

Área de Genética y Genómica

Grupo de Genética y Genómica de Enfermedades raras y Complejas

Métodos y recursos bioinformáticos para el reanálisis de casos con enfermedades raras

Grupo de Bioinformática

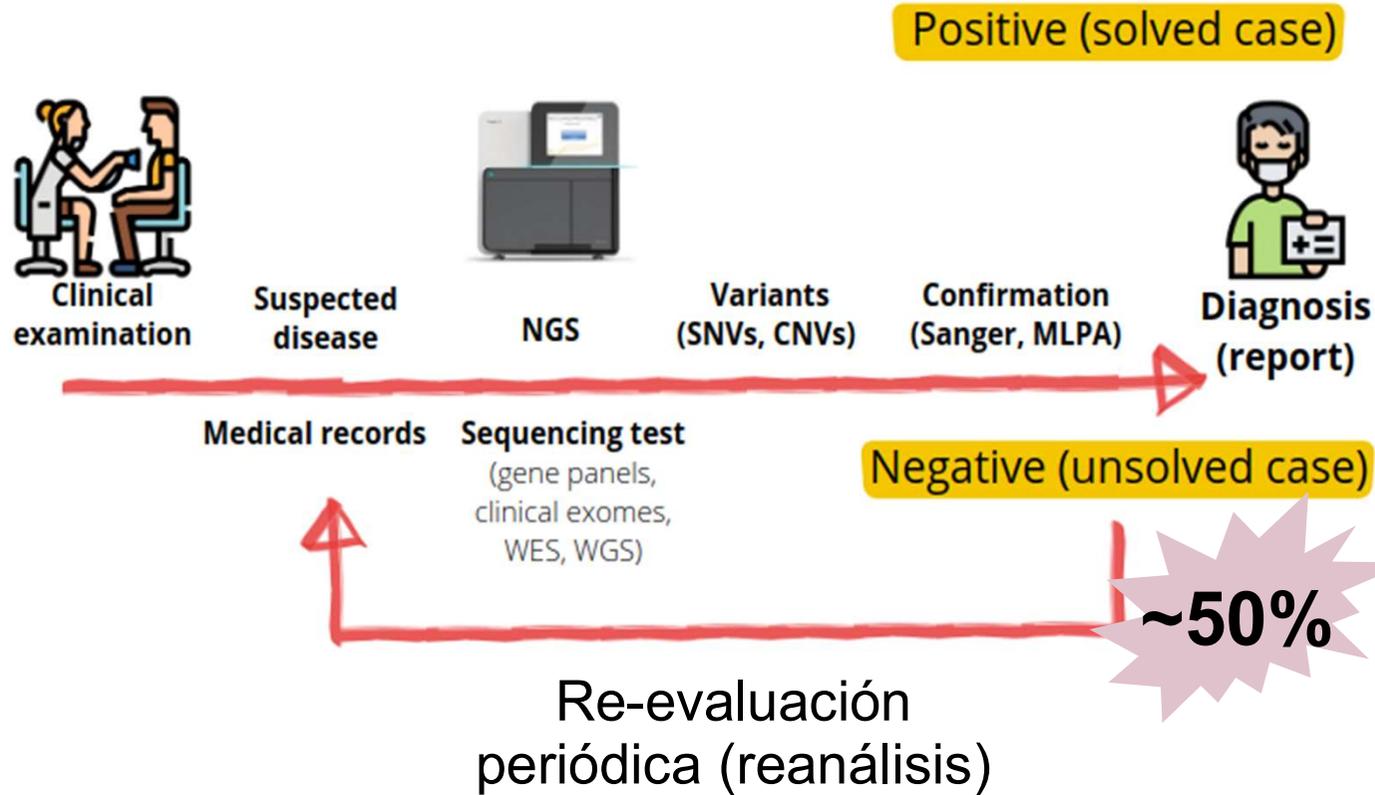
Pablo Mínguez (MSII)

pablo.minguez@quironosalud.es (tlf. 2124), FJD 4ª planta (junto a Immunología)

IV REUNIÓN ANUAL DE ÁREAS Y GRUPOS DEL IIS-FJD
29 de septiembre del 2022



Protocolo de diagnóstico de enfermedades raras



¿Cómo resolver los casos complejos?

- Incrementando la región genómica secuenciada (WES, **WGS**).
- Nuevas técnicas de secuenciación (**long reads**).
- Mejores métodos de detección de variantes que se adapten a las limitaciones de las técnicas (e.g. **CNVs**)
- Disminuir la incertidumbre en la asociación clínica de las variantes (**VUS**).
- Buscar (ayudar a predecir) **nuevos genes** implicados.

Un sistema de reanálisis que integra varias estrategias

Pipeline

SNVs
GATK based

CNVs
Methods overlapping approach

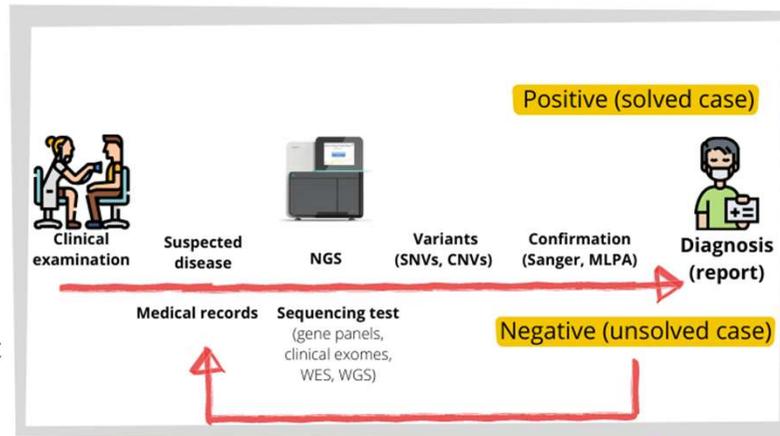
- CE, WE, WGS
- long, short reads

Filtering & prioritization

SNV own cohort frequencies

- Disease specific MAFs
- Pseudocontrols MAFs

Genetic diseases Diagnosis protocol



Reanalysis pipeline

VUS prioritization

- Rules based algorithm
- Base on cohort frequencies
- Next: Machine learning

New candidate genes

- WES analysis
- Network Biology approach

Analysis Tools (PriorR)

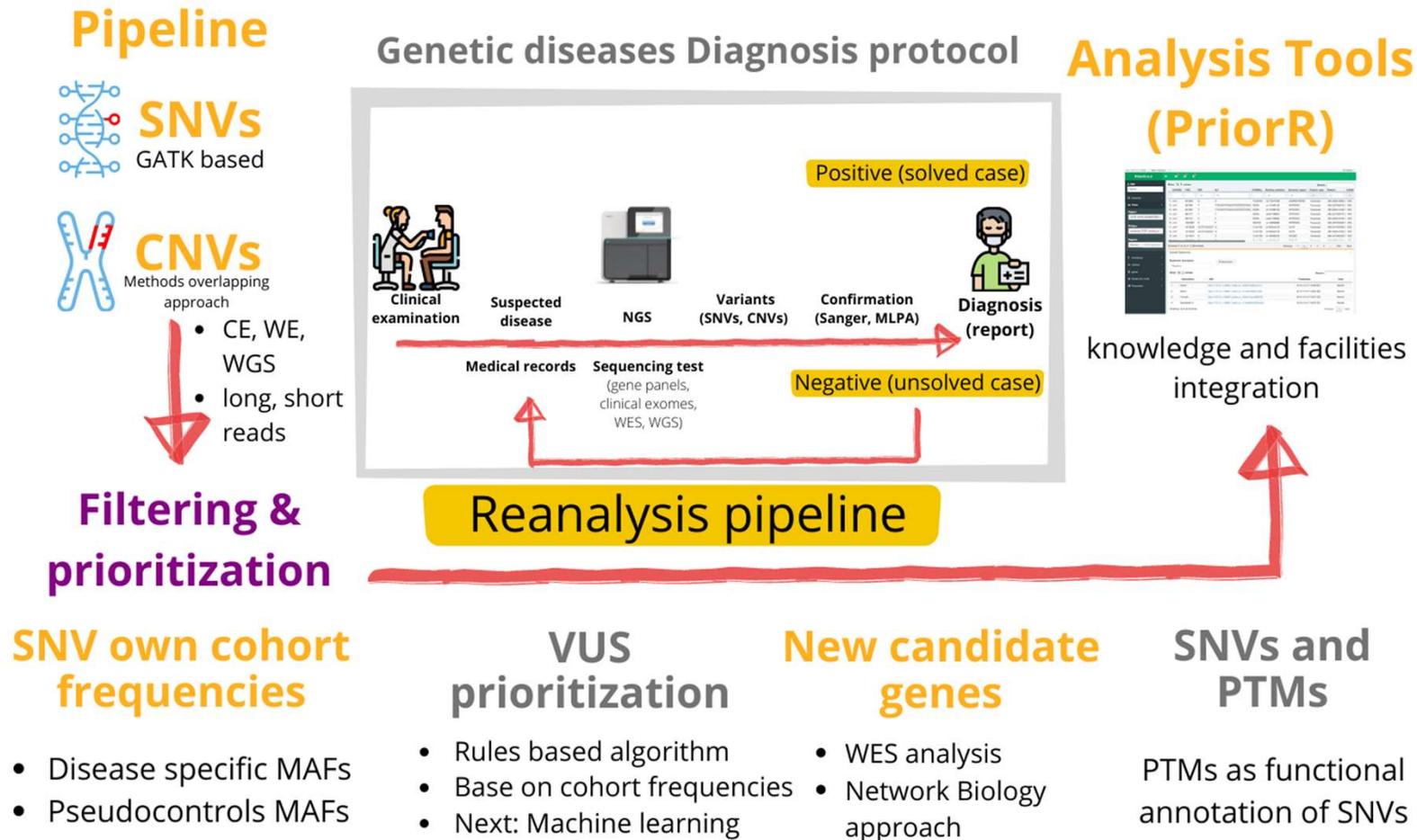


knowledge and facilities integration

SNVs and PTMs

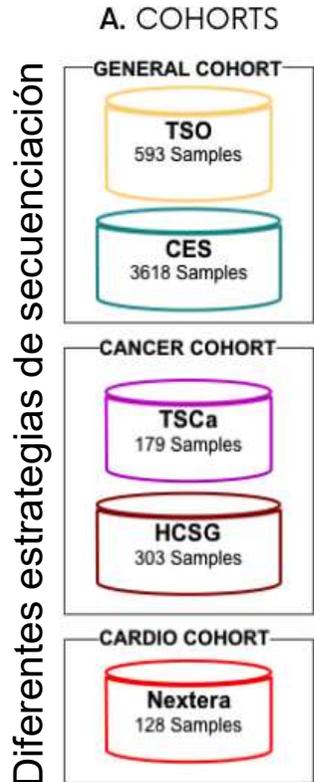
PTMs as functional annotation of SNVs

Un sistema de reanálisis que integra varias estrategias



Detección de SNVs – Es posible complementar el diagnóstico primario

Comparamos la detección de SNVs del software comercial y el nuestro (FJD-pipeline)



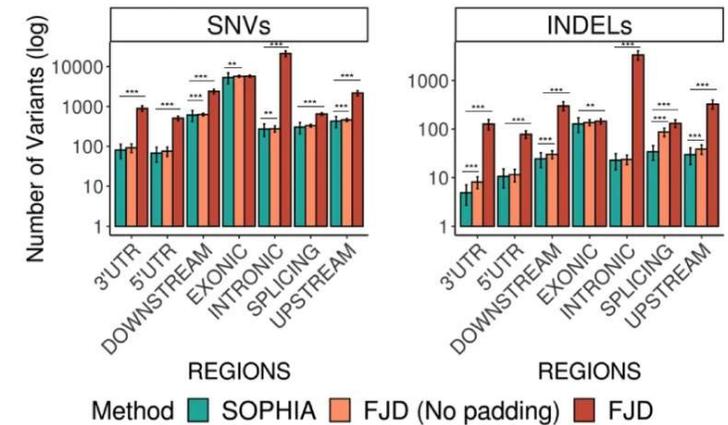
Software comercial
(diagnóstico)



FJD-pipeline
(reanálisis)



- Detectábamos el 99.7% de las variantes causales en casos resueltos.
- Detectábamos mas INDELS y podíamos reportar variantes no exónicas.
- **Algoritmo de reanálisis:** Cogemos las variantes diferentes (~30) y las priorizamos (no exónicas e indels, sig. clínica) (~10 min).
- Aumento en el rendimiento diagnóstico de un **4.4%**



Romero *et al.* NPJ Genome Med 2022

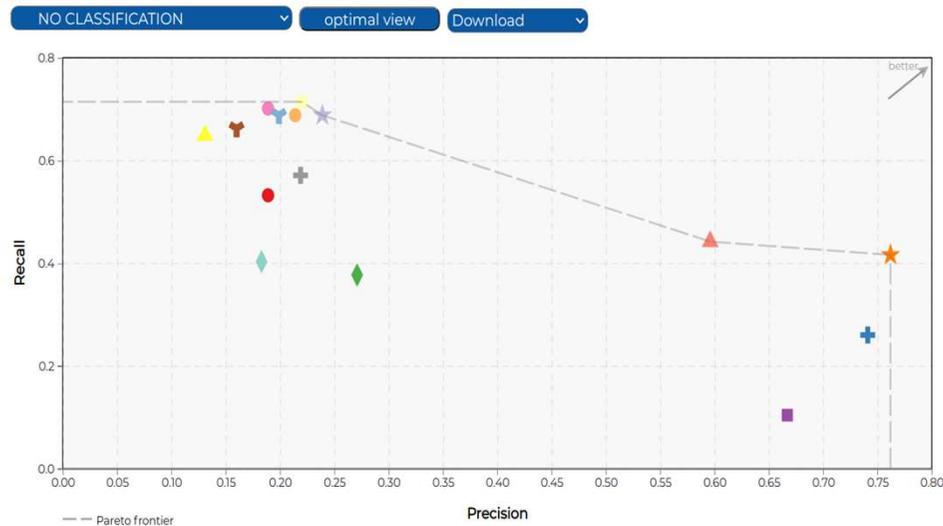
Detección de CNVs – Ejemplo de esfuerzo en comunidad

Detección de CNVs en secuenciación dirigida: Complicada. No estándares de análisis



Scientific Benchmarking Technical Monitoring

<https://openebench.bsc.es>



nextflow

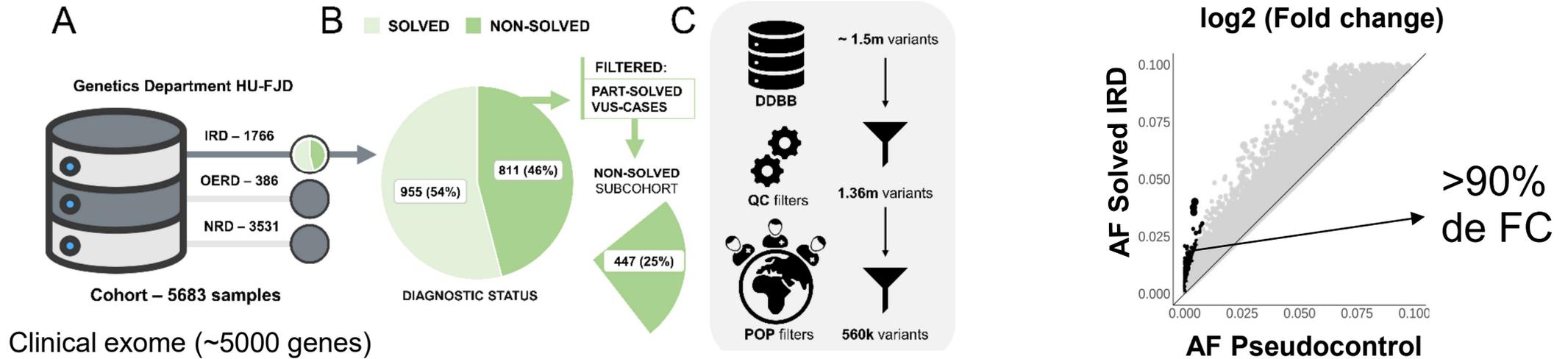
- 15 pipelines evaluadas
- Aplicadas a muestras con CNVs conocidas
- Evaluación de la detección

- Pipelines disponibles para descargar
- Evaluación de su capacidad en el análisis de WES.
- Posibilidad de subir su pipeline y compararse con las demás.
- Adquisición de estándares
- Producción de guías de buenas prácticas.

Otras comunidades: CIBERER, RAREGenomics, IMPACT-G

Priorización de variantes utilizando una base de datos de frecuencias alélicas

Utilizar la información de la cohorte en su conjunto para ayudar a pacientes concretos



Cohorte de pacientes divide en

- Distrofias de retina
- Otras patologías oculares
- No oculares

Enfermedades mendelianas pueden hacer de controles de otras patologías
NO OCULARES:
PSEUDOCONTROLES

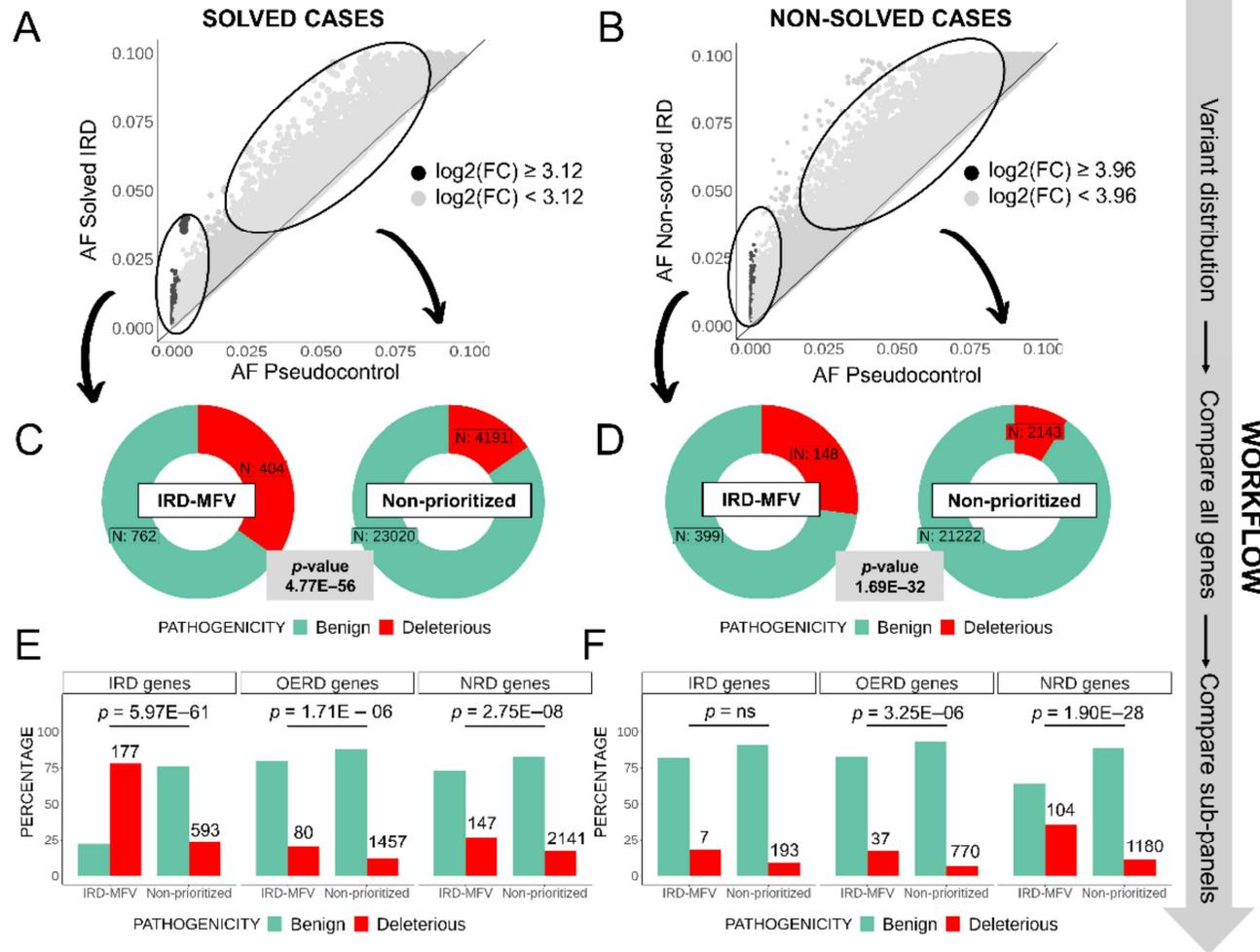
Base de datos de frecuencias alélicas:

- Por enfermedades (su control)
- Casos resueltos/no resueltos

Algoritmo de priorización basado en la comparación de frecuencias en casos y pseudocontroles

MFV (significativas)

Priorización de variantes asociadas a distrofias hereditarias de la retina



IRD-MFV: variantes significativamente más frecuentes en distrofias de retina

- 63 variantes de significado incierto en seguimiento, 10/11 reclasificadas a probablemente patogénicas.
- Reevaluación de casos: 8 casos con variantes asociadas al fenotipo.

Iancu *et al.*, 2022 IJMS

Priorización de genes candidatos asociados a enfermedades raras

Ayudar a la priorización de genes candidatos en análisis de WES y WGS

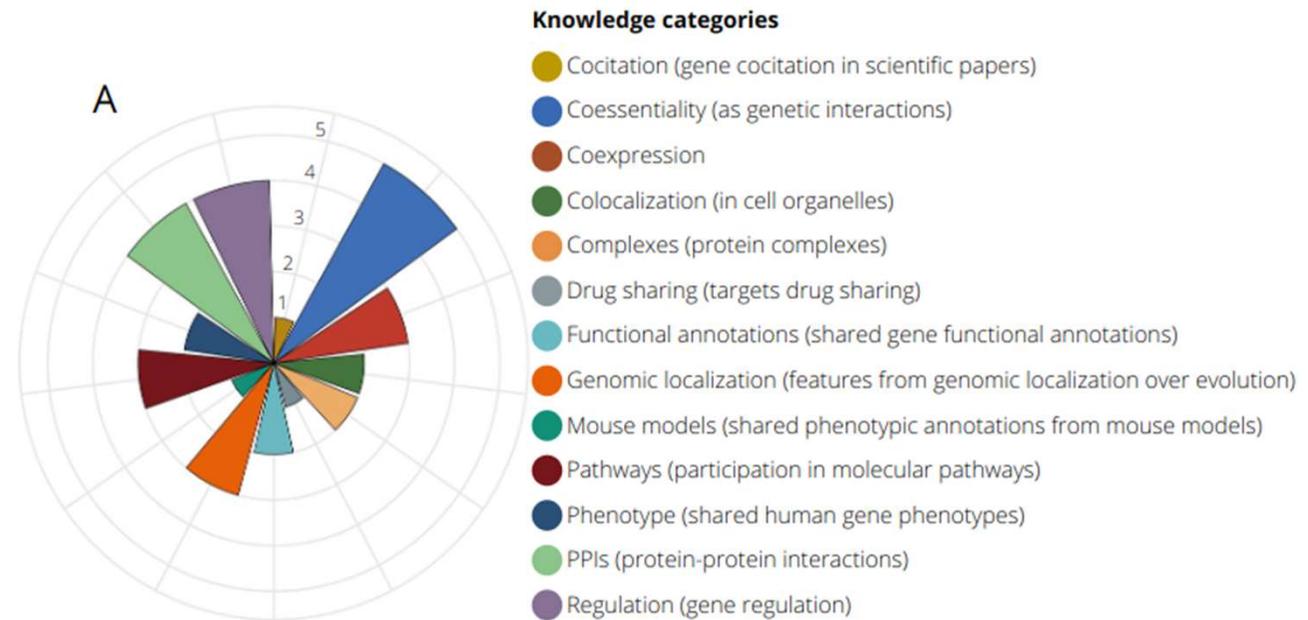
Principios e hipótesis:

1) genes causantes de la misma enfermedad están funcionalmente relacionados;

2) genes funcionalmente relacionados están cerca un interactoma funcional;

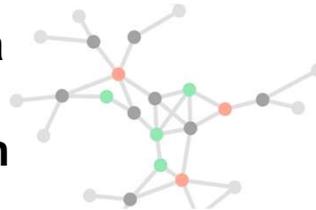
3) la capacidad de diferentes tipos de interactomas funcionales de encontrar genes candidatos varía con la enfermedad

33 redes de diferentes tipos de interacciones funcionales (**Categorías de conocimiento**)



Diferentes categorías funcionales tienen diferente capacidad predictiva

Propagación de la
señal funcional
Random Walk with
Restart



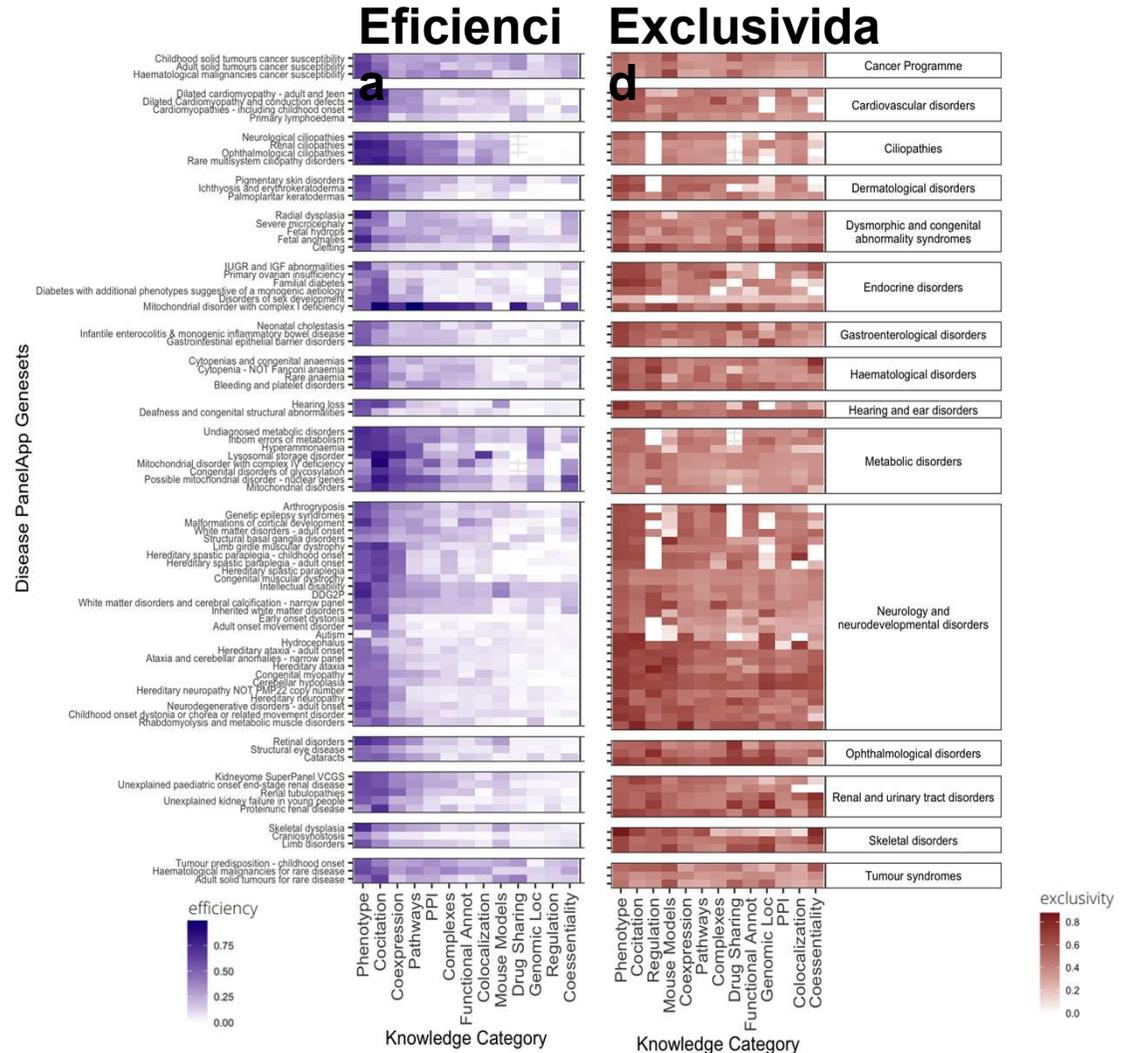
70% entrenamiento
30% validación



Predicciones

90 enfermedades raras Genomics
England PanelApp (seeds)

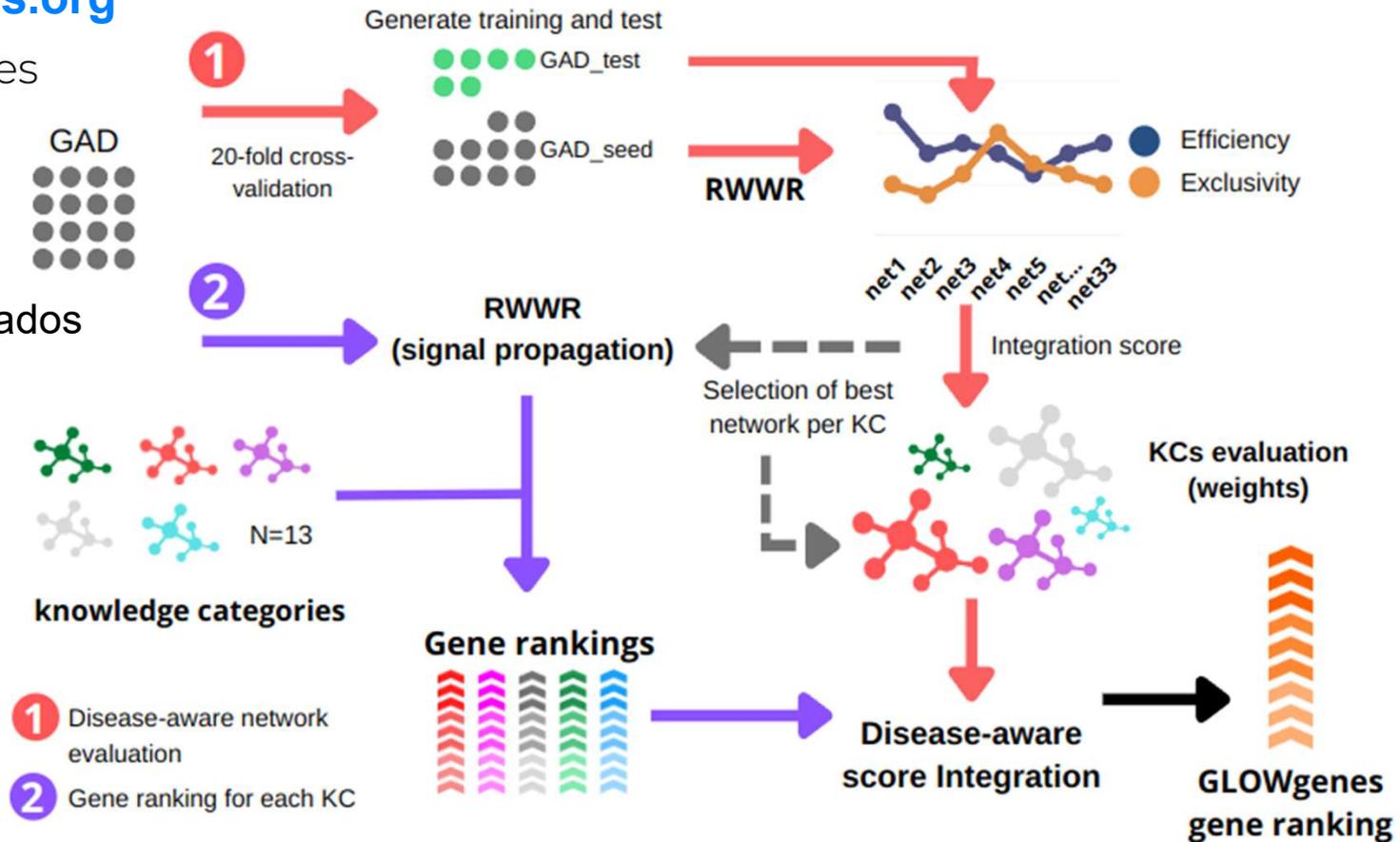
- **Eficiencia (recall):** captura de verdaderos positivos.
- **Exclusividad:** captura de genes de forma única.



GLOWgenes. Un algoritmo para priorizar genes candidatos en el diagnóstico de enfermedades raras

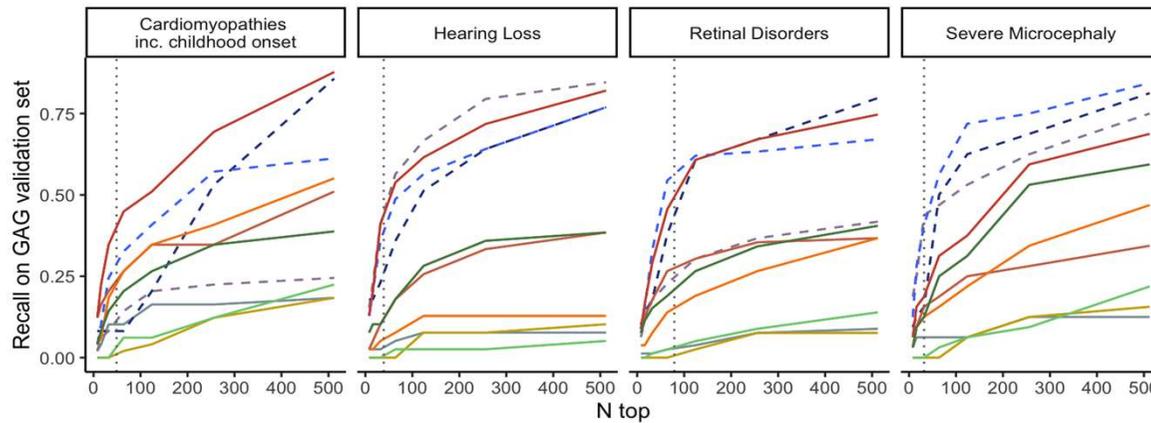


GAD: genes asociados a enfermedad



GLOWgenes predice sobretodo los genes “menos obvios”

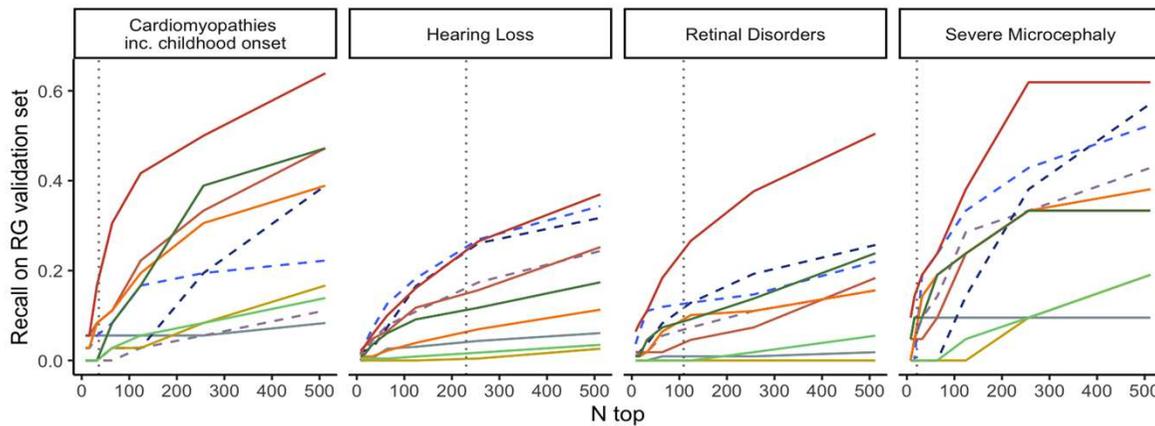
Evaluación en genes asociados a enfermedad (Green/Ambar)



- RED** STOP: not enough evidence for this gene-disease; this gene should not be used for genome interpretation.
- AMBER** PAUSE: moderate evidence for this gene-disease association, and should not yet be used for genome interpretation.
- GREEN** GO: high level of evidence for this gene-disease association, demonstrates confidence that this gene should be used for genome interpretation.

- Method
- GLOWgenes
 - DIAMOnD
 - NetComb-GUILDify
 - ToppGenet - network based
 - ToppGenet - functional similarity
 - Endeavour
 - GeneMANIA
 - DisGeNET - BEFFREE
 - DisGeNET - non curated
 - DISEASES

Evaluación en genes candidatos (menos obvios) (Red)



- Métodos basados en semillas
- ⋯ Métodos de referencia (enfermedades predeterminadas)

PriorR. Programa de análisis que incorpora todas las implementaciones de nuestro reanálisis

PriorR 3.0

User:

SNV CNV

Show 10 entries

Search:

CHROM	POS	REF	ALT	Location	SYMBOL	VARIANT_CLASS	X14.1649_GT	X14.1649_VF
chrX	99662241	T	C	chrX:99662241	PCDH19	SNV	1/1	0.94
chr1	27023375	G	A	chr1:27023375	ARID1A	SNV	0/1	0.55
chr16	88494821	G	A	chr16:88494821	ZNF469	SNV	0/1	0.51
chr3	15495345	T	G	chr3:15495345	COLQ	SNV	0/1	0.47
chr6	144898360	C	A	chr6:144898360	UTRN	SNV	0/1	0.45
chr1	225599128	C	T	chr1:225599128	LBR	SNV	0/1	0.49
chr2	3623234	A	G	chr2:3623234	RPS7	SNV	0/1	0.56
chr14	94844865	G	A	chr14:94844865	SERPINA1	SNV	0/1	0.55
chr10	28903503	C	T	chr10:28903503	WAC	SNV	0/1	0.48
chr1	6528150	T	G	chr1:6528150	PLEKHG5	SNV	0/1	0.45

Showing 41 to 50 of 81 entries

Previous 1 ... 4 5 6 ... 9 Next

Variant Information

VARIANT

chr14- -94844865- -G- -A 0/1
SERPINA1
missense_variant
NM_001127701.2:c.1178C>T
NP_001121173.1:p.Pro393Leu

OMIM Phenotype:
Hemorrhagic diathesis due to antithrombin Pittsburgh, 613490 (3), Autosomal recessive
Emphysema due to AAT deficiency, 613490 (3), Autosomal recessive
Emphysema-cirrhosis, due to AAT deficiency, 613490 (3), Autosomal recessive

Population Frequency

FJD MAF: 0.00018
Gnomadg MAF: 0.000105202
Spanish Frequency: NA

Pathogenicity

CADD Phred: 23.6
Mut-Score: 0.842
ADA Score: NA

Muchas gracias!



Raquel, Gonzalo, Ionut

 @TranslationalB1



Carmen Ayuso y todo el departamento

IV REUNIÓN ANUAL DE ÁREAS Y GRUPOS DEL IIS-FJD
29 de septiembre del 2022

