

Pipeline Escalable y Reproducible para el Análisis de Datos Genómicos y Construcción de Bases de Datos en Glioma Pediátrico

Gisela Pattarone¹, Hebe Durán^{1,2,3,4}, Marina Perona^{2,5}, Irene L. Ibañez^{1,2,3,*}

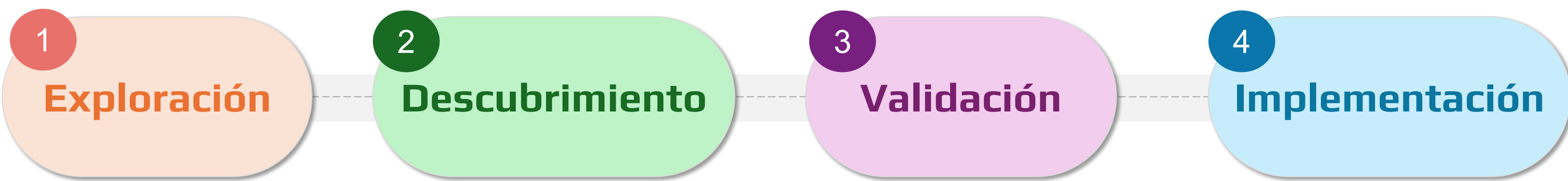
1 Comisión Nacional de Energía Atómica (CNEA), Gerencia de Investigación y Aplicaciones, Subgerencia de Tecnología y Aplicaciones de Aceleradores
2 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
3 Instituto de Nanociencia y Nanotecnología
4 Universidad Nacional de San Martín (UNSAM), Escuela de Ciencia y Tecnología
5 CNEA, Departamento de Radiobiología, División Bioquímica Nuclear

* Contacto: irenuliz@yahoo.com

Objetivos del Proyecto:

- Integración de Datos Genómicos: Desarrollo de un pipeline escalable para el procesamiento de RNA-seq, identificación de variantes y biomarcadores en gliomas pediátricos.
- Análisis de Blancos Terapéuticos y Tratamientos: Meta-análisis de terapias existentes e identificación de objetivos moleculares para respaldar la medicina de precisión.
- Desarrollo y Accesibilidad de Bases de Datos: Implementación de una plataforma de acceso abierto para datos genómicos, clínicos y terapéuticos, fomentando la colaboración en investigación y la toma de decisiones clínicas.

Metodología de Análisis:



Enfoque Bioinformático y Analítico

Análisis Bioinformático de Mutaciones Somáticas

Identificación de alteraciones genéticas utilizando **BioPySRA.py**, Análisis de RNA-seq para la alineación y cuantificación, el Módulo de Perfil Mutacional y el Cargador de Meta-Análisis de Terapias para la extracción de variantes y análisis de terapias.

Análisis de Redes y Vías de Señalización

Construcción de redes de interacción génica utilizando STRING-db y realización de enriquecimiento de vías utilizando herramientas como KEGG, Reactome y Gene Ontology.

Análisis estadístico e implementación

Correlaciones entre expresión génica e impacto de mutaciones realizadas con el método de Spearman para identificar biomarcadores clínicamente significativos en R e implementación en Shiny.

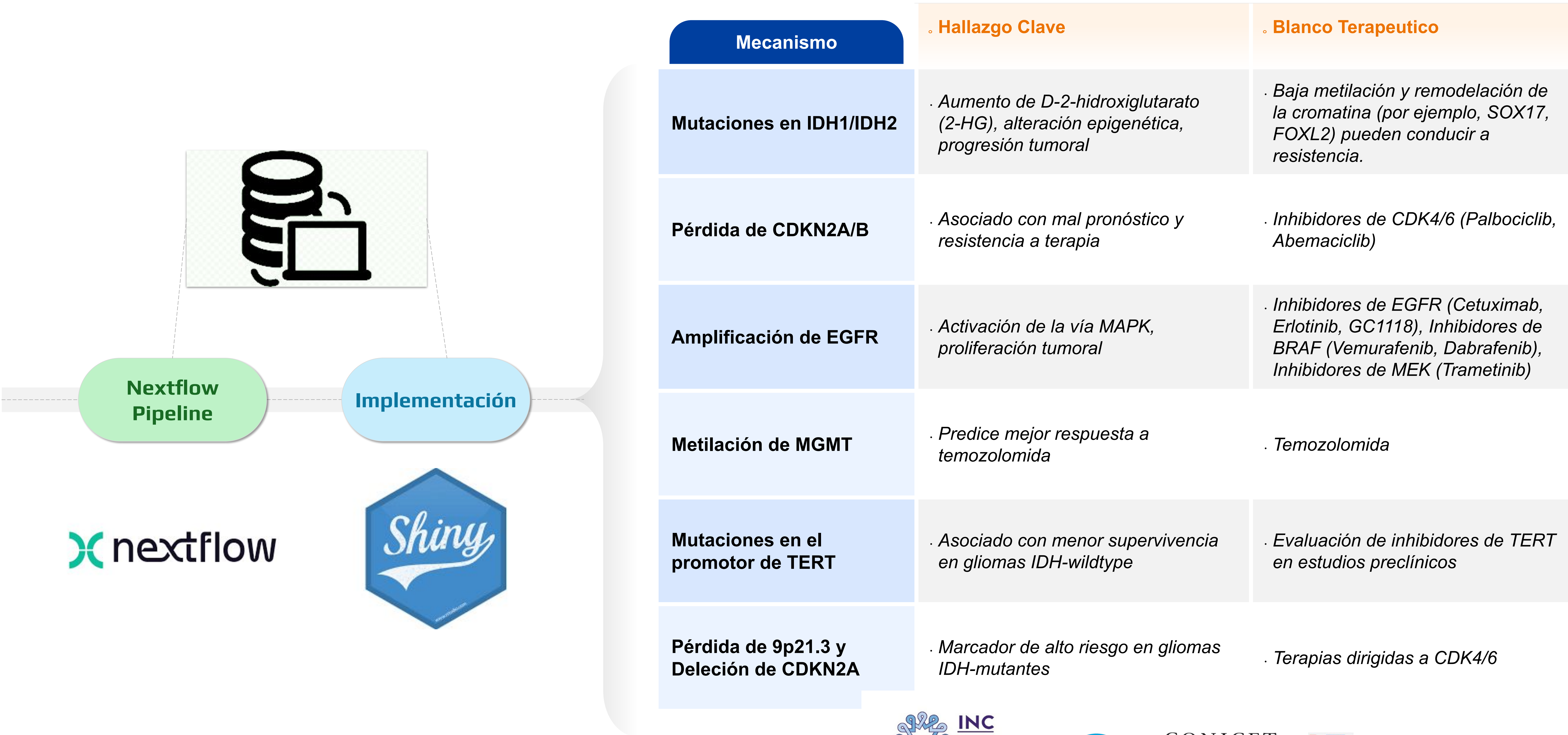
Resultados del Análisis

- Se recopilaron y curaron datos genómicos de glioma pediátrico a partir de SRA y GEO, identificando 6700 muestras de RNA-seq. Se priorizaron 46 muestras bien caracterizadas para garantizar coherencia metodológica y reproducibilidad en el análisis bioinformático.
- El análisis transcriptómico y de variantes en gliomas pediátricos identificó alteraciones clave en **IDH1/IDH2**, **CDKN2A/B**, **EGFR** y **TERT**, entre otros genes. Se observaron firmas moleculares asociadas con progresión tumoral y resistencia a tratamientos, destacando diferencias en perfiles epigenéticos y expresión génica.
- El meta-análisis de tratamientos en gliomas pediátricos permitió evaluar la eficacia de terapias dirigidas, destacando el uso de **inhibidores de IDH1/IDH2**, **CDK4/6**, **BRAF** y **EGFR**, así como la relevancia de biomarcadores como **MGMT** y **TERT** en la respuesta a quimioterapia.
- Se desarrolló una **plataforma escalable y reproducible en R Shiny** para la exploración en tiempo real de datos genómicos y clínicos en **gliomas pediátricos**. El sistema integra **PostgreSQL** y **MongoDB** para una gestión eficiente de datos y permite su implementación en **hospitales**, facilitando el acceso de **médicos** a información clave para la toma de decisiones en tratamientos personalizados.

Pipeline



Esquema de Plataforma de Acceso Público para futura Integración de Hallazgos



Agradecimientos:

Instituto Nacional del Cáncer “Beca de Investigación en Cáncer 2024/2025” (Médica Gisela Pattarone)
Fundación Florencio Fiorini “Beca Estímulo para Investigación en Medicina 2025” (Médica Gisela Pattarone)

