

# SEH<sup>2</sup>Bioinfo

Simposio de Estudiantes Hispanohablantes  
de Bioinformática y Biología Computacional

# LIBRO DE RESÚMENES

02-03 de Junio, 2022



# Índice

<b>Organizadores y Colaboradores</b>	2
<b>Programa</b>	6
<b>Trabajos presentados en SEH<sup>2</sup>Bioinfo 2022</b>	10
<b>Oradora principal: Dra. Fátima Al-Shahrour</b>	17
<b>Orador principal: Dr. Francisco Melo</b>	18
<b>Oradora principal: Dra. Patricia Melin</b>	19
<b>Resúmenes</b>	20

# Patrocinadores y Colaboradores

## Patrocinador Oro:



## Patrocinadores Plata:



## Patrocinadores Bronce:



## Colaboradores:



## **Comité Organizador**

El **I Simposio de Estudiantes Hispanohablantes de Bioinformática y Biología Computacional** está organizado por nueve de los RSGs (Grupos Regionales de Estudiantes) activos en países donde el idioma predominante es el Español (RSG Argentina, RSG Ecuador, RSG España, RSG Venezuela, RSG Colombia, RSG Chile, RSG Peru, RSG Costa Rica, RSG México).

### **Coordinación general**

Sara Monzón Fernández (ISCIH)  
Carla Padilla Franzotti (UNQ-CONICET)

### **Comité de búsqueda de Financiación**

Tamara Hernández (HUNSC)  
Christian Ignacio Olivera Fuentes (UTalca)  
Guillermo Orellana Becerra (PUC)  
Eva Tosco-Herrera (HUNSC)

### **Comité de programa y evaluación de resúmenes**

Agustín Ariel Baricalla (CITNOBA-UNNOBA-CONICET)  
Guillermo Jorge Gorines Cordero (URJC)  
María Florencia Martínez (IBR-CONICET)  
Gabriel J. Olguín-Orellana (CBSM-UTalca)  
Carla Perpiñá-Clérigues (CIPF-UV)  
Luis Alberto Rubio (ITER)  
Karen Sanchez (UFPel)  
Irene Soler (CIPF)  
Eva Suárez Pajés (UI-HUNSC)

### **Comité de Secretaría y Gestión de Asistentes**

Javier Lanillos Manchón (CNIO)  
Zulema Rodríguez Hernández (ISCIH)

## **Comité de Diseño**

Mercedes Didier Garnham (IIB-UNSAM-CONICET)  
Dayana Ticona (GII, UNSA)

## **Comité de actividades de Networking**

Adrián García Moreno (GENYO)  
Maria Clara Miserendino (UNC)  
Ivan Ellson Lancho (GENYO)  
María Soledad Benítez Cantos (UGR)

## **Comité de difusión y Redes sociales**

Monica Cabrera Pasadas (BSC)  
Claudia Silva (UMayor)  
Jennifer Vélez Segura (UNAL)  
Maria Delgado (UCSM)  
José Antonio Ovando Ricárdez (UJAT)

## **Comité de web**

Sebastián Ayala Ruano (USFQ)  
Cleidy Osorio (USP)  
Rafael Puche (IVIC)  
Luis Angel Chicoma Rojas (UNESP)

## **Comité de libro de abstracts**

Guillermo Jorge Gorines Cordero (URJC)  
Tamara Hernández (HUNSC)  
Javier Lanillos Manchón (CNIO)  
Sara Monzón Fernández (ISCIII)  
Zulema Rodríguez Hernández (ISCIII)  
Carla Padilla Franzotti (UNQ-CONICET)

## Glosario de instituciones

**A2B2C:** Asociación Argentina de Bioinformática y Biología Computacional, Argentina

**BSC:** Barcelona Supercomputing Center, España

**CBSM-UTALCA:** Centro de Bioinformática y Simulación Molecular, Universidad de Talca, Chile

**CIPF:** Centro de Investigación Príncipe Felipe, España

**CITNOBA-UNNOBA:** Centro de Investigaciones y Transferencia del Noroeste de la Provincia de Buenos Aires, Argentina

**CNIO:** Centro Nacional de Investigaciones Oncológicas, España

**CONICET:** Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

**GENYO:** Pfizer-University of Granada-Junta de Andalucía Centre for Genomics and Oncological Research, España

**HUNSC:** Hospital Universitario Ntra. Sra. de Candelaria, España

**IBR-CONICET:** Instituto de Biología Molecular y Celular de Rosario - Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

**IIB-UNSAM:** Instituto de Investigaciones Biotecnológicas - Universidad Nacional de San Martín, Argentina

**INB:** Instituto Nacional de Bioinformática, España

**ISCIH:** Instituto de Salud Carlos III, España

**ISCB-SC:** International Society for Computational Biology - Student Council

**IVIC:** Instituto Venezolano de Investigaciones Científicas, Venezuela

**PUC:** Pontificia Universidad Católica de Valparaíso, Chile  
Tachyon Perú

**SCB:** Sociedad Chilena de Bioinformática, Chile

**UCSM:** Universidad Católica de Santa María, Perú

**UMayor:** Universidad Mayor, Chile

**UNAL:** Universidad Nacional de Colombia, Colombia

**UNESP:** Universidade Estadual Paulista, Brasil

**UJAT:** Universidad Juárez Autónoma de Tabasco, México

**UNC:** Universidad Nacional de Córdoba, Argentina

**UFPel:** Universidade Federal de Pelotas, Brasil

**UNQ:** Universidad Nacional de Quilmes, Argentina

**UPCH:** Universidad Peruana Cayetano Heredia, Perú

**URJC:** Universidad Rey Juan Carlos, España

**USFQ:** Universidad San Francisco de Quito, Ecuador

**USP:** Universidade de São Paulo, Brasil

**UTalca:** Universidad de Talca, Chile

## Programa

### 2 de junio, “Jornada de Networking”

**08:00 (México GMT-6) / 09:00 (Ecuador/Perú/Colombia GMT-5) / 10:00 (Venezuela GMT-4) / 11:00 (Argentina/Chile GMT-3) / 16:00 (Madrid GMT+2):**

**Título:** Palabras de bienvenida y presentación de la motivación del simposio SEH<sup>2</sup>Bioinfo.

**Duración:** 10 mins.

**Moderadoras:** Carla Padilla Franzotti y Sara Monzón

**08:10 (GMT-6) / 09:10 (GMT-5) / 10:10 (GMT-4) / 11:10 (GMT-3) / 16:10 (GMT+2):**

**Título:** Video presentación de los RSGs Hispanohablantes.

**Duración:** 35 mins.

**Moderador:** Adrian Garcia Moreno

**08:45 (GMT-6) / 09:45 (GMT-5) / 10:45 (GMT-4) / 11:45 (GMT-3) / 16:45 (GMT+2):**

**Título:** Mesa redonda con Líderes de los RSG: Entendiendo los diferentes retos en la investigación en bioinformática en cada región.

**Duración:** 1 h

**Moderadores:** Adrian Garcia Moreno e Ivan Ellson Lancho - España

**Participantes:** Sebastián Ayala Ruano - Ecuador

Claudia Silva Andrade - Chile

Carla Padilla Franzotti - Argentina

Jennifer Vélez Segura - Colombia

Cleidy Mirela Osorio Mogollón - Perú

Rafael Puche - Venezuela

Gabriel Olguín-Orellana - ET Student Council ISCB

Gonzalo Parra - ET Student Council ISCB

**09:45 (GMT-6) / 10:45 (GMT-5) / 11:45 (GMT-4) / 12:45 (GMT-3) / 17:45 (GMT+2) - Descanso**

**10:00 (GMT-6) / 11:00 (GMT-5) / 12:00 (GMT-4) / 13:00 (GMT-3) / 18:00 (GMT+2):**

**Título:** “CABANA.network: una plataforma para fomentar el desarrollo de la bioinformática en LAC”.

**Duración:** 10 mins. + 5 mins. preguntas

**Moderadora:** Carla Padilla Franzotti

**10:15 (GMT-6) / 11:15 (GMT-5) / 12:15 (GMT-4) / 13:15 (GMT-3) / 18:15 (GMT+2):**

**Título:** Presentación del programa de Mentoreo

**Duración:** 10 mins. + 5 mins. preguntas

**Moderadora:** Clara Miserendino

**10:30 (GMT-6) / 11:30 (GMT-5) / 12:30 (GMT-4) / 13:30 (GMT-3) / 18:30 (GMT+2):**

**Título:** Encuesta en vivo: “Conociendo las costumbres bioinformáticas de nuestra comunidad”

**Duración:** 30 mins.

**Moderadores:** Clara Miserendino, Marisol Benitez Cantos e Ivan Ellson Lancho

**11:00 (GMT-6) / 12:00 (GMT-5) / 13:00 (GMT-4) / 14:00 (GMT-3) / 19:00 (GMT+2):**

**Título:** Palabras finales del primer día y apertura de la Sesión virtual de Pósters y actividades de Networking.

**Duración:** 5 mins.

**Moderadoras:** Sara Monzón y Carla Padilla Franzotti

**11:05 (GMT-6) / 12:05 (GMT-5) / 13:05 (GMT-4) / 14:05 (GMT-3) / 19:05 (GMT+2):**

**Título:** Sesión Virtual de Pósters y Actividades de Networking. - Transición y familiarización con la Plataforma Gather Town.

**Duración:** 2 hs

- **Sesión de Pósters en vivo durante las 2 hs**

**Lugar:** SALA 1, 2 y 3 Gather Town

- **Cronograma de actividades de Networking:**

**11:15 (GMT-6) / 12:15 (GMT-5) / 13:15 (GMT-4) / 14:15 (GMT-3) / 19:15 (GMT+2):**

**Título:** Escape Room Virtual "Encuentra al Hacker"

**Duración:** 45 mins.

**Lugar:** SALA 7 Gather Town - Si no te inscribiste ¡Pasá por la Sesión de pósters!

**Moderadora:** Mónica Cabrera

**12:00 (GMT-6) / 13:00 (GMT-5) / 14:00 (GMT-4) / 15:00 (GMT-3) / 20:00 (GMT+2):**

**Título:** Kahoot sobre Cultura General de los RSGs en el SEH<sup>2</sup>

**Duración:** 20 mins.

**Lugar:** Auditorio Gather Town

**Moderadores:** Marisol Benitez Cantos e Ivan Ellson Lancho

**12:20 (GMT-6) / 13:20 (GMT-5) / 14:20 (GMT-4) / 15:20 (GMT-3) / 20:20 (GMT+2):**

**Título:** ¡Seguir jugando!. Mesas organizadas para: Pinturillo, Gartic Phone y otros.

**Lugar:** Actividad Libre - ¡Pasa por la Sesión de pósters!

**Duración:** 40 mins.

### 3 de junio, "Jornada Científica"

**08:00 (México GMT-6) / 09:00 (Ecuador/Perú/Colombia GMT-5) / 10:00 (Venezuela GMT-4) / 11:00 (Argentina/Chile GMT-3) / 16:00 (Madrid GMT+2)**

**Título:** Metodologías bioinformáticas para oncología de precisión: logros y retos

**Duración:** 45 mins. + 15 mins. preguntas

**Oradora principal:** Dra. Fátima Al-Shahrour - CNIO, España

**Moderadora:** Mónica Cabrera



**09:00 (GMT-6) / 10:00 (GMT-5) / 11:00 (GMT-4) / 12:00 (GMT-3) / 17:00 (GMT+2):**

**Título:** Sesión de charlas relámpago:

- **Yanel Elina Bernardi** – “Descifrando el interactoma de miARN-ARNm necesario para las transiciones epitelio-mesénquima-epitelio (EMT/MET) de las células de la cresta neural embrionaria./ Deciphering the miRNA-miRNA interactome required for epithelial-mesenchymal-epithelial transitions (EMT/MET) of embryonic neural crest cells.”
- **Arturo González Vilanova** – “Evaluación comparativa de los métodos de cuantificación de la expresión de isoformas para los datos de RNA-seq./ Benchmarking isoform expression quantification methods for RNA-seq data.”
- **Jochem Nelen** – “ Optimización del consenso de acoplamiento mediante la incorporación de la similitud de la pose de unión prevista: un estudio de caso de Wee1. /Optimizing consensus docking by incorporating predicted binding pose similarity: a Wee1 case study.”
- **Mauro Gabriel Rey** – “Resolución de un transcriptoma de referencia en *Setaria italica* para análisis de splicing alternativo en condiciones de estrés hídrico./ Resolution of a reference transcriptome in *Setaria italica* for alternative splicing analysis under water stress conditions.”
- **Cristhian N. Rodríguez Silva** - “Diseño *in silico* de inhibidores de ptpa y ptpb como agentes terapéuticos frente a *Mycobacterium tuberculosis*. / *In silico* design of ptpa and ptpb inhibitors as therapeutic agents against *Mycobacterium tuberculosis*. “
- 

**Duración/charla:** 2 mins. + preguntas

**Moderador:** Sebastián Ayala Ruano

**09:30 (GMT-6) / 10:30 (GMT-5) / 11:30 (GMT-4) / 12:30 (GMT-3) / 17:30 (GMT+2) - Descanso**

**09:45 (GMT-6) / 10:45 (GMT-5) / 11:45 (GMT-4) / 12:45 (GMT-3) / 17:45 (GMT+2):**

**Título:** Sesión de comunicaciones orales cortas:

- **Martin de Jesus Loza López** – “Integración no sesgada de réplicas de transcriptomas de células individuales. /Unbiased integration of single cell transcriptomes replicates.”
- **Sara Herráiz Gil** – “Efectos de la inflamación en la alteración del metabolismo de los lípidos en la epidermólisis bullosa distrófica recesiva severa: lipidómica y transcriptómica. / Effects of inflammation on altered lipid metabolism in severe recessive dystrophic dystrophic epidermolysis bullosa: lipidomics and transcriptomics.”
- **Tadeo Enrique Saldaño/Nahuel Escobedo** – “Impacto de la diversidad conformacional de las proteínas en las predicciones de *alphafold*. / Impact of protein conformational diversity on alphafold predictions.”
- **Julieth Andrea López Castiblanco** – “Utilidad del análisis integrativo de microARN para diagnóstico. /Utility of integrative microRNA analysis for diagnostic purposes.”

**Duración/charla:** 10 mins. + 5 mins. Preguntas

**Moderadora:** Cleidy Osorio

**10:45 (GMT-6) / 11:45 (GMT-5) / 12:45 (GMT-4) / 13:45 (GMT-3) / 18:45 (GMT+2)**

**Título:** Estudios de la especificidad del factor de transcripción MarA mediante experimentos de unión in vitro, evolución dirigida in vivo, secuenciación profunda y análisis bioinformáticos

**Duración:** 45 mins. + 15 mins. preguntas

**Orador principal:** Dr. Francisco Melo - Facultad de Ciencias Biológicas, UC, Chile

**Moderadora:** Claudia Silva

**11:45 (GMT-6) / 12:45 (GMT-5) / 13:45 (GMT-4) / 14:45 (GMT-3) / 19:45 (GMT+2) - Descanso**

**12:00 (GMT-6)/ 13:00 (GMT-5) / 14:00 (GMT-4) / 15:00 (GMT-3) / 20:00 (GMT+2):**

**Título:** Sesión de comunicaciones orales cortas:

- **María Victoria Borja Lozano** – “Identificación de polimorfismos de nucleótido simple en cuyes (*Cavia porcellus L*) de las zonas altoandinas del Perú a través de genotipado basado en secuenciamiento (GBS). / Identification of single nucleotide polymorphisms in guinea pigs (*Cavia porcellus L*) from the high Andean zones of Peru through genotyping based on sequencing (GBS).”
- **Matías Carletti** – “Identificación y caracterización de proteínas humanas con repeticiones de unión a ARN. / Identification and Characterization of Human RNA Binding Repeat Proteins.”
- **Karla Verónica Nolívos Rodríguez** – “Comparación de ensambladores en la recuperación de vMAGs de bacteriófagos/ Comparison of assemblers in the recovery of vMAGs from bacteriophages.”
- **Giulliano Nájera Bless/Jaime Tamayo Ramos** – “Diseño in silico de derivados de quercetina con posible actividad dual inhibitoria contra GSK3 $\beta$  y CDK5/P25 para el tratamiento de la enfermedad de Alzheimer. *In silico* design of quercetin derivatives with potential dual inhibitory activity against GSK3 $\beta$  and CDK5/P25 for the treatment of Alzheimer's disease.”
- 

**Duración/charla:** 10 mins. + 5 mins. preguntas

**Moderadora:** María Florencia Martínez

**13:00 (GMT-6)/ 14:00 (GMT-5) / 15:00 (GMT-4) / 16:00 (GMT-3) / 21:00 (GMT+2):**

**Título:** “Modelos Híbridos Inteligentes basados en Redes Neuronales, Lógica Difusa y Algoritmos de Optimización Bioinspirados para aplicaciones en ciencia e ingeniería”

**Duración:** 45 min + 15 min preguntas

**Oradora principal:** Dra. Patricia Melin - Instituto Tecnológico de Tijuana, México

**Moderador:** Rafael Puche

**14:00 (GMT-6)/ 15:00 (GMT-5) / 16:00 (GMT-4) / 17:00 (GMT-3) / 22:00 (GMT+2):**

**Título:** Premios a la mejor comunicación oral y póster. Comunicación y clausura.

**Moderadoras:** Carla Padilla Franzotti y Sara Monzón

## Trabajos presentados en SEH<sup>2</sup>Bioinfo 2022

GatherTown, Sala 1:

N°		Autor Expositor	Título del trabajo
1	P	Lucas Inchausti	Análisis funcionales de perfiles de co-expresión génica de <i>Trypanosoma cruzi</i>
2	P	Ary Lautaro Di Bartolo	Dinámica molecular en sistemas biológicos: implementando variable colectiva en software open-source.
3	P	Silvia Talavera Marcos	Acercamientos computacionales basados en señal filogenética explican las leyes de ensamblaje subyacentes en datos composicionales de comunidades microbianas cultivadas <i>ex situ</i> .
4	P	Lilibeth Torres-Elizalde	Detección <i>in silico</i> de integrones de resistencia en <i>Salmonella enterica</i> en la región andina usando dos herramientas bioinformáticas.
5	P, CR	Yanel Elina Bernardi	Descifrando el interactoma de miARN-ARNm necesario para las transiciones epitelio-mesénquima-epitelio (EMT/MET) de las células de la cresta neural embrionaria.
6	P, CR	Arturo González Vilanova	Comparación de métodos de cuantificación de la expresión de isoformas para datos de RNA-seq.
7	P	Yessica Galicia Pérez	Wikiproyecto bases de datos de biociencias.
8	P	José Francisco Català Senent	Estudio de las bases moleculares afectadas por el sexo en Esclerosis Múltiple a través del transcriptoma de tejido nervioso.
9	P	Cristina Galiana Roselló	Reposicionamiento <i>in silico</i> de Fármacos en la Enfermedad de Alzheimer con Perspectiva de Sexo.
10	P	Sara Herraiz Gil	Efectos de la inflamación en la alteración del metabolismo de los lípidos en la epidermolísis bullosa distrófica recesiva severa: lipidómica y transcriptómica.
11	P	Miguel Carmena-Bargueño	TOLEDO: un nuevo software que puede ejecutar y analizar simulaciones de dinámica molecular de Desmond en clústeres informáticos de alto rendimiento.

12	P	Alejandro Rodríguez-Martínez	ASGARD: Revisión de una herramienta de análisis de dinámica molecular GROMACS simple y automática.
13	P	Jochem Nelen	Optimización del acoplamiento de consenso al incorporar la similitud de pose de unión predicha: un estudio de caso de Weel.
14	P	Sandra Pamela Cangui Panchi	Análisis de variabilidad de genomas del SARS-CoV-2
16	P	Irene Soler Sáez	Atlas funcional de las diferencias de sexo en esclerosis múltiple mediante el análisis de datos transcriptómicos de célula única
17	P	Carlos Martínez Cortés	MetaScreener: un paquete de software enfocado en flujos de trabajo de detección virtual en infraestructuras informáticas de alto rendimiento.
18	P	Marco Orlando Fuel Herrera	Análisis de los subtipos y cadenas de transmisión predominantes en la epidemia de VIH-1 en España durante los años 2011-2020
19	P	Ester Muñoz del Campo	Identificación de nuevos biomarcadores microRNAs en esclerosis múltiple con un perfil diferencial por sexo
20	P	Jorge Martín Arana	Priorización de variantes en la secuenciación de exoma completo de plasma para la caracterización del tumor en la recaída e identificación de nuevas dianas terapéuticas en pacientes con cáncer de colon.
21	P	Helena Gómez Martínez	Nuevas estrategias bioinformáticas para la identificación de biomarcadores en melanoma
22	P	César Eusebio Daniel Rivera Morales	Un algoritmo multimodal mejorado basado en un esquema NSGA-II para la inferencia de árboles filogenéticos
23	P	Sindy Paola Buitrago Puentes	Filodinámica y evolución molecular de SARS-Cov-2 en Colombia
24	P	Andrés Patricio Ibacache Chía	Identificación in silico de moléculas inhibidoras de transportadores de monocarboxilatos y su validación como blancos para la inhibición de la proliferación de células tumorales
25	P	Heliairis Alfonso González	Subclases de IgG en monos del nuevo mundo, ¿un tema de debate?

26	P	Haydé Saracho	Análisis de genomas de plásmidos actinobacterianos a través de redes bipartitas.
28	P	José Alberto Santiago de la Cruz	Diseño computacional de un tricomonida
30	P	Mario Fernando Bustillo López	Integración de realidad aumentada como recurso educativo en la bioinformática.

GatherTown, Sala 2:

N°		Autor Expositor	Título del trabajo
32	P	Patricia Genius	Explorando características cerebrales cuantitativas asociadas a una mayor predisposición genética a la enfermedad de Alzheimer mediante el análisis de datos composicionales
33	P	Rubén Grillo Risco	Firmas genéticas y funcionales de la lesión medular: una revisión sistemática y meta-análisis de estudios transcriptómicos en rata
34	P	Unai Illarregi Insausti	ARN largos no codificantes como predictores de resultados en la leucemia linfoblástica aguda pediátrica de células T
35	P	Eva Maicas Alcaine	Métodos de imputación de la variable sexo a partir de los niveles de expresión en genes.
36	P	Priscila Pagnotta	Uso de bases de datos en la complementación del estudio de las porfirias y la medicina personalizada.
37	P	Fernanda Salgado Farias	Uso de dispositivos en un modelo bioinformático para la atracción de <i>Macroductylus mexicanus</i> (Coleoptera: Melolonthidae).
38	P	Jesica Britez	Predicción y selección de candidatos vacunales contra la coccidiosis aviar mediante herramientas bioinformáticas.
39	CO	Tadeo Enrique Saldaño	Impacto de la diversidad conformacional de proteínas en las predicciones de AlphaFold.

40	P	Noe Yamil Pages Paez	Representación por Docking molecular del efecto de volátiles de hojas de <i>Acaciella angustissima</i> (Fabaceae) en adultos de <i>Macrodactylus mexicanus</i> (Coleoptera: Melolonthidae).
41	P	Karina Elizabeth Ruiz Segura	Variantes de un solo nucleótido asociados a neuropatía diabética en población latinoamericana.
42	P	Jessica Alejandra Zapata García	Análisis de CGH y del transcriptoma revela genes relacionados con supervivencia en leucemia linfoblástica aguda
43	P, CR	Mauro Gabriel Rey	Resolución de un transcriptoma de referencia en <i>Setaria italica</i> para análisis de splicing alternativo en condiciones de estrés hídrico
45	P, CR	Cristhian N. Rodríguez Silva	Diseño in silico de inhibidores de PtpA y PtpB como agentes terapéuticos frente a <i>Mycobacterium tuberculosis</i>
46	P	Vladimir Alfaro	Clasificador de imágenes con patrones de Anticuerpos Antinucleares (ANA)
48	P, CO	María Victoria Borja Lozano	Identificación de polimorfismos de nucleótido simple en cuyes ( <i>Cavia porcellus</i> L) de las zonas altoandinas del Perú a través de genotipado basado en secuenciamiento (GBS)
49	P	Melvin Nadir Rosalez	Comparison of the affinity of acetaminophen and a novel boron analogue on the TRPV channels
50	P	Belén Cuevas López	Evaluación in silico de análogos de poliaminas con potencial inhibidor del transportador TcPOT1 de alta eficiencia de <i>Trypanosoma cruzi</i>
52	CO	Julieth Andrea López Castiblanco	Utilidad del análisis integrativo de microARN para diagnóstico.
53	P	Borja Gómez Cabañes	Revisión de métodos para el resumen y simplificación de los resultados de análisis de enriquecimiento funcional
54	P	Jaime Llera Oyola	Estudio de las diferencias de sexo en la enfermedad de Alzheimer: un meta-análisis de estudios de microRNAs
55	P	Fernando Gordillo González	Caracterización de las diferencias de sexo en Parkinson mediante el análisis masivo de datos de expresión de célula única y su integración.

56	CO	Karla Verónica Nolivos Rodríguez	Comparación de ensambladores en la recuperación de vMAGs de bacteriófagos
57	P	Javier Lanillos Manchón	Interrogando la expresión de PARP1 y la genética de PBRM1 como biomarcadores de respuesta a inmunoterapia en pacientes de cáncer renal
58	P	Anthony Vladimir Campos Segura	Análisis bioinformático para explorar el impacto de las mutaciones somáticas en los genes SAMHD1, GNL1, POLE, MRE11A y ASXL2 como potenciales mini impulsores en Cáncer Colorrectal.
59	P	Álvaro de Jesús Huamani Ortiz	El análisis basado en transcriptomas identifica ASXL1 y MAP4K4 como marcadores de pronóstico en el cáncer gástrico con inestabilidad de microsatélites.
60	P	Carla Perpiñá-Clérigues	Diferencias de sexo en el lipidoma de las vesículas extracelulares tras el consumo masivo de alcohol en la adolescencia.

GatherTown, Sala 3:

Nº		Autor Expositor	Título del trabajo
62	CO	Matías Carletti	Identificación y caracterización de proteínas humanas con repeticiones de unión a ARN.
63	P	María Cecilia Merino	Modelamiento molecular y análisis bioinformático de polimorfismos de la proteasa NS3 del virus Dengue presente en Argentina, Brasil y Paraguay.
64	P	Alicia Lizet Arica Sosa	Búsqueda de inhibidores de PknG de Mycobacterium tuberculosis mediante tamizaje virtual masivo basado en farmacóforo.
65	P	Dayana Rosa Frutos Bottega	Identificación del número de Marcos Abiertos de Lectura en los cromosomas del complejo de especies <i>Cryptococcus neoformans</i> y <i>Cryptococcus gattii</i>
66	P	Montserrat Loredo Guillén	Análisis de marcadores genéticos relacionados con aterosclerosis

68	P	Uriel Sebastian Capcha Jimenez	Influencia de la etnicidad latina y no latina en la asociación entre niveles de mutaciones y expresiones genéticas en pacientes con cáncer de mama.
69	P	Edwin L Bonilla R	Análisis in silico de la selectividad de los inhibidores comerciales de la 5 $\alpha$ -Reductasa.
71	P	Adolfo López Cerdán	Diferencias de sexo en la Enfermedad de Parkinson: metaanálisis de estudios transcriptómicos de tejido cerebral.
72	P	Erika Meyer Kvaalem Soto	Plataforma para la vigilancia genómica de SARS-CoV-2 en España.
73	P	Carolina Monzó	¿Cómo responde el repertorio de receptores de células B a la restricción dietética en ratones durante el envejecimiento?
74	P	Felipe Alonso Maurelia Gaete	Identificación y caracterización de interactores proteicos del extremo c-terminal de osc-espondina durante el desarrollo embrionario
75	P	Maryori Fernanda González Olivero	Análisis y caracterización de la señalización de FGF en el OSC de <i>Gallus gallus</i> en estadio embrionario temprano
76	P	Luis Gonzalo Espinoza Arcos	Identificación del sitio de unión de PIP2 en canales TRP: descifrando los efectos estructurales de la modulación de la familia TRP por fosfolípidos.
77	P	Maria Emilia Baez	Análisis bioinformático de la variabilidad genética de la polimerasa NS5B del virus de la Hepatitis C
79	P	Agustina E. Nardo	Análisis de datos de péptidos bioactivos de proteínas alimentarias para el diseño y construcción de una nueva base de datos
80	P	Juan Mac Donagh	Las proteínas humanas de unión a ARN con repeticiones en tándem evolucionan más lentamente que lo esperado
81	P	Viviana Coliboro Dannich	Análisis bioinformático de la correlación de expresión de NUA1 y G6PD en cáncer de mama, de colon y glioblastoma
82	CO	Martin Loza	Integración imparcial de réplicas de transcriptomas de células individuales.



83	P	Carla Luciana Padilla Franzotti	Aspectos moleculares del mecanismo de unión del antígeno T grande de SV40 a retinoblastoma
84	P	Arley Ricardo Gutarra Inga	Identificación de ligandos bloqueantes de L-selectina durante el proceso de implantación embrionaria como posible anticonceptivo
86	P, CO	Maria Fernanda Hashimoto Vargas y Andrea Alexa Sánchez Sánchez	Búsqueda de potenciales inhibidores del ensamblaje de la cápside del virus de la hepatitis B mediante el tamizaje virtual masivo basado en farmacóforo
87	P	Giulliano Nájera Bless y Jaime Tamayo Ramos	Diseño <i>in silico</i> de derivados de quercetina con posible actividad dual inhibitoria contra GSK3 $\beta$ y CDK5/P25 para el tratamiento de la enfermedad de Alzheimer.
88	P	J. Leonardo Gómez Chávez y Emilio L. Angelina	Estudio del efecto antiinflamatorio de los componentes de <i>C. citratus</i> mediante el Cribado Virtual sobre la red biológica de la enfermedad de Chagas
89	P	Rafael Betanzos San Juan	Hallazgo y validación de genes ortólogos distantes del gen DAF-12 en <i>Meloidogyne incognita</i> mediante el uso de herramientas bioinformáticas.

CO: Comunicación Oral

CR: Charla Relámpago

P: Trabajo con Póster

## Oradora principal: Dra. Fátima Al-Shahrour



Centro Nacional de Investigaciones Oncológicas (CNIO), España

La Dra. Fátima Al-Shahrour obtuvo su doctorado por la Universidad Autónoma de Madrid en 2006, trabajando en la unidad de Bioinformática del Centro Nacional de Investigaciones Oncológicas (CNIO, Madrid) y el Centro de Investigaciones Príncipe Felipe (CIPF, Valencia). Su labor se centró en el desarrollo de herramientas bioinformáticas y métodos computacionales para la

interpretación funcional de datos ómicos. En 2007, comienza una trayectoria internacional en el Broad Institute del Massachusetts Institute of Technology (MIT) y en Harvard para identificar firmas moleculares de genes “drivers” en cáncer a través de metodologías computacionales y datos de expresión genética. Tras cuatro años como bióloga computacional en el MIT, Harvard y el Brigham Women’s Hospital estudiando la biología del cáncer y terapias oncológicas, regresa al CNIO para dirigir la Unidad de Bioinformática Traslacional.

Desde 2017 hasta hoy, dirige la Unidad de Bioinformática del CNIO, que enfoca su investigación en el desarrollo de herramientas bioinformáticas que permitan la integración de datos biológicos y clínicos así como la identificación de biomarcadores de respuesta a terapias oncológicas.

## Orador principal: Dr. Francisco Melo



Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile

El Dr. Francisco Melo, bioquímico de profesión, cuenta con más de 25 años de trayectoria internacional en el campo de la bioinformática, habiendo obtenido su doctorado por la Universidad Notre Dame de la Paix, Bélgica y realizando un postdoctorado en The Rockefeller University, Estados Unidos. En la actualidad, es profesor titular en el Departamento de Genética Molecular y

Microbiología de la Universidad Pontificia de Chile, donde ha desarrollado su labor docente e investigadora desde el año 2000.

Sus investigaciones en el área de la biología computacional estructural se enfocan en las interacciones secuencia-estructura-función de las macromoléculas encargadas de dirigir el flujo de información genética en los seres vivos (las proteínas, el ADN y el ARN). Para conducir sus investigaciones, combina conocimientos en física, química y estadística a metodologías computacionales, que le permiten procesar grandes cantidades de datos experimentales.

## Oradora principal: Dra. Patricia Melin



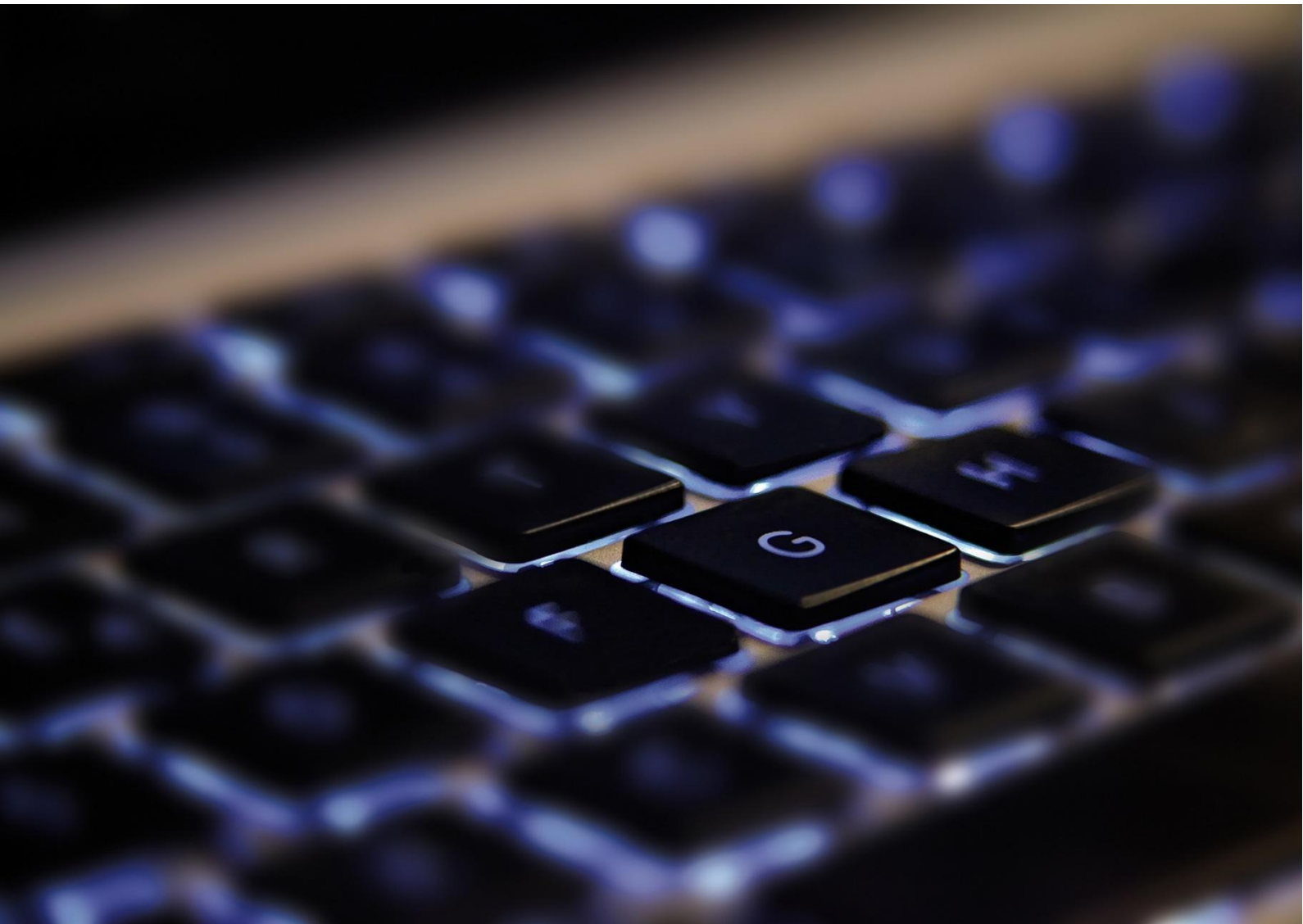
Instituto Tecnológico de Tijuana, México

La Dra. Patricia Melin es doctora en Ciencias Computacionales por la Academia Polaca de Ciencias (PASIFIC), por su trabajo en Sistemas Inteligentes Híbridos para Reconocimiento de Patrones usando “Soft Computing”. Es profesora en Ciencias de la Computación en el Instituto Tecnológico de Tijuana, México desde 1998, y es Directora de los estudios de Grado en Ciencias de la Computación. Desde el año 2000, dirige su grupo de investigación

de Sistemas Híbridos Inteligentes. Mantiene un papel muy activo en múltiples sociedades internacionales e hispánicas (HAFSA) y revistas científicas en los campos de la Inteligencia Computacional y de las Ciencias de la Información (IEEE).

Sus líneas de investigación se relacionan a Redes neuronales modulares, lógica difusa de tipo 2, reconocimiento de patrones, control difuso y enfoques híbridos neuro- y genético-borrosos. Su contribución científica es muy prolífica, con autoría en cerca de 900 publicaciones contando artículos, libros y conferencias.

# Resúmenes



# Análisis funcionales de perfiles de co-expresión génica de *Trypanosoma cruzi*

Lucas Inchausti (1), Álvaro Martín (2), Leticia Pérez-Díaz (1), Beatriz Garat (1), José Sotelo-Silveira (1), Pablo Smircich (1)

1: IIBCE, Facultad de Ciencias UdelaR

2: Facultad de Ingeniería UdelaR

*Trypanosoma cruzi* es un parásito protozoario que causa la enfermedad de Chagas, una enfermedad tropical desatendida que afecta a millones de personas y prolifera en entornos empobrecidos.

*T. cruzi* está caracterizado por tener un ciclo de vida complejo, que implica diferentes etapas de diferenciación.

Los genes *T. cruzi* son expresados de forma policistronica, siendo los mecanismos post-transcripcionales los principales reguladores de la expresión génica.

Los análisis de coexpresión génica son una herramienta valiosa para estudiar los cambios en los niveles de expresión de grupos de genes que interactúan funcionalmente entre sí. El objetivo de este trabajo es caracterizar nuevos mecanismos de expresión a través de la caracterización y análisis funcional de grupos de co-genes expresados utilizando datos transcriptómicos de varias etapas del ciclo de vida de *T. cruzi*.

En este contexto, hemos identificado varios grupos de genes co-expresados mediante el uso de diferentes metodologías de agrupamiento no supervisadas. Actualmente estamos optimizando estos métodos mediante el desarrollo de métricas que evalúen la agrupación funcional de genes conocidos y la consistencia funcional global de la red. Se está llevando a cabo el análisis de las características comunes de los genes en cada grupo.

Esperamos que estos resultados realicen una contribución significativa a la comprensión de la biología molecular de este parásito de gran relevancia.

**PALABRAS CLAVE:** RNA-seq, Co-expression Network

# Dinámica molecular en sistemas biológicos: implementando variable colectiva en software open-source

Ary Lautaro Di Bartolo (1,2)

1: Instituto de Histología y Embriología de Mendoza (IHEM) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de Cuyo (UNCuyo), Mendoza, Argentina

2: Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo (UNCuyo), Mendoza, Argentina

Las simulaciones de Dinámica Molecular (DM) son utilizadas para estudiar sistemas biológicos y estimar cuantitativamente sus propiedades termodinámicas. Dentro de sus limitaciones encontramos que el tiempo de simulación se limita por la capacidad de cómputo disponible. La mayoría de los eventos que se buscan analizar son transiciones que son de corta duración y de rara ocurrencia. Para sortear este inconveniente se diseñan Variables Colectivas (VC) que describen los cambios conformacionales del sistema para sesgar la simulación y acelerar la ocurrencia de estos fenómenos. Las VC son una función de las posiciones tridimensionales de las partículas que conforman el sistema a simular.

PLUMED (<https://www.plumed.org>) es un plugin open-source compatible con los principales motores de DM, como GROMACS o LAMMPS, que permite aplicar algoritmos de muestreo mejorado y analizar las trayectorias generadas por las simulaciones. Tiene la característica distintiva de que es un software creado por la comunidad, por lo que agregar código específico para realizar tareas particulares es sencillo desde su concepción.

Para llevar adelante el trabajo de investigación que terminó publicado en la revista de bandera de la Royal Society of Chemistry, titulado “Synaptotagmin-1 C2B domains cooperatively stabilize the fusion stalk via a master-servant mechanism”, fue necesario implementar una VC en el código fuente de PLUMED que ahora está disponible para ser usado por toda la comunidad científica en cualquiera de los softwares de DM compatibles.

El presente trabajo se centra en dar un panorama general de la dinámica molecular, sus métodos de muestreo mejorado con base en VCs y el proceso de implementar código en un proyecto de código abierto.

**PALABRAS CLAVE:** Dinámica Molecular, Variables Colectivas, PLUMED, Open-source

# Acercamientos computacionales basados en señal filogenética explican las leyes de ensamblaje subyacentes en datos composicionales de comunidades microbianas cultivadas *ex situ*

Silvia Talavera, Daniel Aguirre de Cárcer

Universidad Autónoma de Madrid (España)

**Introducción:** Las comunidades microbianas tienen un gran potencial positivo en, entre otros, la agricultura y la salud humana. Un importante reto de la ecología microbiana es comprender las reglas de su ensamblaje.

**Objetivos:** En este trabajo exploramos la validez de nuestro marco teórico, que considera la señal filogenética preponderante en las comunidades microbianas como factor definidor del ensamblaje.

**Materiales y métodos:** Para ello analizamos datos composicionales preexistentes, procedentes de experimentos de dilución-crecimiento de comunidades naturales complejas en medios con una única fuente de carbono. Utilizando bases de datos de información genómica de acceso libre, llevamos a cabo estudios pangenómicos, modelado metabólico, modelado de la deriva y análisis de redes de interacciones.

**Resultados:** En cada fuente de carbono, la señal filogenética de las comunidades finales quedó delimitada en dos nodos filogenéticos específicos afiliados a las familias Enterobacteriaceae (E) y Pseudomonadaceae (P). El análisis inicial del pangenoma de cada nodo proporcionó indicios para explicar el mayor ratio de abundancias E/P experimental, que se relaciona con una mayor producción de energía por parte de las rutas metabólicas de E, pero ninguna explicación para la presencia persistente de P en las comunidades finales.

El análisis de interacciones indicó que P y E son capaces de intercambiar metabolitos en co-cultivo. Subsecuentes modelizaciones metabólicas de P consumiendo excreciones de E reprodujeron los ratios E/P experimentales. La exploración a fondo reveló claramente las fuerzas motoras del ensamblaje: E fermenta la fuente de carbono inicial a acetato que después es consumido por P en el ciclo de Krebs.

Los análisis de redes indicaron la existencia de parejas de especies con fuertes interacciones inter-nodo, al mismo tiempo que la neutralidad intra-nodo probó ser coherente con las medidas experimentales.

**Conclusión:** Nuestros resultados muestran que el marco teórico propuesto es un acercamiento valioso para descifrar el ensamblaje de comunidades.

**PALABRAS CLAVE:** Microbiology, Community assembly, Metabolic modelling



# **Detección *in silico* de integrones de resistencia en *Salmonella enterica* en la Región Andina utilizando dos herramientas bioinformáticas**

**Lilibeth Torres-Elizalde (1), Karen Loiza (2), Marco Larrea-Álvarez (1), David Ortega-Paredes (3)**

1: School of Biological Science and Engineering, Yachay-Tech, Ecuador.

2: Statens Serum Institut, Bacteria, Parasites and Fungi, Dinamarca

3: Unidad de Investigación en Enfermedades Transmitidas por Alimentos y Resistencia a los Antimicrobianos (UNIETAR), Facultad de Veterinaria-Universidad Central del Ecuador, Ecuador

**Introducción:** La resistencia a los antimicrobianos en *Salmonella enterica* continúa aumentando a nivel mundial y representa un riesgo inminente. Uno de los mecanismos por los cuales este patógeno puede adquirir genes de resistencia antimicrobiana son los integrones. Estas estructuras capturan, transportan e integran numerosos genes de resistencia y tienen la capacidad de asociarse con plásmidos conjugativos que permiten la movilización de toda la estructura entre diferentes poblaciones bacterianas. Por lo tanto, este estudio tuvo como objetivo detectar integrones de resistencia en secuencias de *S. enterica* en algunos países de la Región Andina, utilizando dos herramientas bioinformáticas, I-VIP v1.2 e IntFinder 1.0. Además se determinó la relación entre integrones y los genes de resistencia a los antimicrobianos.

**Metodología:** El conjunto de datos consistió de 1688 aislamientos secuenciados del genoma completo de *S. enterica* que se descargaron de la EnteroBase. Las secuencias procedían de Bolivia, Colombia, Ecuador, Perú y Venezuela y se aislaron desde 1956 hasta 2021. La presencia de integrones de resistencia se investigó mediante IntFinder 1.0 e I-VIP v1.2. Los genes de resistencia a los antimicrobianos se consideraron asociados si se ubicaban dentro de la región variable del integrón.

**Resultados:** IntFinder 1.0 detectó 749 integrones, el 92,4% de clase 2; el 7,6% de clase 1; no se detectaron integrones de clase 3. I-VIP v1.2 predijo 680 integrones, 94,9%; 3,8 % de clase 1; el 94,9 % se clasificó como otras clases de integrones y el 1,3 % eran integrones no funcionales. La mayoría de los integrones detectados estaban relacionados con fuentes animales y ambientales. Se detectaron principalmente en las cepas *S. Infantis*, *S. Paratyphi* y *S. Typhimurium*. Los genes de resistencia a trimetoprima (*dfrA*) fueron los más abundantes. Ambos programas también predijeron genes de resistencia a aminoglucósidos (*aadA*) y betalactámicos (*bla*).

**Conclusiones:** El uso inapropiado de cotrimoxazol en la región podría estar ejerciendo una presión externa que favoreció la alta prevalencia de genes de resistencia a trimetoprima (*dfrA*) entre los integrones detectados. Debido a que los integrones pueden asociarse con elementos móviles, nuestros resultados enfatizan la necesidad de establecer programas de vigilancia para rastrear la evolución de estas plataformas genéticas. Además, esta investigación demostró la utilidad y versatilidad de IntFinder 1.0 e I-VIP v1.2 como herramientas complementarias para estudiar integrones y su relación con genes de resistencia a antibióticos.

**PALABRAS CLAVE:** *Salmonella enterica*, Integrones, Región-Andina

# **Descifrando el interactoma de miARN-ARNm necesario para las transiciones epitelio-mesénquima-epitelio (EMT/MET) de las células de la cresta neural embrionaria**

**Yanel Elina Bernardi, Rocio Marquez, Pablo H. Strobl-Mazzulla**

Laboratorio de Biología del Desarrollo. Instituto Tecnológico de Chascomús (CONICET-UNSAM), Chascomús, Argentina

Durante el desarrollo embrionario de vertebrados, las células de la cresta neural (CCN) delaminan desde la parte dorsal del tubo neural mediante un proceso denominado transición epitelio-mesénquima (EMT). De manera inversa y en ciertas regiones del embrión estas células requieren condensarse para formar ciertos ganglios craneales mediante un proceso similar a la transición mesénquima-epitelio (MET). Durante ambas transiciones, las CCN sufren cambios moleculares y celulares que asemejan mucho a las transiciones que ocurren en las células tumorales durante la metástasis. Particularmente, se ha demostrado que los microARNs (miARNs) juegan un papel crucial en la modulación de la expresión génica durante dichas transiciones tanto en cáncer como durante el desarrollo embrionario normal. En este sentido, nuestro objetivo fue identificar miARNs claves involucrados en las EMT (delaminación) y TME (condensación) de las CCN, y predecir sus genes blanco para abordar experimentalmente la funcionalidad de los mismos en embriones de pollo. Para esto realizamos miRNA-seq y mRNA-seq a partir de CCN obtenidas por FACS a partir de embriones electroporados con reporteros fluorescentes en estadios premigratorio, migratorio y condensantes. El análisis comparativos entre estadios nos lleva a la identificación de dos grupos de microRNAs exclusivos de las CCN durante EMT (miR-140-3p, miR-140-5p y miR-455-5p) y MET (miR-23b-3p, miR-187-3p y miR363-5p). Al comparar mediante una búsqueda in silico de los genes blanco de los miRNAs (predictores TargetScan y miRDB) y los genes downregulados en cada transición, pudimos identificar grupos de genes vinculados en procesos biológicos de adhesión celular, regulación de la comunicación celular, señalización y morfogénesis. Estos estudios nos permitirán tener un mayor entendimiento de la dinámica de las transiciones EMT-MET en un sistema completamente in vivo para poder trazar un futuro paralelismo y las implicancias de estos actores durante la metástasis tumoral.

**PALABRAS CLAVE:** miRNA-seq, RNA-seq, Cresta neural, EMT, MET, microRNAs

# Benchmarking isoform expression quantification methods for RNA-seq data

Arturo González-Vilanova (1), Juan Antonio Carbonell-Asíns (2),  
Tania Fleitas (1), Sheila Zúñiga-Trejos (2)

1: Department of Oncology, Biomedical Research Institute INCLIVA

2: Bioinformatics and Biostatistics Unit, Biomedical Research Institute INCLIVA

**Motivation:** Isoform expression levels are crucial for understanding molecular functions in biological tissues and the progression of disease states. Accurate estimation of expression at the isoform level is suboptimal with current RNA-seq technologies due to different factors such as short read length and sequencing errors among others. Many computational tools have been developed to deal with these problems. This raises the question of which tool is better to use.

**Materials and Methods:** We performed a study using both simulated and experimental data to assess the performance of 5 widely used tools: Salmon, RSEM, kallisto, StringTie and eXpress. We considered the effect of using a reduced annotation by filtering the isoforms that did not code for proteins. Various metrics (Absolute Relative Differences, Spearman Correlation, Coefficient of Variation) were calculated, and statistical methods were applied (beta regression and linear regression) with the aim of comparing the accuracy, sensitivity, specificity, and robustness of the methods.

**Results:** The results indicate that RSEM, Salmon and kallisto performed best and provided similar metrics. On the other hand, StringTie and eXpress have poorer performance with simulated data. The changes in annotation were not indicative of substantial variation in accuracy metrics. Small isoforms like microRNAs and snoRNAs were wrongly estimated by most methods due to differences in the estimated effective transcript length.

**Conclusion:** RSEM, Salmon and kallisto are all excellent options for the purpose of estimating isoform expression. If computational resources are aspects of concern and alignment files are not needed for additional purposes, Salmon and kallisto are the fastest alternatives. The use of a reduced annotation file is not recommended as it does not improve accuracy and yields an incomplete expression profile. Effective transcript length calculation should be an area of concern for methods that aim to improve the performance of published tools.

**PALABRAS CLAVE:** RNA-Seq, Isoform quantification, Benchmarking

# WIKIPROYECTO BASES DE DATOS DE BIOCIENCIAS

**Minerva María Romero Pérez, Yessica Galicia Pérez,  
Paola Castañeda Covarrubias, Layla Michán Aguirre**

Laboratorio de Bioinformación, Departamento de Biología Comparada, Facultad de Ciencias, UNAM

La principal herramienta de los biólogos computacionales son las bases de datos con información en biociencias. Las fuentes de bases de datos para biociencias son catálogos públicos que reúnen y describen estos recursos útiles en la investigación bioinformática, sin embargo, cada colección registra información parcial y con propósitos distintos. De tal manera que, localizar las bases de datos y conocer sus características es tardado y complejo. Wikidata es una base de datos con licencia de dominio público, cuya plataforma es abierta, colaborativa, semántica y multilingüe que permite realizar consultas, generar aplicaciones y generar mapas de conocimiento con la información. Nuestro objetivo es realizar un proyecto colaborativo para localizar, reunir y curar las bases de datos de biociencias en Wikidata con información vigente, completa y ligada.

Hasta el momento, en el Wikiproyecto bases de datos en biociencias ([https://www.wikidata.org/wiki/Wikidata:Bases\\_de\\_Datos\\_Biociencias](https://www.wikidata.org/wiki/Wikidata:Bases_de_Datos_Biociencias)) se han reunido 14 fuentes que incluyen 26,439 registros que servirán para seleccionar únicamente las bases de datos activas para crear y enriquecer la información de Wikidata. Se han identificado 42 propiedades para describir las bases de datos, se han creado más de 350 elementos nuevos y se han curado más de 3000 entidades en Wikidata con información básica. La documentación del proyecto está abierta y se puede consultar la página. En la siguiente etapa se completará para cada base de datos toda la información disponible en las fuentes y la que resulte de la exploración de cada plataforma con datos sobre taxón, tema, objeto de estudio, licencia, tecnologías asociadas, por mencionar las más importantes. Con base en lo anterior, estimamos que estarán disponibles cerca de 5,000 elementos de bases de datos en biociencias localizables, abiertas, interoperables y reusables para consulta, análisis y visualización semántica en Wikidata cuando se complete este proyecto.

**PALABRAS CLAVE:** Bases de datos biológicas, Wikidata, Web semántica, Biocuración, Interoperabilidad

# Estudio de las bases moleculares afectadas por el sexo en Esclerosis Múltiple a través del transcriptoma de tejido nervioso

José Francisco Català-Senent (1), Zoraida Andreu (2), Marta R. Hidalgo (1), Francisco José Roig (1,3), Natalia Yanguas-Casás (4), Almudena Neva-Alejo (1), Adolfo López-Cerdán (5), Irene Soler-Sáez (1), María de la Iglesia-Vayá (5), Francisco García-García (1)

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

3: Faculty of Health Sciences. San Jorge University, 50830, Zaragoza, Spain

4: Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana (IDIPHISA), Grupo de Investigación en Linfomas, C/Joaquín Rodrigo 2, Majadahonda, 28222 Madrid, Spain

5: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

La esclerosis múltiple (EM) es una enfermedad crónica, autoinmune y degenerativa que afecta al sistema nervioso central. Esta enfermedad se comporta de forma diferente entre hombres y mujeres, siendo más prevalente en las mujeres. A pesar de los avances en el estudio de la enfermedad, no se conoce con precisión las causas de las diferencias observadas en función del sexo.

Por este motivo, el objetivo del trabajo ha sido identificar posibles mecanismos moleculares subyacentes a las diferencias de sexo en EM. Con este fin, se llevó a cabo una revisión sistemática de estudios transcriptómicos con muestras de tejido nervioso, con información del sexo de los pacientes disponibles en los repositorios de datos públicos. Los trabajos fueron seleccionados según las normas de la declaración PRISMA y sometidos a metaanálisis para identificar las alteraciones genéticas diferenciales entre mujeres y hombres. Finalmente, se realizó un enriquecimiento funcional de los resultados para caracterizar funcionalmente estas alteraciones.

Como resultado de nuestra revisión sistemática, se seleccionaron 4 estudios de la base de datos Gene Expression Omnibus que contenían un total de 224 pacientes (137 hembras y 87 machos). A nivel genético se detectaron 13 genes significativos alterados entre mujeres y hombres, mientras que a nivel funcional únicamente una ruta KEGG mostró diferencias significativas. Interesantemente, las respuestas inmunes y las células asociadas difieren por sexo. Estos resultados podrían abrir la puerta a nuevas investigaciones que lleven a tratamientos más específicos para cada sexo.

**PALABRAS CLAVE:** Esclerosis múltiple, Biomarcadores, Metaanálisis, Diferencias de sexo, Neurodegeneración

# Reposicionamiento in silico de Fármacos en la Enfermedad de Alzheimer con Perspectiva de Sexo

**Cristina Galiana-Roselló (1), Rubén Grillo-Risco (1), José Francisco Català-Senent (1), Adolfo López-Cerdán (1,2), Isabel del Pino (3), Francisco García-García (1)**

1: Unidad de Bioinformática y Bioestadística, Centro de Investigación Príncipe Felipe (CIPF), Valencia, 46012, España.

2: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, España.

3: Laboratorio de Plasticidad Neuronal, Centro de Investigación Príncipe Felipe (CIPF), Valencia, 46012, España.

La enfermedad de Alzheimer (EA) es la causa más común de demencia, situándose entre las 10 primeras causas de muerte en todo el mundo. Al igual que otras enfermedades neurodegenerativas, la EA presenta un evidente dimorfismo sexual en su prevalencia, pronóstico, y progresión. A pesar de estos indicadores epidemiológicos y clínicos, el tratamiento sigue siendo sintomático. Además, el desarrollo de nuevos fármacos es limitado por una alta tasa de fallo terapéutico en los ensayos clínicos, en los cuales no siempre se proporciona un desglose por sexo de la seguridad y eficacia.

En vista de la carga epidemiológica, la ausencia de tratamientos efectivos y el dimorfismo sexual de esta enfermedad, este trabajo tiene como objetivo identificar potenciales estrategias terapéuticas frente a la EA, considerando el factor sexo, mediante el reposicionamiento o reutilización in silico de fármacos basado en el principio de reversión del transcriptoma: identificar fármacos que reviertan las firmas de expresión génica asociadas con estados de enfermedad.

Con esta finalidad, se han analizado más de 5000 perfiles transcriptómicos de células progenitoras neuronales, neuronas y motoneuronas tratadas con más de 3000 compuestos, a diferentes dosis y tiempos, generados por el consorcio Cmap-LINCS. La librería de compuestos incluye tanto fármacos aprobados por las agencias reguladoras o en ensayos clínicos, así como nuevas entidades moleculares evaluadas en estudios preclínicos. El análisis de correlación posterior de estos perfiles con firmas consenso de hombres y mujeres con EA, a partir de diferentes algoritmos basados en GSEA y métricas de similitud, nos proporcionará una lista de priorización de fármacos ordenados por sus puntajes de conectividad. Esta lista de priorización permitirá postular nuevas estrategias terapéuticas, así como elucidar si los fármacos presentan un patrón diferencial de correlación entre los perfiles de expresión de hombres y mujeres, y en su caso, aproximarnos a un tratamiento personalizado.

**PALABRAS CLAVE:** Reposicionamiento de fármacos, Alzheimer, Diferencias de Sexo

# Efectos de la inflamación en la alteración del metabolismo de los lípidos en la epidermólisis bullosa distrófica recesiva severa: lipidómica y transcriptómica.

Herraiz-Gil S (1), Arriba MC (1), Martínez-Santamaría L (1), Sacedón R (2), Maseda R (3), Capellades J (4), Junza S (4), Tell-Martí G (5), Potrony M (5), Badenas C (5), Jiménez E (2), Quintana L (3), Illera N (1), García M (1), Lwin SM (6), McGrath JA (6), Vicente MA (2), Yanes O (4), Dopazo J (7), Puig S (5), Puig JA (5), de Lucas R (3), del Río M (1), León C (1), Escámez MJ (1)

1: Department of Bioengineering, UC3M; CIBERER U714; Regenerative Medicine and Tissue Engineering Group, IIS-FJD; CIEMAT, Madrid, Spain.

2: Department of Cell Biology, Faculty of Medicine, UCM, Madrid, Spain.

3: Department of Dermatology, La Paz University Hospital, Madrid, Spain.

4: Universitat Rovira i Virgili, Department of Electronic Engineering, IISPV, CIBERDEM, Tarragona, Spain.

5: Melanoma Unit, Hospital Clinic & IDIBAPS (Institut d'Investigacions Biomèdiques Agustí Pi i Sunyer), CIBERER (U726), Universitat de Barcelona, Barcelona, Spain.

6: King's College London, London, United Kingdom.

7: Clinical Bioinformatics Area, Fundación Progreso y Salud, CDCA; Functional Genomics Node, INB-ELIXIR-es, FPS; Bioinformatics in Rare Diseases (BiER-U715), CIBERER, FPS, Hospital Virgen del Rocío, Sevilla, Spain.

**Introducción:** La epidermólisis bullosa distrófica recesiva (EBDR) es una genodermatosis de baja prevalencia, causada por mutaciones en COL7A1, que cursa con fragilidad mucocutánea y otras complicaciones que merman la calidad de vida y aumentan la morbimortalidad. Se han descrito genes modificadores y factores epigénéticos, asociados a inflamación, que contribuyen a la severidad clínica y repercutirían en el diseño de terapias efectivas.

**Objetivos:** Profundizar, mediante tecnologías ómicas y herramientas bioinformáticas, en las bases moleculares de la EBDR en una cohorte con alta prevalencia de hipolipidemia, que correlaciona con su estado inflamatorio.

**Materiales y métodos:** La cohorte incluye 8 pacientes pediátricos reclutados en Mesensistem-EB (NCT04153630). Se realizó un estudio transcriptómico mediante RNA-Seq, y se integró con un análisis lipidómico empleando UPLC-MS (Q-Exactive). Los datos se preprocesaron en R, y posteriormente, se integraron usando Cytoscape para el modelado de redes de interacción. Finalmente se realizó un análisis de enriquecimiento funcional para la interpretación biológica empleando STRING y LipidMap.

**Resultados:** El análisis transcriptómico reveló la desregulación de 255 genes (FDR<0.05) en la piel de los pacientes. En cuanto al metabolismo lipídico, se encontró una baja expresión en el gen de la leptina (FDR=0.0065) y otros de su red de interacción, relacionados con los procesos de inflamación. El análisis lipidómico mostró la disminución de 162 lípidos circulantes (FDR<0.05) en el suero de los pacientes. Entre las clases de lípidos más representadas se encuentran las carnitinas, ácidos grasos, esteroides y glicerofosfolípidos, sugiriendo una desregulación lipídica en EBDR asociada al proceso de inflamación sistémica.

**Conclusiones:** Este estudio identifica, por primera vez, los perfiles transcriptómico cutáneo y metabólico sistémico en pacientes con EBDR y dilucida los eventos moleculares subyacentes a la hipolipidemia clínicamente observada. Estos resultados contribuyen a la comprensión de las bases moleculares de EBDR sugiriendo una relación entre la inflamación crónica y el metabolismo lipídico.

**PALABRAS CLAVE:** Epidermólisis bullosa, Análisis de datos, Lipidómica, Transcriptómica, Inflamación

# **TOLEDO: a new software that is able to run and analyze Desmond Molecular Dynamics simulations on high-performance computing clusters.**

**M. Carmena-Bargueño, C. Martínez-Cortés, A. J. Banegas-Luna, H. Pérez-Sánchez**

Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), UCAM Universidad Católica San Antonio de Murcia, 30107 Murcia, Spain

Molecular Dynamics (MD) simulations predict the dynamical evolution of biological systems. Using MD in conjunction with high performance computing (HPC) architectures, we can study the potential interactions and their dynamical evolution between a ligand library against one protein with the purpose to find a drug which can influence on a protein target to cure a disease. A previous MD implementation on GPUs by our group with the rCUDA software [1] showed a 4x throughput improvement with respect to classical MD (~500 ns/day vs ~125 ns/day). However, our previous study [1] focused only on GROMACS; targeted to experienced users, due to the necessity of configuring a lot of parameters via command line. Whereas, the MD software Desmond, executed via Maestro Suite ([www.schrodinger.com](http://www.schrodinger.com)) allows to setup, run and analyze MD with a graphical interface, allowing to be used by non-expert users. Maestro-Desmond is a software exploited via Graphical User Interface (GUI), however using the GUI, users can typically run only one MD calculations at the same time. In this work, we provide a tool called TOLEDO (Throughput Optimization of Ligand-Protein systems Exploration through Dynamics simulation in Optimized HPC systems) to overcome such limitation and run effortlessly several MDs at the same time. First, we setup the protein-ligand complex via Maestro GUI, next we run several MD in a supercomputer such as SAGA ([https://documentation.sigma2.no/hpc\\_machines/saga.html](https://documentation.sigma2.no/hpc_machines/saga.html)) or Turgalium of CIEMAT (<https://www.ceta-ciemat.es/>) and finally, show the results with graphics that are very simple to interpret. In summary, TOLEDO is a tool that allows researchers to run MDS with higher throughput than the Maestro GUI.

**References:** [1] DOI: [10.1177/1094342019857131](https://doi.org/10.1177/1094342019857131).

**PALABRAS CLAVE:** Molecular Dynamics, High Performance Computing, Drug Discovery



# ASGARD: Review of a simple and automatic GROMACS Molecular Dynamics analysis tool

Alejandro Rodríguez-Martínez (1), Jochem Nelen (2), Miguel Carmena-Bargueño (2), Carlos Martínez-Cortés (2), Irene Luque (1), Horacio Pérez-Sánchez (2)

1: Departamento de Química Física, Instituto de Biotecnología y Unidad de Excelencia de Química aplicada a Biomedicina y Medioambiente Universidad de Granada. Campus Fuentenueva s/n 18071 Granada, Spain

2: Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), UCAM Universidad Católica de Murcia (UCAM), 30107 Murcia, Spain

**Background:** Molecular Dynamics (MD) has become one of the most popular techniques to understand molecular structure and interactions [1]. Among the different programs used to perform MD simulations, GROMACS is an open-source and free software and one of the most widely used for this purpose [2].

**Objectives:** The main objective is to obtain a tool that can generate an PDF analysis report with one single command from Molecular Dynamics simulation files obtained by GROMACS

**Methods:** ASGARD is developed in Python and bash languages, using libraries like Numpy and Matplotlib, and using GROMACS command and other protein-ligand interaction programs such as PoseView.

**Results:** We have created ASGARD, a tool that allows via an automated MD workflow to perform MD protein or protein-ligand complex simulations and to easily generate the corresponding analyses. Thus, this tool generates a set of analyses after completing simulation. First, several graphs about the system stability are generated. Also, it generates a system flexibility analysis with calculations as RMSD Fluctuation and Distribution. Finally, a dynamic analysis is performed with a SASA analysis, DSSP method graphs and different interaction analyses.

**Conclusions:** This tool allows the user to run a MD simulation with a single command line instead of using the following GROMACS programs individually. In addition, after finishing the simulation, it automatically creates an analysis report to understand the different molecular interactions and structural changes.

**References:** [1] DOI: [10.1016/s0959-440x\(02\)00308-1](https://doi.org/10.1016/s0959-440x(02)00308-1); [2] DOI: [10.1016/j.softx.2015.06.001](https://doi.org/10.1016/j.softx.2015.06.001)

**PALABRAS CLAVE:** Structural Bioinformatics, Molecular modelling, Molecular Dynamics, GROMACS

# Optimizing consensus docking by incorporating predicted binding pose similarity: a Wee1 case study

J. Nelen (1), M. Carmena-Bargueño (1), C. Martínez-Cortés (1), A. Rodríguez-Martínez (1), J.M. Villalgorido (2), H. Pérez-Sánchez (1)

1: Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), Computer Engineering Department, UCAM Universidad Católica de Murcia, Murcia, Spain

2: Eurofins Villapharma Research, Murcia, Spain

**Introduction:** Molecular docking is a structure-based drug design technique that aims to predict both the binding pose and affinity of a ligand against a chosen protein target<sup>1</sup>. It is used extensively as a virtual screening (VS) technique, where it aids in the identification of potential active compounds present in large chemical libraries. The success of a docking-based VS campaign can vary substantially depending on various factors, including the protein target and docking algorithm. One way to reduce this variability is to make use of consensus docking<sup>2</sup>, where multiple docking algorithms are used, and only compounds which show good binding affinities in both calculations are predicted to be active.

**Objectives:** For this research, the aim was to both automate and further optimize a consensus docking protocol by including an additional metric which quantifies the similarity in predicted binding pose.

**Materials and Methods:** A Python script was developed to automatically calculate the root-mean-square deviation (RMSD) between binding poses that were predicted by the different Docking algorithms. Subsequently, MetaScreener<sup>3</sup> was used to perform docking-based VS of the DrugBank<sup>4</sup> library for the protein Wee1, using both the AutoDock Vina<sup>5</sup> and the LeadFinder<sup>6</sup> algorithms. Drugbank contains ~10000 compounds, 8 of which are listed as Wee1 inhibitors. To finally validate the RMSD metric, the script was used on the MetaScreener consensus docking results, to see if it could further improve them.

**Results and Conclusion:** Using the regular consensus docking approach, 4 out of the 8 known Wee1 inhibitors were identified in the top 50 compounds. However, when using the RMSD metric, it was possible to identify these 4 compounds in the top 10. Thus in this case, it managed to substantially reduce the amount of false positives. To conclude, the RMSD metric looks like a promising enrichment factor, which will be further explored and utilized at the BIO-HPC group in several drug discovery projects.

**PALABRAS CLAVE:** *In silico* drug design, Virtual Screening, Molecular Docking, Consensus Docking

# ANÁLISIS DE VARIABILIDAD DE GENOMAS DEL SARS-COV-2

Canguí, Sandra; Fuel, Marco

Universidad de Málaga - España

**Introducción:** Las investigaciones sobre el genoma del coronavirus SARS-CoV-2 son de gran importancia para poder conocer su variabilidad y la implicación de esta en la transmisión y patología de la enfermedad. La secuenciación del genoma del virus mediante NGS permite determinar las mutaciones existentes para conocer la evolución del virus, dar un seguimiento a la enfermedad y encontrar posibles dianas terapéuticas.

**Objetivos:** En este trabajo se realizó un análisis bioinformático de los cebadores empleados en dos protocolos muy utilizados de las empresas ARTIC y Entebbe por su particularidad en ofrecer mejorar la calidad y cobertura de los genomas proporcionando menos sitios gaps en su fase de secuenciación. También se incluyó una investigación bibliográfica y bioinformática de la variabilidad génica y las mutaciones existentes en los sitios de interfaz e interacción entre los dímeros Nsp10-Nsp14 y Nsp10-Nsp16 del sitio activo del complejo replicasa-transcriptasa (RTC) del SARS-CoV-2 en estudios de cuasiespecies virales para conocer el espectro mutante del virus y determinar estos sitios como posibles dianas terapéuticas para futuros estudios de mutagénesis letal.

**Materiales y métodos:** En el análisis de la variabilidad genómica se trabajó con un total de 7462 secuencias obtenidas de las bases de datos GenBank y GISAID. En los análisis bioinformáticos se hizo uso de softwares como BLAST y MEGA-X.

**Resultados y conclusión:** Mediante estos análisis bioinformático se pudo determinar que algunos de los cebadores empleados en el proceso de secuenciación no eran específicos de SARS-CoV-2. En cuanto al análisis de la variabilidad de genomas del virus, se identificó una frecuencia baja de mutación entre  $3 \times 10^{-4}$  y 0.01 mutaciones de sustitución por secuencia analizada y se encuentran en los sitios de interacción de los complejos proteicos Nsp10-Nsp14 y Nsp10-Nsp16, esto indica que los sitios de interés pueden ser usados como dianas para el desarrollo de péptidos que bloqueen las interacciones entre estos dímeros y puedan ser empleados como alternativa terapéutica frente al SARS-CoV-2.

**PALABRAS CLAVE:** SARS-CoV-2, Genoma, Cebadores, Variabilidad génica

# Atlas funcional de las diferencias de sexo en esclerosis múltiple mediante el análisis de datos transcriptómicos de célula única

Irene Soler-Sáez(1), Zoraida Andreu(2), José Francisco Català-Senent(1), Rubén Grillo-Risco(1), Adolfo López-Cerdán(3), Almudena Neva-Alejo(1), Héctor Carceller(3), María de la Iglesia-Vayá(5), Marta R. Hidalgo(1), Francisco García-García(1)

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

3: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

La esclerosis múltiple (EM) es una enfermedad crónica, autoinmune y neurodegenerativa del sistema nervioso central. Pese a que su etiología es desconocida, esta enfermedad constituye la primera causa de discapacidad no asociada a accidentes entre la población adulta joven. Se han identificado diferencias de sexo en EM a nivel clínico y epidemiológico. Sin embargo, los mecanismos moleculares subyacentes implicados en dichas diferencias continúan siendo un enigma para la comunidad científica.

El objetivo de este trabajo es caracterizar a nivel de tipo celular, el perfil funcional que diferencia a hombres y mujeres que sufren EM. El abordaje bioinformático in silico realizado a tal fin consta de los siguientes pasos: 1) revisión sistemática para la identificación y posterior descarga de los datos a analizar, procedentes de células o núcleos individuales; 2) procesamiento de datos, 3) control de calidad, 4) normalización, 5) reducción de la dimensionalidad, 6) anotación del tipo celular, 7) análisis de expresión diferencial y, 8) caracterización funcional, dónde se evaluaron funciones biológicas y rutas de señalización.

Como resultado de la revisión sistemática se identificaron tres estudios con datos procedentes de muestras de tejido nervioso o de sangre. La anotación de tipos celulares reveló la existencia de astrocitos, microglía, neuronas, oligodendrocitos, y precursores de oligodendrocitos en tejido nervioso; y linfocitos B, linfocitos T CD4+, linfocitos T CD8+, linfocitos T, linfocitos NK, células dendríticas y monocitos en sangre. Finalmente, la caracterización funcional de cada tipo celular permitió conocer la existencia de funciones biológicas y rutas de señalización alteradas de forma característica en un único tipo celular, mientras que otras diferencias son compartidas entre distintos tipos celulares. Con todo ello, estos resultados pueden orientar la potencial búsqueda de biomarcadores dependientes del sexo en la enfermedad; situación que ayudará a promover la medicina personalizada para la EM.

**PALABRAS CLAVE:** Esclerosis múltiple, Diferencias de sexo, scRNA-seq, snRNA-seq, Medicina personalizada

# MetaScreener: a software suite focused on Virtual Screening workflows on High Performance Computing infrastructures

Carlos Martínez Cortés, Miguel Carmena Bargueño , Antonio Jesús Banegas Luna, Alfonso Pérez Garrido, Horacio Pérez Sánchez

Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), Universidad Católica de Murcia (UCAM), Murcia, España

**Introducción:** MetaScreener (MS) es una colección de scripts desarrollados en diferentes lenguajes de programación (Shell script, Python, Java, C, ...), usado para llevar a cabo cálculos de descubrimientos de fármacos a gran escala. MS integra docking, cribado farmacóforo y otros enfoques de modelado molecular, cuyas ejecuciones son realizadas en supercomputadores mediante distintos sistemas de gestión de tareas y clústeres. Además los resultados pueden consensuarse, obteniendo un resumen de los mejores compuestos de múltiples ejecuciones. El grupo BIO-HPC la ha usado en más de 30 publicaciones y varias patentes, algunas licenciadas.

**Objetivos:** El desarrollo, test y lanzamiento de la herramienta para su uso por toda la comunidad científica. Durante su etapa de mantenimiento, el objetivo será aumentar el número de métodos y softwares soportados por la herramienta.

**Materiales y métodos:** MS recibe del usuario una línea de comandos y, tras una serie de preparaciones, manda al supercomputador los jobs indicados, que obtienen el resultado del método y software requeridos. Opcionalmente se puede hacer una fase de análisis, que optione más información y gráficos de utilidad, así como una sesión Pymol con el resultado en forma gráfica y las interacciones de los distintos residuos.

**Resultados y conclusiones:** La herramienta está disponible de forma gratuita en github: <https://github.com/bio-hpc/metascreeener>. Algunos casos de uso reciente en [1,2].

**References:** [1] DOI: [10.1016/j.jff.2021.104479](https://doi.org/10.1016/j.jff.2021.104479); [2] DOI: [10.3390/cancers13040861](https://doi.org/10.3390/cancers13040861)

**PALABRAS CLAVE:** Molecular modeling, Virtual Screening, Molecular docking, Similarity

# Análisis de los subtipos y cadenas de transmisión predominantes en la epidemia de VIH-1 en España durante los años 2011-2020

Marco Fuel (1), Sandra Cangui (2)

1: Centro de Investigación Biomédica (CIBM). Universidad de Granada, 18100 Granada, España.

2: Facultad de Ciencias. Universidad de Málaga. 29010 Málaga, España

**Introducción:** La vigilancia de la variabilidad molecular del VIH-1, el riesgo de transmisión de la resistencia a los antirretrovirales y la evolución de nuevas variantes, constituyen un problema de salud pública debido a la presencia de subtipos y nuevas formas recombinantes circulantes (CRFs) resistentes que limita las opciones terapéuticas. La metodología estándar en su identificación es la filogenia, sin embargo, en la clínica por su rapidez y facilidad se emplean herramientas automáticas. Éstas presentan limitaciones, para la correcta identificación de algunos subtipos no-B, CRFs y formas recombinantes únicas (UFRs).

**Objetivos:** El objetivo fue analizar los subtipos predominantes del VIH-1 en España y analizar su transmisión de resistencia.

**Métodos:** Las secuencias del gen pol del VIH-1 se obtuvieron de bases de datos. Para el subtipado se emplearon tres herramientas automatizadas (Rega, Comet y Geno2Pheno) y se comparó los resultados con la filogenia, determinando su sensibilidad. Por último, se determinó las mutaciones de resistencia.

**Resultados:** De las 1546 secuencias, los subtipos más prevalentes corresponden a los B y F1, seguido de las formas recombinantes CRF02\_AG y CRF14\_BG. La sensibilidad de las herramientas automatizadas para el subtipo B fue del 90% y para los no-B del 60%. Respecto a la resistencia a los antirretrovirales, se evidenció una mayor frecuencia de mutaciones a la familia de los inhibidores de la transcriptasa inversa (ITIAN e ITINAN), seguido de los inhibidores de la proteasa (IP) y a los inhibidores de la integrasa (INI). Las mutaciones más frecuentes para los ITIAN fueron M184V/I, T215Y y M41L, para los ITINAN, K103N, Y181C y para los IP, M46I/L.

**Conclusión:** Este estudio permitió tener un panorama general sobre la prevalencia de las variantes circulantes del VIH-1 en España, así como de las mutaciones más prevalentes, lo que ayudaría a la toma de decisiones en el diagnóstico y tratamiento.

**PALABRAS CLAVE:** VIH-1, Farmacorresistencia viral, Filogenia

# Identificación de nuevos biomarcadores microRNAs en esclerosis múltiple con un perfil diferencial por sexo

**Ester Muñoz del Campo, José Francisco Català-Senent, Francisco García-García**

Bioinformatics and Biostatistics Unit, Príncipe Felipe Research Center (CIPF), 46012, Valencia, Spain

La esclerosis múltiple es una enfermedad que conlleva la desmielinización y neurodegeneración del sistema nervioso central. Sus síntomas presentan variaciones en cuanto al inicio, progreso y manifestaciones clínicas entre pacientes; además hay determinados aspectos que difieren entre hombres y mujeres. En concreto, las mujeres presentan mayor riesgo de desarrollar la enfermedad, mientras que en los hombres la progresión es más rápida y severa.

Para estudiar y comprender mejor estas diferencias asociadas al sexo, se realiza un estudio que incluye una revisión sistemática de datos de microRNAs, y un posterior metaanálisis de estos. Ello permitirá identificar las posibles diferencias a nivel de regulación de la expresión génica entre hombres y mujeres que padecen esclerosis múltiple. De esta manera se pretende conocer mejor los mecanismos moleculares de esta patología, así como dar pie a nuevas herramientas para su detección y tratamiento; permitiendo su abordaje mediante medicina personalizada.

**PALABRAS CLAVE:** Esclerosis Múltiple; microRNAs; Metaanálisis; Diferencias de sexo

# PRIORIZACIÓN DE VARIANTES EN LA SECUENCIACIÓN DE EXOMA COMPLETO DE PLASMA PARA LA CARACTERIZACIÓN DEL TUMOR EN LA RECAÍDA E IDENTIFICACIÓN DE NUEVAS DIANAS TERAPÉUTICAS EN PACIENTES CON CÁNCER DE COLON

**Jorge Martín-Arana (1,2), Francisco Gimeno-Valiente (1), Roberto Tébar-Martínez (1,2), Santiago Moragón (1), Marta Tapia (1), Marisol Huerta (1), Carolina Martínez-Ciarpaglini (3), Clara Alfaro-Cervello (3), Manuel Cabeza-Segura (1), Blanca García-Micó (1), Elena Duréndez-Sáez (1), Valentina Gambardella (1,4), Pilar Rentero-Garrido (1,2), Sheila Zuñiga (1,2), Tania Fleitas (1), Susana Roselló (1,4), Josefa Castillo (4,5), Desamparados Roda (1,4), Andrés Cervantes (1,4), Noelia Tarazona (1,4)**

1: Unidad de Oncología, Instituto de Investigación Sanitaria INCLIVA, Universidad de Valencia, Valencia, España.

2: Unidad de Medicina de Precisión, Instituto de Investigación Sanitaria INCLIVA, Valencia, España.

3: Departamento de Patología, Instituto de Investigación Sanitaria INCLIVA, Valencia, España.

4: Instituto de Salud Carlos III, CIBERONC, Madrid, España.

5: Departamento de Bioquímica y Biología Molecular, Universidad de Valencia, Valencia, España.

**Introducción:** Alrededor del 80% de los pacientes con cáncer de colon (CC) con DNA tumoral circulante (ctDNA) positivo en plasma no negativizan después de la quimioterapia. La secuenciación del exoma completo (WES) puede ayudar a la caracterización de la enfermedad, pero aporta una gran cantidad de información mutacional que debe ser filtrada.

**Objetivos:** Implementación de un proceso de priorización de variantes en el estudio pareado del plasma en el momento del diagnóstico y la recaída mediante WES en pacientes con CC para una mejor comprensión de los cambios dinámicos del tumor e identificar potenciales dianas accionables.

**Metodología:** Se realizó WES del tumor primario, sangre periférica y plasma de la recaída de 25 pacientes con CC. En 12 de ellos se dispuso de muestras de plasma al diagnóstico. Se llevó a cabo un proceso de priorización de variantes mediante la anotación con bases de datos internas y públicas como COSMIC y OncoKB para la identificación de mutaciones oncogénicas.

**Resultados:** En el 92% de los pacientes se detectó al menos 1 mutación oncogénica en la recaída. Se comprobó que las muestras de tejido presentaban la limitación de la heterogeneidad intratumoral, por lo que el plasma fue elegido para la comparación entre diagnóstico y recaída. Se identificaron un 26.5% de mutaciones nuevas en la recaída relacionadas con vías de respuesta a estrés celular, la vía de señalización de WNT; y un enriquecimiento mutacional significativo en vías de evasión inmune. Se detectaron potenciales dianas terapéuticas en el 80% de los pacientes.

**Conclusiones:** La priorización de variantes en el estudio del perfil mutacional del plasma mediante WES en pacientes con CC permite evaluar la evolución dinámica del tumor, superando la barrera de la heterogeneidad intratumoral, así como la detección de potenciales dianas de tratamiento que mejoren la respuesta obtenida en terapias convencionales.

**PALABRAS CLAVE:** WES, ctDNA, Cáncer de colon, Oncogénico, OncoKB



# Nuevas estrategias bioinformáticas para la identificación de biomarcadores en melanoma

Helena Gómez-Martínez (1), Adolfo López Cerdán (1,2), Francisco García-García (1)

1: Unidad de Bioinformática y Bioestadística, Centro de Investigación Príncipe Felipe (CIPF), 46012 Valencia, España.

2: Unidad de Imagen Biomédica FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012 Valencia, España.

**Introducción:** El melanoma es uno de los tipos de cáncer de piel más letales. Esto es debido principalmente a que su detección temprana se dificulta por la aparición tardía de síntomas y a la falta de un tratamiento eficaz. Sin embargo, a pesar de que su incidencia y el interés científico en la enfermedad han ido en aumento en los últimos años, todavía existe un desconocimiento sobre biomarcadores eficaces que favorezcan la detección temprana y la predicción de la evolución de la enfermedad en los pacientes. Por tanto, la identificación de biomarcadores mediante herramientas bioinformáticas que integren la información conocida, podría mejorar la comprensión del desarrollo del melanoma y la predicción del diagnóstico del paciente.

**Objetivos:** El objetivo principal de este trabajo fue la identificación de miRNAs expresados diferencialmente que permitan tanto un diagnóstico temprano, como la predicción del avance de la enfermedad, mediante una aproximación computacional.

**Materiales y métodos:** En primer lugar, se realizó una revisión sistemática de estudios de melanoma en diferentes repositorios públicos y, en base a criterios previos, se seleccionaron sólo aquellos que incluían datos de microarrays o de secuenciación masiva de miRNAs. Tras la descarga de los datos, se realizó un análisis exploratorio y la normalización de estos. Posteriormente, se evaluó individualmente cada estudio y se llevó a cabo un metaanálisis global integrando todos los resultados individuales obtenidos considerando la información específica de los pacientes.

**Resultados:** El estudio permitió la identificación de los miRNAs expresados diferencialmente entre tejido de melanoma y nevus, en un escenario de diagnóstico de la enfermedad.

**Conclusión:** Los hallazgos generados en este estudio han contribuido a una mejor comprensión de los mecanismos moleculares y los procesos de desarrollo del melanoma, mediante el desarrollo de recursos novedosos para su detección y tratamiento.

**PALABRAS CLAVE:** Melanoma, Expresión de miRNAs, Medicina de precisión, Metaanálisis

# An improved multi-modal algorithm based on an NSGA-II scheme for phylogenetic tree inference

César Rivera(1), Eduardo Kessi-Pérez(2), Mario Inostroza-Ponta(1) and Manuel Villalobos-Cid(1)

1: Departamento de Ingeniería Informática, Facultad de Ingeniería, Laboratory of Artificial Intelligence applied to Bioinformatics, Universidad de Santiago de Chile

2: Centro de Estudios en Ciencia y Tecnología de Alimentos (CECTA), Universidad de Santiago de Chile

Phylogenetic inference methods allow proposing evolutionary hypotheses using molecular or morphological data, often representing the evolutionary relationships as phylogenetic trees. The optimisation-based inference methods attempt to find the best topologies according to a specific criterion among all the possible topologies inferred for the same dataset. It has been classified as an NP-hard problem.

The literature has proposed different inference criteria: minimum evolution, least-squares, parsimony, and likelihood. Several multi-objective optimisation strategies have been published during the last decade to reduce the bias associated with the dependency on the selected criterion, improving the quality of the solutions regarding the single-optimisation approaches considering dominance. The state-of-the-art multi-objective strategies search for solutions using operators and metrics based on the objective space optimising pair of inference criteria. However, recent works concluded that the topological features of the trees (decision space) and the objective space are not related in the multi-objective phylogenetic inference context, becoming phylogeny in a multimodal optimisation problem.

This work deals with this problem by incorporating additional information from the decision space over the searching process, designing a new metric based on a topological tree distance by proposing a new multimodal multi-objective optimisation inference algorithm. This work aims to improve the diversity of the solutions according to the decision space, obtaining good solutions in terms of the optimised inference criteria.

We compare the performance of the proposed strategy with state-of-the-art algorithms using classical datasets from the literature. Additionally, we thoroughly analyse a study case using yeast *Saccharomyces cerevisiae* aligned sequences. Results show that our proposal increases the diversity of topologies, improving the quality of solutions in terms of hypervolume and the optimised inference criteria.

Our current research is focused on (1) reducing the consumption time regarding our proposed algorithm and (2) exploring different topological distances as part of the multimodal operator.

**PALABRAS CLAVE:**Phylogenetic inference, Multi-objetive,  
Multi-modal optimisation

# FILODINÁMICA Y EVOLUCIÓN MOLECULAR DE SARS-COV-2 EN COLOMBIA

Sindy P. Buitrago(1,2,3\*), Jhoan S. Beltran(1,4), Diego Garzón-Ospina(1,2,3)

1: PGAME - Population Genetics And Molecular Evolution, Fundación Scient

2: GEBIMOL, Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

3: GEO, Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

4: Estudiante del programa de Biología, Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

**Introducción:** SARS-CoV-2, el agente causal de la Covid-19, ha causado tasas de mortalidad y carga socioeconómica significativas en todo el mundo. La dinámica del virus en Colombia es evaluada por el Instituto Nacional de Salud (INS) a partir de los casos sintomáticos diarios (enfoque empírico) usando un modelo SIR (del inglés, Susceptible, Infectious, and Recovered/Removed). Sin embargo, este enfoque podría estar sesgado por pacientes asintomáticos, personas con síntomas leves que no reportan la enfermedad, o la capacidad de detectar casos positivos. No obstante, las secuencias genómicas pueden ser útiles brindando información sobre epidemias a través de análisis filodinámicos.

**Objetivos:** Evaluar la dinámica y la evolución del SARS-CoV-2 en Colombia durante 22 meses utilizando datos del genoma completo del virus.

**Materiales y métodos:** Se analizaron 1975 genomas completos de SARS-CoV-2 de diferentes regiones de Colombia disponibles en GISAID. Estas secuencias se alinearon utilizando el método MAFFT y parámetros epidemiológicos fueron obtenidos mediante el modelo SIR, empleando el programa BEAST. Estos fueron comparados con los parámetros obtenidos por el enfoque empírico del INS. Adicionalmente, la diversidad genética y los mecanismos evolutivos que modulan la diversidad observada fueron evaluados con diferentes programas bioinformáticos tales como DnaSP, HyPhy/Datamonkey, RDP, entre otros.

**Resultados:** El análisis filodinámico estimó los valores de  $R_0$  para Colombia (3.1), el cual fue ligeramente superior al obtenido por el INS. Adicionalmente, el  $R_e$  obtenido por este análisis mostró un panorama más preciso respecto a la dinámica del SARS-CoV-2, que el obtenido por el enfoque empírico. Se encontraron señales de selección positiva y de recombinación en genomas virales de Colombia, principalmente en aquellos recuperados del 2021.

**Conclusión:** Los resultados acá obtenidos indican que el uso de datos de secuencia y los análisis retrospectivos por métodos coalescentes y de evolución molecular son una fuente alternativa de información sobre las epidemias.

**PALABRAS CLAVE:** Virus, SARS-CoV-2, Análisis filodinámicos, modelo SIR, Colombia

# Identificación in silico de moléculas inhibidoras de transportadores de monocarboxilatos y su validación como blancos para la inhibición de la proliferación de células tumorales

Andrés Ibacache Chía (1,2), Jimena Sierralta (2), Andreas Schüller (1)

1: Pontificia Universidad Católica de Chile

2: Universidad de Chile

**Background:** El cáncer es la segunda causa de muerte a nivel mundial, alcanzando anualmente más de 10 millones de muertes. Las células tumorales basan su proliferación en lactato, que difunde a través de las membranas gracias a los transportadores de monocarboxilatos (MCTs). En humanos existen 14 MCTs, pero solo MCT1 y MCT4 transportan lactato en cáncer, convirtiéndose en atractivos blancos farmacológicos. El desarrollo de inhibidores de MCTs ha sido obstaculizado por la escasez de estructuras cristalográficas para estas proteínas, que en el caso de MCT4 han imposibilitado además el desarrollo de inhibidores selectivos y la identificación de residuos relevantes para su actividad. El modelamiento in silico de proteínas y el screening virtual de ligandos surgen como estrategias para identificar inhibidores de ambos transportadores. La evaluación in vitro de los hits usando líneas celulares de cáncer permitirá verificar el efecto de cada molécula.

**Objetivos:** (1) Generar un modelo tridimensional de MCT4, caracterizando los residuos potencialmente importantes para la inhibición de su actividad. (2) Identificar in silico moléculas inhibidoras de la actividad de ambos MCTs. (3) Determinar la capacidad inhibitoria de las moléculas sobre la proliferación y el transporte de lactato en líneas celulares de cáncer expresando MCT1 y MCT4.

**Metodología:** Basándonos en la estructura cristalográfica de MCT1 y mediante modelamiento por homología construimos la estructura de MCT4. Con un screening virtual identificamos candidatos que fueron validados con ensayos in vitro usando células transfectadas con un sensor de lactato (laconic).

**Resultados:** Algunos hits afectaron la proliferación de las líneas de cáncer SiHa y MDA, pero compuestos con scaffolds moleculares de cumarina fueron más versátiles inhibiendo el transporte de lactato.

**Conclusión:** Usando una plataforma de evaluación computacional e in vitro identificamos ligandos que disminuyen la viabilidad de líneas celulares de cáncer, que además afectaron el transporte de lactato mediado por estos MCTs.

**PALABRAS CLAVE:** Cáncer, Lactato, Transportadores de monocarboxilatos

# SUBCLASES DE IGG EN MONOS DEL NUEVO MUNDO, ¿UN TEMA DE DEBATE?

**Helairis Alfonso-González (1), Jehymin D. Parra-Montaño (1), Kimberly C. Mateus-Rincon (1), Juliana V. Aranguren-Borrás (1), Mary Medrano-Robayo (1), Alejandro Figueredo-López (1), Laura M. González-Amaya (1), Juan D. Vega-Valderrama (1), Luisa F. González-Bautista (1), Andrea L. Becerra-Embus (1), Yury Aponte-Rubio (1), Sindy P. Buitrago (2,3,4), Diego Garzón-Ospina (2,3,4)**

1: Estudiantes laboratorio de Genética I, Programa de Biología de la Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

2: PGAME - Population Genetics And Molecular Evolution, Fundación Scient

3: GEBIMOL, Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

4: GEO, Facultad de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia - UPTC

**Introducción:** La inmunoglobulina G (IgG) es un anticuerpo esencial durante la inmunidad adaptativa observándose una expansión diferencial del gen que codifica la IgG (IGHG) en mamíferos. En humanos y modelos biomédicos como ratones y macacos, se observan 4 genes, cada uno codificando una subclase. Sin embargo, los datos para los monos del nuevo mundo (NWM) parecen ser controvertidos. Algunas publicaciones defienden la existencia de un solo gen mientras, un artículo reciente sugiere la presencia de subclases en algunas especies. **Objetivo:** Evaluar la presencia de subclases de IgG en especies de NWM. **Materiales y métodos:** Las secuencias de genes IGHG de 13 especies de NWM recuperadas de datos genómicos y generadas por PCR-clonación disponibles en GenBank fueron descargadas y alineadas junto con secuencias de IGHG de Homo sapiens, Pan troglodytes, Pongo abelii, Macaca mulata, M. nemestrina, y Prolemur simus (grupo externo) usando el método MUSCLE. A partir del alineamiento se estimaron las distancias genéticas entre las secuencias y se construyeron árboles de máxima verosimilitud y bayesianos para identificar las relaciones filogenéticas. **Resultados:** Se observó un resultado desconcertante para las secuencias de Aotus nancymae y Cebus albifrons y para las especies del género Ateles y Alouatta. Mientras que los datos de las secuencias del genoma mostraron un valor de distancias genéticas superior a 0,07 entre estas especies, las secuencias generadas por PCR mostraron valores inferiores a 0,008. Los análisis filogenéticos mostraron que varias secuencias no se agruparon en el taxón esperado, probablemente debido a contaminación cruzada durante los procedimientos de laboratorio. Finalmente, varias secuencias reportadas como subclases mostraron ser 100% idénticas en los dominios inmunoglobulina y entre especies. **Conclusión:** Los datos aquí presentados sugieren que no hay pruebas suficientes para establecer la presencia de subclases de IgG en NWM y, por lo tanto, la IgG es codificada por un solo gen en estas especies.

**PALABRAS CLAVE:** IgG, Primates del nuevo mundo, Distancias genéticas, Relaciones filogenéticas

# Análisis de genomas de plásmidos actinobacterianos a través de redes bipartitas.

Haydé Saracho, Daniel G. Kurth

Planta Piloto de Procesos Microbiológicos Industriales (PROIMI) - CONICET

**Introducción:** Los plásmidos son moléculas de ADN extracromosómico circulares o lineales que se replican de forma autónoma. Se sabe que los plásmidos permiten la transferencia horizontal de genes (HGT) entre microorganismos de diferente especie y que esta es la fuerza fundamental que impulsa la evolución bacteriana. El uso de redes ayuda a dilucidar la dinámica de especies en la HGT.

**Objetivos:** Ampliar el conocimiento sobre plásmidos actinobacterianos mediante la comparación y el análisis de sus genomas.

**Materiales y métodos:** Para determinar los límites de HGT estudiamos la distribución del genoma codificado por plásmidos actinobacterianos con AcCNET, una herramienta bioinformática basada en redes bipartitas que permite el análisis simultáneo de genomas accesorios. Analizamos 833 plásmidos y construimos una red bipartita con dos tipos de nodos: Unidades Genómicas (GUs, plásmidos en nuestro caso) y Clusters de Proteínas Homólogas (HPC, proteínas codificadas por los plásmidos). Cuando el genoma de un plásmido codifica un miembro de HPC, ambos nodos se unen por un borde. Los archivos de salida fueron procesados con R, visualizados y analizados mediante el programa de visualización de gráficos Gephi.

**Resultados:** AcCNET extrajo el proteoma codificado por los plásmidos y los organizó en 60601 grupos de HPC con 4154 diferentes tipos de proteínas, siendo las más abundantes las proteínas de función desconocida. Los nodos de la red están conectados por 107333 bordes y se agrupan en grupos correspondientes a conjuntos de plásmidos con alta similitud general en la composición de los proteomas por el algoritmo ForceAtlas2 de Gephi.

**Conclusiones:** La aplicación del software AcCNET facilita el análisis de nuestro conjunto de datos. Los plásmidos actinobacterianos se organizan en grupos genotípicamente coherentes homogéneos y heterogéneos. La transmisión de plásmidos está restringida por límites taxonómicos, pero son lo suficientemente permeables como para sostener redes de intercambio de genes.

**PALABRAS CLAVE:** Genómica, Transcriptómica, Plásmidos y Redes.

## Diseño computacional de un tricomonocida

José Alberto-Santiago (1), Juan Carlos Gómez-Verjan (2), María Elizabeth Alvarez-Sánchez(1)

1: Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México.

2: Departamento de Ciencias Básicas, Dirección de Investigación, Instituto Nacional de Geriátría (INGER).

La tricomoniasis es la infección de transmisión sexual (ITS) no viral más común a nivel global, llegando a infectar a unos 276.4 millones de personas anualmente en todo el mundo. El agente etiológico causante de esta parasitosis es *Trichomonas vaginalis*. Actualmente, los nitroimidazoles como el metronidazol y tinidazol constituyen la clase de medicamentos que se administra con mayor frecuencia para tratar la tricomoniasis. Sin embargo, se ha observado que el parásito presenta resistencia a estos medicamentos, por tanto, es necesario buscar nuevos compuestos con actividad tricomonocida. El objetivo de este trabajo es realizar una aproximación al mundo del diseño de fármacos asistidos por computadora, realizando una búsqueda de moléculas pequeñas (< 500 Daltons) que puedan actuar como fármaco contra la tricomoniasis. Primeramente, realizamos una búsqueda de moléculas con potencial terapéutico contra *T. vaginalis*, usamos como blanco una metaloproteinasa, TvMP50, que es un factor de virulencia del parásito e involucrada en la citotoxicidad de este hacia la célula del hospedero. Posteriormente se desarrolló un screening de alto rendimiento basado en un análisis por descriptores moleculares; las moléculas resultantes se procesaron mediante un análisis de docking molecular para identificar aquellas moléculas que interactúan con el sitio activo de TvMP50. Finalmente, se generaron fragmentos de las moléculas más activas que fueron ensamblados en nuevas entidades moleculares, que se volvieron a evaluar contra la proteína blanco. De las más 34 moléculas iniciales encontradas en la literatura se logró filtrar seis moléculas candidatas que cumplen con reglas de Lipinski y tuvieron interacción con el sitio activo de TvMP50, las cuales fueron fragmentadas y ensambladas en quimeras, donde una de ellas presentó un acoplamiento más óptimo hacia TvMP50.

**Conclusión:** Mediante estos enfoques computacionales en el presente proyecto fuimos capaces de diseñar un compuesto que in silico pareciera tener una mayor especificidad contra TvMP50 de *T. vaginalis*, este compuesto tendrá que ser sometido a pruebas experimentales de laboratorio para conocer sus verdaderas propiedades ADME, antiparasitarias y toxicológicas.

**PALABRAS CLAVE:** *Trichomonas vaginalis*, TvMP50, Diseño de fármacos asistidos por computadora, Descriptores moleculares, Fragmentación de fármacos

# INTEGRACIÓN DE REALIDAD AUMENTADA COMO RECURSO EDUCATIVO EN LA BIOINFORMÁTICA

**Mario Bustillo, Liliana Ferrer, Silvina Videla, Gabriela Ohanian, Sergio Vardaro**

Facultad de Ingeniería, Universidad Nacional de Cuyo. Mendoza, Argentina

La Realidad Aumentada brinda a los estudiantes la posibilidad de ver una molécula desde todos sus ángulos, visualizar cómo se organizan los átomos en un elemento, hasta comprender conceptos químicos más abstractos.

Durante el proceso de enseñanza de la Química Orgánica, se comprobó que los estudiantes de ingeniería, ciencias químicas y ciencias biológicas presentan dificultad con ciertos temas básicos que son fundamentales para poder llevar la continuidad de la materia. Uno de los problemas detectados fue la complicada visualización o imaginación de las estructuras químicas de las moléculas químicas orgánicas, lo cual es fundamental para determinar propiedades físicas-químicas e información relevante sobre la molécula.

A partir de ello, en el marco de innovación y aplicación de nuevas tecnologías para la educación científica, en especial para el área de Química Orgánica, se desarrolla AUMENTED, plataforma educativa que permite la visualización de estructuras tridimensionales de moléculas químicas mediante el uso de realidad aumentada.

Dicha plataforma mejoró las condiciones de aprendizaje y enseñanza en esta área con métodos disruptivos y modernos para la identificación de la nomenclatura de compuestos orgánicos, determinación de propiedades físicas- químicas y sus respectivas fórmulas.

A través de la implementación de la Realidad Aumentada en la educación científica, se espera que esta tecnología pueda ser vista como una herramienta que promueva, facilite y enriquezca la comprensión de temas y conceptos propios de las ciencias, además de estar a la vanguardia de los procesos educativos. Se reporta una aceptación por parte de los estudiantes que fue incrementándose en los dos años de uso. Basado en los resultados, se propone considerar la incorporación de TIC en educación no sólo como un nuevo insumo, sino como un elemento de innovación disruptiva, es decir, que obliga al cambio de las prácticas educativas y, en definitiva, a un cambio importante de los sistemas escolares.

**PALABRAS CLAVE:** Realidad Aumentada, Química Orgánica, Educación Científica, Innovación Educativa, Innovación Disruptiva



# Explorando características cerebrales cuantitativas asociadas a una mayor predisposición genética a la enfermedad de Alzheimer mediante el análisis de datos composicionales

**Patricia Genius (1,2), Juan D. Gispert (2,3,4,5), Grégory Operto (2,3,6), Manel Esteller (7,8,9,10), Arcadi Navarro (1,2,11,12,13), Roderic Guigó (1,11), Malu Calle (14), Natalia Vilor-Tejedor (1,2,11,15)**

- 1: Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain.
- 2: BarcelonaBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain.
- 3: IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.
- 4: Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER- BBN), Madrid, Spain.
- 5: Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain.
- 6: Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES), Madrid, Spain.
- 7: Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain.
- 8: Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Spain.
- 9: Physiological Sciences Department, School of Medicine and HealthSciences, University of Barcelona (UB), Barcelona, Catalonia, Spain.
- 10: Centro de Investigación Biomédica en Red Cancer (CIBERONC), Madrid, Spain.
- 11: Pompeu Fabra University, Barcelona, Spain.
- 12: Institute of Evolutionary Biology (CSIC-UPF), Department of Experimental and Health Sciences, Universitat Pompeu Fabra, 08003, Barcelona, Spain.
- 13: Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
- 14: Biosciences Department, Faculty of Sciences and Technology, University of Vic-Central University of Catalonia, Carrer de la Laura, 13, 08500 Vic, Spain.
- 15: Erasmus MC University Medical Center, Department of Clinical Genetics, Rotterdam, Netherlands

**Introducción:** Los estudios de asociación genética en el campo de la neuroimagen tienen como objetivo analizar cómo la información genética influye en la estructura y función cerebral combinando características cerebrales y datos genéticos del mismo individuo. La mayoría de estudios se centran en métodos univariados que, en un contexto multidimensional, se traduce en una reducción del poder estadístico. Es por ello que proponemos una nueva aproximación para explorar la asociación entre la predisposición genética a la enfermedad de Alzheimer (EA) y estructuras cerebrales específicas a través del análisis de datos composicionales.

**Objetivo:** El objetivo principal del estudio es evaluar los determinantes genéticos que afectan a la variación volumétrica conjunta de las subregiones del hipocampo y ofrecer nuevas claves para mejorar el conocimiento sobre la EA.

**Material y métodos:** La muestra fue definida por 1071 participantes de mediana edad (45-75 años) cognitivamente sanos que forman parte del estudio ALFA, la mayoría de ellos descendientes de pacientes con Alzheimer. Todos ellos poseen información genética y datos de neuroimagen. Se utilizó la selección de balances (algoritmo Selbal) para evaluar el cambio volumétrico conjunto de los componentes cerebrales seleccionados (volúmenes de las subregiones del hipocampo) relacionados con una mayor predisposición genética a la EA.

**Resultados:** Los resultados mostraron una modulación volumétrica conjunta de subregiones específicas del hipocampo asociadas significativamente con una mayor predisposición genética a la EA. También se observaron diferencias entre mujeres y hombres en la modulación conjunta de estas estructuras.

**Conclusiones:** Este trabajo ofrece una perspectiva innovadora para la modelización conjunta de datos genéticos y de neuroimagen. En específico, enfatiza la necesidad de explorar el cerebro como una composición de estructuras, ofreciendo una aproximación que puede ser más cercana a los cambios volumétricos conjuntos que los individuos con un mayor riesgo genético a la EA pueden presentar.

**PALABRAS CLAVE:** Análisis multivariante; Análisis de datos composicionales; Algoritmo Selbal; Alzheimer; Neurogenética

# Firmas genéticas y funcionales de la lesión medular: una revisión sistemática y meta-análisis de estudios transcriptómicos en rata

Rubén Grillo-Risco (1), Marta Hidalgo-García (1), Beatriz Martínez-Rojas (2),  
Victoria Moreno-Manzano (2), Francisco García-García (1)

1: Bioinformatics & Biostatistics Unit, Centro de Investigación Príncipe Felipe, E-46012 Valencia, Spain.

2: Neuronal and Tissue Regeneration Laboratory, Centro de Investigación Príncipe Felipe, E-46012 Valencia, Spain.

La lesión medular (LM) es una patología con aproximadamente 11000 nuevos casos cada año tan solo en los Estados Unidos. Los daños en la médula espinal, sean cual sea su origen, desencadenan una cascada de eventos patofisiológicos que culminan en disfunciones del sistema motor, sensorial y autónomo, con consecuencias devastadoras para la calidad de vida de los pacientes. Actualmente no existe un tratamiento efectivo para la LM debido a la gran complejidad de los eventos fisiopatológicos inducidos tras la lesión, modulados por múltiples tipos celulares, rutas de señalización y genes. Para presentar nuevos tratamientos que permitan una recuperación funcional es importante identificar los mecanismos moleculares que subyacen a cada una de las etapas de la LM, y que todavía no son completamente conocidos.

Por ello, el objetivo de este trabajo es la caracterización de las diferentes fases y severidades de la LM a nivel de gen y función e identificación de biomarcadores. La estrategia empleada se compone de los siguientes pasos: 1) revisión sistemática y selección de estudios transcriptómicos; 2) análisis individuales de los estudios, que consta de pre-procesamiento, análisis exploratorio y análisis de expresión diferencial; 3) metaanálisis a nivel de gen; 4) análisis de enriquecimiento funcional; 5) identificación de biomarcadores.

En la revisión sistemática se seleccionaron un total 14 estudios y 273 muestras que cumplieran los criterios de inclusión. Cada muestra fue clasificada en base a la severidad de la lesión (daño moderado, severo), la fase (fase aguda subaguda, intermedia o crónica) o controles (sin lesión), dando lugar a 9 grupos experimentales. El metaanálisis de genes y el análisis funcional arrojaron un gran número de resultados significativos. Para facilitar a la comunidad consultar nuestros resultados y recopilar información se ha construido una herramienta web que puede consultarse en <https://rubbmetasci.shinyapps.io/metaSCI/>.

**PALABRAS CLAVE:** Spinal cord Injury, Transcriptomics, Meta-analysis, Biomarkers, Functional analysis

# Long non-coding RNAs as outcome predictors in T-cell pediatric Acute Lymphoblastic Leukemia

Unai Illarregi (1), Anastasija Jakjimovska (1), Ángela Gutierrez-Camino (2),  
Idoia Martin-Guerrero (1), Elixabet Lopez-Lopez (3), Nerea Bilbao-Aldaiturriaga (1).

1: Department of Genetics, Physical Anthropology & Animal Physiology, Science and Technology Faculty, IIS Biocruces Bizkaia, UPV/EHU, Leioa, Spain;

2: Division of Hematology-Oncology, Research Center, Sainte-Justine University Health Center, Montreal, Canada;

3: Department of Biochemistry & Molecular Biology, Science and Technology Faculty, IIS Biocruces Bizkaia, UPV/EHU, Leioa, Spain.

**INTRODUCTION:** T-cell Acute Lymphoblastic Leukemia (T-ALL) remains an important cause of death from disease in children. Therefore, the identification of new outcome predictors and therapeutic targets is needed. In this context, long non-coding RNAs (lncRNAs), an important fraction of the functional genome poorly studied to date, could be novel candidates with great potential. **OBJECTIVE:** We aimed to identify new biomarkers of T-ALL relapse, analyzing expression differences in lncRNAs through a high throughput RNA-seq approach.

**MATERIAL AND METHODS:** Total RNA of 9 pediatric T-ALL patients treated following the same protocol in two Spanish hospitals was extracted from tumor samples at diagnosis. RNA libraries were generated with TruSeq Stranded Total RNA Library Prep Kit and sequenced on NovaSeq 6000 System (Illumina), with a depth of ~150 million paired-reads. Alignment to hg19 human genome was performed with HISAT2, and featureCounts was used to quantify all lncRNAs annotated in LNCipedia v5.2. Differential expression analysis was performed using DESeq2, and Kaplan-Meier method was used to perform 5-year event-free survival (EFS) analyses. Patients were divided into three groups for survival analyses: “High expression”, “Low expression”, and “No expression”.

**RESULTS AND DISCUSSION:** From 56,846 lncRNAs annotated in LNCipedia, 7 were differentially expressed ( $p_{adj} < 0.01$ ;  $\log_2FC > 2 / < -2$ ) comparing relapsed ( $n=3$ ) vs. non-relapsed patients ( $n=6$ ). From these 7 lncRNAs, 6 showed significant differences ( $p < 0.05$ ) in Kaplan-Meier survival analyses. Notably, lower expression of all these lncRNAs was associated to worst survival, except for lnc-ODF2L-1, the one with the most statistically significant result ( $p=0.0033$ ). These results point to an important role of lncRNA expression in treatment response, and warrant further downstream studies to understand their role in the pathogenesis of the disease and as potential new therapeutic targets.

**CONCLUSION:** Our high throughput approach identified 6 lncRNAs associated with relapse and EFS, suggesting that lncRNA expression pattern could serve as prognostic predictor in pediatric T-ALL.

**PALABRAS CLAVE:** RNA-seq; T-ALL; lncRNA; Survival; Leukemia

# Métodos de imputación de la variable sexo a partir de los niveles de expresión en genes

Eva Maicas Alcaine, Marta R. Hidalgo, Francisco García-García

Unidad de Bioinformática y Bioestadística. Centro de Investigación Príncipe Felipe (CIPF), Valencia, 46012, España

Se han descrito más 600 enfermedades humanas donde existen diferencias por sexo, en su frecuencia, evolución y tratamiento. Las estrategias bioinformáticas basadas en el reanálisis de datos transcriptómicos disponibles en los repositorios públicos, constituyen una potente herramienta para generar nuevo conocimiento que permita una mejor caracterización de estas diferencias. Desafortunadamente, en numerosos estudios biomédicos no se especifica el sexo de los sujetos o animales que forman parte del experimento, lo que supone un importante problema en el desarrollo del proyecto, puesto que se reduce el número de datos disponibles de alguno de los sexos, o incluso se desconoce esta información en las muestras. Así pues, se establece como objetivo del trabajo suplir la falta de información acerca del sexo de los individuos, proporcionando una herramienta que permita imputar la variable sexo a partir de los niveles de expresión de ciertos genes incluidos en los cromosomas sexuales. Para ello, se han evaluado los diferentes métodos de imputación desarrollados hasta el momento y que están disponibles en los paquetes de R: MetaIntegrator y massiR. Su aplicación y comparación se ha llevado a cabo en dos grupos de estudios: 1) una serie sobre accidentes cerebrovasculares (ACV) y 2) otra serie sobre adenocarcinoma de pulmón (NSCLC). El siguiente paso incluye el diseño y la implementación de una aplicación web en R con el paquete shiny, que permita predecir el sexo de las muestras pertenecientes a un estudio o experimento.

Paralelamente, también se ha estudiado la posibilidad de predecir el sexo a partir de los niveles de expresión de microRNAs, haciendo uso de estudios sobre melanoma. Dado que hasta el momento no se dispone de procedimientos que lleven a cabo esta imputación, el primer paso ha sido identificar microRNAs específicos del sexo mediante un análisis integrativo de los resultados de la expresión diferencial en el conjunto de estudios evaluados.

El desarrollo de nuevos métodos y herramientas que faciliten la imputación de la variable sexo supondrá una importante mejora en el input empleado en numerosos estudios que tienen como objetivo la caracterización de las diferencias de sexo en enfermedades humanas mediante abordajes computacionales.

**PALABRAS CLAVE:** Transcriptómica, Genes, microRNAs, Diferencias sexo, Imputación

# Uso de bases de datos en la complementación del estudio de las porfirias y la medicina personalizada

Priscila Pagnotta (1, 2), María Laura Buscalia (3), Johanna Zuccoli (3), Viviana Melito (2, 3), Victoria Parera (3), Ana María Buzaleh (2, 3)

1: Instituto de Biología y Medicina Experimental (IBYME)

2: Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales (FCEN, UBA);

3: Centro de Investigaciones sobre Porfirinas y Porfirias (CIPYP)

**Introducción:** Las Porfirias son enfermedades metabólicas producidas por alteraciones en la biosíntesis del Hemo. La Porfiria Cutánea Tardía (PCT), debido a alteración de la Uroporfirinógeno decarboxilasa, puede ser hereditaria o adquirida, existiendo elevada asociación PCT-VIH. La Porfiria Aguda Intermitente (PAI), causada por una Porfobilinógeno deaminasa deficiente, se hereda con carácter autosómico dominante, siendo la mutación no suficiente para su manifestación.

**Objetivos:** Utilizar bases de datos para complementar el estudio de Porfirias en relación con portación de variantes genéticas.

**Materiales y métodos:** Bases de Datos: 1000Genomes ([browser.1000genomes.org](http://browser.1000genomes.org)), Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>), Scielo (<https://scielo.org/es/>), PharmGKB (<https://www.pharmgkb.org/>), Gen Expression Omnibus (<https://www.ncbi.nlm.nih.gov/gds>, GSE44228), UniProt (<https://www.uniprot.org/>), Ensembl (<http://ensembl.org/>), GenBank ([www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)). Se realizó un metaanálisis (PRISMA) comparando grupos control. Se analizó toxicidad de tratamientos para VIH y drogas contraindicadas en PAI en relación a variantes genéticas.

**Resultados :** Datos experimentales del grupo control coincidieron con 1000Genomes y el meta análisis. Algunas terapias para VIH aumentan la probabilidad de toxicidad en relación a variantes genéticas: Efavirenz y Nelfinavir (TT, c.3435C>T, *ABCB1*), Atazanavir (TT, c.2677G>T/A, *ABCB1*), Nevirapina (CC, c.3435C>T, *ABCB1*; Presencia, *GSTM1*). Individuos tratados con inhibidores de proteasas comparados con inhibidores no nucleósidos de transcriptasa reversa, presentaron: Menor expresión de *ABCB1* (FC=0,83; p adj<0,05) y expresión diferencial de 19 transportadores ABC (68,4% sobreexpresados), 25 CYPs (72% sobreexpresados) y 13 GSTs (53,8% subexpresados). Se encontraron casos de toxicidad asociada a drogas contraindicadas para PAI y variantes genéticas: *ABCB1*: rs1045642 (12), rs1128503 (2) y rs1128503 (2); *GSTT1* nulo (3); *GSTM1* nulo (8); *GSTP1*: rs1695 (7).

**Conclusión:** Mediante bioinformática se corroboraron datos experimentales y realizaron búsquedas complementarias. Los resultados aumentaron las evidencias sobre el rol de variantes genéticas en la asociación PCT-VIH y la manifestación de la PAI. Se demostró la relevancia de la expresión genética. Esto refuerza la importancia de avanzar hacia la medicina personalizada considerando el contexto genético y metabólico de cada individuo.

**PALABRAS CLAVE:** Porfirias, Medicina personalizada, Variantes genéticas, Bases de datos

# Uso de dispositivos en un modelo bioinformático para la atracción de *Macrodactylus mexicanus* (Coleoptera: Melolonthidae)

Salgado-Farias Fernanda (1), Contreras-González Carolina (2), Romero-López Angel Alonso (1)

1: Maestría en Ciencias Biológicas, Benemérita Universidad Autónoma de Puebla, Ciudad Universitaria, Boulevard Valsequillo y Av. San Claudio, Edificio 119A, Col. Jardines de San Manuel, Puebla, México, C.P. 72570.

2: Dinámica no lineal y Sistemas Complejos, S. Lorenzo 290, Col. Del Valle Sur, Benito Juárez 03100 Ciudad de México, CDMX, Universidad Autónoma de la Ciudad de México.

Se creó un modelo bioinformático en el software NetLogo, en el cual se representa virtualmente el empleo de dispositivos de atracción y patrones de comportamiento relacionados con la respuesta de adultos de los coleópteros *Macrodactylus mexicanus* Burmeister, hacia atrayentes sexuales. Con base en una revisión bibliográfica exhaustiva, se establecieron los parámetros requeridos para el diseño y construcción del modelo, como es el caso de la planta hospedera *Quercus grahamii* (Benth), la cantidad de machos y hembras de este insecto, además del número de dispositivos de atracción por incluir. Se consideraron 30 machos y 30 hembras, así como cinco dispositivos de atracción diseñados con base en Arakaki et al. (2003), conteniendo los principales componentes del atrayente sexual (ácido n-hexanoico y ácido oleico). La simulación virtual mostró una disminución constante en el número de machos de *M. mexicanus* al utilizarse los dispositivos de atracción, en contraste con los datos que se proyectan con los machos de esta especie en la naturaleza. En el caso de las hembras, el modelo resultó en mínima atracción hacia los dispositivos y por ende, un número menor de individuos con respecto a los machos, sobre todo cuando ya habían transcurrido alrededor de 30 días. Se sugiere que este tipo de modelado basado en agentes es una herramienta idónea para predecir el efecto de dispositivos de atracción preparados con atrayentes sexuales sobre el comportamiento de ambos sexos de *M. mexicanus*, reflejándose en actividades de monitoreo más pertinentes.

**PALABRAS CLAVE:** Modelo bioinformático, Dispositivos de atracción, *Macrodactylus mexicanus*, Atrayentes sexuales, Monitoreo

# Predicción y selección de candidatos vacunales contra la coccidiosis aviar mediante herramientas bioinformáticas

Britez, J.D. (1, 2), Rodriguez, A.E. (1), Tomazic, M.L. (1, 3)

1: Instituto de Patobiología Veterinaria, INTA-CONICET. Hurlingham, Buenos Aires, Argentina.

2: Universidad de Buenos Aires. Facultad de Ciencias Veterinarias. Cátedra de Microbiología. Ciudad Autónoma de Buenos Aires, Argentina

3: Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Cátedra de Biotecnología, Ciudad Autónoma de Buenos Aires, Argentina

La coccidiosis es una parasitosis de alto impacto económico en la producción avícola. Es causada por parásitos del género *Eimeria*. Existen siete especies aviares y generan inmunidad especie-específica. Actualmente, se emplean vacunas vivas que poseen ciertas desventajas. Las vacunas recombinantes están en desarrollo y aún no están comercialmente disponibles. Para ello, la búsqueda de nuevos antígenos es fundamental. Se demostró experimentalmente que las proteínas con anclaje a glicosilfosfatidilinositol (GPI) son buenas candidatas vacunales ya que generan respuesta inmunogénica. El objetivo de este trabajo fue predecir y seleccionar potenciales candidatos vacunales de dos de las especies más frecuentes *E. tenella* y *E. acervulina* a través de una estrategia secuencial bioinformática para el futuro desarrollo de una vacuna recombinante. Se comenzó con la predicción del secretoma, mediante Signal-P. Se usaron 4 programas de predicción de la señal de anclaje a GPI. Se descartaron proteínas con más de un dominio transmembrana o con localización subcelular. Además, se usó UniProt, ToxoDB, BLAST y NetMHCpan para ampliar el análisis. Para el estudio de conservación se comenzó con 2 genes de *E. tenella*, *Etsag1* y *Etsagta4*. Se diseñaron oligonucleótidos específicos, se amplificaron los genes de 2 aislamientos locales y una cepa de referencia propagada experimentalmente, y se secuenciaron. Las secuencias se editaron con el programa BioEdit y se alinearon mediante Clustal Omega regiones conservadas. De 6864 proteínas de *E. acervulina* y 8627 de *E. tenella* se preseleccionaron 11 y 23 respectivamente, con anclaje a GPI, localización superficial, transcripción en el primer estadio infectivo asexual, rol en la invasión y sitios de unión a MHC I. *Etsag1* y *Etsagta4* demostraron 100 % y 99.8 % de conservación respectivamente, hallando una mutación con cambio de sentido en *Etsagta4*. Concluyendo, se logró predecir y seleccionar proteínas con características relevantes para considerarse como candidatos vacunales.

**PALABRAS CLAVE:** Coccidiosis, Proteínas-GPI, Candidatos vacunales, Conservancia

# Impact of protein conformational diversity on AlphaFold predictions

**Tadeo Saldaño (1, 3), Nahuel Escobedo (1, 2), Julia Marchetti (1, 2), Juan Mac Donagh (1, 2), Nicolás Palopoli (1, 2), Gustavo Parisi (1, 2)**

1: National University of Quilmes

2: National Scientific and Technical Research Council (CONICET)

3: Ministry of Science, Technology and Innovation (MinCyT)

After the outstanding breakthrough of AlphaFold in predicting protein 3D models, new questions appeared and remain unanswered. The ensemble nature of proteins, for example, challenges the structural prediction methods because the models should represent a set of conformers instead of single structures. The evolutionary and structural features captured by effective deep learning techniques may unveil the information to generate several diverse conformations from a single sequence. Here we address the performance of AlphaFold2 predictions obtained through ColabFold under this ensemble paradigm.

Using a curated collection of apo-holo pairs of conformers, we found that AlphaFold2 predicts the holo form of a protein in ~70% of the cases, being unable to reproduce the observed conformational diversity with the same error for both conformers. More importantly, we found that AlphaFold2's performance worsens with the increasing conformational diversity of the studied protein. This impairment is related to the heterogeneity in the degree of conformational diversity found between different members of the homologous family of the protein under study. Finally, we found that main-chain flexibility associated with apo-holo pairs of conformers negatively correlates with the predicted local model quality score pLDDT, indicating that pLDDT values in a single 3D model could be used to infer local conformational changes linked to ligand binding transitions.

**PALABRAS CLAVE:** AlphaFold predictions, Conformational Diversity



# Representación por Docking molecular del efecto de volátiles de hojas de *Acaciella angustissima* (Fabaceae) en adultos de *Macrodactylus mexicanus* (Coleoptera: Melolonthidae)

Pages-Paez N.Y. (1), Carrasco-Carballo A. (2), Romero-López A.A. (1)

1: Facultad de Ciencias Biológicas, Benemérita Universidad Autónoma de Puebla, Boulevard Valsequillo y Av. San Claudio, Col. Jardines de San Manuel, Puebla, México, C.P. 72570.

2: Laboratorio de Elucidación y Síntesis en Química Orgánica, Benemérita Universidad Autónoma de Puebla, Boulevard Valsequillo y Av. San Claudio, Col. Jardines de San Manuel, Puebla, México, C.P. 72570.

Los compuestos orgánicos volátiles (COVs) son sustancias producidas a partir del metabolismo secundario de las plantas y en el contexto donde coexisten árboles de *Acaciella angustissima* (Mill.) y adultos de *Macrodactylus mexicanus* Burmeister, con base en registros de comportamiento previos se ha evidenciado que la interacción que se establece entre ambas especies está mediada por COVs. Para complementar estos registros con evidencias basadas en técnicas de acoplamiento molecular, en el presente estudio se propuso el uso del Docking para modelar y simular la actividad biológica entre ligandos (COVs) y proteínas de unión a olores (OBPs, por sus siglas en inglés) que forman parte de los quimiorreceptores antenales de los insectos. Primero, se extrajeron los COVs por medio del sistema Clevenger modificado, para posteriormente generar el perfil químico correspondiente por cromatografía de gases y espectrometría de masas. Enseguida, se seleccionaron cuatro OBPs (1OW4, 3OGN, 3Q8I y 2ERB) de la base de datos RSCB PDB y un control ( $\beta$ -ionona) con la afinidad óptima para las cuatro proteínas. Por último, se prepararon los ligandos y proteínas con el software AutoDockTools y se realizó el acoplamiento con AutoDockVina. En el cromatograma se identificaron el  $\beta$ -mirceno, el  $\alpha$ -pineno y grasas bajas de origen vegetal inmersas en el volátil líquido. Con relación al Docking, se representaron en diagramas 2D interacciones de naturaleza hidrófoba y con residuos de aminoácidos, además de una mayor energía de acoplamiento del  $\beta$ -mirceno respecto al  $\alpha$ -pineno en las cuatro OBPs. El Docking nos permitió tener una aproximación de cómo se lleva a cabo la recepción de los COVs, con qué aminoácidos interactúa, el sitio de acople al que llegan los COVs y observando sus niveles de energía, indicando valores altos en términos de acoplamiento que pueden traducirse a una atracción en esta interacción.

**PALABRAS CLAVE:** Docking molecular, COVs, *Acaciella angustissima*, *Macrodactylus mexicanus*, OBPs

# Variantes de un solo nucleótido asociados a neuropatía diabética en población latinoamericana

**Karina Elizabeth Ruiz Segura (1), Michelle Guadalupe Casillas Gallegos (1), Juan Carlos Ávila Gutiérrez (1), Alejandro Bernal Rivera (1), Antonio de Jesús Casillas Navarro (1), Dulce Méndez (1), Laura Yareni Zuñiga (1), María Guadalupe Sánchez Parada (1), Arieh Roldan Mercado Sesma (1), Raúl Cuauhtemoc Baptista Rosas (1, 2).**

1: Centro de Investigación Multidisciplinaria en Salud, Centro Universitario de Tonalá, Universidad de Guadalajara  
2: Jal. Hospital General de Occidente, Secretaría de Salud Jalisco

**Introducción:** Se estima una prevalencia mundial de diabetes mellitus (DM) de 425 millones, donde México ocupa el 4to lugar. De cada 10 adultos con DM entre 1-6 desarrollará neuropatía diabética (ND). El riesgo de amputación es 20 veces mayor en ND. La neuropatía es una complicación crónica que afecta nervios sensitivos periféricos de las extremidades, principalmente inferiores, presentándose en 75% de los sujetos, impactando en la economía y calidad de vida.

**Objetivo:** Identificar variantes genómicas asociadas a ND y su frecuencia en población mexicana.

**Material y métodos:** Las variantes asociadas a ND se obtuvieron de los artículos de Politi et al. 2016, Prabodha et al. 2018 y Jankovic 2021. Se utilizaron: la base de datos BioMe AMP T2D GWAS, para identificar la frecuencia de polimorfismos en individuos con ND y la base de datos GWAS SIGMA, para obtener frecuencia de polimorfismos en mexicanos; criterios de búsqueda de variantes genéticas, Type 2 Diabetes como fenotipo,  $P < 0.05$  y radio de probabilidades  $< 1$ . Se exploró la frecuencia alélica del efecto y frecuencia del genotipo en heterocigotos y homocigotos de mexicanos (base de datos 1000 Genomes, Ensembl).

**Resultados:** Se identificaron 61 variantes en la base de datos según los criterios establecidos; 23 polimorfismos se encuentran en población Mexicana.

**Conclusiones:** Las muestras de BioMe AMP T2D GWAS se incluyen en la base de datos "DIAMANTE (European) T2D GWAS", compuesta por 13,034 individuos, estadounidenses mixtos (41.5%), afroamericanos (38%), europeos (19.8%) y asiáticos (0.006%). La base de datos GWAS SIGMA, Proyecto de la Iniciativa Slim en Medicina Genómica para las Américas, base genómica de DM en mexicanos y latinoamericanos, cuenta con 3,848 individuos con DM y 4,366 individuos controles. En prospectiva, se explorarán cambios identificados en proteínas e importancia clínica, buscando asociaciones (bases de datos especializadas, ClinVar y OMIM).

**PALABRAS CLAVE:** Neuropatía diabética, Polimorfismos, BioMe AMP T2D GWAS, GWAS SIGMA, Frecuencia

# **Análisis de CGH y del transcriptoma revela genes relacionados con supervivencia en leucemia linfoblástica aguda**

**Jessica Alejandra Zapata-García (1, 2), Luis Felipe Jave-Suárez (2), Adriana Aguilar-Lemarroy (2)**

1: Programa de Doctorado en Ciencias Biomédicas, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara  
2: División de Inmunología, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco

La inestabilidad cromosómica es una de las principales causas del desarrollo de la leucemia; el estudio de estas anomalías ha permitido identificar genes que actualmente son útiles para el diagnóstico, pronóstico o terapia dirigida; es por esto que el objetivo de este trabajo fue identificar regiones de pérdida o ganancia cromosómica frecuentes en muestras derivadas de médula ósea de individuos con leucemia linfoblástica aguda vírgenes de tratamiento y en las líneas celulares JURKAT y CEM, así como evaluar la expresión a nivel del transcriptoma de los genes presentes en las regiones de pérdida o ganancia establecidas usando RNA seq y correlacionar los mismos con la supervivencia global o libre de eventos. Para lo anterior se realizó hibridación genómica comparativa (CGH) y se calcularon los valores de expresión del genoma, además se usaron secuenciaciones derivadas del proyecto TARGET fase II para realizar el análisis a nivel del transcriptoma y para la elaboración de las curvas de supervivencia global y libre de eventos se usó la herramienta en línea Survival Genie. Se encontraron cromosomas que mantuvieron zonas de pérdida (2 y 9) o ganancia (14,17 y 22) comunes en seis o más muestras. En total se identificaron 65 genes presentes en las zonas de pérdida y 22 en las zonas de ganancia y se estableció que la baja expresión de los genes: *IGKV3-20*, *IFNB1*, *SNORD39*, *IFNA16*, *KLHL9* y la alta expresión de: *RAB2B*, *JUP* y *COMT*, se correlacionan con una supervivencia global y supervivencia libre de eventos pobre en individuos con ALL. En conclusión, los hallazgos de este trabajo brindan información valiosa acerca de ocho nuevos genes que podrían ser útiles en la clínica debido a su relación con el mal pronóstico para los individuos con LLA; además nuestros resultados constituyen un fértil campo de investigación; ya que algunos de los genes identificados aún no han sido estudiados en el contexto del cáncer o LLA.

**PALABRAS CLAVE:** RNA seq, CGH, Leucemia Linfoblástica Aguda, Supervivencia global, Supervivencia libre de eventos

# Resolución de un transcriptoma de referencia en *Setaria italica* para análisis de splicing alternativo en condiciones de estrés hídrico

Mauro Rey (1), Luisa Bermúdez (2), Magdalena Rossi (3), Nicolas Bellora (1)

1: Instituto de Energías Nucleares para la Salud, INTECNUS-CONICET, Bariloche, Argentina

2: IABIMO (Instituto de Agrobiotecnología y Biología Molecular), CICVyA - INTA; CONICET, Argentina.

3: Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Brasil.

**Marco Teórico:** Los sistemas agrícolas se encuentran cada vez más amenazados a cambios climáticos extremos como sequías, afectando el rendimiento del cultivo. Las plantas superiores poseen una alta elasticidad en su desarrollo y crecimiento, permitiéndose adaptar a diversas condiciones ambientales, maximizando así sus posibilidades de supervivencia y reproducción. El splicing alternativo, que permite generar diferentes isoformas por gen, es un importante modulador en la expresión génica aumentando la diversidad del proteoma y generando isoformas no-codificantes con potencial regulatorio.

**Objetivos:** Se buscó definir un transcriptoma *de novo* para *S. Italica* que refleje la complejidad estructural por locus. Las anotaciones existentes no consideraban la diversidad ni la cantidad de genes e isoformas de *S. Italica* para ser utilizadas a posteriori en análisis de eventos de splicing alternativo.

**Materiales y métodos:** Se cultivaron muestras de *S. Italica* bajo tres condiciones de estrés hídrico, control, moderado y severo. Se extrajo el ARN total con eliminación del ARN ribosomal y se armaron bibliotecas pair-end, obteniendo así secuenciaciones de alta calidad. Estas secuenciaciones fueron mapeadas con STAR al genoma de referencia y anotaciones de *S. Italica* obtenidas en Phytozome. A los archivos BAM obtenidos se los corrió en Scallop2 para el primer ensamblado de transcriptomas por condición, que luego fueron procesados con GffCompare para generar un único transcriptoma. Finalmente se realizó una última corrida con GffCompare utilizando este último transcriptoma y el transcriptoma original de Phytozome para evitar pérdida de isoformas que no se hayan expresado en nuestras condiciones.

**Resultados y Conclusión:** La comparación del transcriptoma obtenido y las anotaciones previas revelan un aumento considerable de 9628 (12.8%) en genes no anotados, además de duplicar las variantes de splicing asociadas: +59085 (137%) isoformas totales y de +1.12 (90%) isoformas por gen. El transcriptoma obtenido será utilizado para inferir el splicing alternativo en condiciones de estrés hídrico.

**PALABRAS CLAVE:** *Setaria*, Transcriptoma, RNA-seq, Estrés hídrico

# Diseño *in silico* de inhibidores de PtpA y PtpB como agentes terapéuticos frente a *Mycobacterium tuberculosis*

Cristhian N. Rodríguez-Silva (1), Terecita E. Carrión Z (1), Andrés Reyes-Chaparro (2),  
Andrés Alvarado-Salazar (3)

1: Facultad de Farmacia y Bioquímica, Unidad de Posgrado en Farmacia y Bioquímica, Universidad Nacional de Trujillo, Perú.

2: Centro de Física Aplicada y Tecnología Avanzada, Universidad Nacional Autónoma de México, México.

3: Área Farmacéutica, Química Farmacéutico Biológica, Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México, México.

La tuberculosis es una de las epidemias con más muertes en el mundo, por ello es necesario sea atendida con nuevas estrategias farmacológicas de mayor eficacia y seguridad. Asimismo, las proteínas tirosina fosfatasa de *Mycobacterium tuberculosis* PtpA y PtpB representan un factor de virulencia al disrumpir la desfosforilación-fosforilación en los macrófagos del huésped evadiendo la respuesta inmunitaria; representando un potencial blanco terapéutico. Por lo cual, surge la necesidad de investigar nuevos tratamientos antituberculosos, en este sentido, se han reportado derivados de chalconas y moléculas con núcleo de benzofuroxano con actividad antituberculosa. Nuestro objetivo fue diseñar y evaluar *in silico* compuestos híbridos de benzofuroxano y chalconas con potencial actividad sobre las proteínas PtpA y PtpB de *Mycobacterium tuberculosis*. Se realizó un cribado virtual con diversas herramientas quimioinformáticas como Way2Drug, molinspiration, OSIRIS Property Explorer y pkCSM-pharmacokinetics, para obtener una quimioteca 250 moléculas bioisostérica con alta eficacia y baja o nula toxicidad, así mismo obtener la predicción de la actividad biológica, parámetros fisicoquímicos, de toxicidad y farmacocinéticos, respectivamente. En la última etapa se realizó el estudio de acoplamiento molecular con el software Molegro Virtual Docker 6.0 de las 4 mejores moléculas híbridas, así como, de los medicamentos isoniazida y etambutol en el sitio catalítico de las proteínas PtpA y PtpB de *M. tuberculosis*. Las moléculas BC-11, 12, 24 y 236 presentaron mejores energías de afinidad en comparación con los fármacos de referencia isoniazida y etambutol en ambas proteínas, PtpA (-6.7, -6.6, -6.5, -6.5, -5.8 y -6.4, respectivamente) y PtpB (-8.1, -8.1, -8.1, -8.2, -6.3 y -7.4, respectivamente). Por medio de un estudio quimioinformático basado en tamizaje virtual y docking molecular se obtuvieron 4 moléculas de benzofuroxano-chalconas con óptimas características fisicoquímicas, farmacocinéticas, farmacológicas, de nula toxicidad, por lo cual, es posible que dichas moléculas presenten una adecuada actividad terapéutica anti-*M. tuberculosis*.

**PALABRAS CLAVE:** Antituberculosos, Diseño de fármacos, Chalconas, Benzofuroxanos, Quimioinformática

# Clasificador de imágenes con patrones de Anticuerpos Antinucleares (ANA)

Vladimir Alfaro

Universidad Simón Bolívar

**Introducción:** Los anticuerpos antinucleares (ANA) son un tipo de auto-anticuerpo que puede generar patrones específicos y que ayudan al diagnóstico y seguimiento de enfermedades autoinmunes. Su clasificación es una tarea subjetiva que depende de la opinión de expertos en el área, es por esto que se propone la creación de un modelo de aprendizaje profundo para poder clasificar dichos patrones en diferentes niveles.

**Marco teórico:** Actualmente hay un consenso mundial que acuerda la existencia de hasta veintinueve patrones para las imágenes ANA. Estos patrones están divididos en 3 grandes grupos; Nuclear, Citoplasmático y Mitótico. Estos grupos tienen subcategorías que se encuentran documentadas en la página oficial del consenso internacional para patrones ANA (<https://www.anapatterns.org>). En el área del procesamiento automático de imágenes se ha vuelto tendencia el uso de redes neuronales convolucionales, ya que su eficiencia ha quedado más que demostrada en muchos trabajos previos, es por esto que se usaron estos modelos para crear un clasificador de imágenes de patrones ANA que sea capaz de identificar los 3 grandes grupos.

**Objetivos:** Crear un modelo capaz de clasificar patrones ANA en los tres grandes grupos principales.

**Métodos:** El desarrollo de la red se hizo en Google Colab para hacer uso de su servicio de GPU y acelerar el entrenamiento de la red. Para entrenar la red se usó una base de datos pública usada en trabajos previos.

**Resultados:** Por los momentos no se tienen resultados satisfactorios para el conjunto de pruebas, por lo que se pretende implementar distintos tipos de regularización ya que se tiene un problema de sobre-entrenamiento de la red. Precisión = 71%. Puntuación F1 = 76%

**Conclusión:** Se creó un modelo basado en redes neuronales convolucionales que es capaz de clasificar patrones ANA en sus tres grandes grupos de forma automática con una precisión hasta del 71%.

**PALABRAS CLAVE:** Anticuerpos Antinucleares, Aprendizaje Automático, Aprendizaje Profundo, Redes Neuronales Convolucionales

# Identificación de polimorfismos de nucleótido simple en cuyes (*Cavia porcellus* L) de las zonas altoandinas del Perú a través de genotipado basado en secuenciamiento (GBS)

María Victoria Borja Lozano (1, 3), Bianca Vigil Santillan (1), Manuel J. More Montoya (2), Jonathan A. Morón Barraza (3), Aura Liz García-Serquén (1), Gustavo Gutiérrez Reynoso (1, 3), Claudia E. Yalta-Macedo (1)

1: Laboratorio de Biología Molecular y Genómica, Dirección de Recursos Genéticos y Biotecnología, Instituto Nacional de Innovación Agraria (INIA). Av. La Molina 1981, 15024 Lima, Perú.

2: Facultad de Ciencias Agrarias, Universidad Nacional de San Antonio Abad del Cusco. Av. de la Cultura 733, Cusco, Perú.

3: Instituto de Bioquímica y Biología Molecular y Facultad de Zootecnia, Universidad Nacional Agraria La Molina (UNALM), Av. La Molina s/n, 15024 Lima, Perú

El cuy es un mamífero nativo sudamericano, que tiene gran capacidad de adaptación a diversas condiciones climáticas y de crianza para el consumo humano por su fuente proteica. Por ello, su caracterización genómica es importante para su manejo y conservación. Las tecnologías de secuenciamiento de alto rendimiento han sido utilizadas en especies nativas del Perú como en alpacas, pero a la fecha no se ha empleado en el cuy nativo. El objetivo del presente trabajo es identificar polimorfismos de nucleótido simple (SNPs) para estudiar la diversidad genómica de cuyes. Se extrajo ADN genómico de 37 muestras de pelos de cuyes de zonas altoandinas peruanas y 3 de razas, las cuales fueron digeridas usando *NotI* III. El secuenciamiento (PE 150 pb) fue realizado en un Illumina®HiSeq. El filtrado de lecturas y el alineamiento al genoma de referencia *Cavia porcellus* (Cavpor3) se realizó con los programas Trimmomatic v0.32 y BWA v0.7 respectivamente. La identificación de variantes fue realizada con el programa STACKS v2.6, considerando las variantes presentes en al menos el 30% de la población y con una frecuencia de alelo menor (MAF)  $\geq 0.01$ . El control de calidad de las variantes se realizó con el programa VCFtools v0.1.16. La Heterocigosidad observada ( $H_o$ ), Heterocigosidad esperada ( $H_e$ ), MAF y el coeficiente de endogamia  $F$ , fueron calculados con los paquetes adegenet, hierfstat y snpReady en el programa R. Se identificaron 39,232 SNPs válidos para el análisis de diversidad genética. El promedio de MAF,  $H_o$  y  $H_e$  fueron 0.31, 0.1216 y 0.3918 respectivamente. El bajo valor de  $H_o$  podría atribuirse a un efecto de endogamia ( $F=0.69$ ) o a un efecto de muestreo. Un mayor número de muestras serán incluidas en el análisis final. Estos son los primeros resultados de la aplicación de GBS como una tecnología alternativa para estudios de diversidad genómica en especies nativas poco estudiadas.

**PALABRAS CLAVE:** Genotyping by sequencing, Cuy, Diversidad genómica, SNP

# Comparison of the affinity of acetaminophen and a novel boron analogue on the TRPV channels

Melvin Nadir Rosalez, Celina de los Á. Castillo Figueroa, Marvin A. Soriano Ursua

Department of Physiology, Escuela Superior de Medicina, Instituto Politécnico Nacional, Plan de San Luis and Diaz Miron S/N, Mexico City, 11340, Mexico.

Pain is the main reason why people attend hospitals and physicians. The most used over-the counter drug to treat pain is acetaminophen, nevertheless, due to its easy access and overuse it has become the main cause of acute liver failure. Studies have shown that boron counterparts to known pharmaceutical drugs show an increased efficiency, potency and affinity for the same protein targets. One of the main targets for the analgesic action of acetaminophen has been correlated to the activation of transient receptor potential vanilloid channel 1, or TRPV1, a non-selective cation channel, located in the central nervous system. The main objective of the study was to analyze and compare if acetaminophen with its metabolites and a boron analogue and the proposed metabolites would bind to the vanilloid pocket or active site of the channel. For the study a computer with an AMD12 CPU, 12 GB of RAM and 1 TB of memory was used. The molecules were designed using ChemsSketch®, they were optimized using Gaussview® and Gaussian 9.0®, while the PDB file for the channel was found in the PDB database, and the docking files were prepared using Autodock Tools 1.5.7. The docking was run using Fedora Workstation v32, and the analysis and visualization of the free energies and bonds formed was done using Autodock Tools and Discovery Studio Visualizer v20.1.0.19295. The Results showed that the molecules did bind to the vanilloid pocket albeit with the boron compounds showing a better free energy in comparison with acetaminophen. Thus, it can be said that the boron analogue in study had a better affinity for the TRPV1 channel than acetaminophen nevertheless further studies are needed to assess the in vivo prospect of an analgesic effect of the new molecule.

**PALABRAS CLAVE:** Pain, Acetaminophen, Boron, TRPV1



# Evaluación *in silico* de análogos de poliaminas con potencial inhibidor del transportador TcPOT1 de alta eficiencia de *Trypanosoma cruzi*

Belén Cuevas-López (1), Carlos Eliel Maya-Ramírez (1), Dulce Estefanía Nicolás-Álvarez (1), Andrés Reyes-Chaparro (2).

1: Instituto Politécnico Nacional. Escuela Nacional de Ciencias Biológicas. Av. Luis Enrique Erro S/N, UPALM, Alcaldía Gustavo A. Madero, C.P. 07738, CDMX.

2: Centro de Física Aplicada y Tecnología Avanzada (CFATA). Universidad Nacional Autónoma de México. Blvd. Juriquilla 3000, Querétaro, México.

**Introducción:** La enfermedad de Chagas o Tripanosomiasis americana, es una parasitosis endémica de América Latina ocasionada por el parásito protozoario *Trypanosoma cruzi*. Esta enfermedad afecta a más de 1 millón de mexicanos, esto debido a que las condiciones climatológicas, posición geográfica y condiciones de vida propician el desarrollo del artrópodo transmisor, los triatomíneos. *Trypanosoma cruzi* tiene la capacidad de crecer en su hospedero vertebrado, gracias a la captación de moléculas llamadas poliaminas. Estas poliaminas son producidas por casi todos los seres vivos, sin embargo, el parásito no tiene las enzimas necesarias para sintetizarlas, por lo que es necesario emplear al transportador de membrana TcPOT1 para su incorporación, y promover su reproducción e infección a las células hospederas.

**Objetivo:** Modelar fármacos para el tratamiento de la Tripanosomiasis a través de ensayos *in silico* con la estructura tridimensional del transportador TcPOT1 y de 7 análogos de poliaminas, los cuáles pretenden bloquear el transportador para así evitar el ingreso de poliaminas provenientes del hospedero hacia el interior del parásito, provocando un efecto negativo en la viabilidad de este.

**Materiales y métodos:** Se realizó el modelado tridimensional del transportador TcPOT1 mediante la secuencia de aminoácidos proveniente del GenBank, mediante el software SwissMODEL y se realizó la validación de este mediante el servidor Procheck. Los análogos de poliaminas se realizaron en el software ChemsSketch para su posterior minimización con el servidor Avogadro para su utilización en los estudios de Docking molecular. El acoplamiento se realizó en tres diferentes softwares, AutoDock Vina, AutoDock4 y SwissDock.

**Resultados:** Los derivados de putrescina (PM2) y espermidina (PM3) mostraron una alta afinidad (-11.4 y -13.48  $\Delta G$ , respectivamente) por la porción intracelular del transportador TcPOT1 encontrando que la molécula PM3 presenta la más alta afinidad con el receptor, siendo esta molécula el mejor candidato para emplearse contra la tripanosomiasis. La molécula PM3 se compone por 8 farmacóforos, dos de los cuáles que pueden donar hidrógenos, además de dos sitios de enlace de aceptores de Hidrógeno y tres sitios de unión de fuerzas hidrofóbicas que le permite hacer interacción en el sitio activo del transportador.

**Conclusiones:** En este trabajo se propone al transportador TcPOT1 como potencial blanco farmacológico. Se obtuvo y validó el modelo del transportador TcPOT1 para utilizarlo en estudios de diseño de fármacos. El análogo PM1.1 presenta la mayor energía de afinidad hacia el transportador TcPOT1. El análogo PM1.1 presenta propiedades ADMET adecuadas para ser un potencial fármaco para el tratamiento de la Tripanosomiasis Americana.

**PALABRAS CLAVE:** Enfermedad de Chagas, *Trypanosoma cruzi*, Docking molecular

# Utilidad del análisis integrativo de microARN para diagnóstico

**Julieth López-Castiblanco (1), David Niño-Torres (1), Liliana López-Kleine (1),  
Adriana Rojas (2), Litzy Bermúdez (2)**

1: Departamento de estadística, Facultad de ciencias, Universidad Nacional de Colombia, Bogotá, Colombia

2: Instituto de genética humana, Facultad de medicina, Pontificia Universidad Javeriana, Bogotá, Colombia

El cáncer del pulmón es el tercer cáncer con mayor incidencia (22.4) en el mundo siendo el cáncer de células no pequeñas (NSCLC) el más común, sin embargo, es el cáncer con la mayor tasa de mortalidad (18.0) según la Organización mundial de la salud (2020). En la mayoría de los casos el diagnóstico se realiza en etapas tardías, por ello, investigaciones que permitan la detección temprana son cruciales para disminuir la tasa de mortalidad y mejorar el éxito de los tratamientos.

El objetivo del estudio fue aplicar un flujo de trabajo para el análisis bases de datos secundarias de exosomas provenientes de biopsia líquida, que permitiera el análisis integrativo de las bases seleccionadas para detectar posibles biomarcadores para NSCLC.

La ruta de trabajo que se estableció consistió en una búsqueda de bases de datos en el **NCBI** (GSE71661 y GSE111803). Luego, se realizó un análisis de control de calidad en R y un análisis de expresión diferencial utilizando **DESEQ2** para detectar microARNs comunes diferencialmente expresados. Para ver su papel biológico, se realizó la identificación de genes blanco usando **MiRNet 2.0** que permitió generar una red de interacción que se analizó en **Cytoscape**. Posteriormente, se realizó un análisis de enriquecimiento con la herramienta **DAVID** en donde se detectaron procesos involucrados con los blancos seleccionados. De aquí se realizó un análisis biológico y se revisó en la literatura para entender cómo influyen los microARNs encontrados en la enfermedad.

Se encontraron 5 microARNs en común, todos sobre-expresados: hsa-miR-1246, hsa-miR-144-3p, hsa-miR-16-2-3p, hsa-miR-194-5p, hsa-miR-505-5p y 293 genes blanco con los que la anotación dio 398 términos biológicos. A través de la ruta de trabajo que involucró alrededor de 4 plataformas se encontraron resultados con sentido y sustento biológico y que cumplen con el objetivo de hallar potenciales biomarcadores.

**PALABRAS CLAVE:** MicroARNs, cáncer de pulmón de células no pequeñas (NSCLC), potenciales biomarcadores, análisis integrativo

# Revisión y propuestas de métodos para resumir y mejorar los resultados de los análisis de enriquecimiento funcional en GO

**Borja Gómez Cabañes, Francisco García-García, Marta R. Hidalgo**

Bioinformatics and Biostatistics Unit, Príncipe Felipe Research Center (CIPF), 46012, Valencia, Spain

Los análisis de enriquecimiento con términos GO (Gene Ontology) son utilizados para caracterizar funcionalmente los resultados de diferentes análisis bioinformáticos. GO presenta una alta **interdependencia** entre sus términos debido a su estructura jerárquica, donde estos se distribuyen de menor a mayor especificidad, y los genes anotados en un término concreto se anotan también en todos sus ancestros. Esto genera una gran **redundancia**, pues la mayoría de los análisis de enriquecimiento (p. ej. Over-Representation Analysis (ORA), y Gene Set Enrichment Analysis (GSEA)) analizan cada término de forma independiente, sin tener en cuenta las relaciones existentes entre ellos. Esto da como resultado largas listas de términos GO donde la gran mayoría tienen relación de parentesco, aumentando la redundancia, y siendo solo unos pocos relevantes e informativos.

El objetivo de este trabajo es la revisión de las principales estrategias empleadas para disminuir y simplificar los resultados de un análisis de enriquecimiento, facilitando su interpretación. Actualmente existen dos ramas principales para tratar la redundancia y sesgos del enriquecimiento en Gene Ontology:

- Actuar durante el enriquecimiento funcional, modificando los análisis para tener en cuenta las relaciones existentes entre los términos, tratándolos como un conjunto y no de forma independiente (p. ej. los algoritmos weight o elim).
- Actuar posteriormente al análisis, utilizando distintas herramientas y algoritmos (p. ej. similitud semántica o weighted set cover) para filtrar, detectar y eliminar los términos no informativos.

Con todo ello se pretende señalar cuáles son las mejores aproximaciones para corregir uno de los principales sesgos de la Gene Ontology, así como detectar las limitaciones, necesidades y mejoras necesarias en este campo de cara a un futuro próximo, donde la importancia del enriquecimiento funcional es cada vez mayor dado el avance y progreso de las tecnologías de alto rendimiento.

**PALABRAS CLAVE:** Enriquecimiento funcional, Gene Ontology, Redundancia de resultados, Revisión sistemática

# **Estudio de las diferencias de sexo en la enfermedad de Alzheimer: un meta-análisis de estudios de microRNAs**

**Jaime Llera Oyola (1), Zoraida Andreu (2), Irene Soler-Sáez (1),  
Fernando Gordillo González (1), Héctor Carceller (3,4), María de la Iglesia-Vayá (3),  
Marta R. Hidalgo (1), Francisco García-García (1)**

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

3: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

4: Institute of Biotechnology and Biomedicine (BIOTECMED), Universitat de València, 46100, Burjassot, Spain

La enfermedad de Alzheimer (EA) supone, actualmente, la principal causa de demencia en la población, con previsiones de que su incidencia aumente de forma conjunta con el envejecimiento de la población. Se han descrito diferencias en base al sexo en el desarrollo de la enfermedad, pero existen incógnitas sobre cómo afecta el sexo biológico a nivel genético y molecular. La identificación de perfiles de expresión y elementos reguladores específicos de sexo, como los microRNAs, podrían resultar de gran ayuda en la elaboración de métodos dirigidos de diagnóstico y terapia para la enfermedad.

El objetivo del presente trabajo es identificar microRNAs que se encuentren expresados diferencialmente en función del sexo en pacientes de EA. El abordaje utilizado consiste en una revisión sistemática, acorde a los criterios PRISMA, de los estudios de expresión de microRNAs con información de sexo en EA publicados en las principales bases de datos. Los estudios seleccionados serán analizados individualmente incorporando la variable sexo, y seguidamente se realizará el meta-análisis del conjunto de estudios encontrados. De este modo, esperamos identificar biomarcadores de microRNAs en relación a la variable sexo que mejoren el conocimiento sobre la EA. A través de la revisión sistemática, se han seleccionado 27 estudios de microRNA relacionados con la EA, recogidos en los repositorios públicos de datos: GEO y ArrayExpress. Tras la aplicación de diferentes filtros de calidad, la selección final incluyó 6 estudios de expresión con muestras de tejido cerebral y de sangre que utilizan las tecnologías de microarray y RNAseq. Se realizaron análisis de expresión de cada estudio y un meta-análisis global para determinar potenciales microRNAs marcadores específicos del sexo.

**PALABRAS CLAVE:** Alzheimer, microRNA, Transcriptómica, Diferencias de sexo, Metaanálisis

# Caracterización de las diferencias de sexo en Parkinson mediante el análisis masivo de datos de expresión de célula única y su integración

**Fernando Gordillo González (1), Zoraida Andreu (2), Irene Soler-Sáez (1), Jaime Llera Oyola (1), Héctor Carceller (3,4), María de la Iglesia-Vayá (3), Marta R. Hidalgo (1), Francisco García-García (1)**

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

3: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

4: Institute of Biotechnology and Biomedicine (BIOTECMED), Universitat de València, 46100, Burjassot, Spain

La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más frecuente en la población, así como el trastorno del movimiento más común asociado al envejecimiento. Su causa aún es desconocida, siendo su diagnóstico principalmente clínico y no existiendo un tratamiento para frenar el proceso neurodegenerativo. La variable sexo tiene gran importancia en la EP, siendo determinante en la prevalencia, edad de aparición, mortalidad y el desarrollo y progresión de síntomas. Sin embargo, los mecanismos moleculares responsables de estas diferencias aún no están completamente definidos.

El objetivo del presente trabajo es profundizar en estos mecanismos subyacentes a las diferencias de sexo en la EP mediante la integración de estudios de secuenciación masiva de célula única en muestras cerebrales de EP. En primer lugar, se ha realizado una revisión sistemática de los estudios de célula única desarrollados hasta la fecha con información de sexo, siguiendo los criterios PRISMA. Tras la realización de un análisis exploratorio e individual de cada estudio seleccionado se procederá a la integración de los resultados obtenidos en un único meta-análisis, que revelará genes por tipo celular analizado y características funcionales que podrían estar implicados en las diferencias de sexo observadas en esta enfermedad.

Inicialmente, en la revisión sistemática se han encontrado 49 estudios, de los cuales, tras aplicar los criterios PRISMA, se seleccionaron 3 estudios de scRNA-seq con un total de 46 pacientes (16 controles y 30 casos). El análisis de dichos estudios contribuirá a un mejor conocimiento de cómo el factor sexo influye en la EP, tanto dando robustez a mecanismos previamente descritos como hallando posibles nuevas rutas implicadas.

**PALABRAS CLAVE:** Parkinson, scRNA-seq, Diferencias de sexo, Meta-análisis

# Comparación de ensambladores en la recuperación de vMAGs de bacteriófagos

**Karla Nolivos (1), Estefanía Tisalema (1), Jorge Reyes (1,2), Daniel Garzón (2), Fausto Cabezas (2)**

1: Facultad de Ciencias Químicas, Universidad Central del Ecuador (UCE), Ciudadela Universitaria –Avenida América, Quito, Pichincha, Ecuador

2: Instituto de Microbiología, Colegio de Ciencias Biológicas y Ambientales (COCIBA), Universidad San Francisco de Quito (USFQ), Diego de Robles y Vía Interoceánica, Quito, Ecuador

El estudio del genoma de los bacteriófagos se ha visto limitado por la deficiente recuperación y escasa información disponible de las bases de datos actuales. La población de fagos es extensa, dinámica y genéticamente muy diversa por lo que, las secuencias recuperadas no muestran una similitud significativa con ninguna secuencia almacenada, lo que dificulta, en principio, su ensamblaje y posterior anotación. Con el desarrollo de nuevas herramientas bioinformáticas y la mejora de las existentes, se han propuesto varias opciones con distintos enfoques para intentar recuperar genomas virales ensambladas a partir de metagenomas (vMAGs, por sus siglas en inglés). El objetivo de este estudio fue comparar el desempeño de tres ensambladores (metaviralSPAdes<sup>1</sup>, MEGAHIT<sup>2</sup> y Unicycler<sup>3</sup>) para recuperar genomas fágicos a partir de lecturas generadas mediante secuenciamiento de Illumina tipo “shotgun” de cuatro muestras concentradas de fagos aislados de aguas residuales con actividad en *Klebsiella pneumoniae* resistente a carbapenémicos. Finalmente se evaluó parámetros de calidad e integridad de los vMAGS recuperados mediante checkV<sup>4</sup>. Se ensamblaron cuatro genomas fágicos de Taipeivirus spp. con discretas diferencias de longitud y profundidad. Unicycler y metaviralSPAdes ensamblaron 4 vMAGs, mientras que MEGAHIT solo 3. Los genomas recuperados fueron de alta calidad y estaban completos >94%, asumidos por la detección de repeticiones terminales directas (DTRs) o por reconocimiento de patrones virales mediante modelos ocultos de Márkov (HMM). Por tanto, sugerimos Unicycler y metaviralSPAdes para la recuperación y ensamblaje de genomas fágicos.

**PALABRAS CLAVE:** Bacteriófagos, vMAGs, Ensamblaje, Aguas residuales

# Interrogando la expresión de *PARP1* y la mutación en *PBRM1* como biomarcadores de respuesta a tratamiento con atezolizumab y bevacizumab o sunitinib en pacientes con carcinoma renal de células claras

Javier Lanillos, María Santos-Romero, Carlos Valdívía, Alberto Mora,  
Javier de Nicolás, Cristina Rodríguez-Antona

Laboratorio de Cáncer Endocrino Hereditario, Centro Nacional de Investigaciones Oncológicas (CNIO)

**Introducción:** La expresión del gen *PARP1* en conjunto con el estado mutacional del gen *PBRM1* se ha asociado a la respuesta a inmunoterapia en pacientes con carcinoma de células renales de células claras (CCRcc). Sin embargo, este hallazgo requiere validación en otras cohortes clínicas.

**Materiales y métodos:** Se utilizaron los datos ómicos (RNA-seq y DNA-seq) y datos clínicos de 794 pacientes de cáncer renal disponibles en el ensayo clínico Immotion151 (EGAS00001004353) tratados con la combinación de atezolizumab y bevacizumab o el grupo control con sunitinib. Se evaluó la Supervivencia Libre de Progresión (SLP) teniendo en cuenta la histología (CCRcc), el grupo pronóstico (criterio de riesgo según el MSKCC), las firmas de expresión del Immotion151, la expresión de *PARP1* (alta/baja) y las mutaciones en *PBRM1* (mutado/no mutado).

**Resultados y discusión:** Se observó una SLP más corta en el tumores con *PARP1*-alta tanto en el grupo tratado con atezolizumab más bevacizumab como en el de sunitinib ( $p=0,003$ ,  $p<0,001$ , respectivamente). Esta asociación se mantuvo estadísticamente significativa tras corregir por el grupo pronóstico (MSKCC). Cuando se combinaron *PARP1* y *PBRM1*, el subgrupo *PARP1*-alta/*PBRM1*-no mutado mostró diferencias significativas solamente en el brazo de sunitinib ( $p<0,001$ ). En el brazo de sunitinib, el grupo *PARP1*-alta/*PBRM1*-no mutado mostró la SLP más corta con respecto al resto y a *PARP1*-baja/*PBRM1*-no mutado ( $p<0,01$ ). Comparando esos dos subgrupos de *PBRM1*-no mutado, se observaron diferencias en la firma de angiogénesis ( $p<0,01$ ). Además, se observaron mayores porcentajes de las firmas de linfocitos T efectores y proliferativas en *PARP1*-baja/*PBRM1*-no mutado.

**Conclusión:** La expresión de *PARP1* junto a las mutaciones en *PBRM1* se muestran como un potencial biomarcador de SLP en la cohorte clínica del Immotion151, en pacientes tratados con antiangiogénicos con o sin combinación con inmunoterapia. Otras cohortes clínicas servirán para revalidar estas observaciones, incluyendo pacientes tratados solamente con inmunoterapia.

**PALABRAS CLAVE:** Inmunoterapia, Secuenciación, Ómicas, Carcinoma de células renales

# **Análisis bioinformático para explorar el impacto de las mutaciones somáticas en los genes *SAMHD1*, *GNL1*, *POLE*, *MRE11A* y *ASXL2* como potenciales mini impulsores en Cáncer Colorrectal.**

**Anthony Vladimir Campos Segura (1, 3), Alexis German Murillo Carrasco (2, 3)**

1: Universidad Nacional Federico Villarreal, Escuela profesional de Biología, Lima-Perú

2: Centro de Investigação Translacional em Oncologia (LIM24), Instituto do Cancer do Estado de Sao Paulo (ICESP), Hospital das Clinicas da Faculdade de Medicina da Universidade de Sao Paulo (HCFMUSP), Sao Paulo CEP 01246-000, SP, Brasil

3: Grupo de Investigación en Inmunología y Cáncer. INMUCA

**Introducción:** La mayoría de las investigaciones en cáncer identifican las mutaciones somáticas más importantes. Sin embargo, muchas de las mutaciones encontradas no son impulsoras principales (drivers), algunas mutaciones tienen efectos promotores de tumores relativamente débiles y a este tipo de mutaciones se les denomina “mini impulsores”.

**Objetivos:** Explorar el posible efecto mini impulsor de los genes *SAMHD1*, *GNL1*, *POLE2*, *MRE11* y *ASXL2* en cáncer colorrectal.

**Métodos:** Realizamos un análisis de características clínicas utilizando el portal web del cBioPortal (<https://www.cbioportal.org/>). Revisamos datos de expresión génica por RNA-seq en muestras de cáncer colorrectal y tejidos saludables incluidos en el proyecto TCGA (The Cancer Genome Atlas Program) a través del servidor online UALCAN (<http://ualcan.path.uab.edu/analysis.html>). Obtuvimos datos de expresión de los cinco genes propuestos en muestras de pacientes y controles e información sobre genes coexpresados. Usando el lenguaje de programación R descargamos la data de enriquecimiento de los módulos KEGG, Reactome y Gene Ontology del servidor Enrichr, y, por último, generamos dotplots y cnetplots que muestran la implicancia de nuestros genes en diferentes vías metabólicas.

**Resultados:** Tres de los 5 genes evaluados (*GNL1*, *POLE* y *MRE11A*) mostraron sobreexpresión en pacientes con cáncer colorrectal en relación a los controles ( $p < 0.01$ ). En relación a los genes coexpresados *GNL1*, *POLE* y *MRE11* mostraron 3689, 1550 y 3529 genes coexpresados ( $R=0.7$ ) respectivamente asociados a vías de respuesta inmunológica, control del ciclo celular, factores de transcripción y mecanismos de reparo de DNA.

**Conclusiones:** *GNL1*, *MRE11A* y *POLE* fueron descritos una lista de genes con potencial de impulsar el cáncer. Los resultados de nuestros análisis muestran la viabilidad de estos genes en cáncer colorrectal a través del control del proceso de transcripción, mecanismos de reparo y transición entre las fases del ciclo celular (*POLE* y *MRE11A*), así como la intervención del tumor en el sistema inmune del individuo (*GNL1*).

**PALABRAS CLAVE:** Cáncer colorrectal, Expresión génica, Genes mini impulsores



# Transcriptome-based analysis identifies ASXL1 and MAP4K4 as prognostic markers in microsatellite instability gastric cancer.

Álvaro de Jesús Huamani Ortiz (1,2,8), Anthony Vladimir Campos Segura (3,8),  
Kevin Jorge Magano Bocanegra (4,8), Obert Marín Sánchez (5,8),  
Alexis Germán Murillo Carrasco (6,8), César Alexander Ortiz Rojas (7,8,9)

- 1: Universidad Nacional Mayor de San Marcos, Escuela profesional de Tecnología Médica, Lima-Perú
- 2: Universidad Nacional Mayor de San Marcos, Sociedad Científica de Estudiantes de Tecnología Médica en el área de Laboratorio Clínico y Anatomía Patológica, Lima-Perú
- 3: Universidad Nacional Federico Villarreal, Escuela profesional de Biología, Lima-Perú
- 4: Universidad Nacional Mayor de San Marcos, Escuela profesional de Genética y Biotecnología, Lima-Perú
- 5: Departamento Académico de Microbiología Médica, Facultad de Medicina, Universidad Nacional Mayor de San Marcos
- 6: Centro de Investigação Translacional em Oncologia (LIM24), Instituto do Cancer do Estado de São Paulo (ICESP), Hospital das Clinicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP), São Paulo CEP 01246-000, SP, Brasil
- 7: Centro de Terapia Basada en Células, Universidad de São Paulo, Ribeirão Preto, Brasil
- 8: División de Hematología, LIM31, Facultad de Medicina, Universidad de São Paulo, São Paulo, Brasil
- 9: Grupo de Investigación en Inmunología y Cáncer. INMUCA

**Background:** Microsatellite instability (MSI) tumors are identified by alteration in mismatch repair (MMR) genes. In gastric cancer (GC), The Cancer Genome Atlas (TCGA) study and the Asian Cancer Research Group (ACRG) identified that MSI tumors characterize patients with a high probability to respond to the treatment. Nevertheless, the exact pathways involved in this are unknown. By this, we analyzed individual expression profiles that are associated with prognosis from TCGA and ACRG transcriptome data.

**Objective:** To identify genes whose expression profile is consistently associated with overall (OS) and disease-free survival (DFS) in two MSI-GC cohorts.

**Methods:** The TCGA (n=68) and ACRG (n=68) cohorts were retrieved from the Firebrowse database ([www.firebrowse.org](http://www.firebrowse.org)) and the Gene Expression Omnibus portal (GEO, <https://www.ncbi.nlm.nih.gov/geo/>, GSE66229), respectively. Patients were dichotomized into groups of high and low gene expression, for each gene, according to the optimal cut-point calculated from the ROC curve related to the OS. The association of genes with OS rates was analyzed using Cox regression and area under the curve (AUC). Additionally, we used expression levels of previously identified genes for predicting cell fractions in the tumor microenvironment by the TIMER 2.0 web resource (<http://timer.cistrome.org/>). All calculations were performed with R 4.1.1 software ([www.r-project.org](http://www.r-project.org)).

**Results:** We identified two genes (ASXL1 and MAP4K4) consistently associated with OS and DFS (HR>1, p<0.05, AUC>0.5). Multivariate Cox regression analysis and analysis with clinical variables confirmed that the ASXL1 and MAP4K4 association with prognosis is not influenced by relevant clinical variables such as age and tumor stage, highlighting the importance of our results. Additionally, we found an enrichment of tumor-associated fibroblasts and follicular T cells in tumors with elevated expression of MAP4K4 and ASXL1, respectively.

**Conclusion:** The high expression of ASXL1 and MAP4K4 is associated with lower survival rates and contributes to the enrichment of specific immune cells in MSI-GC patients.

**PALABRAS CLAVE:** transcriptome, microsatellite instability (MSI), overall survival, disease-free survival

# Diferencias de sexo en el lipidoma de las vesículas extracelulares tras el consumo masivo de alcohol en la adolescencia

Carla Perpiñá-Clérigues (1,2), José F. Català-Senent (1) , Francesc Ibáñez (3),  
Consuelo Guerri (3), María Pascual-Mora (2), Francisco García-García (1)

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Department of Physiology, School of Medicine and Dentistry, University of Valencia (UV), 46003, Valencia, Spain

3: Cellular and Molecular Pathology of Alcohol, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

Las vesículas extracelulares (VE) tienen un papel crucial en procesos patológicos, incluida la inflamación. Su contenido lipídico no solo es crucial a nivel estructural sino que participan en su biogénesis, liberación e internalización en las células diana y como mensajeros. La capacidad de las VEs de atravesar la barrera hematoencefálica, las convierte en buenos biomarcadores no invasivos, siendo el objetivo de este estudio evaluar si la intoxicación por alcohol altera el lipidoma de las VEs plasmáticas de los adolescentes y su posible perfil diferencial por sexo. Para ello, analizamos VEs plasmáticas de adolescentes humanos de ambos sexos tras una intoxicación aguda por alcohol y controles. Además, para ver si la respuesta inmune está asociada con el receptor TLR4 se han utilizado dos modelos murinos (WT y TLR4-KO), pues estudios previos han demostrado su papel en la respuesta neuroinflamatoria tras el consumo de alcohol.

El abordaje bioinformático realizado consta de 1) procesamiento de los datos, 2) análisis de abundancia diferencial, 3) anotación por clases, 4) análisis de enriquecimiento de clases.

Los resultados muestran diferencias significativas en el perfil lipídico entre las VEs plasmáticas en los diferentes grupos, y estos difieren significativamente entre sexos tanto en humanos como en ratones. Por otro lado, en la caracterización por clases detectamos diferencias en lípidos relacionados tanto con la actividad exosomal como con la neuroinflamación. Destacar un mayor número de lípidos significativos en hembras adolescentes que pertenecen a clases relacionadas con la respuesta inmune inflamatoria, lo que respalda resultados anteriores, que demuestran que las adolescentes son más vulnerables a los efectos del alcohol que los machos.

Por lo tanto, los lípidos de las VEs plasmáticas pueden ayudarnos a explicar los mecanismos que subyacen a la respuesta neuroinflamatoria tras intoxicación aguda por alcohol, siendo potenciales biomarcadores de la enfermedad.

**PALABRAS CLAVE:** Vesículas extracelulares, Lipidómica, Alcohol, Diferencias de sexo, Biomarcadores

# Identification and Characterization of Human RNA-Binding Repeat Proteins

**Matías Carletti (1,2), Agustín Ormazábal (1,2), Julia Marchetti (1,2),  
Martin González Buitrón (1,2) , Tadeo Saldaño (1,3) and Nicolás Palopoli (1,2)**

1: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires, Argentina

2: Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

3: Ministerio de Ciencia, Tecnología e Innovación (MinCyT), Argentina

**Introduction and Goals:** Tandem Repeat Proteins (TRPs) are of great scientific interest due to their structural and functional characteristics [1]. It has been previously reported that TRPs are enriched in functions related to RNA binding like transcription, RNA splicing or stabilization, and multiprotein complexes assembly [2]. However, a large-scale analysis of RNA Binding Proteins (RBPs) and are predicted to have tandem repeats (TRs) remains pending. Our main objectives are to identify and characterize the structure, function and evolutionary history of these repetitive human proteins.

**Materials and methods:** Our work is based on preliminary studies performed during a collaborative hackathon organized by the Biocomputing Group at the University of Padova. To build our TR-RBPs dataset we collected annotated entries from different resources such as CoDNAS-RNA, RBPbase, IntAct, PDB and UniProt, plus the intrinsically disordered RBPs from Zhao et al [3]. We defined consensus TR regions based on residue level predictions from RepeatsDB-Lite, CE-Symm, REP2, Tally, TAPO and TRAL. Annotation and statistical description of the resulting TR-RBPs was performed by analysis of the integrated data. Manual curation was assisted by inspection of AlphaFoldDB models.

**Results and Conclusion:** We found 5244 UniProt entries that likely interact with RNA, of which 3093 exhibit TR regions identified by two or more predictors. In a subset of 59 repetitive disordered proteins, 57% of the RNA binding residues overlap with repetitive regions. Also, in our highest-quality dataset of 132 proteins with strong evidence of RNA binding and presence of TRs, 94 are putative novel TRPs not yet included in RepeatsDB. 107 of these proteins were correctly modelled by AlphaFold2, adding to a large dataset of experimental TR-RBP structures. At least 25 TR-RBPs are conserved in mice, flies, nematodes and yeasts. Overall, we have curated a novel dataset of TR-RBPs and were able to provide preliminary annotations for many of its proteins.

**References:** [1] DOI: [10.1016/j.jsb.2020.107608](https://doi.org/10.1016/j.jsb.2020.107608); [2] DOI: [10.3390/genes11040407](https://doi.org/10.3390/genes11040407); [3] DOI: [10.1016/j.jmb.2021.167229](https://doi.org/10.1016/j.jmb.2021.167229).

**PALABRAS CLAVE:** Tandem Repeat Proteins (TRP), RNA Binding Proteins (RBP), AlphaFold2 (AF2)

# **Modelamiento molecular y análisis bioinformático de polimorfismos de la proteasa NS3 del virus Dengue presente en Argentina, Brasil y Paraguay**

**Merino M. Cecilia (1,3), Elio R. Castillo (2), Karina Salvatierra (3)**

1: FCEQyN, Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN-UNaM), Posadas, Misiones, Argentina

2: IBS, FCEQyN, Laboratorio de Genética Evolutiva Dr. Claudio J. Bidau. Instituto de Biología Subtropical (CONICET-UNaM), Posadas, Misiones, Argentina

3: PMMM, FCEQyN, Instituto de Materiales de Misiones (CONICET-UNaM), Posadas, Misiones, Argentina

El Dengue es una enfermedad infecciosa endémica de países tropicales y subtropicales. Es causada por el virus dengue, perteneciente al género Flavivirus. El genoma consiste en un ARN monocatenario de polaridad positiva. La proteína NS3 cumple funciones enzimáticas en el procesamiento de la poliproteína, en la replicación del ARN viral y está involucrada en procesos que permiten evadir la respuesta inmune del hospedador. Debido a la importancia de las funciones que cumple, la proteasa de NS3 es considerada un sitio blanco clave para la acción de los antivirales. La presencia de mutaciones en el sitio catalítico de la proteasa de NS3 puede modificar la relación de afinidad respecto a los fármacos, ya sea en aumento o en decrecimiento. El objetivo del presente trabajo fue identificar polimorfismos en las secuencias de la proteasa NS3 de aislados clínicos de Argentina, Brasil y Paraguay, realizar el modelamiento tridimensional de la proteína y analizar las interacciones resultantes del acoplamiento con antivirales. Se emplearon programas bioinformáticos para el análisis de secuencias virales: BioEdit, PyMol y Geneious para la edición de secuencias y análisis de sustituciones, Swiss Model para el modelamiento de la proteína, AutoDock Tools para el análisis de acoplamiento con antivirales y Amber para las dinámicas moleculares. Se registraron sustituciones aminoacídicas en la secuencia de la proteasa NS3: Argentina 1,7% (DENV1), Brasil 1,8% (DENV1), 4,1% (DENV2), 1,2% (DENV3), 0,4% (DENV4), y Paraguay 1,2% (DENV3). Los cambios aminoacídicos detectados no se encuentran próximos al sitio catalítico de la proteína, sin embargo interfieren en la afinidad por los antivirales ensayados. Los resultados obtenidos permiten concluir que los polimorfismos influyen directamente en la estructura de la proteína y en la afinidad de interacción con antivirales.

**PALABRAS CLAVE:** Antivirales, Dengue, NS3

# Búsqueda de inhibidores de PknG de *Mycobacterium tuberculosis* mediante tamizaje virtual masivo basado en farmacóforo

Arica-Sosa A., Alcántara R., Quiliano M.

Grupo de Desarrollo e Innovación de Medicamentos, Laboratorio de Biomoléculas, Facultad de Ciencias de la Salud, Universidad Peruana de Ciencias Aplicadas (UPC), 15023, Lima, Perú

La Tuberculosis (TB) es una enfermedad infecciosa causada por *Mycobacterium tuberculosis* (Mtb). Según el último informe de la OMS, en el año 2021 se estima que más de 1.9 millones de personas enfermaron de TB provocando cerca de 1.3 millones de fallecimientos. A pesar de ser curable y prevenible, la TB es la segunda causa de muerte infecciosa después del COVID-19, que, sumado al incremento de resistencia antibiótica se ha convertido en un grave problema de salud pública. Mtb es un patógeno capaz de sobrevivir dentro de los macrófagos causando una infección latente a largo plazo debido a su serina/treonina proteína quinasa G (PknG) quien se encarga de participar dentro de procesos celulares, así como del bloqueo de la fusión fagosoma-lisosoma evitando la degradación.

Estudios han demostrado que la inhibición de PknG evita que Mtb entre en la etapa de latencia, convirtiéndolo así en una diana terapéutica de valioso interés. En este trabajo, se tamizaron 1578014 millones de moléculas utilizando un modelo farmacofórico dentro de una plataforma de tamizaje virtual masivo implementada en la nube (Amazon Web Service) logrando identificar 689 candidatos, posteriormente estos candidatos fueron filtrados nuevamente mediante acoplamiento molecular (docking). El acoplamiento molecular detectó 62 compuestos con afinidades de unión estimadas de  $-7,54$  kcal/mol (d.e. =  $0,77$  kcal/mol). Finalmente, se seleccionaron 14 compuestos para experimentos in vitro considerando las actividades biológicas y la disponibilidad comercial reportadas previamente. Los ensayos in vitro de la actividad de PknG mostraron que CHEMBL3237561 inhibe la actividad de la quinasa de manera similar a AX20017, un inhibidor conocido. El efecto inhibitor es dependiente de la dosis con una IC50 relativa de  $4,4 \pm 1,1$   $\mu$ M. En conjunto, nuestro estudio indica que CHEMBL3237561 es un potencial candidato a fármaco antituberculoso debido a su actividad inhibitoria en PknG y sus excelentes parámetros farmacocinéticos informados.

**PALABRAS CLAVE:** Tuberculosis, *Mycobacterium tuberculosis*, Proteína quinasa G, Farmacóforo, Drug Discovery

# Identification of the number of Open Reading Frames in chromosomes of *Cryptococcus neoformans* and *Cryptococcus gattii* species complex

Frutos Bottega D., Semhan Franco P., Miño O., Salvatierra K.

Laboratory of Applied Molecular Biology (LaBiMAP), Faculty of Exact, Chemical and Natural Sciences, National University of Misiones (UNaM)  
Materials, Modeling and Metrology Research Group (PMMM), Misiones Materials Institute (IMAM), National University of Misiones (UNaM).  
National Institute of Agricultural Technology (INTA), EEA Cerro Azul, Misiones

**Introduction:** The *Cryptococcus* species complex includes two species of encapsulated yeasts from the phylum Basidiomycota. *Cryptococcus neoformans*, characterized as an opportunistic pathogen that affects immunosuppressed patients, and *Cryptococcus gattii*, a primary pathogen, causing infections in immunocompetent individuals. Both are haploid, yeast-like, heterothallic fungi characterized by having a polysaccharide capsule. In contrast to fungi that have been widely studied due to their biological importance as pathogens for humans, there is no update and information regarding gene annotations, as well as a line dedicated to maintaining and providing said information in the species of the *Cryptococcus* complex. ORF are sequences that have the ability to be translated, starting at a start codon and ending at a stop. However, the number of ORFs present in each of the chromosomes of these fungi is unknown. **Objectives:** To identify the number of open reading frames of chromosomes one and seven of the *Cryptococcus* species complex of the genomes published in the database, using different bioinformatic tools for ORF search. **Methods:** The sequences corresponding to chromosome one and seven of the *Cryptococcus* species complex were obtained, which are published by the International Nucleotide Sequence Database Collaboration. The ORF search was performed using the GetORF, OrfM and OrfFinder programs. Two search window sizes (600pb and 300pb) were used, and Fisher's analysis of variance and a one-way ANOVA analysis were performed with the InfoStat program and it was determined if there were significant differences between the results obtained. **Results and conclusion:** For both chromosomes, a total of 12 strains of *C. neoformans* and 3 of *C. gattii* were obtained. The ORF numbers of 300 and 600 base pairs were obtained for each chromosome and from the statistical analysis it was concluded that there are no significant differences between the programs, therefore it is possible to obtain reproducible results.

**PALABRAS CLAVE:** *C. neoformans*, *C. gattii*, Open reading frames (ORF), chromosome

# **Análisis de marcadores genéticos relacionados con aterosclerosis**

**Montserrat Loredo-Guillén, Manuel Rivera**

Laboratorio de medicina translacional y biología de sistemas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México

La aterosclerosis cardiovascular pertenece a las enfermedades con mayor índice de mortalidad y morbilidad a nivel mundial como en México. Se caracteriza por ser un trastorno crónico inflamatorio asociado a la acumulación de lípidos dentro de las arterias. Actualmente, se sabe que el inicio del desarrollo de esta placa está ligada a otras patologías como diabetes Mellitus tipo 2, obesidad, resistencia a la insulina, altos niveles de lipoproteína de baja densidad, hipertensión arterial e hipertrigliceridemia las cuales participan como factores de riesgo. Su desarrollo ocurre en un lapso de 10 a 40 años sin presentar síntomas, lo que promueve que su diagnóstico sea en una etapa tardía. Actualmente, la generación de modelos de aprendizaje de máquina y la asignación de un valor de riesgo poligenético son pilares fundamentales en la predicción de enfermedades complejas. Este trabajo tiene como objetivo analizar la participación de marcadores biológicos en aterosclerosis identificados a través de estrategias de inteligencia artificial tales como redes neuronales. Se generó una estrategia basada en modelos de aprendizaje de máquina a partir de SNPs depositados en GWAS asociados a aterosclerosis así como a los factores de riesgo genéticos relacionados con la patología. El algoritmo desarrollado permitió realizar la selección, procesamiento, transformación y minería de datos especializadas de biomarcadores con el propósito de obtener SNPs putativos. Por último, se desarrolló una función de riesgo poligenético para obtener valores de riesgo de los SNPs seleccionados. En conclusión los modelos generados en este trabajo tienen capacidad de relacionar SNPs depositados en diversas bases de datos con los principales factores de riesgo para desarrollar aterosclerosis.

**PALABRAS CLAVE:** Aterosclerosis, Aprendizaje de máquina, SNPs, Valor de riesgo poligenético, biomarcadores

# Influencia de la etnicidad latina y no latina en la asociación entre niveles de mutaciones y expresiones genéticas en pacientes con cáncer de mama

Uriel S. Capcha-Jimenez (1,2,3), Jesus F. Pasache Juarez (1,2,3),  
Alexis Germán Murillo Carrasco (1,4)

1: Grupo de Investigación en Inmunología y Cáncer (INMUCA)

2: Facultad de Medicina, Universidad Nacional Mayor de San Marcos (UNMSM), Lima, Perú

3: Sociedad Científica de San Fernando (SCSF), Lima, Perú

4: Centro de Investigação Translacional em Oncologia (LIM24), Instituto do Cancer do Estado de Sao Paulo (ICESP), Hospital das Clínicas da Faculdade de Medicina da Universidade de Sao Paulo (HCFMUSP), Sao Paulo CEP 01246-000, SP, Brasil

**Introducción:** El cáncer de mama (CaMa) es la neoplasia más común en mujeres presentando una alta mortalidad (15,5%). La aparición de mutaciones es una característica del proceso tumoral que permite clasificar pacientes en subtipos moleculares. Muchas veces, estas mutaciones tienen un impacto en los niveles de expresión génica, la mayoría de los estudios se centraron en los genes *BRCA1* y *BRCA2* dejando de lado genes importantes involucrados en la reparación del ADN. Adicionalmente se han reportado estas mutaciones principalmente para la población caucásica, judía y asiática, con pocos informes para la población latina (PL).

**Objetivos:** Hallar el nivel de asociación entre las alteraciones y expresiones genéticas de *BRCA1*, *BRCA2*, *ATM*, *TP53*, *CHEK2*, *BARD1*, *MLH1*, *PALB2*, *BRIP1*, *MSH2*, *MSH6*, *NBN* y *PMS2* en pacientes latinos y no latinos.

**Materiales y métodos:** Utilizamos el repositorio cBioPortal (<https://www.cbioportal.org/>) para buscar conjuntos de datos con información sobre ancestralidad, niveles de expresión (por RNA-seq) y perfil de mutación para los genes más analizados en cáncer de mama (*BRCA1*, *BRCA2*, *ATM*, *TP53*, *CHEK2*, *BARD1*, *MLH1*, *PALB2*, *BRIP1*, *MSH2*, *MSH6*, *NBN* y *PMS2*). Descargamos los datos y los analizamos usando el software R v4.1.3 (paquetes ggplot2, ggmisc y ggpubr) para realizar gráficos comparativos y calcular el valor de p usando el Test de Mann-Whitney.

**Resultados:** En los genes *BRCA2*, *ATM*, *CHEK2*, *BARD1*, *MLH1*, *MSH2* y *NBN* mostraron una asociación ( $p > 0.05$ ) entre la presencia de mutaciones y sus niveles de expresión en la población total ( $n=797$ ) y en el segmento de individuos no latinos ( $n=763$ ). Sin embargo, esta asociación no fue vista en individuos autodeclarado como latinos ( $n=34$ ,  $p > 0.05$ )

**Conclusión:** La pérdida de asociación entre la presencia de mutaciones y niveles de expresividad en la PL abre nuevos nichos de investigación útiles para la estratificación molecular y desarrollo de terapias dirigidas más efectivas.

**PALABRAS CLAVE:** Cáncer de mama, ancestralidad, expresión génica, alteraciones genéticas



# **Análisis in silico de la selectividad de los inhibidores comerciales de la 5 $\alpha$ -reductasa**

**Bonilla R Edwin L, Diana Alexandra Torres Sánchez, Elkin Sanabria-Chanaga**

Universidad de Pamplona, Facultad de Ciencias Básicas, Programa de Química

La 5 $\alpha$ -reductasa (5- $\alpha$ R) es una familia de isoenzimas encargadas de transformar la testosterona en dihidrotestosterona. Con base en esto, esta proteína se convierte un blanco de interés farmacéutico, debido a que se ha demostrado que, muchas patologías como; hiperplasia prostática benigna (HPB) y cáncer de próstata (CaP), son afecciones fuertemente correlacionados con altos niveles de dihidrotestosterona, producto de esta actividad enzimática. En el presente trabajo se estudió la selectividad de algunos inhibidores comerciales de la enzima 5 $\alpha$ -reductasa. Para esto, se realizaron estudios de acoplamiento molecular y dinámica molecular con las isoformas I y II de la proteína de interés. Con base en los resultados obtenidos, se concluyó que la selectividad de cada inhibidor está fuertemente relacionada con el mecanismo de inhibición covalente que exhiben los compuestos ensayados, lo cual reduce la selectividad a dos características claves, primero, la conformación del inhibidor y segundo, su distancia con el cofactor con el cual reacciona.

**PALABRAS CLAVE:** Hiperplasia prostática benigna, Inhibidores 5 $\alpha$ R, Selectividad, 5 $\alpha$ -reductasa

# Diferencias de sexo en la Enfermedad de Parkinson: metaanálisis de estudios transcriptómicos de tejido cerebral

**Adolfo López-Cerdán (1,2), Zoraida Andreu (2), Marta R. Hidalgo (1), Rubén Grillo-Risco (1), José Francisco Català-Senent (1), Irene Soler-Sáez (1), Almudena Neva-Alejo (1), María de la Iglesia-Vayá (2), Francisco García-García (1)**

1: Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

2: Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

3: Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

La Enfermedad de Parkinson (EP) es una de las enfermedades neurodegenerativas más comunes en las sociedades con mayor esperanza de vida. El principal factor de riesgo de esta enfermedad es la edad, con una mayor prevalencia en la población más envejecida. No obstante, el sexo constituye también un factor diferencial tanto en incidencia como en desarrollo de síntomas. A pesar de los avances en el estudio de la enfermedad, no se conoce con precisión los mecanismos moleculares que subyacen a estas diferencias de sexo.

Por este motivo, el objetivo del trabajo ha sido la identificación de posibles patrones de expresión génica relacionados con diferencias sexuales. Para ello, se llevó a cabo una revisión sistemática, en bases de datos públicas, de estudios transcriptómicos con muestras de las tres regiones cerebrales más afectadas por la EP (Sustancia negra, Estriado y Corteza). Todos los estudios seleccionados fueron filtrados en base a la disponibilidad de información del sexo de los pacientes. Los trabajos fueron seleccionados según las normas de la declaración PRISMA y sometidos a metaanálisis para identificar las alteraciones genéticas diferenciales entre mujeres y hombres. Finalmente, se realizó un enriquecimiento funcional de los resultados para caracterizar funcionalmente estas alteraciones.

Como resultado de nuestra revisión sistemática, se seleccionaron 7 estudios que contenían un total de 267 pacientes (132 controles y 135 enfermos), de los cuales el 41% eran mujeres y el 59% hombres. A nivel genético se detectaron alteraciones significativas, entre mujeres y hombres, en dos de las tres regiones (1 gen alterado en el estriado y 237 en la sustancia negra). Interesantemente, la mayoría de los genes alterados están vinculados con las principales funciones biológicas vinculadas con esta enfermedad: respuesta inflamatoria, disfunción mitocondrial y estrés oxidativo. Estos resultados podrían abrir la puerta a nuevas investigaciones que lleven a tratamientos más específicos para cada sexo.

**PALABRAS CLAVE:** Enfermedad de Parkinson, Transcriptómica, Metaanálisis, Diferencias de sexo, Tejido cerebral

# Plataforma para la vigilancia genómica de SARS-CoV-2 en España

**Erika Meyer Kvaem Soto (1), Alberto Lema Blanco (1), Sarai Varona Fernández (1), Luis Chapado (1), Luis Aranda Lillo (1), María de la Montaña Iglesias Caballero (2), Sonia Vázquez Moron (2), Francisco Pozo Sánchez (2), Sara Monzón Fernández (1), Inmaculada Casas Flecha (2), Isabel Cuesta De La Plaza (1)**

1: Unidad de Bioinformática, Unidades Científico-técnicas, Instituto de Salud Carlos III

2: Laboratorio de Referencia de Gripe y Virus Respiratorios, Centro Nacional de Microbiología, Instituto de Salud Carlos III

Para hacer frente a la pandemia de SARS-CoV-2, en España se crea una red de laboratorios de microbiología a nivel nacional, bajo el nombre de RELECOV. Esta red, coordinada por el Centro Nacional de Microbiología del Instituto de Salud Carlos III (CNM-ISCIII), tiene la finalidad de monitorizar la evolución viral y el efecto de las variantes a través de la secuenciación del genoma de SARS-CoV-2 realizada en las respectivas Comunidades Autónomas (CCAA).

El objetivo es desarrollar una plataforma que permita integrar y analizar las secuencias genómicas junto con el metadata asociado, para el seguimiento a nivel nacional de la evolución de variantes de interés, linajes, mutaciones y análisis filogenéticos. Información que se comunicará a las autoridades de salud pública, para así facilitar la toma de medidas adecuadas de acuerdo a la situación epidemiológica del momento.

Las secuencias genómicas obtenidas en los laboratorios a través de diferentes plataformas (Illumina, Nanopore e Ion Torrent) son analizadas con el pipeline nfcov/viralrecon. Los resultados del análisis y los metadatos recogidos son procesados en la plataforma de RELECOV.

El procesamiento de los datos incluye la transferencia de datos siguiendo un protocolo sftp, el tratamiento, la validación y subida a bases de datos públicas (ENA y GISAID) así como la ejecución del análisis y el almacenamiento de los resultados. Para la gestión de los metadatos se definen esquemas json anotados con ontologías (GENEPIO) siguiendo el modelo empleado en Public Health Alliance for Genomic Epidemiology (PH4AGE) y las recomendaciones FAIR. Todo ello se realizará usando las herramientas desarrolladas "relecov-tools". La plataforma almacenará las secuencias generadas durante estos 2 años de pandemia de forma retrospectiva, y las secuencias virales de forma prospectiva de acuerdo a la incidencia de la infección.

La plataforma web aportará información que facilitará llevar a cabo una vigilancia genómica coordinada a nivel nacional.

**PALABRAS CLAVE:** SARS-CoV-2, Vigilancia genómica, RELECOV, nfcov/viralrecon, Plataforma

# ¿Cómo responde el repertorio de receptores de células B a la restricción dietética en ratones durante el envejecimiento?

**Carolina Monzó (1,2), Lisonia Gkioni (1), Sebastian Grönke (1), Andreas Beyer (2),  
Dario R. Valenzano (1), Linda Partridge (1,3)**

1: Max Planck Institute for Biology of Ageing, Cologne, Germany

2: CECAD Research Center, Cologne, Germany

3: Institute of Healthy Aging, University College of London, United Kingdom

La restricción dietética, definida como una reducción del consumo nutricional sin pasar hambre, es una de las estrategias de antienvjecimiento más eficaces, ya que mejora la salud y alarga la duración de la vida. Sin embargo, se desconocen los procesos a través de los cuales la restricción dietética interactúa con el proceso de envejecimiento para mejorar la salud en la vejez. Entre los rasgos relacionados con el envejecimiento, el sistema inmunitario adaptativo experimenta una inmunosenescencia asociada al envejecimiento, que se traduce en una menor protección frente a patógenos y eficacia de las vacunas. Utilizamos la secuenciación del repertorio de receptores de células B (BCR) en los bazos de ratones jóvenes, adultos, y envejecidos, alimentados ad libitum o con restricción dietética para estudiar si las células B del sistema inmunitario adaptativo desempeñan un papel en la mejora de los fenotipos asociados a la restricción dietética en ratones hembra. También calculamos un índice de morbilidad que resume el estado de salud de los ratones. En los ratones alimentados ad libitum a lo largo del envejecimiento, se observó el patrón de repertorio BCR tradicional, con una disminución de la variabilidad intraindividual, un aumento de la abundancia clonal y de la disimilitud interindividual, y un repertorio ingenuo reducido. Encontramos que la restricción dietética retiene a los ratones en una condición más homeostática y más joven, asociada a un índice de morbilidad menor, llevando un repertorio BCR más estable en el envejecimiento, donde las métricas evaluadas se alivian. Así, en la vejez, los ratones con restricción dietética crónica tienen una mayor capacidad para reconocer y responder a nuevos antígenos, lo que sugiere que la modulación del sistema inmunitario adaptativo es uno de los elementos que contribuyen a la mejora de la salud y la vida de estos ratones.

**PALABRAS CLAVE:** Inmunidad adaptativa, Repertorio de receptores de células B, AIRR-Seq, Restricción dietética, Envejecimiento

# Identificación y caracterización de interactores proteicos del extremo C-terminal de OSC-espondina durante el desarrollo embrionario

**Felipe Maurelia (1), Maryori Gonzalez (1), Jaime Aguayo (1),  
Vania Sepúlveda (1), Carlos Farkas (2), Teresa Caprile (1)**

1: Departamento de Biología Celular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile

2: Laboratorio de Investigación en Ciencias Biomédicas, Departamento de Ciencias Básicas y Morfología, Facultad de Medicina, Universidad Católica de la Santísima Concepción, Concepción, Chile

**Introducción:** El desarrollo del sistema nervioso central comienza con la aparición del tubo neural, una estructura hueca cuyas paredes corresponden a células neuroepiteliales que encierran líquido cefalorraquídeo embrionario (LCRe). Las células neuroepiteliales darán origen a todas las células del sistema nervioso, siendo una de las primeras regiones en diferenciarse el órgano subcomisural, estructura especializada desde la cual se liberan factores proteicos hacia el líquido cefalorraquídeo embrionario (LCRe). Uno de estos factores corresponde a OSC-espondina (OSCsp), una glicoproteína de gran tamaño y con múltiples dominios capaces de interactuar con otras proteínas y modular sus efectos. Análisis de Rna-seq realizados en nuestro laboratorio identificaron diversas isoformas de OSCsp, entre las cuales destaca una conformada por el extremo C-terminal de OSCsp (C-term OSCsp). Esta isoforma presenta dominios tipo matricelular, por lo que hipotetizamos que posee una función moduladora de los factores neurogénicos presentes en el LCRe. **Objetivo:** Analizar la capacidad de la isoforma C-term OSCsp de modular factores presentes en el LCRe, regulando la diferenciación y proliferación del neuroepitelio. **Método:** Se produjo de manera heteróloga la proteína C-term OSCsp, la cual será obtenida desde el sobrenadante de cultivo, para estudiar su efecto en cultivo de neuroepitelio. Dicha isoforma será además incubada con el LCRe y analizada por Co-Inmunoprecipitación y WesternBlot. Para analizar el efecto neurogénico in vivo de C-term OSCsp, un plásmido codificante para la proteína será electroporado en embriones de Gallus gallus. **Resultados:** La proteína fue producida de manera exitosa en células de mamífero, y actualmente se está probando la interacción in-vitro con los factores del LCRe. Además, se ha detectado la expresión de la proteína in vivo en los embriones de Gallus gallus tras la electroporación in ovo. **Conclusión:** Se espera determinar los factores del LCRe que interactúan con la nueva variante de OSCsp recientemente descrita y avalar su rol neurogénico.

**Financiación:** Fondecyt 1191860.

**PALABRAS CLAVE:** OSC-espondina, Neurogénesis, Proteína matricelular, Líquido cefalorraquídeo embrionario, RNA-seq

# Análisis y caracterización de la señalización de FGF en el OSC de *Gallus gallus* en estadio embrionario temprano

Maryori González (1), Felipe Maurelia (1), Carlos Farkas (2), Estefanía Tarifeño (3),  
Jaime Aguayo (1), Vania Sepulveda (1), Teresa Caprile (1)

1: Departamento de Biología Celular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile

2: Laboratorio de Investigación en Ciencias Biomédicas, Departamento de Ciencias Básicas y Morfología, Facultad de Medicina, Universidad Católica de la Santísima Concepción, Concepción, Chile

3: Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile

**Introducción:** El sistema nervioso central se origina del tubo neural, una estructura cilíndrica hueca rellena de líquido cefalorraquídeo y recubierta por las células neuroepiteliales. Durante etapas tempranas de desarrollo, en la zona cefálica del tubo neural comienza la aparición de las vesículas cerebrales, las cuales continuarán regionalizándose hasta formar el cerebro adulto. Los cambios morfogenéticos locales de las vesículas cerebrales son comandados por los centros organizadores secundarios (COS), localizados en el límite entre las vesículas encefálicas. Los COS se caracterizan por secretar señales difusibles como los factores de crecimiento de fibroblastos (FGF), responsables de regular eventos fundamentales como proliferación, diferenciación y migración celular.

El órgano subcomisural (OSC) es una glándula cerebral que se desarrolla de forma muy temprana en el límite de la vesícula diencefálica y mesencefálica. Diversos estudios han implicado al OSC en la diferenciación, supervivencia y guía axonal. Análisis recientes han demostrado la expresión de receptores de FGF en el OSC, posicionando esta glándula como un posible COS.

**Objetivo:** Analizar y caracterizar la señalización de FGF en el OSC durante el desarrollo embrionario.

**Materiales y métodos:** Se analizaron muestras de tejido de OSC embrionario mediante RNA-seq, se validaron por qPCR y ensayos de western-blot. Además, se caracterizó la localización de los FGF en el OSC mediante análisis de inmunofluorescencia.

**Resultados y conclusiones:** Los análisis transcriptómicos demuestran que en el OSC se expresan de forma temprana gran cantidad de miembros de la familia de FGF, tanto ligandos como receptores. Además, se logró detectar y caracterizar la presencia de isoformas de receptores de FGF. Los análisis in situ sugieren que el OSC secreta FGFs al líquido cefalorraquídeo donde podría estar señalizando de forma paracrina al neuroepitelio cercano y de forma autocrina debido a la presencia de receptores en el mismo OSC en la zona apical.

**Financiación:** Fondecyt 1191860.

**PALABRAS CLAVE:** RNA-seq, OSC, FGF, Morfogénesis

# **Identificación de the PIP2 binding site in TRP Channels: Deciphering the structural effects of modulation of the TRP family by phospholipids**

**L.G. Espinoza Arcos, Ariela Vergara-Jaque, Mariela González, Matias Zuñiga, Horacio Poblete**

Center for Bioinformatics and Molecular Simulation, Universidad de Talca, 2 Norte 685, Talca-Chile.

Phosphatidylinositol 4,5-bisphosphate (PIP2) has been recognized as an important modulator of Transient Receptor Potential Receptor (TRP) channels, as well as rectifier channels and signaling proteins. PIP2 acts as an anchor phospholipid for the recruitment of proteins to the plasma membrane for subsequent activation and cellular signal transduction. TRP channels are one of the most recently discovered members of the ligand-activated ion channel superfamily and current evidence indicates that they are ubiquitously expressed. TRPs can be modulated by endogenous stimuli such as: temperature, pH, mechanical stimuli, endogenous and synthetic ligands, among others. PIP2 has been shown to be crucial for the recovery from desensitized states of TRPV1 and TRPM8 channels, two of the most studied members of the Thermo-TRP channel subgroup. Detailed knowledge of the function and modulation of TRPs has been a focus of interest for their potential use as a therapeutic target for pain-related diseases. In this work, coarse-grained molecular simulations were used to elucidate possible PIP2 binding sites in TRP channels. Protein backmapping was performed to study TRPs in atomic models, in which affinity of these channels for positively charged amino acids was found. Additionally, by means of sequence and structure information in literature, the effect of key amino acids (arginines and lysines) in the interaction between PIP2 and TRPV1, TRP1, TRPM8, TRPC6 and Kir 2.2 channels was analyzed, which by means of sequence alignments allow us to raise new questions for the study of these ion channels.

**PALABRAS CLAVE:** TRPV1, PIP2, molecular dynamics, multiscale simulation

# Análisis bioinformático de la variabilidad genética de la polimerasa NS5B del virus de la Hepatitis C

María Emilia Baez (1), Elio Rodrigo Castillo (2), Karina Alejandra Salvatierra (3)

1: Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones (FCEQyN- UNaM), Argentina

2: Laboratorio de Genética Evolutiva Dr. Claudio J. Bidau. Instituto de Biología Subtropical (IBS) CONICET - UNaM, Argentina

3: Grupo de Investigación Materiales, Modelización y Metrología (PMMM), Instituto de Materiales de Misiones (IMaM), FCEQyN-UNaM-CONICET, Argentina

El virus de la Hepatitis C representa la principal causa de carcinoma hepatocelular. Posee un ARN monocatenario de polaridad positiva que traduce una única poliproteína. Esta poliproteína, es procesada en proteínas estructurales (Core, E1 y E2) y no estructurales (vp7, NS2, NS3-NS4A, NS4B, NS5A, NS5B). Con el conocimiento del ciclo viral, la polimerasa NS5B, se convirtió en diana para el desarrollo de antivirales de acción directa. Debido a la alta variabilidad genética del virus, las variaciones que ocurren en su genoma pueden modificar la composición de aminoácidos de *NS5B* afectando la interacción con el antiviral y, por lo tanto, generando resistencia. Nos propusimos analizar la secuencia nucleotídica del gen completo de la polimerasa *NS5B*, obtenidos de aislados virales de un paciente infectado con el virus, mediante secuenciación de nueva generación (NGS). Identificar las variaciones asociadas a resistencia (VAR), modelar la proteína y analizar el acoplamiento molecular con un antiviral. Se utilizaron diferentes herramientas bioinformáticas para edición (BioEdit, Geneious, Chimera), detección de variaciones (BMA), obtención de modelo estructural de la polimerasa NS5B con VAR (Swiss Model) y docking molecular (Autodock). Se detectaron 13 VAR a INN (inhibidores no nucleotídicos), sin encontrarse VAR a IAN (inhibidores nucleotídicos). El modelo estructural de la NS5B con VAR fue confiable presentando una cobertura del 90% y el acoplamiento con el antiviral no fue efectivo, por lo que se podría inducir que este fármaco no sería útil para su tratamiento. En conclusión, observamos que el aislado viral presentaba al menos una VAR y polimorfismos de efecto desconocido. Pese a que no produjeron un cambio conformacional significativo en la estructura proteica, se produjo fallo en la interacción con el antiviral. Finalmente, destacar la importancia de la detección de VAR, para predecir la reacción viral a diferentes drogas y escoger el tratamiento más efectivo para cada paciente.

**PALABRAS CLAVE:** Hepatitis C, Antivirales, Variaciones, Acoplamiento



# Análisis de datos de péptidos bioactivos de proteínas alimentarias para el diseño y construcción de una nueva base de datos

Manuel Lonigro (1), Virginia González (2), Sebastián Bassi (2),  
Nicolás Palopoli (1,3), Agustina Nardo (1,4,5)

1: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Bs. As, Argentina

2: Toyoko LLC, Berkeley, California, EE. UU

3: Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

4: Laboratorio de Investigación, Desarrollo e Innovación en Proteínas Alimentarias, CIDCA, La Plata, Bs. As, Argentina

5: Departamento de Ciencias Biológicas, Facultad de Cs. Exactas – UNLP, La Plata, Bs. As, Argentina

**Introducción:** Los péptidos bioactivos (PB) derivados de proteínas alimentarias son secuencias encriptadas en proteínas de distinto origen que, luego de su liberación in vitro y/o in vivo, son capaces de ejercer uno o más efectos fisiológicos (actividades) en humanos, dependiendo de sus características fisicoquímicas y estructurales. La caracterización e identificación de PB resulta un área de gran interés para el desarrollo en el área de alimentos y/o la industria farmacéutica.

**Objetivos:** Existe la necesidad de sistematizar la información disponible de PB derivados de proteínas alimentarias y en particular desarrollar una base de datos versátil, que permita optimizar el estudio y la caracterización de nuevos PB para facilitar análisis cuantitativos de relación estructura-actividad.

**Metodología y Resultados:** Se realizó un relevamiento de las bases de datos de PB. Se seleccionaron aquellas con información de PB derivados de proteínas alimentarias: BIOPEP-UWM (Minkiewicz, 2019), PlantPepDB (Das, 2020) y FermFooDb (Chaudhary, 2021). Se analizó el número de entradas (péptidos) registrados, las actividades biológicas en las cuales se agrupan, parámetros de cuantificación de la actividad (IC50/EC50), distribución de longitudes y composición aminoacídica de todos los registros y por actividad. Se compararon los registros y se construyó el dataset inicial. Éste fue extendido con información estructural disponible en PDB y obtenida por simulaciones exhaustivas de dinámica molecular de las estructuras de tri- y penta-péptidos ([https://registry.opendata.aws/short\\_peptides/](https://registry.opendata.aws/short_peptides/)).

**Conclusiones:** Nuestro trabajo nos permitió el diseño y construcción de una novedosa base de datos de péptidos bioactivos con anotaciones funcionales y estructurales. La misma se pondrá a disposición a través de una página web pública con datos que serán de utilidad académica e industrial.

**Referencias:** [1] DOI: [10.3390/ijms20235978](https://doi.org/10.3390/ijms20235978); [2] DOI: [10.1038/s41598-020-59165-2](https://doi.org/10.1038/s41598-020-59165-2); [3] DOI: [10.1016/j.heliyon.2021.e06668](https://doi.org/10.1016/j.heliyon.2021.e06668).

**PALABRAS CLAVE:** Péptidos bioactivos, Base de datos, Estructura de proteínas

# Las proteínas humanas de unión a ARN con repeticiones en tándem evolucionan más lentamente que lo esperado

Juan Mac Donagh (1,2), Julia Marchetti (1,2), Martín González Buitrón (1,2), Agustín Ormazábal (1,2), Matías Carlett (i1,2), Tadeo Saldaño (1,3), Gustavo Parisi (1,2), Nicolás Palopoli (1,2)

1: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires, Argentina

2: Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

3: Ministerio de Ciencia, Tecnología e Innovación (MinCyT), Argentina

**Introducción:** Las proteínas con repeticiones en tándem (TRs) son de gran interés científico debido a sus características estructurales singulares. Se ha reportado que muchas de ellas están enriquecidas en funciones de unión a ARN. Nuestra hipótesis es que dichas características estructurales y de interacción impactan diferencialmente en su velocidad de evolución.

**Objetivos:** En este trabajo utilizamos un dataset curado de proteínas repetitivas que interactúan con moléculas de ARN con el objetivo de realizar estudios evolutivos a nivel genómico y evaluar el aporte de distintos parámetros secuenciales para explicar sus perfiles de velocidad de evolución.

**Materiales y métodos:** Nuestro dataset consiste en 3056 proteínas humanas con repeticiones en tándem que interactúan con ARN. Para estas secuencias, calculamos la relación entre sustituciones no sinónimas y sinónimas a nivel genómico, comparando las secuencias humanas con sus homólogas de cuatro especies de mamíferos cercanos (*Cercocebus atys*, *Macaca fascicularis*, *Pan troglodytes* y *Rhinopithecus bieti*). Sumamos a este análisis los valores de distintos parámetros condicionantes de las tasas evolutivas, sobre la base de estudios previos, disponibles en distintas bases de datos públicas.

**Resultados:** Encontramos que las proteínas repetitivas que unen ARN evolucionan significativamente más lento que un conjunto de proteínas humanas de referencia. Además, poseen tasas de expresión comparativamente más altas y grados de desorden secuencial significativamente más altos respecto de las referencias.

**Conclusión:** Nuestros resultados preliminares sugieren que las características estructurales de proteínas repetitivas con unión a ARN condicionan la tasa de cambio evolutivo. Estos resultados concuerdan con el supuesto general que proteínas con altos niveles de expresión evolucionan con tasas lentas.

**Referencias:** DOI: [10.1006/jsbi.2001.4392](https://doi.org/10.1006/jsbi.2001.4392); DOI: [10.3390/genes11040407](https://doi.org/10.3390/genes11040407);  
DOI: [10.1073/pnas.0504070102](https://doi.org/10.1073/pnas.0504070102).

**PALABRAS CLAVE:** Proteínas con repeticiones en tándem, Proteínas de unión a ARN, Tasa de evolución, Niveles de expresión

# **Análisis bioinformático de la correlación de expresión de NUA1 y G6PD en cáncer de mama, de colon y glioblastoma**

**Viviana Coliboro Dannich, Ariel Castro Alma**

Laboratorio de Transducción de Señales y Cáncer, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción, Chile

**Introducción:** La patología del cáncer es una problemática de salud a nivel mundial y por ello, su estudio es fundamental. Las células cancerígenas comparten características metabólicas que las llevan a un ambiente de elevado estrés oxidativo, perjudicial para su sobrevivencia y proliferación. A raíz de ello estas células desarrollan estrategias para la mantención del balance redox, entre las cuales se han estudiado a proteínas como NUA1. Esta quinasa se expresa ampliamente en cáncer, relacionada con un peor pronóstico y como participante de la homeostasis redox. En nuestro laboratorio se ha estudiado su rol en la inhibición de la enzima fosfofructoquinasa, lo que promovería la vía de las pentosas fosfato (PPP) y la generación de NADPH. Además, se ha demostrado su relación con hnrNPK, que participa en el splicing de glucosa-6-fosfato deshidrogenasa (G6PD), sugiriendo un rol de NUA1 en el control de la expresión de G6PD, mecanismo que también podría contribuir a la estimulación de PPP en la generación de NADPH.

**Objetivos:** Determinar si existe correlación entre la expresión de NUA1 y G6PD en muestras de tejido de pacientes con cáncer de mama, colon y glioblastoma.

**Métodos:** Se utilizaron las bases de datos bioinformáticas GEPIA, UALCAN, R2, MiPanda y GEO, para analizar si existe una relación entre la expresión de NUA1 y G6PD en cáncer de mama, colon y glioblastoma (importantes en incidencia, mortalidad o pronóstico).

**Resultados y Conclusión:** A pesar de las limitaciones encontradas en los análisis, se observó que existe correlación positiva en la expresión de ambas proteínas entre los estadios II y III de adenocarcinoma de colon; y cierto grado de correlación en glioblastoma (y glioblastoma recurrente), siendo estos resultados útiles para la elección de modelos celulares al momento de estudiar esta relación a nivel experimental a futuro.

**Financiación:** FONDECYT 1201215, adjudicado por el Dr. Ariel Castro Alma.

**PALABRAS CLAVE:** Cáncer, NUA1, G6PD, Correlación, Database

# Unbiased integration of single cell transcriptome replicates

Martin Loza (1,2), Shunsuke Teraguchi (1,3,4), Daron M Standley (1,3), Diego Diez (1)

1: Immunology Frontier Research Center, Osaka University, Suita 565-0871, Japan

2: Laboratory of Functional Analysis in silico, The University of Tokyo, Minato-ku 108-8639, Japan

3: Research Institute for Microbial Diseases, Osaka University, Suita 565-0871, Japan

4: Faculty of Data Science, Shiga University, Hikone 522-8522, Japan

Single cell transcriptomic approaches are becoming mainstream, with replicate experiments commonly performed with the same single cell technology. Methods that enable integration of these datasets by removing batch effects while preserving biological information are required for unbiased data interpretation. Here, we introduce Canek for this purpose. Canek leverages information from mutual nearest neighbor to combine local linear corrections with cell-specific non-linear corrections within a fuzzy logic framework. Using a combination of real and synthetic datasets, we show that Canek corrects batch effects while introducing the least amount of bias compared with competing methods. Canek is computationally efficient and can easily integrate thousands of single-cell transcriptomes from replicated experiments.

**PALABRAS CLAVE:** Single-cell-transcriptomics, Machine-learning, R-package, Batch-effects

# Aspectos moleculares del mecanismo de unión del antígeno T grande de SV40 a retinoblastoma

Carla Luciana Padilla Franzotti, Nicolas Palopoli, Gustavo Pierdominici-Sottile

Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires, Argentina  
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

**Introducción:** Las proteínas de bolsillo humanas, como retinoblastoma (pRb), son reguladores negativos del ciclo celular eucariota. Mediante interacciones proteína-proteína inactivan factores transcripcionales, reprimen la expresión de ciertos genes y son por ello consideradas supresores tumorales. SV40 (Simian Virus 40) es un virus oncogénico que codifica dos oncoproteínas no estructurales. Entre ellas, el antígeno T grande (LTSV40) contiene una región intrínsecamente desordenada (ID) en la cual se encuentra el motivo lineal (SLiM) de unión a pRb caracterizado por el patrón secuencial LxCxE. Este SLiM más sus residuos flanqueantes inactivan a pRb, provocando la modificación del ciclo celular en oncogénesis.

**Objetivos:** Mediante técnicas de simulación computacional buscamos comprender los aspectos moleculares del mecanismo de unión entre LTSV40 y pRb, caracterizando los determinantes secuenciales y conformacionales de esta interacción.

**Metodología y Resultados:** Utilizamos diferentes estructuras conocidas (PDB ID: 1GH6, 1Q1T y 4GDF) para obtener la conformación inicial del complejo pRb-LTSV40. Realizamos Dinámicas Moleculares (DM) de este sistema en las que se pudo observar que el SLiM interactúa fuertemente y con movimiento restringido. Las regiones flanqueantes al motivo poseen movilidad sensiblemente mayor, alternando entre interactuar con pRb y estar libres interactuando con el solvente. Estudios de Umbrella Sampling nos sugirieron que en el evento de separación de las proteínas, el SLiM de LTSV40 es la última región que deja de interactuar con pRb, lo que implica que actúa en el proceso de reconocimiento para la unión. Luego caracterizamos las conformaciones y flexibilidad de LTSV40 y pRb cuando se hallan libres mediante DM. Notamos que el bolsillo de pRb es relativamente rígido, mientras que en LTSV40 la región ID posee un vasto conjunto de conformaciones posibles, todas compatibles con la unión a pRb. Por último empleamos análisis de componentes principales y pudimos observar que tanto el SLiM como toda la región ID poseen movimientos globales coordinados y de gran amplitud. Estos podrían permitirle al sistema establecer una cinética de interacción relativamente rápida.

**Conclusión:** Estos hallazgos aportan información valiosa para comprender las razones energéticas y estructurales de unión entre LTSV40 a pRb. Además, podrían contribuir al conocimiento general de la interacción proteica mediada por regiones desordenadas.

**PALABRAS CLAVE:** Antígeno T grande de SV40, Supresor tumoral pRb, Dinámica molecular

# Identification of blocking ligands for I-selectin during embryo-implantation process as a possible contraceptive

Arley Gutarra Inga, Ana Paula Blanco Ortiz, Santiago Gutiérrez Quiquia,  
Aracelly Ameri Fierro, Tyfani Vera Rodríguez

Facultad de Ciencias, Universidad Nacional Agraria La Molina, Av. La Molina, Lima 23, Perú,

**Introduction:** Nowadays, the use of contraceptives is widespread among young women, particularly the use of Morning-after pill as an emergency contraceptive to avoid pregnancy. However, these types of drugs cause complications due to hormonal dysregulation caused by levonorgestrel (the active compound) and besides, they are not 100% effective.

**Objective:** In view of the above, here we aimed to identify a novel contraceptive with reduced side-effects and enhanced efficacy. For this purpose, we performed a virtual screening and molecular docking approach using ligands from a public database and targeting L-Selectin protein to block blastocyst implantation in the decidual lining.

**Methods:** We used a total of 172 ligands recovered from ZINC15 database and the 3D-structure of human L-Selectin from RCSB (PDBID: 3CFW). The ligands were prepared using a python script and OpenBabel. Virtual screening was performed with Autodock Vina (num\_modes=1, exhaustiveness=6) and a gridbox covering the known binding site of L-Selectin. The ligands were ranked according to their free energy from the simulation using a python script. From the screening, the top 6 molecules were selected to further perform molecular docking with Vina (num\_modes=10, exhaustiveness=8). The resulting poses from the docking protocol were analyzed in python and Drug Discovery Studio to visualize and measure protein-ligand interactions that may stabilize the complex. Moreover, dissociation constants (kd) were calculated according to the formula  $\Delta G = RT \ln(kd)$ .

**Results:** The 6 ligands mentioned before got all a better affinity to L-Selectin than Sialyl Lewis-x and, when considering the affinity with PGEL (full glycoprotein ligand), Lifitegrast ligand still has a greater affinity, which is explained mainly by three strong H-Bonds (Lys84, Arg97 y Lys111).

**Conclusion:** Our results point Lifitegrast as a potential contraceptive that may be used for future commercial contraceptives with less side-effects.

**PALABRAS CLAVE:** Drug-discovery, Molecular docking, Virtual screening, Contraceptives, L-Selectin blockage

# **Búsqueda de potenciales inhibidores del ensamblaje de la cápside del virus de la hepatitis B mediante el tamizaje virtual masivo basado en farmacóforo**

**Hashimoto Vargas Maria Fernanda, Sánchez Sánchez Andrea Alexa, Quiliano Meza Miguel**

Grupo de Desarrollo e Innovación de Medicamentos, Laboratorio de Biomoléculas, Facultad de Ciencias de la Salud, Universidad Peruana de Ciencias Aplicadas (UPC)

**Introducción:** La hepatitis B es un problema de salud global por su alto riesgo de progresión de la enfermedad crónica a cirrosis hepática, insuficiencia hepática y carcinoma hepatocelular. Los fármacos disponibles en el mercado para el tratamiento de la infección crónica por el virus de la hepatitis B (VHB) no son curativos, son costosos y están limitados por su toxicidad y resistencia. La cápside resulta una diana farmacológica atractiva debido a su relevancia en el ciclo de replicación del virus y a que no existen medicamentos aprobados actualmente con este mecanismo de acción. Dentro de los moduladores del ensamblaje de la cápside que están siendo estudiados destaca la familia Heteroarildihidropirimidina (HAP), resultando una alternativa terapéutica atractiva.

**Objetivos:** Buscar potenciales inhibidores del ensamblaje de la cápside del virus de la hepatitis B mediante el uso de una plataforma de tamizaje virtual masivo basado en farmacóforo.

**Metodología:** Se realizó un estudio de tipo experimental usando herramientas In silico para el diseño de modelos farmacofóricos basado en el ligando y en el receptor, los cuales fueron refinados y validados. Posteriormente se realizó el tamizaje virtual masivo usando un clúster virtual en bibliotecas virtuales de un total de 2 680 904 moléculas para seleccionar los potenciales nuevos inhibidores de la cápside de VHB.

**Resultados:** Ambos modelos fueron validados teóricamente y demostraron tener un buen desempeño (sensibilidad, especificidad, AUC-ROC). Producto del tamizaje virtual masivo de 2 680 904 moléculas se obtuvieron 3 329 hits. Hasta la fecha se ha encontrado 1 potencial inhibidor del ensamblaje de la cápside de VHB.

**Conclusiones:** El tamizaje virtual masivo basado en farmacóforo usando un clúster virtual permitió en forma rápida y eficiente el descubrimiento de 1 potencial inhibidor del ensamblaje de la cápside del VHB hasta la fecha.

**PALABRAS CLAVE:** Farmacóforos, Tamizaje virtual, Hepatitis B, Inhibidores de la cápside

# Diseño *in silico* de derivados de quercetina con posible actividad dual inhibitoria contra GSK3 $\beta$ y CDK5/p25 para el tratamiento de la enfermedad de Alzheimer

Nájera Giulliano (1), Tamayo Jaime (1), Latorre Alessandra (1), Zavaleta Juan(2)

1: Universidad Nacional Mayor de San Marcos, Perú

2: Universidad Privada Antenor Orrego, Perú

La enfermedad de Alzheimer (EA), es un desorden neurodegenerativo progresivo que afecta a millones de personas a nivel mundial. El enfoque actual para el desarrollo de nuevos tratamientos de la EA se basa en el diseño de fármacos multidiana, los cuales actúan sobre diversas vías de la fisiopatología de la EA. Las principales causas de la EA son la acumulación de placas de proteína  $\beta$ -amiloide e hiperfosforilación de la proteína Tau, relacionado a la activación de enzimas cinasas como GSK3 $\beta$  y CDK5/p25, por lo cual la búsqueda de inhibidores duales es una novedosa estrategia terapéutica para el tratamiento de la EA.

El esqueleto de flavonoides ha sido base de diseño de nuevos derivados como candidatos a fármacos contra la EA. Estudios con sustituciones en la posición C8 han mostrado buenos resultados, sin embargo, estos son escasos. Por ello, el objetivo del presente trabajo es el diseño *in silico* de nuevos derivados de quercetina sustituidos en la posición C8 como inhibidores duales de GSK3 $\beta$  y CDK5/p25.

Se diseñaron 613 bioisómeros usando la plataforma Swisbioisostere. El primer tamizaje se realizó en base a criterios de actividad inhibitoria de cinasas, toxicidad, reglas de Lipinski - Veber y farmacocinética, empleando las plataformas Way2Drug, DataWarrior, SwissADME, respectivamente. 25 bioisómeros se optimizaron con Avogadro y finalmente el docking molecular dirigido a los sitios activos respectivos se realizó empleando Autodock4.

10 bioisómeros resultaron con mejores energías de afinidad a GSK3 $\beta$  y CDK5/p25 en comparación con los controles quercetina (-9,8; -10.62 kcal/mol) e indirubina (-9.26; -10.43 kcal/mol), destacando los compuestos QT94-1 (-12.36; -13.06 kcal/mol) y QT112-2 (-12.23; -13.81 kcal/mol). En conclusión, se demostró que los derivados de quercetina sustituidos en C8 son posibles inhibidores duales de GSK3 $\beta$  y CDK5/p25 para el tratamiento de EA.

**PALABRAS CLAVE:** Quercetina, CDK5/p25, GSK3 $\beta$ , Molecular docking, Enfermedad de Alzheimer



# Estudio del efecto antiinflamatorio de los componentes de *C. citratus* mediante el Cribado Virtual sobre la red biológica de la enfermedad de Chagas

J. Leonardo Gómez Chávez, German A. Conti, Emilio L. Angelina, Nelida M. Peruchena

Lab. Estructura Molecular y Propiedades, IQUIBA-NEA, Universidad Nacional del Nordeste, CONICET, FACENA, Av. Libertad 5470, Corrientes 3400, Argentina.

**Introducción:** La enfermedad de Chagas (ECh) causada por *Trypanosoma cruzi* dispone de tratamientos poco efectivos durante la fase crónica, etapa donde extractos de *Cymbopogon citratus* (CC) han demostrado ser efectivos.

**Objetivo:** Comprender los mecanismos moleculares de la acción antiinflamatoria del extracto de CC en la miocardiopatía chagásica, mediante docking molecular sobre los Hubs de las redes de interacciones proteína-proteína (PPI) asociadas a la ECh.

**Materiales y métodos:** Se utilizó el dataset [GSE41089](#) constituido por datos de expresión génica en corazón de ratones infectados con *T.cruzi*. El procesado, expresión diferencial, enriquecimiento de Ontología génica (GO) y pathways e interacciones proteína-proteína (PPI) fueron obtenidos con Oligo, LIMMA, TopGO, SPIA y STRINGdb respectivamente, disponibles en [Bioconductor](#). Las estructuras fueron obtenidas de [Protein Data Bank](#) y [AlphaFold DB](#), mientras que los compuestos de CC mediante revisión bibliográfica. El docking molecular fue realizado con Autodock-vina y las dinámicas moleculares (DM) con AMBER16.

**Resultados:** El enriquecimiento funcional de genes sobreexpresados del dataset GSE41089 ( $\log\text{Fold-Change} > 1$  y  $p\text{-valor ajustado} < 0,05$ ) priorizó los términos GO asociados con respuesta inmune innata y las vías biológicas relacionadas con la liberación de citoquinas. La PPI construida con genes de los pathways más enriquecidos priorizó blancos moleculares sobre los cuales se realizó docking inverso de componentes de CC. Entre los complejos compuesto-proteína, destaca Nerolidol unidos a PTGS2, con una  $\Delta G_{\text{bind}}$  similar a inhibidores específicos como diclofenac e ibuprofeno.

**Conclusiones:** La combinación de farmacología en red con docking molecular para el cribado virtual de datos ómicos resulta una estrategia prometedora para comprender la acción de extractos vegetales contra ECh y permitirá buscar tratamientos más efectivos para la misma.

**PALABRAS CLAVE:** Enfermedad de chagas, Etnomedicina, Análisis de expresión diferencial, Pathways biológicos, Docking

# Hallazgo y validación de genes ortólogos distantes del gen DAF-12 en *Meloidogyne incognita* mediante el uso de herramientas bioinformáticas

Rafael Betanzos San Juan (1,2), Claudio David Schuster (1,2), Carlos Modenutti (1,2)

1: Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. II (CE1428EHA), Buenos Aires, Argentina

2: Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales (IQUIBICEN) CONICET. Ciudad Universitaria, Pab. II (CE1428EHA), Buenos Aires, Argentina

**Introducción:** *Meloidogyne incognita* es un nemátodo parasítico que daña cultivos y genera grandes pérdidas económicas. En la búsqueda de blancos moleculares para controlar otros nemátodos como *Strongyloides stercoralis* se descubrió que la interacción entre el receptor DAF-12 y los ácidos dafacrónicos (AD) es importante porque regula su ciclo de vida. Trabajos previos fallaron en identificar a DAF-12 en *M. incognita* debido a los métodos utilizados, pero mediante nuestra estrategia basada en minería de datos, filogenia y análisis estructural pudimos hallar posibles ortólogos de DAF-12 en *M. incognita*. La posterior validación experimental confirmó la eficiencia de esta estrategia.

**Objetivos:** Desarrollar una estrategia de búsqueda de genes ortólogos para hallar el gen DAF-12 en el parásito *M. incognita*.

**Métodos:** La identificación de potenciales ortólogos se hizo mediante una combinación de minería de datos con HMMER sobre el proteoma de *M. incognita* y un análisis filogenético de las proteínas obtenidas y ortólogos conocidos de DAF-12 para obtener candidatos validables. Las estructuras de los candidatos se modelaron con AlphaFold2 y se refinaron con simulaciones de dinámica molecular, se realizaron alineamientos estructurales de los candidatos y DAF-12 y se identificó el posible sitio activo mediante Fpocket. Además, ya que DAF-12 une compuestos de tipo AD, realizamos simulaciones de docking molecular entre los candidatos y diversos AD para analizar la energía de interacción, las formas y los modos de unión del ligando.

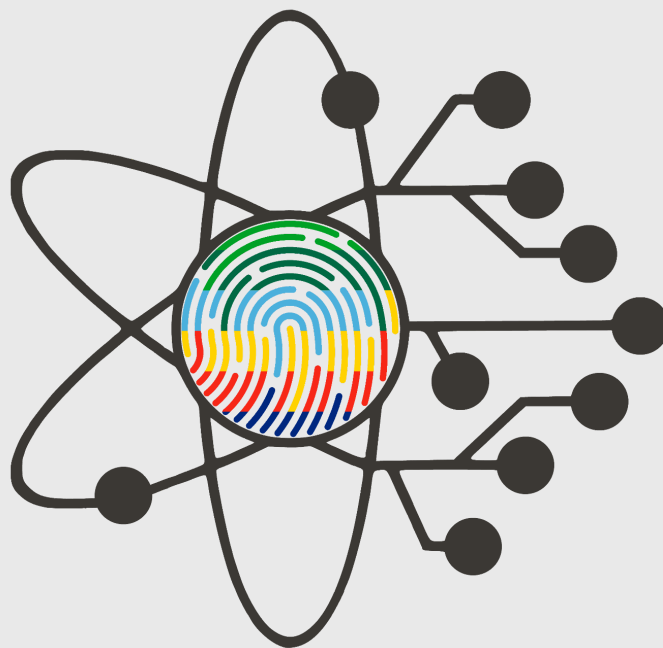
**Resultados:** Se identificaron 3 posibles ortólogos de DAF-12 en *M. incognita* con similitudes estructurales significativas en el plegado en comparación con estructuras cristalográficas de DAF-12 de otros organismos; con características fisicoquímicas similares en sus sitios activos; que pueden unir DA; y que el compuesto C24 funciona como antagonista de DAF-12 en *M. incognita*.

**Conclusión:** La estrategia antes descrita para identificar ortólogos permitió llegar desde el proteoma a 3 candidatos de DAF-12 en *M. incognita*, a su modelado estructural e incluso la validación experimental.

**PALABRAS CLAVE:** *M. incognita*, DAF-12, HAMMER, Docking, Dinámica molecular

# LIBRO DE RESÚMENES

I Simposio de Estudiantes  
Hispanohablantes de Bioinformática  
y Biología Computacional



**SEH<sup>2</sup>Bioinfo**