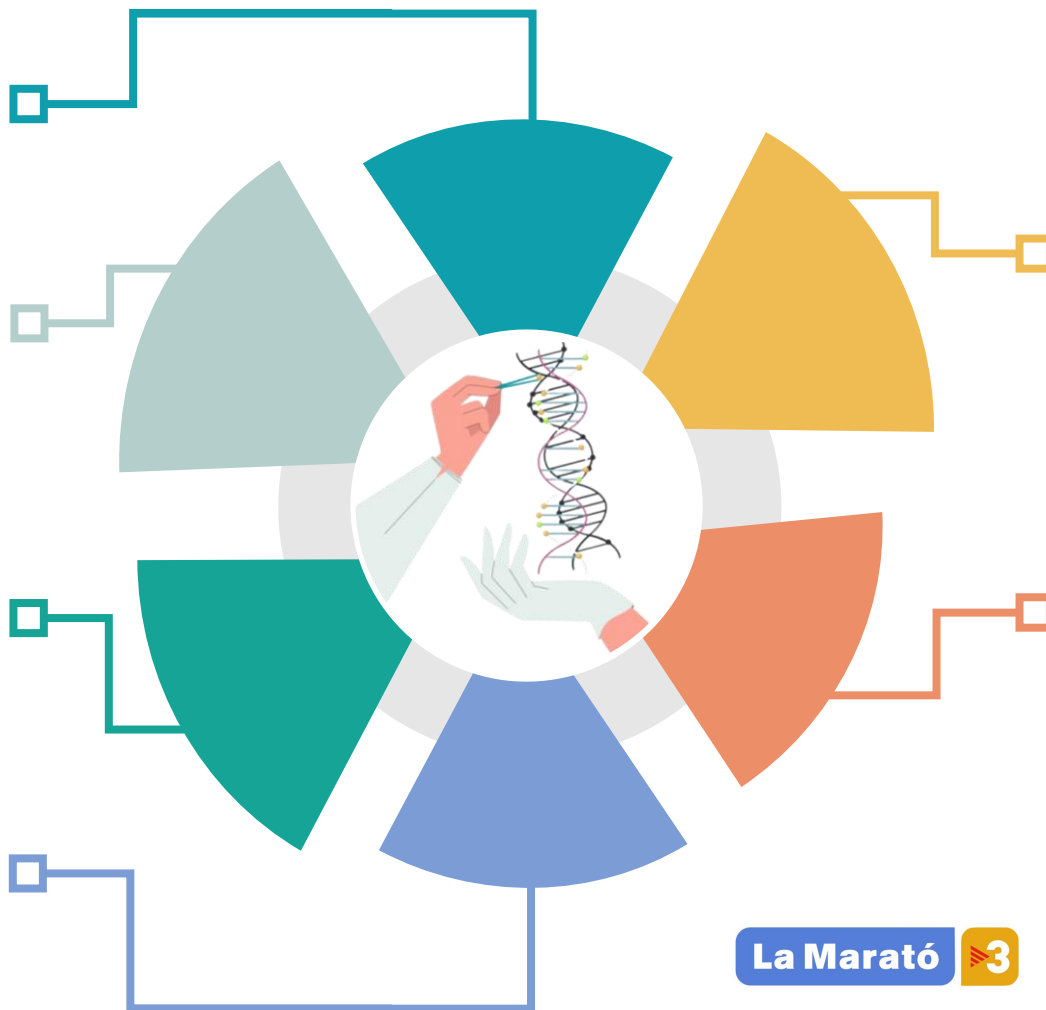
BUMPHUNTER + OUTLIER TEST

★ **MANOVA**
Análisis multivariante
de la varianza

MLM
Modelo lineal
multivariante

iForest
Isolation forest

MAH-DIST
Distancia de
Mahalanobis



QUANTILE ★
Distribución de
cuantiles

BETA
Distribución de
betas

OUTLIERS + CLUSTERING

La Marató 3

brge

ISGlobal

MÉTODOS DE DETECCIÓN

BUMPHUNTER + OUTLIER TEST

★ **MANOVA**
Análisis multivariante
de la varianza

MLM
Modelo lineal
multivariante

iForest
Isolation forest

MAH-DIST
Distancia de
Mahalanobis



La Marató 3

LB brge

ISGlobal

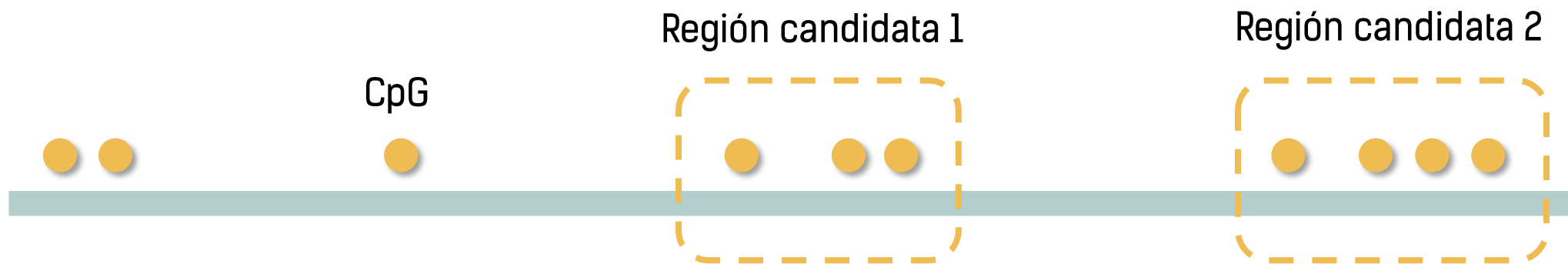
MÉTODOS DE DETECCIÓN

BUMPHUNTER + OUTLIER TEST

01

IDENTIFICACIÓN DE REGIONES CANDIDATAS

Regiones con al menos 3 CpGs separados por menos de 1 kb



~ 450,000 CpGs → ~ 40,000 regiones candidatas

La Marató 3

brge

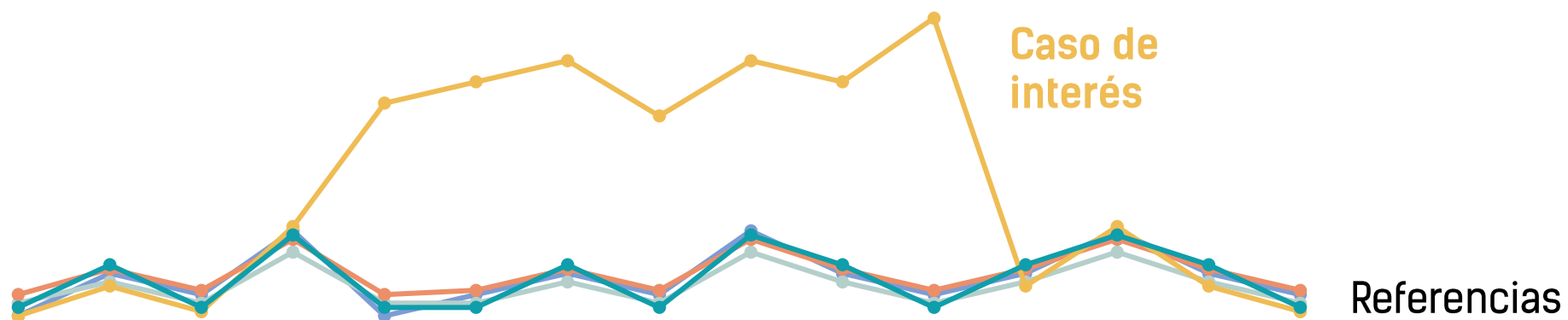
ISGlobal

MÉTODOS DE DETECCIÓN

BUMPHUNTER + OUTLIER TEST

02 IDENTIFICACIÓN DE REGIONES DIFERENCIALMENTE METILADAS

Compara la muestra de interés con el panel de referencia e identifica todas aquellas regiones que están diferencialmente metiladas en la muestra.



MÉTODOS DE DETECCIÓN

BUMPHUNTER + OUTLIER TEST

03

OUTLIER TEST

Usando uno de los 4 métodos posibles, se evalúa cada región por separado y se identifican aquellas que son significativas.

MANOVA

Análisis multivariante
de la varianza

MLM

Modelo lineal
multivariante

iForest

Isolation forest

MAH-DIST

Distancia de
Mahalanobis

MANOVA Y MODELO LINEAL MULTIVARIANTE

ANOVA:

$$\text{cp}g_i = \alpha + \beta x_i$$

$$i = 1 \dots n$$

x = caso/control

MANOVA:

$$c(\text{cp}g_i, \text{cp}g_{i+1}, \text{cp}g_{i+2}, \dots, \text{cp}g_{i+k}) = \alpha + \beta x$$

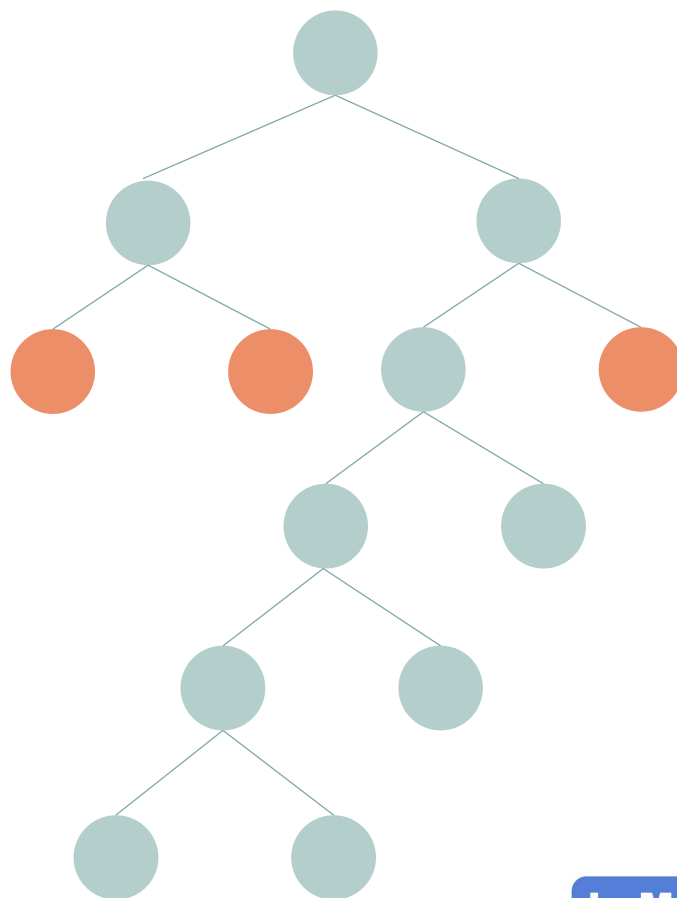
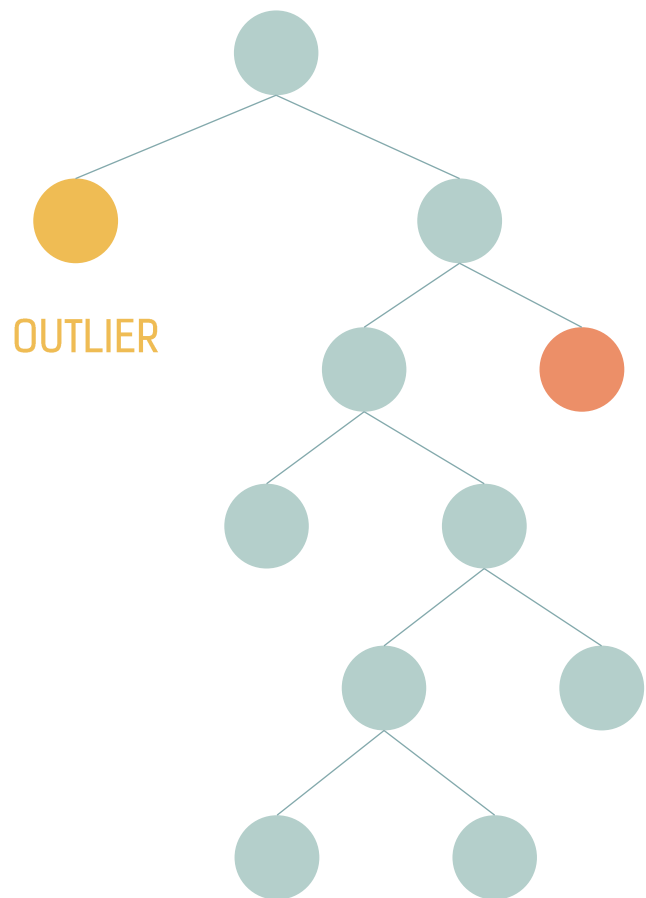


Asume distribución normal multivariada

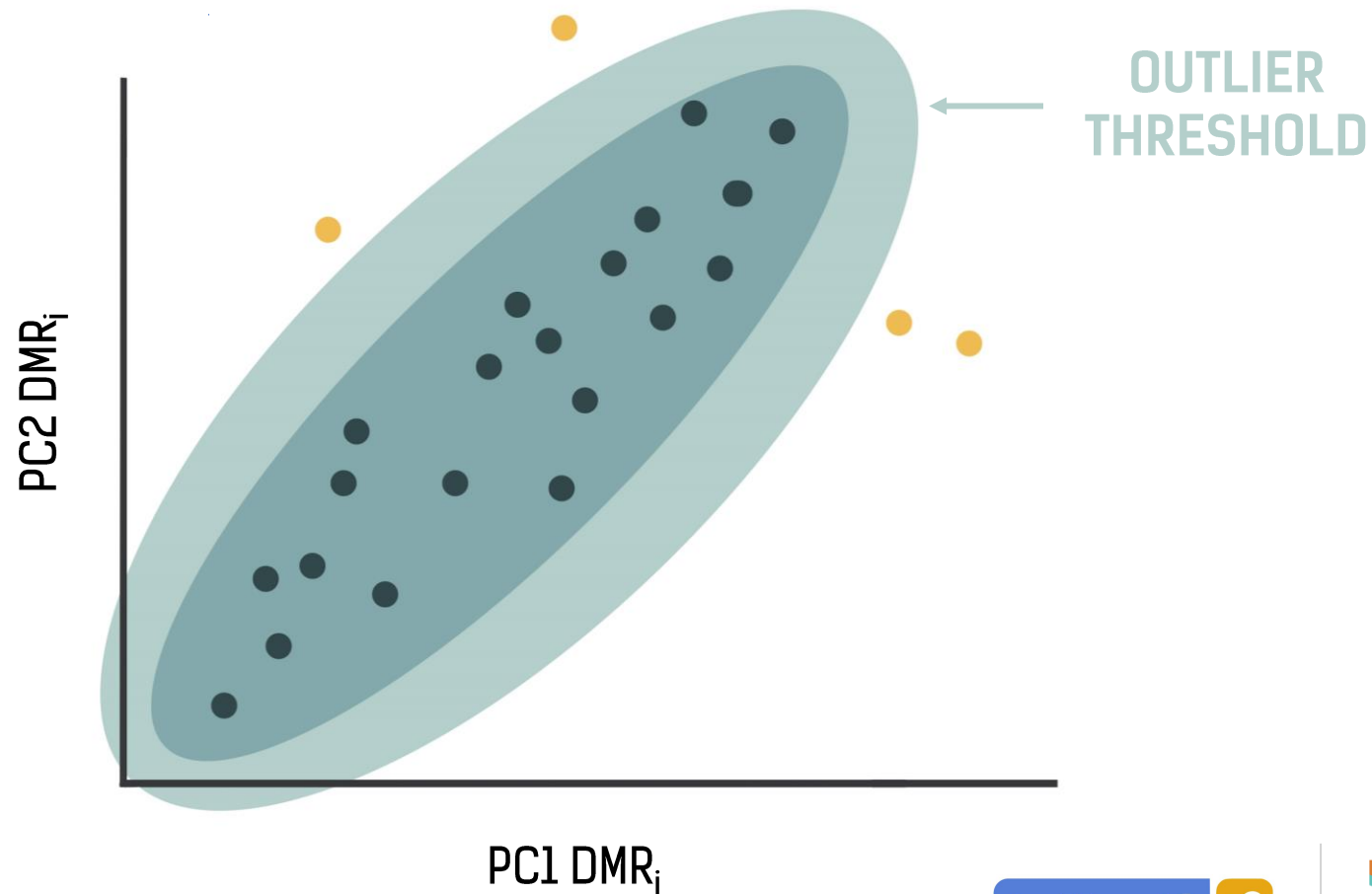


MLM: Método robusto y no paramétrico

ISOLATION FOREST



DISTANCIA DE MAHALANOBIS



MÉTODOS DE DETECCIÓN

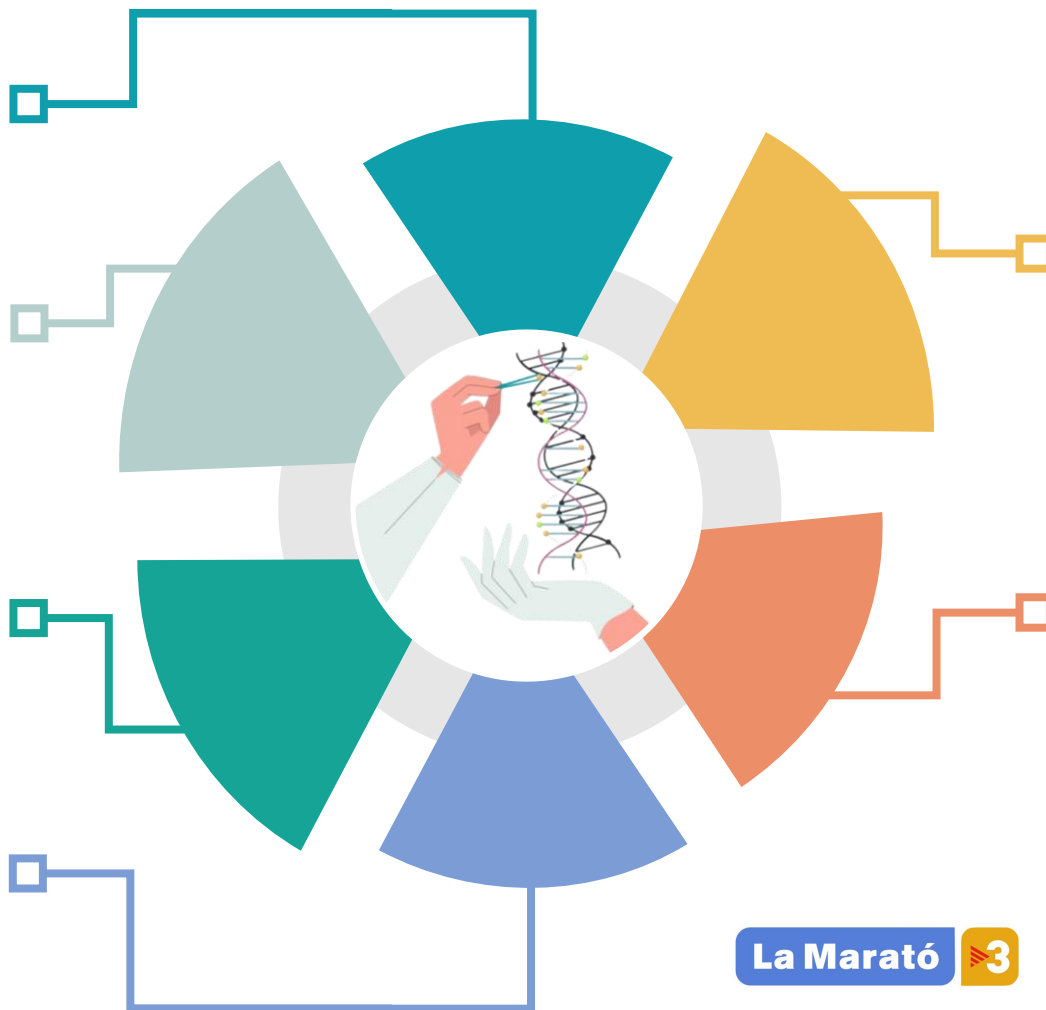
BUMPHUNTER + OUTLIER TEST

★ **MANOVA**
Análisis multivariante
de la varianza

MLM
Modelo lineal
multivariante

iForest
Isolation forest

MAH-DIST
Distancia de
Mahalanobis



QUANTILE ★
Distribución de
cuantiles

BETA
Distribución de
betas

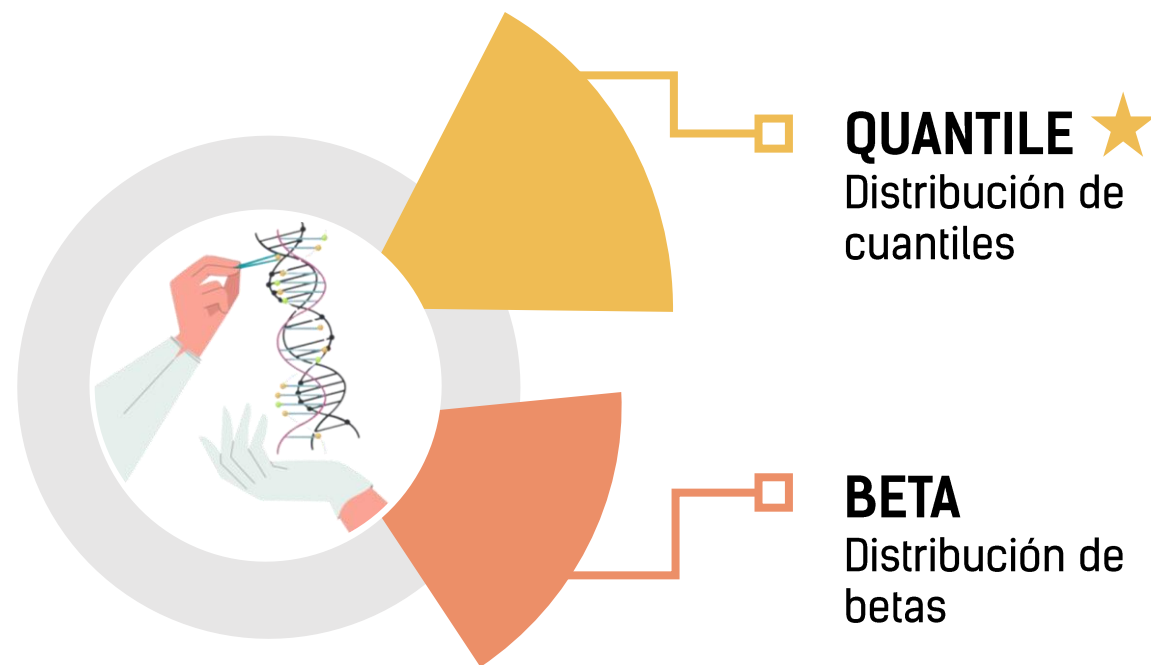
OUTLIERS + CLUSTERING

La Marató 3

brge

ISGlobal

MÉTODOS DE DETECCIÓN



OUTLIERS + CLUSTERING

MÉTODOS DE DETECCIÓN

OUTLIERS + CLUSTERING

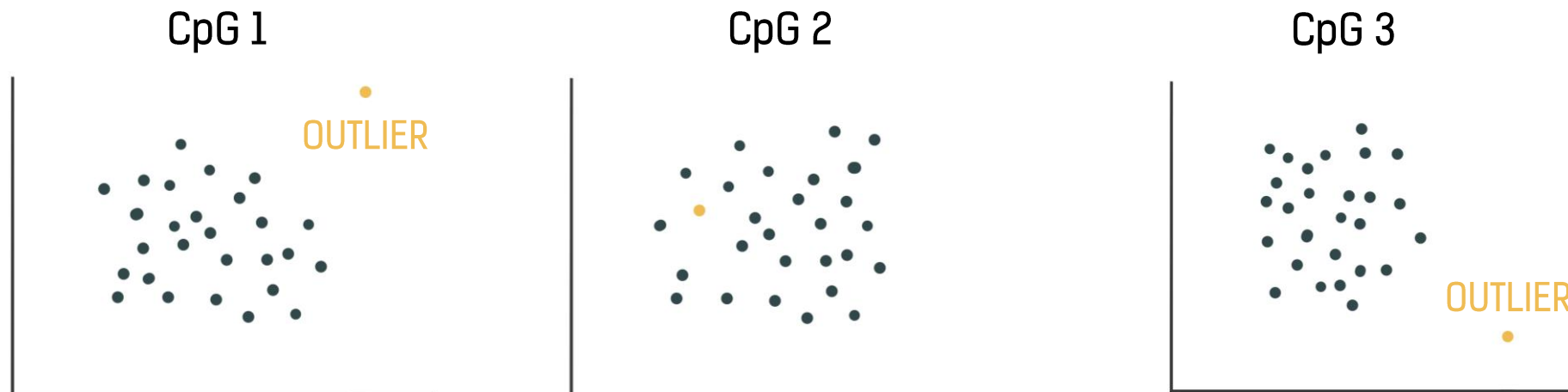
01

IDENTIFICACIÓN DE OUTLIERS PARA CADA CpG

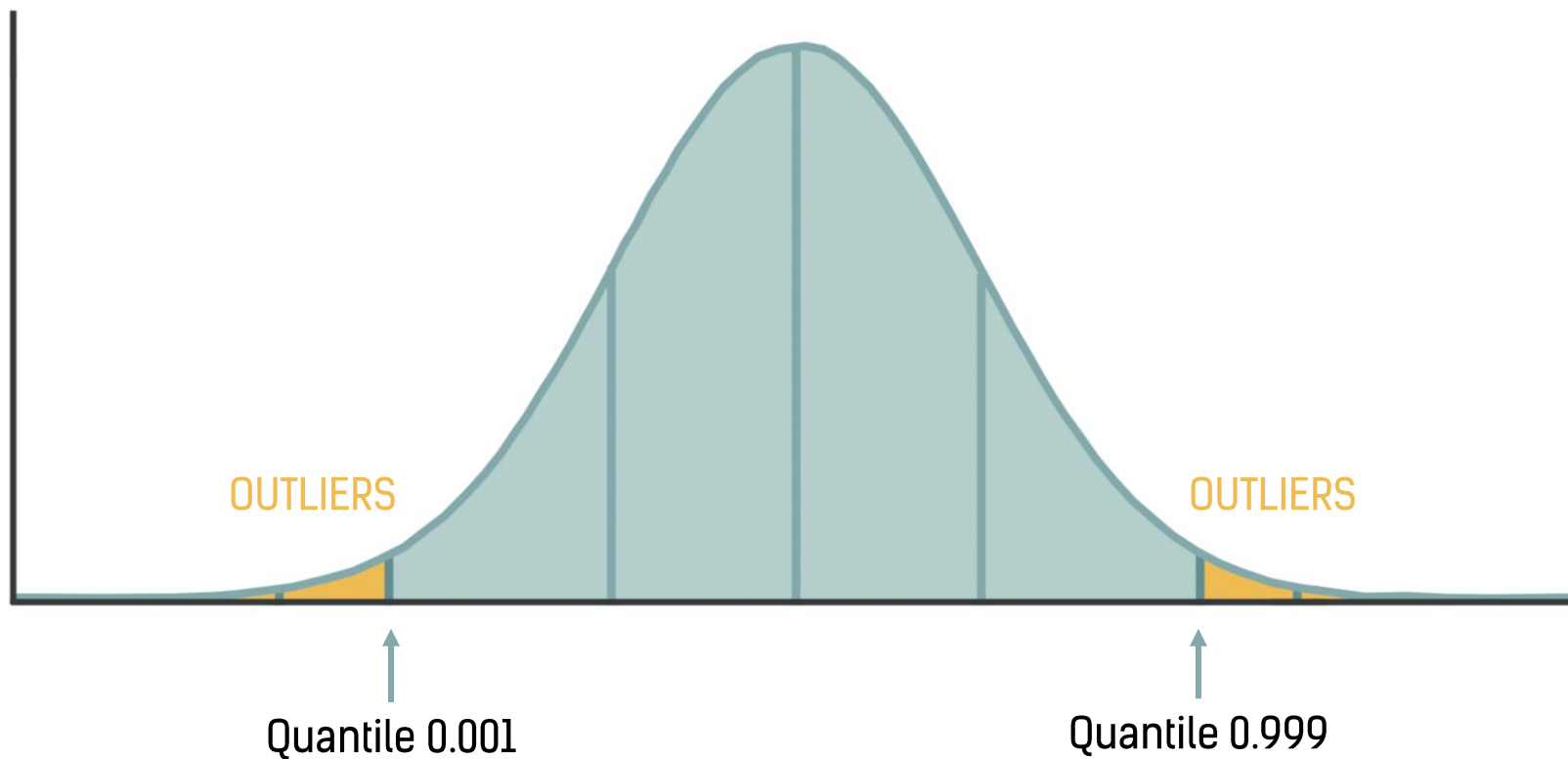
Para cada CpG, se analiza si la muestra tiene una metilación anormal comparado con el panel de referencia.

QUANTILE
Distribución de
cuantiles

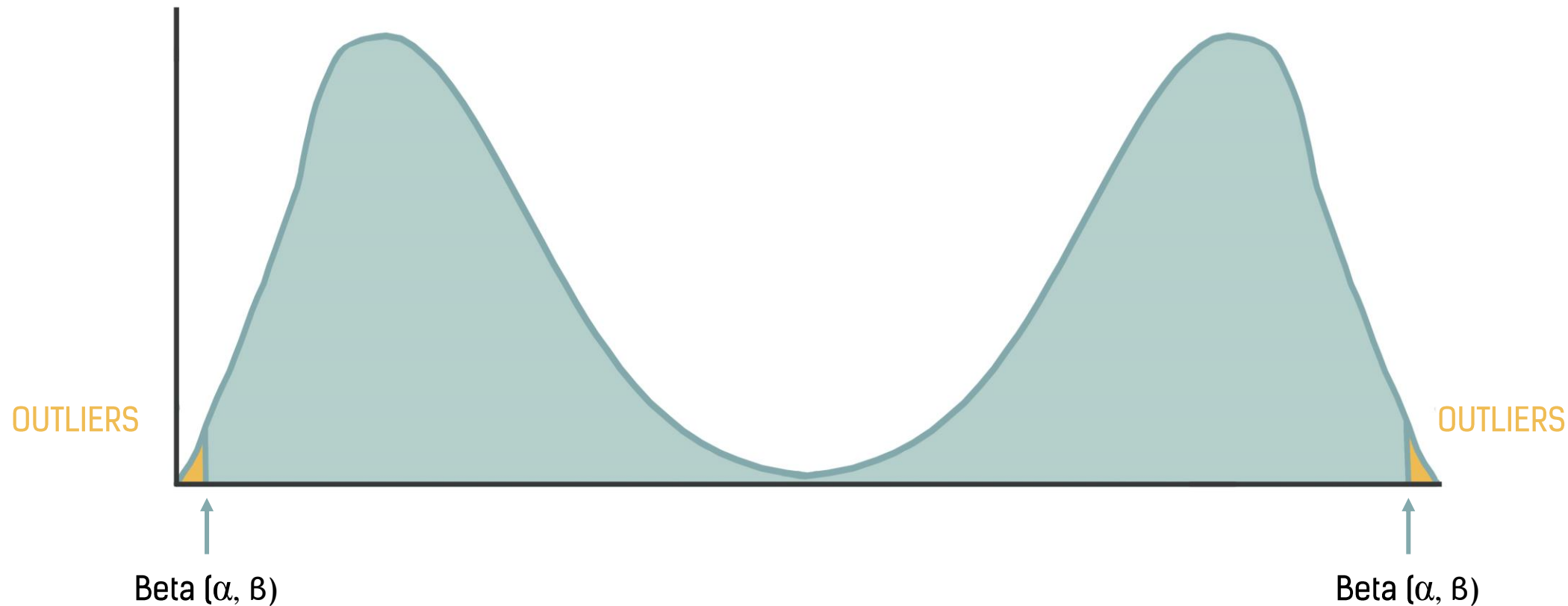
BETA
Distribución de
betas



DISTRIBUCIÓN DE CUANTILES



DISTRIBUCIÓN DE BETAS





MÉTODOS DE DETECCIÓN

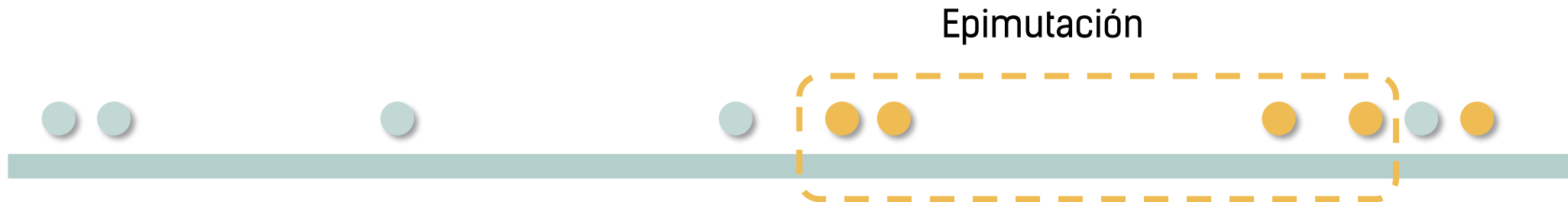
OUTLIERS + CLUSTERING

02

IDENTIFICACIÓN DE EPIMUTACIONES

Cuando encontramos 3 o más CpGs *outliers* consecutivos con menos de 1kb entre ellos, se considera una epimutación.

- OUTLIER
- METILACIÓN NORMAL

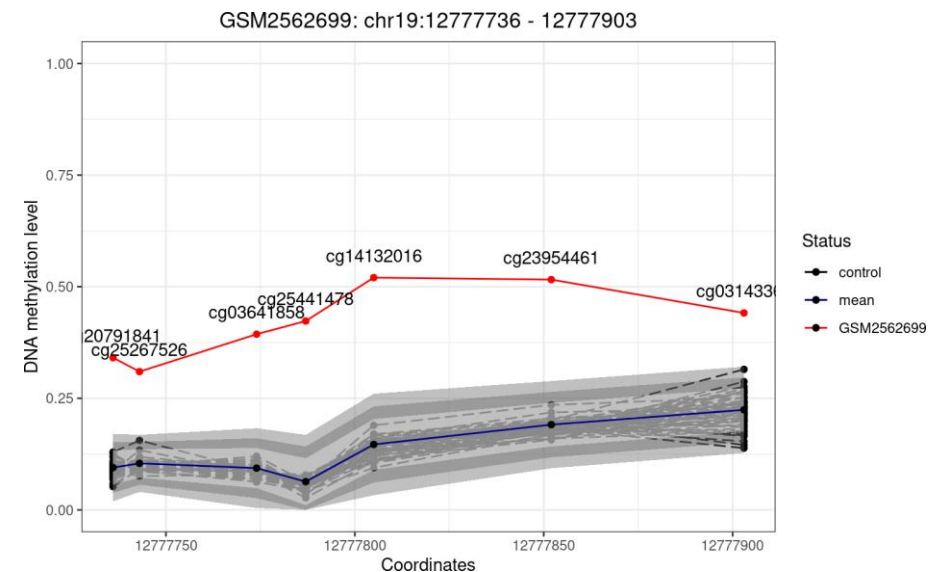


PAQUETE *EPIMUTACIONES*

OUTPUT

Sample	Chr	Start	End	Size	Num. CpGs
GSM256269	chr19	12777736	12777903	167	4
CLT1236766	chr8	89554	89634	80	3

CpGs	Direction	P value	P value adjusted	GENCODE
cg20791841,cg25267526, cg03641858,cg25441478	hypermethylation	$3.44 \cdot 10^{-9}$	$2.17 \cdot 10^{-6}$	<i>MSRA, TNKS, RP1L1, RP1L1</i>
cg65267526,cg13641858, cg890374639	Hypomethylation	$2.03 \cdot 10^{-13}$	$6.11 \cdot 10^{-11}$	<i>IL2, ADAD1, IL2</i>



CONSIDERACIONES

- El método que da mejores resultados es **Quantile**.
- Los **controles** deberían estar hechos con el **mismo array** y ser del **mismo sexo** que la muestra de interés.
- En la medida de lo posible, usar el método **caso-control** y no One-leave-Out.
- Para priorizar epimutaciones, es recomendable comprobar si la epimutación afecta la **expresión** de algún gen cercano.



¡MANOS A LA OBRA!

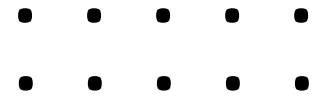
<https://cursopal.isglobal.org>

Github link: https://github.com/isglobal-brge/course_epimutations

La Marató 3

brge

ISGlobal



KAHOOT

<https://kahoot.it/>

