



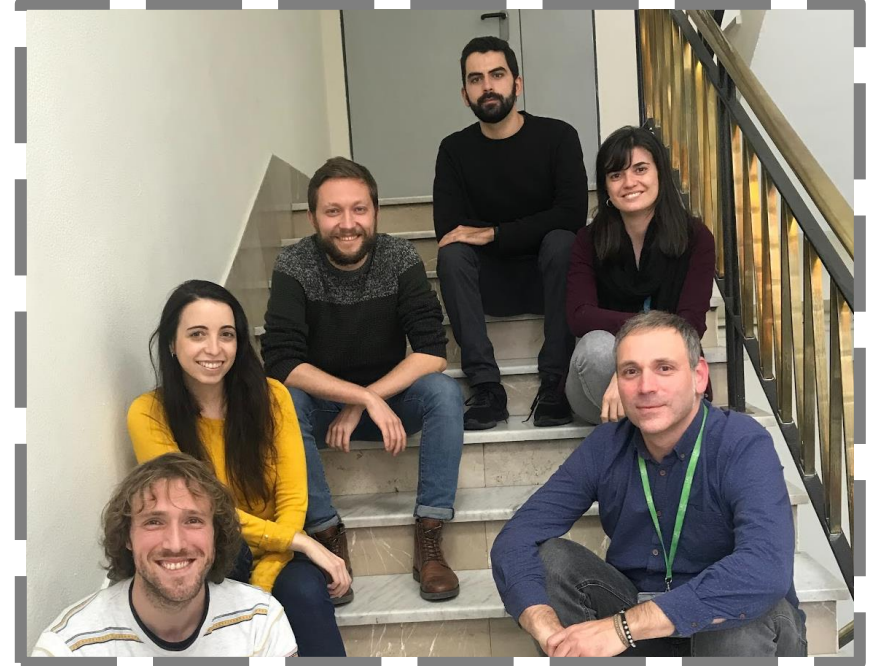
Unidad de Bioinformática del IIS-FJD

Pablo Mínguez & Gonzalo Núñez - IIS-FJD

pablo.minguez@quironosalud.es

18/05/2022

Anti-apoptotic/
pro-proliferation genes



<https://bit.ly/UnidadBioinformaticallIS-FJD>

- Pablo Mínguez (Miguel Servet II)
- Raquel Romero (Postdoc CAM)
- Ionut Iancu (Doctorando, Ayud. Inv. CAM, RAREGenomics, IMPACT)
- Gonzalo Núñez Moreno (Ayud. Inv. CAM)

- **Pablo Mínguez (coordinador)**
- **Gonzalo Núñez Moreno (bioinformático) Ayud. Inv. CAM.**

RECURSOS HUMANOS

- ✓ **AES2018.** Contratos Técnicos Bioinformáticos, ISCIII.
- ✓ **Ene 2021.** Contrato Ayudante investigación CAM.
- ✓ **Mar 2022.** Contrato IMPACT 8 meses.
- ✓ **Solicitado AES2022.** Contratos Técnicos Bioinformáticos, ISCIII.

Recursos computacionales

- ✓ **Cluster de computación científica:** 4 servers Intel Xeon Gold 6138F, 3 PowerEdge R440, 448GB RAM y 280 cores.

NodoBio en el Centro de Supercomputación de la Universidad Autónoma de Madrid (CCC-UAM). Compartido con CBMSO e IMDEA alimentación.



CIVP18A3862, CP16/00116,
PI18/00579

- ✓ Servidor en Telefónica SA, 64GB de RAM y 16 cores.
- ✓ Almacenamiento de 24TB en Telefonica SA.

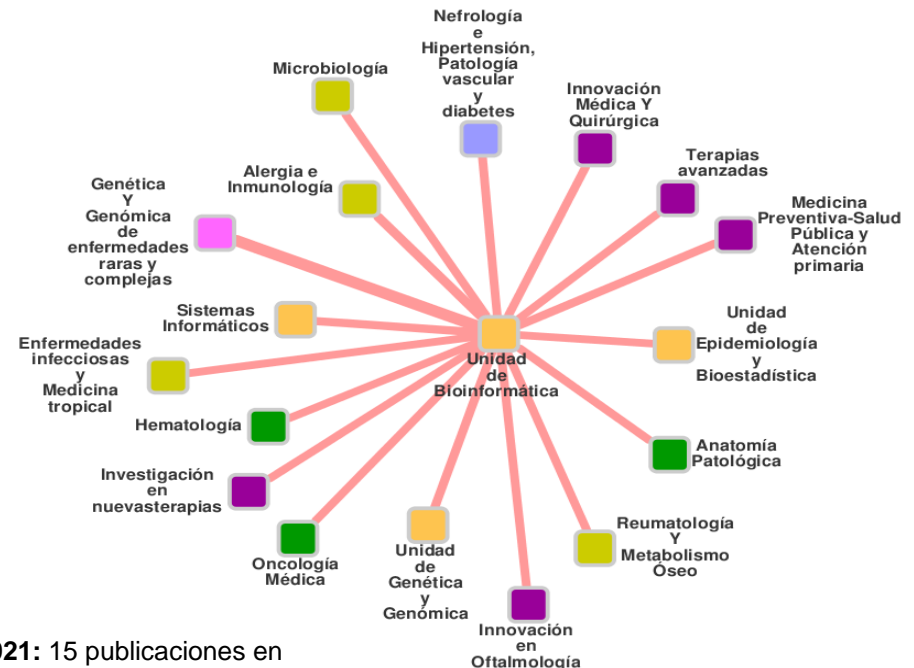


- ✓ Servidor DELL PowerEdge T440 160GB RAM y 13TB almacenamiento.
- ✓ Varias estaciones de trabajo.



Objetivos de Unidad de Bioinformática del IIS-FJD

- Apoyo al diseño y análisis de **experimentos ómicos** (Transcriptómica, Genómica, Proteómica) a grupos del IIS-FJD (y externos) en régimen de colaboración.
 - *También AI en datos clínicos.*
- Diseño, implementación y mantenimiento de **software bioinformático**.
- Asistencia al IIS-FJD en **proyecto institucionales** con necesidades bioinformáticas y de compartición de datos.
 - Dos proyectos integrados (PIE13/00051 y PIE15/00048), proyecto IMPACT, TransBioNet, unCOVER.
- Transferencia de conocimiento.
 - Definición de **estándares** en la comunidad científica.
- **Formación** (cursos, TFM/Gs, clases en Másteres).
- Asesoramiento en la **compartición de datos ómicos** provenientes de investigación.



2021: 15 publicaciones en colaboración con 17 grupos/unidades IIS-FJD

Análisis de experimentos ómicos - Transcriptómica

RNASeq (mRNA, miRNA).

- Descripción de clases de muestras (PCA, clustering, heatmaps).
- Firmas moleculares - Expresión diferencial con diferentes diseños: caso-control, series temporales.
- Clasificadores/predictores (aprendizaje automático).
- Caracterización funcional (GSEA, STRING, Biología de redes – SNOW).
- Seguimiento: escritura, propuesta de nuevos análisis, etc.

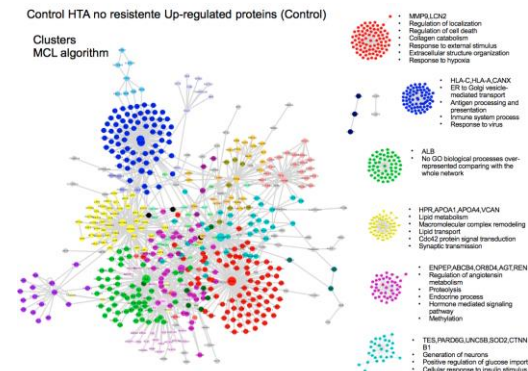
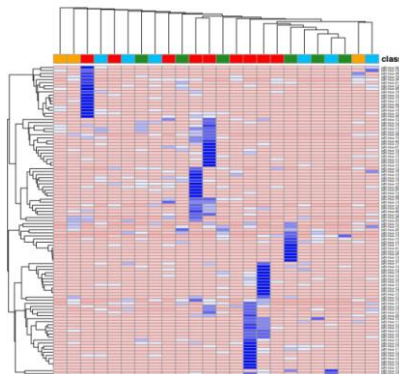
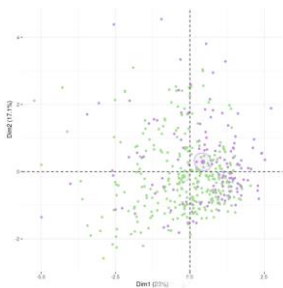
Normalización

Extracción de firmas moleculares
(DE, coexpresión)

Caracterización funcional
de las firmas moleculares

→ Seguimiento del proyecto

Descripción de clases (PCA, heatmaps, clustering).



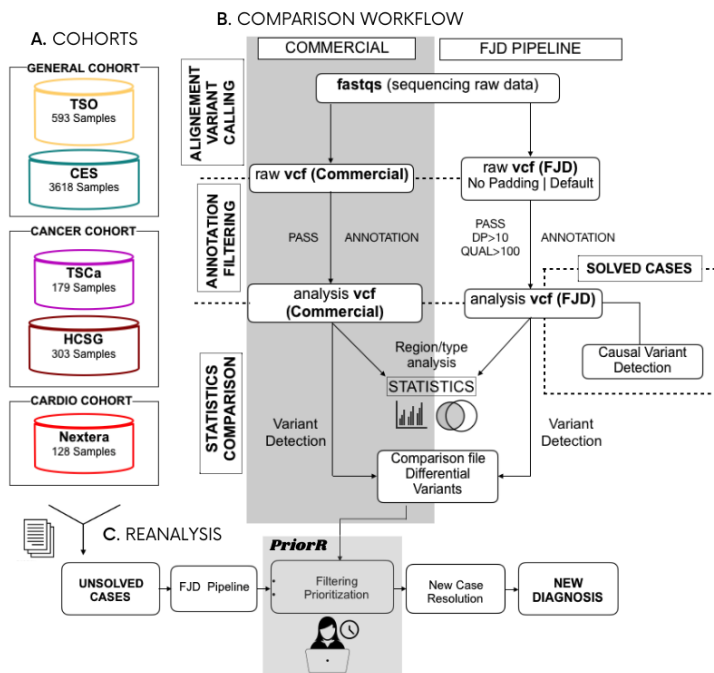
- Cuantificación de isoformas utilizando tecnología Nanopore (secuencias largas), junto a la Unidad de Genómica/Dpto. Genética, **Mini-IsoQLR**.

Análisis de experimentos ómicos - Genómica

DNASeq

- Detección y anotación de variantes **germinales** (SNVs, indels, SVs-CNVs), **FJD-pipeline**, junto al Dpto. de Genética. **Paneles de genes, WES, WGS.**
- Detección y anotación de variantes **somáticas** (SNVs, indels), junto a Anatomía Patológica y Oncología.
- Detección de **variantes en mosaico**, junto al Dpto. de Genética.
- Análisis del espectro mutacional de virus (SARS-CoV-2), para Microbiología.

EJEMPLO DE TRASLACIÓN A LA PRÁCTICA CLÍNICA



Reanálisis de casos sin diagnóstico concluyente de enfermedades genéticas (Dpto. Genética).

Sistema de detección y filtrado de variantes que aumenta en un 4% la tasa de diagnóstico sin necesidad de resecuenciar.

<https://github.com/TBLabFJD/VariantCallingFJD>



Available on **batchX**

Otros recursos en Genómica

PriorR. Herramienta de diagnóstico genético - Filtrado y priorización de variantes

The screenshot displays the PriorR 2.0 web application interface. The top navigation bar is green with the logo 'PriorR 2.0' and a 'Publish' button. The left sidebar contains a 'User' section with the name 'Rachel', a 'Columns' section, a 'Filters' section with 'Region' (3UTR, 5UTR, DOWNSTREA) and 'Biotype' (antisense, CTCF_binding_si) dropdowns, and a 'Regions' section with a 'Browse...' button. The main content area shows a table of variants with columns: CHROM, POS, REF, ALT, SYMBOL, Existing_variation, Genomic_region, Feature_type, Feature, and CANOI. The table lists 11 variants on chromosome 1, including genes like PLEKHN1, AGRN, and C1orf159. Below the table is a 'Saved Sessions' section with a 'Bookmark description' field containing 'Proyect c' and a 'Bookmark...' button. Below that is another table with columns: Description, URL, Timestamp, and User, listing 4 saved sessions. The interface includes search bars, pagination controls, and a 'Show 10 entries' dropdown.

CHROM	POS	REF	ALT	SYMBOL	Existing_variation	Genomic_region	Feature_type	Feature	CANOI	
11	chr1	914605	G	A	PLEKHN1	rs771673458	DOWNSTREAM	Transcript	XM_005244809.1	YES
12	chr1	957967	T	TTGTAGTCTGACCTGTGGTCTGAC	AGRN	rs141489152	INTRONIC	Transcript	NM_001305275.1	YES
13	chr1	957967	T	TTGTAGTCTGACCTGTGGTCTGAC	AGRN	rs141489152	INTRONIC	Transcript	XM_005244748.1	YES
14	chr1	987517	C	T	AGRN	rs561796600	INTRONIC	Transcript	NM_001305275.1	YES
15	chr1	987517	C	T	AGRN	rs561796600	INTRONIC	Transcript	XM_005244748.1	YES
16	chr1	1007987	C	T	RNF223	rs142960669	INTRONIC	Transcript	NM_001205252.1	YES
17	chr1	1018029	GCTCTGAGGT	G	C1orf159	rs764555513	3UTR	Transcript	NM_001330306.1	YES
18	chr1	1018029	GCTCTGAGGT	G	C1orf159	rs764555513	3UTR	Transcript	XM_005244764.1	YES
19	chr1	1019431	G	C	C1orf159	rs139486403	EXONIC	Transcript	NM_001330306.1	YES
20	chr1	1019431	G	C	C1orf159	rs139486403	EXONIC	Transcript	XM_005244764.1	YES

Description	URL	Timestamp	User
1 Gene1	http://127.0.0.1:6688/?_state_id=b9df7b2e64ac141f	2019-10-01T13:59:42Z	Rachel
2 Gene1	http://127.0.0.1:6688/?_state_id=015a192fd82103fe	2019-10-01T14:00:18Z	Rachel
3 FamilyA	http://127.0.0.1:6688/?_state_id=6dfc21bacd899783	2019-10-01T14:01:15Z	Rachel
4 Sample56/13	http://127.0.0.1:6688/?_state_id=77b9d9a502883a48	2019-10-01T14:02:30Z	Rachel

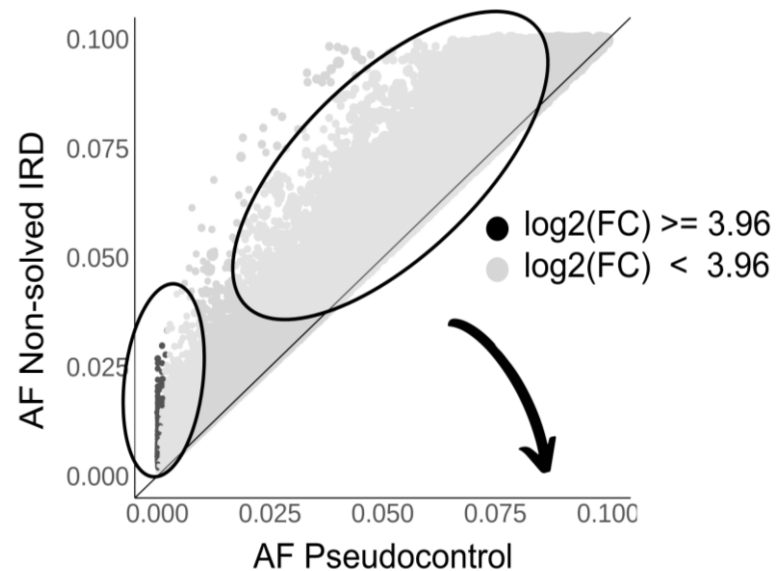
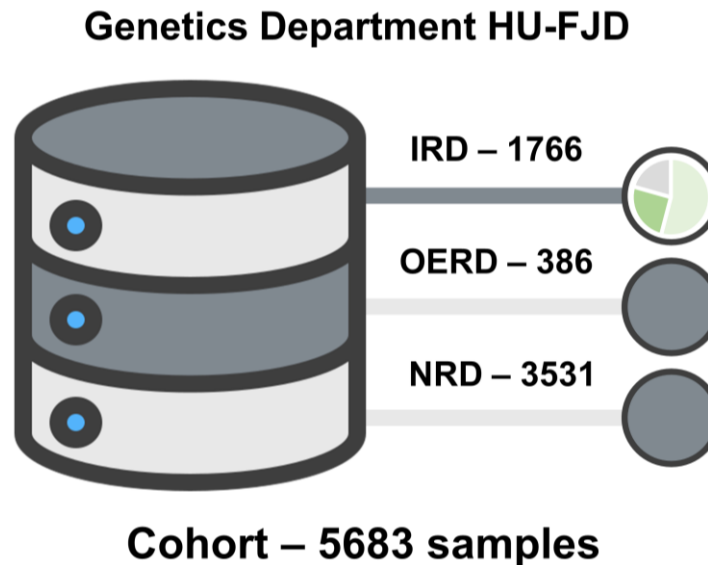
Romero et al NPJ Genom Med. 2022

<https://github.com/TBLabFJD/PriorR>

Otros recursos en Genómica

Base de datos de frecuencias alélicas en la cohorte de pacientes con enfermedades genéticas del Servicio de Genética del HU-FJD

- Frecuencias alélicas → **Filtrado de variantes raras.**
- Comparación frecuencias en patología vs controles (otras enfermedades) → **Priorización de variantes patogénicas.**



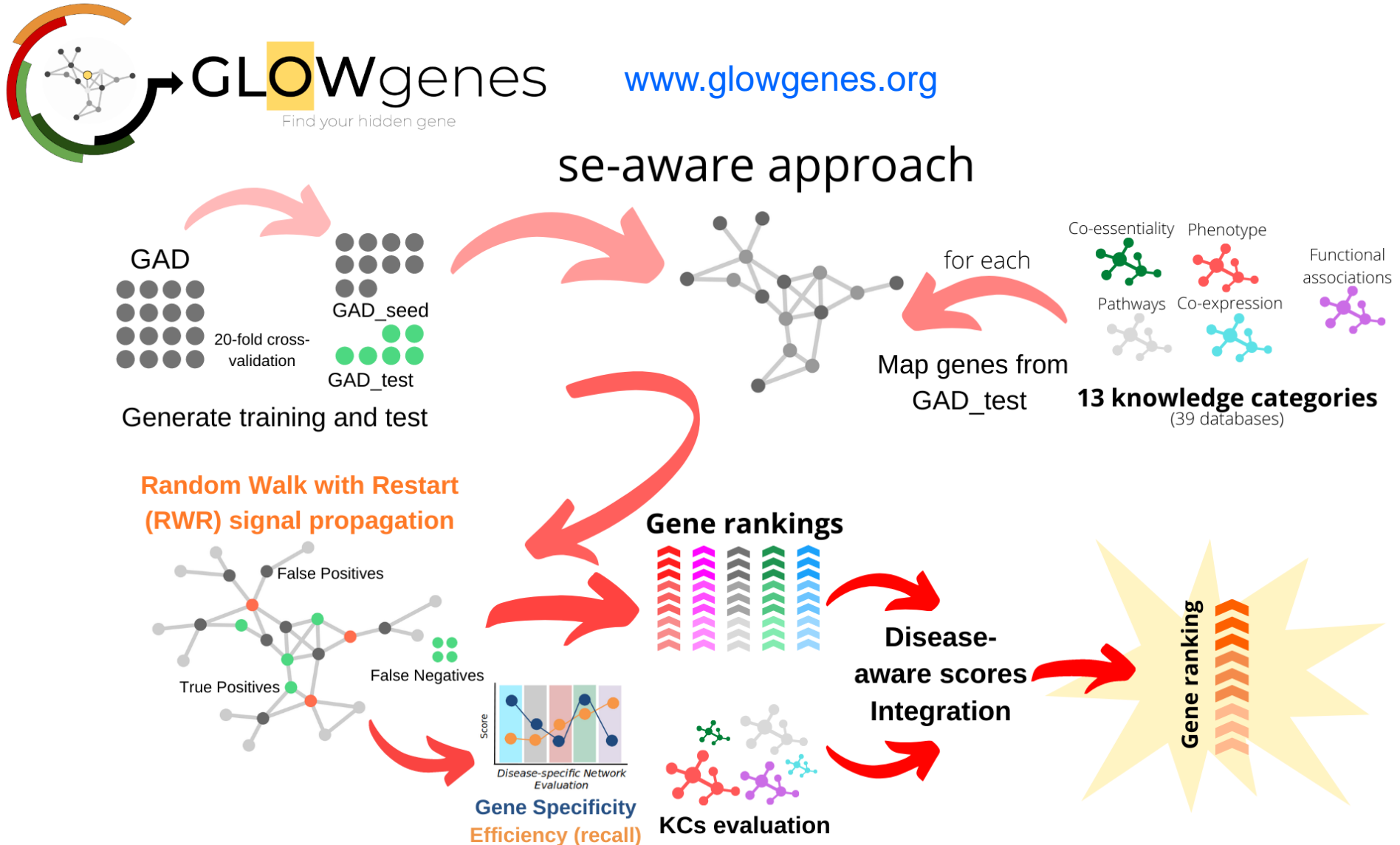
Collaborative Spanish Variant server <http://csvs.babelomics.org>



>700 individuos (cohorte Dpto. Genética HU-FJD) contribuyendo a la base de datos de frecuencias en población española.

Otros recursos en Genómica (y otras ómicas)

Predicador de nuevas asociaciones gen-enfermedad de una firma molecular a genes candidatos.



Análisis de experimentos ómicos - Proteómica

Proteómica

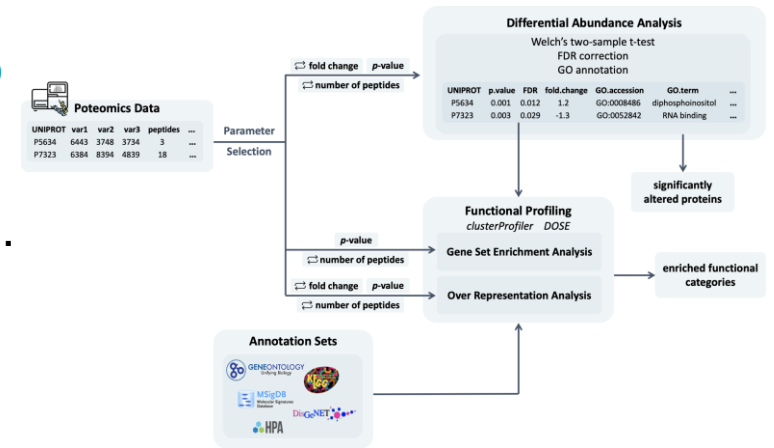
- Caracterización funcional (GSEA, Biología de redes – SNOW).
- **EnrichProt**, estadística y análisis funcional para elegir número de péptidos que identifican una proteína, junto al laboratorio de Inmunoalergia y Proteómica.



An integral tool for proteomics

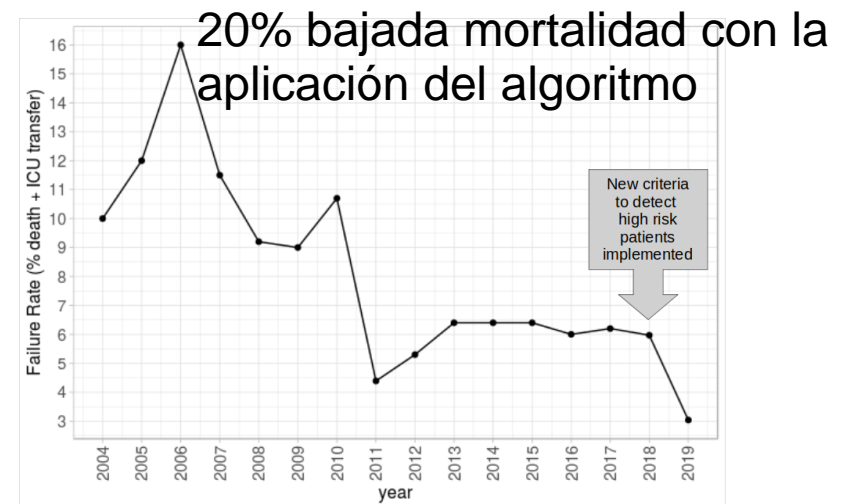
<https://github.com/fatimardlb/EnrichProt>

TFM Fátima Ruiz de la Bastida
Co-dirigido con Gloria Álvarez Llamas.



Análisis de datos clínicos mediante Aprendizaje Automático

- Aprendizaje automático (redes neuronales, random forest, otros).
- Métodos de balanceo de clases.
- Colaboración con UCIR/Neumología (Sarah Heili).
- Estudios en hematología y UCI.



Actividad en la comunidad bioinformática

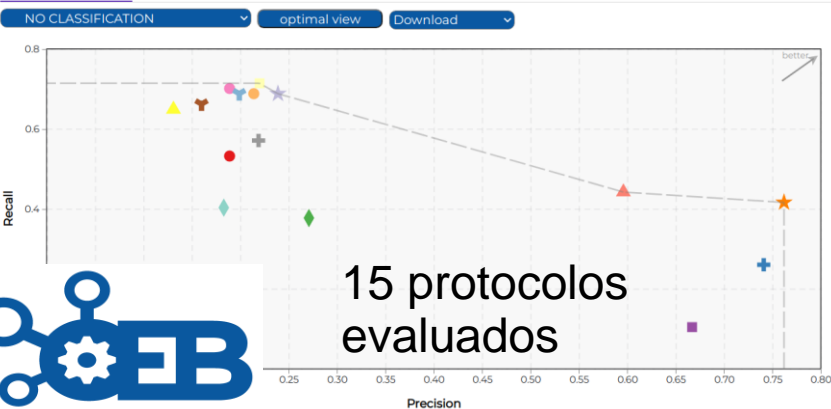


RED2018-102404-T

<https://inb-elixir.es/transbionet>

- **WG1. Standards and Benchmarking**
 - Good practices guides for genomics analysis (Genetic diagnosis).
- **WG2. Sustainability models.**
 - Learn from the experienced groups.
- **WG3. Training**
 - Cover training needs from scientists.

Coordinamos el grupo de trabajo de **evaluación de protocolos de detección de CNVs** (OpenEBench)



Actividad en la comunidad bioinformática



Definición de estándares para los protocolos bioinformáticos en el diagnóstico genético.

Grupos de trabajo

1. Análisis primario de DNaseq
2. Análisis de variantes en regiones codificantes.
3. Análisis de genoma
4. Splicing
5. Mosaicismo
- 6. Variantes estructurales (coordinación)**
- 7. Biología de sistemas (coordinación)**
8. Compartición de datos genómicos.
9. Materiales de referencia
10. Validación funcional de variantes
11. Tecnología long-reads
12. Traslación a la práctica clínica.
13. Flujos de trabajo

Guías clínicas, validación de herramientas, cursos, publicaciones



Proyectos IMPACT (Cohortes, **Genómica y Data**)

Genómica – Infraestructura para la Medicina Genómica – 2000 WGS

Aspectos éticos, selección de pacientes, secuenciación, análisis, uso secundario de los genómicos y compartición de datos.

- Coordinando el WP de **compartición de datos** de enfermedades raras
- WP que definirá el **análisis bioinformático**

Actividad en la comunidad bioinformática

IMPACT

DATA



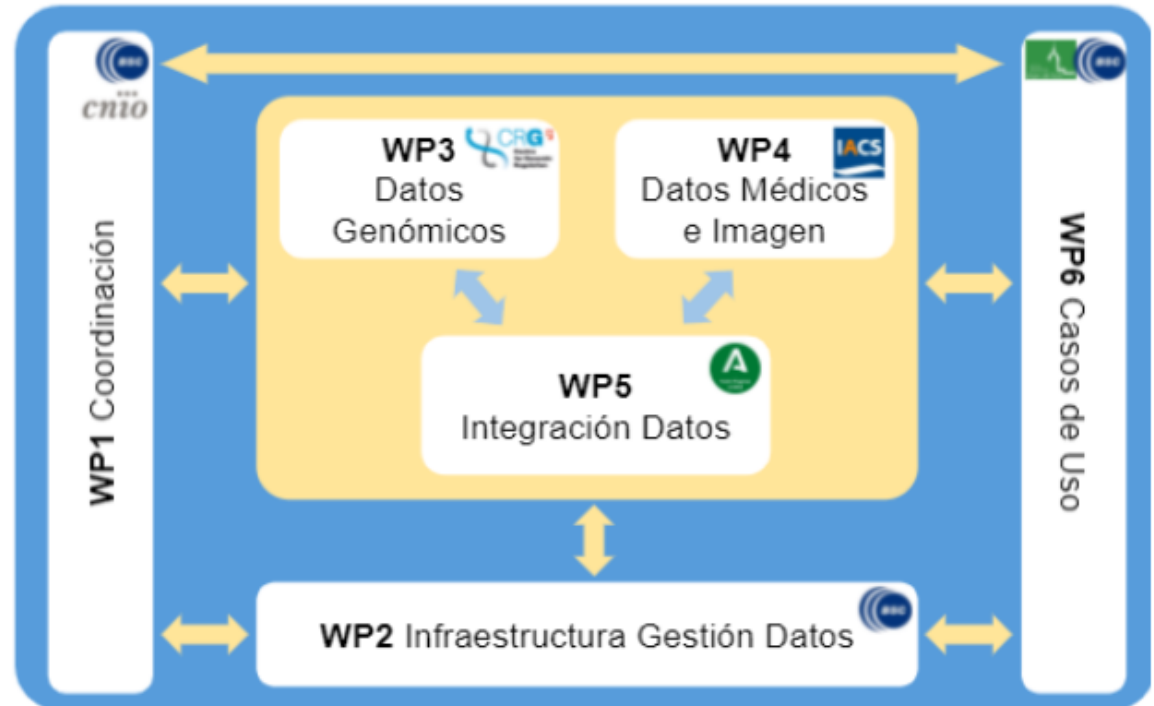
Infraestructura bioinformática y manejo de datos

Junto al Departamento de BigData

Nube de computación con servicios bioinformáticos



Formas de conectar información clínica, genómica y de imagen médica.



Registro de software en bio.tools (Elixir)
<https://bio.tools/>

1+MillionGenomes

Mirador español del WG5 Infraestructura compartición y manejo de datos.

Formación

- Clase en el Máster en Bioinformática aplicada a la Medicina Personalizada y Salud (ISCIII). Seminarios en otros dos Másteres en Bioinformática (UPM, UAM).
- Acogemos TFGs (6) y TFMs (8) incluyendo **compartidos** (Anatomía Patológica, Inmunoalergia y Proteómica, Oncología médica).
- Co-organizamos jornadas BioinfoCAM (IV edición en 2022).
- Curso “Introducción al análisis de datos ómicos 2022.
- Encuesta <https://bit.ly/cursoBioinfoIISFJD2022>

Introducción al análisis de datos ómicos - 2022

Desde la Unidad de Bioinformática del IIS-FJD se realizará un curso de formación en análisis bioinformáticos e interpretación de resultados de experimentos ómicos (genómica/transcriptómica/proteómica).

El curso está orientado a investigadores con o sin experiencia previa en el análisis e interpretación de datos de experimentos ómicos.

Con el fin de adaptar y adecuar los contenidos a las necesidades e intereses de los asistentes hemos creado la presente encuesta. Rellenen los campos que consideren

Incluirá



Compartición de datos en EGA/otros

Preparando **documento sobre compartición de datos** para asesoramiento.

Curso EGA (compartición datos genómicos)

<https://youtu.be/RI9wCs7JVKw>



INSTITUTO DE INVESTIGACIÓN SANITARIA FUNDACIÓN JIMÉNEZ DÍAZ

Gracias!



stay linked

@TranslationalB1

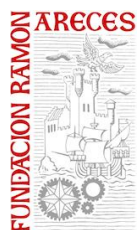
Contact: pablo.minguez@quironsalud.es

Encuesta curso: bit.ly/cursoBioinfoIISFJD2022



Departamento de Genética

Gracias a todos los colaboradores



PTMcode 2

