LIBRO DE RESÚMENES

1era Reunión Anual SCB

Sociedad Chilena de Bioinformática

12, 13 Y 14 DE ENERO DE 2022

www.scbioinfo.cl

scbioinfo

SCB

Sociedad Chilena de Bioinformática

1era Reunión Anual SCB

Sociedad Chilena de Bioinformática

12, 13 Y 14 DE ENERO DE 2022

Organiza:

SCB Sociedad Chilena de Bioinformática

ISCB Regional Student Group Chile

Auspicia:

ISCB Student Council

International Society for Computational Biology

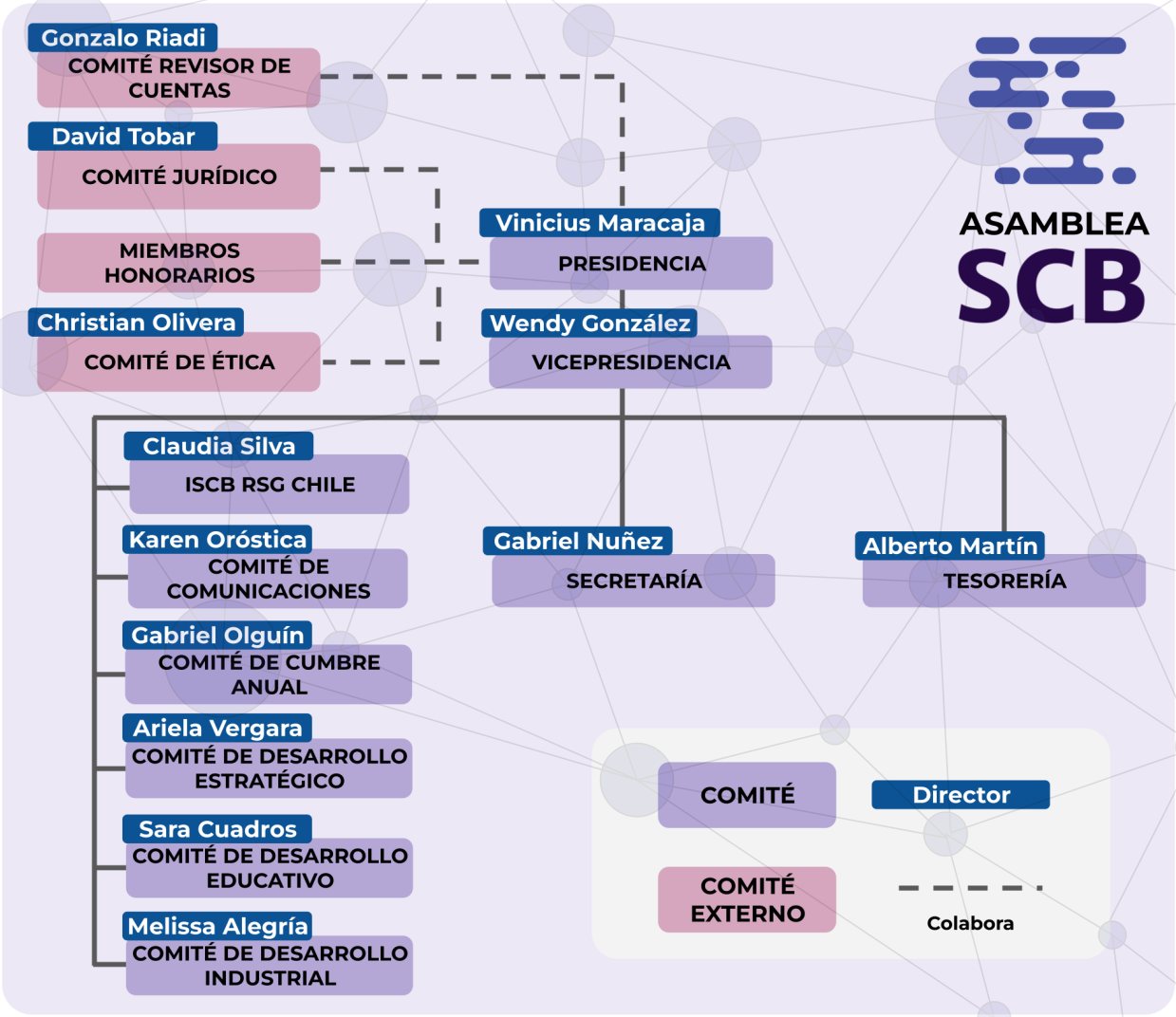
AUSTRAL - omics

ISCB International Society for Computational Biology

www.scbioinfo.cl

scbioinfo

DIRECTIVA 2021-2022



SCB

PROGRAMA

1era Reunión Anual SCB

Sociedad Chilena de Bioinformática

12, 13 Y 14 DE ENERO DE 2022

www.scbioinfo.cl

scbioinfo

SCB

Sociedad Chilena de Bioinformática

DÍA 1
12 DE ENERO DE 2022

09h00 - 09h15 **BIENVENIDA**

09h15 - 09h30 **VINICIUS MARACAJA-COUTINHO**
Universidad de Chile
Presentación de la Sociedad Chilena de Bioinformática

09h30 - 9h45 **WENDY GONZÁLEZ**
Universidad de Talca
Homenaje a los primeros Miembros Honorários de la SCB

09h45 - 10h30 **DAVID HOLMES**
Fundación Ciencia & Vida
How I became a bioinformatician and my role in the early development of bioinformatics in Chile

10h30 - 11h30 **COFFE-BREAK - SESIÓN DE POSTER 1**

11h30 - 12h00 **KAREN ORÓSTICA**
Instituto de Salud Pública
Mutational signatures and transmissibility of SARS-CoV-2 Delta and Gamma variants

12h00 - 12h30 **DANTE TRAVISANY**
Universidad de Las Américas
Hacia una plataforma bioinformática de los genomas Chilenos de SARS-CoV-2

12h30 - 13h00 **CARLOS LAGOS**
Universidad San Sebastián
Identificación in silico de moduladores alostéricos de la ARN polimerasa dependiente de ARN del virus SARS-CoV-2

13h00 - 14h00 **ALMUERZO**

14h00 - 16h00 **WORKSHOP 1**
David Ramírez, Universidad de Concepción
KNIME: Farmacoinformática para el diseño de fármacos

16h00 - 18h00 **ASAMBLEA SCB**

SCB

DÍA 2 - PARTE 1
13 DE ENERO DE 2022

09h00 - 09h30 **DIEGO MIRANDA**
Explora Maule
Proyectos de divulgación científica para aumentar el conocimiento público del naturalista Juan Ignacio Molina y temáticas de biodiversidad

09h30 - 10h00 **NICOLE TREFAULT**
Universidad Mayor
Simbiosis en ambientes extremos y el microbioma de las esponjas antárticas

10h00 - 10h30 **VERÓNICA MOLINA**
Universidad de Playa Ancha
Microorganismos desde el "ID" a su rol en ecosistemas acuáticos diversos

10h30 - 11h30 **COFFE-BREAK - SESIÓN DE POSTER 2**

11h30 - 11h45 **SELECTED TALK 1 - ABSTRACT 33**
Patricio Tapia-Reyes, Universidad Andrés Bello
Ensamblaje y anotación del genoma de Cistanthe longiscapa, una mirada genómica al desierto florido

11h45 - 12h00 **SELECTED TALK 2 - ABSTRACT 37**
Rodrigo Maldonado, Universidad Andrés Bello
Seasonal protein-coding gene expression and identification of non-coding RNAs in the Chilean Altiplano fish Orestias ascotanensis.

12h00 - 12h15 **SELECTED TALK 3 - ABSTRACT 43**
Carolina González, Fundación Ciencia y Vida
Integrative genomics shed light on evolutionary forces shaping acidophilic lifestyle

12h15 - 12h20 **SHORT BREAK**

DÍA 2 - PARTE 2
13 DE ENERO DE 2022

12h20 - 13h30 **CONVERSATORIO 1**

Retrospectiva Histórica de la Bioinformática en Chile

Modera: Melissa Alegria, Universidad de Las Américas
Panelistas: Alejandro Maass, Universidad de Chile
Angélica Fierro Huerta, PUC Chile
Danilo González, Universidad Andrés Bello
Wendy González, Universidad de Talca

13h30 - 14h30 **ALMUERZO**

14h30 - 14h45 **SELECTED TALK 4 - ABSTRACT 51**

Diego Cortez, Fundación Ciencia & Vida

A large-scale genome-based survey of acidophilic Bacteria suggests that genome streamlining is an adaption for life at low pH

14h45 - 15h00 **SELECTED TALK 5 - ABSTRACT 15**

Claudia Silva Andrade, Universidad Mayor

Comparative genomics of Clostridium baratii reveals strain-level diversity in toxin abundance

15h00 - 15h15 **SELECTED TALK 6 - ABSTRACT 21**

Ivana Orellana, Universidad Mayor

Mechano-Phisiology of titin

15h15 - 15h40 **COFFE-BREAK - SESIÓN DE POSTER 3**

15h40 - 16h00 **CHARLA DE LA INDÚSTRIA**

Andrea Silva, Austral-omics

16h00 - 18h00 **CONVERSATORIO 2**

Postgrados en Bioinformática y Biología Computacional: una mirada nacional

Modera: Christian Olivera, Universidad de Talca
Panelistas: Cristina Muñoz, Universidad Andrés Bello
Daniela Araya, Universidad de Talca
Evelyn Sánchez, Universidad Mayor
Javiera Cortés, PUC Chile
Joaquín Jensen, Universidad San Sebastián
Rúben Herzog, Universidad de Valparaiso
Sebastián Urquiza, Universidad de Chile

SCB

DÍA 3 - PARTE 1
14 DE ENERO DE 2022

09h00 - 09h30 **ALEX DI GENOVA**
Universidad de O'Higgins
Ensamblaje híbrido de genomas humanos: Algoritmos y Aplicaciones

09h30 - 10h00 **JULIO CABALLERO**
Universidad de Talca
LigRMSD: A web server for comparing similar compounds in protein-ligand docking

10h00 - 10h30 **CAROL MORAGA**
CNRS, Francia
Desarrollo de nuevos algoritmos para caracterizar el rol de los RNAs pequeños no codificantes en especies no modelo

10h30 - 11h30 **COFFE-BREAK - SESIÓN DE POSTER 4**

11h30 - 11h45 **SELECTED TALK 7 - ABSTRACT 36**
Maira Rivera, PUC Chile
The fold-switch energy profile of the metamorphic protein KaiB

11h45 - 12h00 **SELECTED TALK 8 - ABSTRACT 11**
Nicolas Espinosa, Universidad Adolfo Ibañez
Neo4GeNet: A platform for the study of contextualized gene regulatory networks and its application in SARS-CoV-2 infection

12h00 - 12h15 **SELECTED TALK 9 - ABSTRACT 13**
Alejandro Valdés Jiménez, Universidad de Talca
Parallel Algorithm for discovering and comparing three-dimensional proteins patterns

12h15 - 12h20 **SHORT BREAK**

SCB

DÍA 3 - PARTE 2
14 DE ENERO DE 2022

12h20 - 12h35 **SELECTED TALK 10 - ABSTRACT 54**
Wladimir Corrales, Universidad de Chile
Sex-biased impact of chronic stress on m6A modification machinery and neuroplastic pathways in the rat dorsal hippocampus

12h35 - 12h50 **SELECTED TALK 11 - ABSTRACT 25**
Valentina Opazo, Universidad de Concepción
Discovering the potential role of lncRNA in expression dynamics of POMC neurons during obesity

12h50 - 13h05 **SELECTED TALK 12 - ABSTRACT 48**
Kevin Aguilar-Valdés, Universidad de Talca
Identification of experimental and clinical characteristics for severity and pulmonary sequelae in Chilean COVID-19 patients using Artificial intelligence

13h05 - 13h30 **CLAUSURA & PREMIACIONES**

13h00 - 14h00 **ALMUERZO**

14h00 - 18h00 **WORKSHOP 2**
Laboratorio Nacional de Computación de Alto Rendimiento - NLHPC
Curso de introducción al HPC (Computación de Alto rendimiento)

16h00 - 18h00 **WORKSHOP 3**
Francisca Guzmán, Universidad Mayor
Modelando la Microescala

SCB

CHARLA MIEMBRO HONORARIO

How I became a bioinformatician and my role in the early development of bioinformatics in Chile

David Holmes

Fundación Ciencia & Vida, Chile

SCB

10

CHARLA PRINCIPAL 1

Mutational signatures and transmissibility of SARS-CoV-2 Delta and Gamma variants

Karen Oróstica

Instituto de Salud Pública - ISP, Chile

En este trabajo, utilizamos los datos recolectados por el programa de vigilancia genómica del ISP para cuantificar la transmisibilidad de las variantes del SARS-CoV-2 predominantes en Chile durante el año 2021. Encontramos que en el primer semestre hubo un cambio en las variantes predominantes en el país y determinamos que las variantes P.1 (Gamma) y C.37 (Lambda) son más transmisibles que las variantes originales, siendo estas las dos responsables del último peak con cuarentenas en Chile. Hallamos nuevas mutaciones que aparecieron en las variantes Gamma y Lambda al avanzar la vacunación, y encontramos una relación entre el número de mutaciones en la proteína Spike y la transmisibilidad de las nuevas variantes.

CHARLA PRINCIPAL 2

Hacia una plataforma bioinformática de los genomas Chilenos de SARS-CoV-2

Dante Travisany

Universidad de Las Américas, Chile

Los primeros casos de COVID-19 en Chile se detectaron en marzo de 2020. Desde entonces, la comunidad científica chilena ha contribuido al esfuerzo de seguimiento y mitigación de los efectos de la enfermedad, y se han organizado equipos multidisciplinarios para abordar diversas tareas desde el diagnóstico hasta el desarrollo de dispositivos médicos. Un esfuerzo especial es la secuenciación del genoma, que ha demostrado ser crucial para las autoridades sanitarias y los responsables de la toma de decisiones de muchos países para llevar a cabo la vigilancia de variantes, monitorear brotes y rutas de transmisión, mantener métodos de prueba sólidos y verificar la eficacia de la vacuna.

CHARLA PRINCIPAL 3

Identificación in silico de moduladores alostéricos de la ARN polimerasa dependiente de ARN del virus SARS-CoV-2

Carlos Lagos

Universidad San Sebastián, Chile

En este trabajo se discute el uso de herramientas de bioinformática estructural y enfoques combinados de cribado virtual basado en ligando y estructura, para la identificación de moduladores alostéricos de la enzima ARN polimerasa dependiente de ARN del virus SARS-CoV-2 (RdRp). Se utilizaron bibliotecas químicas estandarizadas para identificar posibles modos de unión de compuestos en sitios alostéricos en los cuales se predice una alta afinidad por la RdRp, y que además presentan un perfil de características fisicoquímicas que indican buen perfil farmacéutico. Estudios celulares se encuentran en curso para realizar la evaluación biológica de los compuestos seleccionados y confirmar su eficacia como inhibidores de la replicación del virus SARS-CoV-2.

CHARLA PRINCIPAL 4

Proyectos de divulgación científica para aumentar el conocimiento público del naturalista Juan Ignacio Molina y temáticas de biodiversidad

Diego Miranda

Explora Maule, Chile

La ponencia tiene como objetivo difundir experiencias memorables interactivas de divulgación científica, en formato libro, que incluyen diseños atractivos y herramientas tecnologías para incentivar el proceso lector y el conocimiento sobre el patrimonio natural y cultural de la Región del Maule.

CHARLA PRINCIPAL 5

Simbiosis en ambientes extremos y el microbioma de las esponjas antárticas

Nicole Trefault

Universidad Mayor, Chile

La simbiosis es el proceso en el que dos especies se relacionan de manera íntima y permanente. Las esponjas antárticas pueden llegar a ocupar hasta el 80% de las superficies bentónicas disponibles y cumplen roles ecosistémicos clave. Estos animales albergan comunidades diversas y abundantes de simbiontes microbianos, los que aportan muchos beneficios a sus hospederos, entre los que destacan la provisión de nutrientes, protección frente a patógenos y ciclaje biogeoquímico. Resultados obtenidos mediante metagenómica centrada en genes y en genomas del microbioma de diversas especies de esponjas antárticas indican que parte del éxito de estos animales primitivos en el continente antártico se puede deber a la compleja relación simbiótica que establecen con microorganismos de los tres dominios de la vida.

CHARLA PRINCIPAL 6

Microorganismos desde el "ID" a su rol en ecosistemas acuáticos diversos

Verónica Molina

Universidad de Playa Ancha, Chile

Los microorganismos juegan un papel clave en la biogeoquímica y sustentación de los ecosistemas ya que son los principales transformadores de la materia, movilizando nutrientes y energía. La caracterización de comunidades microbianas considerando su aporte al ecosistema es fundamental para comprender cómo se regulan los procesos microbianos y su posible respuesta a estresores o perturbaciones tanto naturales como antrópicas. En este seminario se planteará aspectos metodológicos de la caracterización molecular de los microorganismos, incluyendo grupos funcionales activos del ciclo del nitrógeno mayormente mediante 16S rRNA y qPCR, desde el cDNA y DNA, en combinación con estudiosbiogeoquímicos que han permitido avanzar en el conocimiento de los microbiomas de laboratorios naturales en Chile, incluyendo humedales de altura (3.800m.s.n.m.) como Salar de Huasco y series de tiempo oceanográficas de zonas desurgencia costera frente a Concepción y Valparaíso.

CHARLA PRINCIPAL 7

Ensamblaje híbrido de genomas humanos: Algoritmos y Aplicaciones

Alex Di Genova

Universidad de O'Higgins, Chile

El ensamblaje de novo de genomas humanos es el único método que permite caracterizar completamente la variación genética. Desarrollamos Wengan, un nuevo ensamblador que implementa algoritmos eficientes para mejorar la completitud y calidad de las secuencias construidas. Aplicamos Wengan en el ensamblaje de genomas humanos y líneas celulares de cáncer a partir de datos de secuenciación generados con las tecnologías ONT Promethion, MGI y BioNano. Los genomas ensamblados por Wengan poseen una alta calidad, un bajo nivel de error y requieren pocos recursos computacionales. En particular, Wengan reconstruyó los genomas tumorales con alta calidad lo que permitió, mediante la comparación del genoma tumoral y normal, descubrir un alto grado de reorganización genómica tumoral con miles de variantes estructurales. Esto último es difícil de detectar cuando se utilizan métodos alternativos.

CHARLA PRINCIPAL 8

LigRMSD: A web server for comparing similar compounds in protein-ligand docking

Julio Caballero

Universidad de Talca, Chile

En este trabajo, presentamos LigRMSD, un servidor web gratuito para la coincidencia automática y los cálculos de RMSD entre compuestos químicos idénticos o similares. El núcleo de LigRMSD se escribió en lenguaje Python y se implementaron algunas bibliotecas críticas de RDkit. La interfaz del servidor web es muy intuitiva, lo que permite al usuario obtener el RMSD para comparaciones entre un par de moléculas o contra una base de datos.

CHARLA PRINCIPAL 9

Desarrollo de nuevos algoritmos para caracterizar el rol de los RNAs pequeños no codificantes en especies no modelo

Carol Moraga

CNRS, Francia

Los microRNAs (miRNAs) son pequeños RNAs no codificantes que tienen un rol clave a nivel de la regulación post-transcriptcional con diversas aplicaciones biotecnológicas. Muchas veces la caracterización de miRNAs se ve truncada debido a la falta de un genoma de referencia, ya que los métodos actuales se basan en esta información para generar sus predicciones. Nosotros desarrollamos BrumiR, un algoritmo de novo que es capaz de identificar miRNAs directa y exclusivamente a partir de datos de sRNA-seq, sin usar el genoma de referencia. Las predicciones de BrumiR son mejores que los métodos existentes, y a la vez es el método más rápido (20-100X). Además, BrumiR es fácil de usar proporcionando herramientas adicionales para explorar los resultados y maximizar la interpretación biológica. En resumen, BrumiR es un método versátil que implementa nuevas ideas algorítmicas para el estudio de los miRNAs enfocado en especies no modelo.

ABSTRACT #1

Increased frequency of mutations in PIK3CA and TP53 and reduced prevalence of mutations in ESR1 in metastatic hormone receptor-positive breast cancer patients from Chile

Carla Araya, Barbara Miño, Fancy Gaete, Patricio Le Cerf and Daniel Carvajal-Hausdorf

Introduction: metastatic hormone receptor-positive breast cancer (mHR+BC) composes the largest group among advanced breast cancers, with increased burden on health care and limited treatment options. Genomic profiling of these tumors has become standard after the introduction of alpelisib, a small molecule targeting PIK3CA. However, the prevalence of oncogenic mutations in Latin American patients with mHR+BC is largely unknown. Here, studied the frequency of deleterious mutations in key driver genes in a Chilean cohort of mHR+BC patients, and correlated these results with survival. Methods: we used retrospective FFPE tissues from 84 patients with mHR+BC from 3 tertiary hospitals from Chile (n=90 samples). We performed DNA sequencing using a targeted NGS panel including whole-gene analysis for PIK3CA, TP53, PTEN, GATA-3, AKT1, and ESR1, according to NCI MATCH and ACMG standards. Results: metastatic tissues were collected from bone (18%, 16/90), liver (17%, 15/90), lung and pleura (21%, 19/90), lymph node (12%, 11/90), skin (14%, 13/90) and other (18%) including CNS, stomach, ovary, soft tissue, mediastinum and peritoneum. The frequency of detected oncogenic mutations was the following: PIK3CA, 49% (exons 7, 9 and 20; 44/90), TP53, 38% (34/90), PTEN, 10% (9/90), GATA-3, 10% (9/90), AKT1, 6% (5/90), ESR1, 6% (5/90). Mutations in TP53 were associated with decreased overall survival in patients with an HR+ HER2- IHC pattern in their metastases (p=0.0043, HR=2.331, 95% CI: 1.084-5.014) Mutational status for PIK3CA was not associated with overall survival (p=0.64). Discussion: Almost 50% of Chilean mHR+BC patients show oncogenic mutations in PIK3CA, together with increased frequency of TP53 mutations and reduced number of ESR1 mutations, compared to published series. This highlights the need of health policies aiming to incorporate molecular diagnostics as reflex after all mHR+BC diagnoses and access to targeted therapy. More studies are prompted to molecularly characterize breast cancer in Latin America.

SCB

ABSTRACT #2

Study of subjects with Alzheimer's disease using traditional and individualized co-expression networks

Verónica Latapiat, Alberto J Martín Martín, Inti Pedroso Rovira and Mauricio Sáez

Alzheimer's disease is the most prevalent form of dementia and an increasing public health concern in aging societies. A key factor in the failure of treatments for this disease is the assumption that Alzheimer's disease patients are a homogeneous group, obscuring the existence of subgroups with potential differential sensitivity to therapies. In genomics, system malfunction is usually studied by the correlation between the expression of pairs of genes in many samples as coexpression networks, but the traditional approach display averages, erasing the heterogeneity of each individual. However, personalized coexpression networks can identify gene associations for each individual.

Our objective is to identify differential modules in non-cognitive impairment and Alzheimer's disease that can allow the stratification of subjects. To achieve this, we applied the traditional and individualization methods on we used a large publicly available dataset to identify differentiated biological processes between diagnostics groups and Alzheimer's disease diagnosis.

These methods permit to measure the coexpresion preservation among different samples and clinical metadata. Our results show that individualized networks detect novel asociative patterns linked to known classifications and identify the genes that drive these connections. Using spectral clustering strategies, we also found disease-related modules that differ in terms of module numbers per sample and their number of genes using individualized coexpression networks. These findings promise to identify specific changes in molecular interactions for individuals with Alzheimer's disease based on modules, and contribute to integrating different approaches in the study of heterogeneity in conditions on a genomic level of other complex diseases.

SCB

ABSTRACT #3

Atlas: automatic modeling of regulation of bacterial gene expression and metabolism using rule-based languages

Rodrigo Santibáñez, Daniel Garrido and Alberto Martin

Cells are complex systems composed of hundreds of genes whose products interact to produce elaborated behaviors. To control such behaviors, cells rely on transcription factors to regulate gene expression, and gene regulatory networks (GRNs) are employed to describe and understand such behavior. However, GRNs are static models, and dynamic models are difficult to obtain due to their size, complexity, stochastic dynamics and interactions with other cell processes.

We developed Atlas, a Python software that converts genome graphs and gene regulatory, interaction and metabolic networks into dynamic models. The software employs these biological networks to write rule-based models for the PySB framework. The underlying method is a divide-and-conquer strategy to obtain sub-models and combine them later into an ensemble model. To exemplify the utility of Atlas, we used networks of varying size and complexity of Escherichia coli and evaluated in silico modifications, such as gene knockouts and the insertion of promoters and terminators. Moreover, the methodology could be applied to the dynamic modeling of natural and synthetic networks of any bacteria.

Code, models and tutorials are available online (https://github.com/networkbiolab/atlas).

SCB

ABSTRACT #4

Generation of Chloroplast Molecular Markers to Differentiate Sophora toromiro and Its Hybrids as a First Approach to Its Reintroduction in Rapa Nui (Easter Island)

Ignacio Pezoa, Javier Villacreses, Miguel Rubilar, Valentina Provoste, Carolina Pizarro, Maria Jesus Galleguillos, Víctor Polanco, Jaime Espejo and Carolina Sanchez

Sophora toromiro is an endemic tree of Rapa Nui with religious and cultural relevance that despite being extinct in the wild, still persists in botanical gardens and private collections around the world. The authenticity of some toromiro trees has been questioned because the similarities among hybrid lines leads to misclassification of the species. The conservation program of toromiro has the objective of its reinsertion into Rapa Nui, but it requires the exact genotyping and certification of the selected plants in order to efficiently reintroduce the species. In this study, we present for the first time the complete chloroplast genome of S. toromiro and four other Sophora specimens, which were sequenced de-novo and assembled after mapping the raw reads to a chloroplast database. The length of the chloroplast genomes ranges from 154,239 to 154,473 bp. A total of 130–143 simple sequence repeats (SSR) loci and 577 single nucleotide polymorphisms (SNPs) were identified.

ABSTRACT #5

A molecular dynamics simulation analysis of how the ionic liquid media affects interactions between the active site of β-glucosidase and the flavonoid glycoside

Samira Hozhabr Araghi, John Amalraj and Mohammad Sadegh Sadeghi Googheri

i) Background: In enzyme-assisted extraction techniques of flavonoids, β-glucosidase (BGL) is used for deglycosylation of flavonoid glycosides [1, 2]. Since the ionic liquid (IL) media is an appropriate environment for the enzymatic extraction of flavonoid glycosides from natural resources [3, 4], a detailed understanding of the effect of IL concentration on enzyme function is critical. Some findings demonstrated that BGL activity decreases in IL aqueous solutions, but they did not go into detail about why. In order to acquire a comprehensive molecular perspective, MD simulations were used to analyze the influence of IL on BGL–substrate interactions.

ii) Methods: The best docking pose for Quercetin 4'-O-β-D-glucopyranoside (Q-4'), a favored substrate of BGL, was determined using molecular docking methods. Then, the molecular perspective of the [BMIM][Cl] IL concentration effect on BGL–Q-4'complex and its interactions were simulated using 100 ns all-atom MD simulations.

iii) Results: The overall interactions and involved residues in the active site region were observed to be altered by changing the IL concentration. It changed hydrogen bond (HB) and aromatic interactions between Q-4' and both active site subsites, glycone and aglycone subsites. It had a stronger influence on the aglycone subsite, which was attributed to the aglycone subsite's greater solvent accessibility. Furthermore, it had a negative impact on the enzymatic hydrolysis process by influencing the formation of HB with Glu193 as well as the nucleophilic attack of Glu402. It was also observed that an accumulation of IL molecules at the active site entrance could reduce enzyme activity by reducing the active site's accessibility to the substrate.

iv) Conclusions: By providing a molecular perspective on the BGL–substrate complex behavior in the IL systems, our findings lead to a deeper understanding of the flavonoid glycoside hydrolysis process in these useful solvents.

ABSTRACT #6

CTCF binding prediction using genomic and epigenomic features

Camilo Villaman, Mauricio Saez and Alberto Martin

CTCF is the most relevant insulator protein in vertebrates and it is involved in the establishment and maintenance of topologically associated domains (TAD), self-interacting domains with a higher interaction ratio of elements inside the domain than outside it. Aberrant CTCF binding leads to a loss of TAD boundaries and a disrupted transcriptional landscape due to abnormal interactions inside and outside the TAD, and it has been linked with several diseases such as cancer and neuropathies. To explore the role of CTCF in disease, we built a predictor of the binding state of CTCF binding sites (CBS) using Random Forests. The Random Forest Predictor can predict with over 0.8 precision and recall in all of the four tested cell lines.

SCB

ABSTRACT #7

Pan-genomic approach of MDR Salmonella Infantis strains isolated from a poultry farm in Chile

Coral Pardo-Esté, Juan Castro-Severyn, Gabriel Krüger, Marcia Suarez, Nicolas Galleguillos, Phillipi Zepeda and Claudia Saavedra

The emergence of new Salmonella serotypes is a great concern in the food industry; in particular, the serotype Infantis is associated with critical profiles of Multiple Drug Resistance (MDR) that threaten to be challenging in terms of treatment and epidemiology control. Therefore, we aimed to characterize bacterial strains identified as Salmonella Infantis from a poultry farm in Santiago de Chile and determine possible associations between the strains and the unique genes contributing to adaptation and persistence within the production lines despite rigorous disinfection protocols. The analysis was carried out following the anvi'o pangenomic workflow, eliminating all contigs with less than 200 nucleotides, and hierarchically clustering the gene (based on their distribution) and the genomes (based on the gene clusters they share); generating a comprehensive anvi'o Pan-DataBase for downstream analyses. We found that 53.9% of the identified gene clusters belong to the core-genome most of which are single-copy core genes, and only 3.7% were unique genes (the remaining 42.3% belong to the disposable genome). We also detected antimicrobial resistance genes and virulence elements using ABRicate and the ARG-ANNOT, VFDM, and PlasmidFinder databases. The analysis of all genes and their variations among the sampled strains shows no associations between a particular strain and its isolation site, suggesting widespread contamination throughout the facility. Furthermore, the 181 unique genes identified are distributed between ten strains with frequencies from 1 to 140 unique genes per strain, some with functions associated with virulence and stress resistance. We conclude that disinfection protocols are ineffective in eliminating Salmonella from the poultry meat and this work contributes with knowledge for improving food safety. ANID-FONDECYT Regular 1210633 ECOS-ANID 170023 FONDECYT N ◦ 1191019

SCB

ABSTRACT #8

Análisis genómico e identificación de genes involucrados en resistencia a antibióticos desde aislados de Salmonella enterica aisladas de una granja de pollos: una predicción fenotípica

Gabriel I. Kruger, Coral Pardo-Este, Nicolás A. Galleguillos, Phillipi Zepeda, Marcia Suarez, Juan Castro-Severyn and Claudia P. Saavedra

Las enfermedades transmitidas por alimentos (ETA) son un problema de preocupación a nivel de salud pública y también para la industria alimenticia. Una de las bacterias consistentemente relacionadas con las ETAs pertenecen a la especie Salmonella enterica. Durante brotes asociados a Salmonella se han descrito serovares emergentes multirresistentes a antimicrobianos. Este trabajo se enfoca en la búsqueda de los determinantes genéticos de resistencia a antimicrobianos de cepas de Salmonella enterica aisladas desde una granja de pollo ubicada en la Región Metropolitana de Santiago de Chile. 90 cepas, aisladas desde 2018 a la fecha, fueron secuenciadas (Illuminna), genoserotipificandolas, y sus genomas fueron ensamblando y anotando. La búsqueda utilizando el sofware ABRicate permitió encontrar 18 elementos de resistencia a antimicrobianos, tales como genes de resistencia a aminoglucósidos, beta-lactámicos, quinolonas, sulfamidas y tetraciclinas entre otros. Además, encontramos que el 70% de los perfiles genéticos presenta resistencia a 2 o más antimicrobianos. Se realizó antibiogramas para correlacionar la información genotípica y la fisiológica. Por otro lado, encontramos plásmidos de 6 kb a 300 kb donde pudimos describir replicones del tipo Col, IncFIB, IncI1-I, IncH1A y Col3M. El serotipo Infantis se destacó por presentar un megaplásmido IncFIB que presenta además otros elementos móviles como transposones e integrones que llevan asociados elementos de resistencia a antimicrobianos. En consecuencia, los serotipos de Salmonella Infantis presentaron los perfiles de multirresistencia más amplios con más de 11 genes de resistencia por cepa. De este estudio es posible lograr identificar una correlación entre el genoma de las bacterias y su perfil de fenotípico de resistencia.

ANID-FONDECYT Regular 1210633

ABSTRACT #9

IDENTIFICATION OF MULTI-DRUG BIOACTIVE COMPOUNDS IN PARKINSON'S DISEASE USING CHEMOINFORMATICS

Victoria Oviedo, Tabata Barbosa, Carlos Peña and David Ramírez

Advances in drug development for the effective treatment of Parkinson's disease (PD) have been limited because the classic approach to design new drugs is based on the premise one drug - one target. This paradigm does not provide an effective therapeutic alternative, especially for this type of complex/multifactorial disease. PD is characterized by the degradation process that dopaminergic neurons undergo, the origin of this process could be mitochondrial dysfunction, alpha-synuclein aggregation, altered autophagy, endoplasmic reticulum stress or dysregulation of intracellular calcium homeostasis. The important challenge for drug development programs is the identification of molecular therapeutic targets relevant to such diseases as PD. Therefore, designing drugs that simultaneously target multiple pathological mechanisms responsible for the onset and progression of PD could be the best option for better therapeutic solutions. In this work, we developed an automated workflow in KNIME to construct a protein-protein interaction network for PD (PPI-PD) that allowed the identification of the key protein (expected to have a positive effect on PD). With this PPI-PD we identified the most significant targets in the network and investigated the current multitarget compounds deposited in databases such as ChEMBL. The idea is to find experimental information to develop novel programs of multitarget drug design using chemoinformatics tools, not only for PD, but also for other complex diseases.

ABSTRACT #10

COMPOUNDS' DESIGN WITH A POLYPHARMACOLOGICAL PROFILE TO TREAT ALZHEIMER'S DISEASE USING PHARMACOINFORMATICS

Alexandra Cigna, Tabata Barbosa, Carlos Peña and David Ramirez

Alzheimer's disease (AD) is one of the most representative forms of neurodegenerative diseases. Current treatments can only alleviate symptoms; however, they do not solve the etiology of the disease. This condition is a public health problem that would affect 50% of people over 85 years in a near future, and despite advances, there has been limited success to develop and repurpose drugs for an effective treatment. This shortfall derives from the fact that the process of designing drugs to target AD was based on a reductionist "1 target -1 drug" model. So far, this approach has not proven to be an effective therapeutic alternative for the multifactorial nature of this pathology. The synergy between theoretical and computational methods to design new drugs is one of the most efficient manners to develop different compounds with potential therapeutic activity against multiple targets to fight neurodegenerative diseases. Strategies using computational and theoretical approaches such as molecular simulations and chemoinformatics, are all framed within the rational design of drugs targeting AD. Based on the aforementioned, the present project will seek the possibility of identifying useful multitarget compounds to treat Alzheimer's disease by using chemoinformatics tools. In this path, common pharmacophores will be identified among the key targets to be modulated simultaneously and create an effective therapeutic effect for the treatment of Alzheimer's disease. To carry out this approach, an AD-related protein-protein interaction (PPI) network to identify key targets was created using a KNIME workflow designed in Dr. Ramirez's research group. Subsequently, a PPI-AD enrichment of heterogeneous data from multiple databases was performed. Then, using Cytoscape and R-studio, the analysis of the PPI-AD network was performed by identifying topological and functional modules. These are expected to have a favorable effect on AD in order to find new compounds, which can simultaneously modulate multiple targets.

SCB

PRESENTACIÓN ORAL

ABSTRACT #11

Neo4GeNet: A platform for the study of contextualized gene regulatory networks and its application in SARS-CoV-2 infection

Nicolas Espinosa, Alberto J. Martin and Gonzalo Ruz

Background: The increasing use of RNAseq technique to study gene expression and regulatory networks has generated large amount of data, and with these results comes the need to the develop of new platforms that allow researchers to relate their experimental results with existing information in a fastest way, this in order to find relevant conclusions, and at the same time to give them the possibility to compare their findings with those of other researchers in a more reliable way.

Methods: The building of regulatory network was made using curated Gene Regulatory Network as reference and then contextualized for every sample using RNAseq result of gene expression. The database uses nodes as entities that represent genes and these nodes have edges or relationships with other genes in the form of regulations, they can also be connected to the samples if those genes are expressed according to the RNAseq expression results, also we have linked these results with Disgenet, a gene-disease association reference.

Results: We use public RNAseq datasets of SARS-CoV-2 infection in other to generate Gene Regulatory Networks (GRNs) and upload them into our Neo4j database, in where we can compare different samples and determinate differences in regulation quickly, also deliver information about association with diseases of interest and gen ontology.

Conclusions: The comparison between samples and the relationship between genes and diseases using this platform allows us to highlight target genes that have a possible underlying relationship between SARS-CoV-2 and diabetes, cardiac dysfunction or obesity, which are the main comorbidities associated with the most severe cases of covid-19 disease and for which there is currently no good evidence in literature.

ABSTRACT #12

Reconstruction of miRNA-mRNA regulatory network in idiopathic pulmonary fibrosis

José A. Ovando-Ricárdez and Yalbi I. Balderas-Martínez

Idiopathic Pulmonary Fibrosis (IPF) is a chronic, progressive interstitial lung disease, characterized by the incessant production of matrix type I collagen in pulmonary interstitium. The lung tissue becomes rigid and scarred, affecting the correct gas exchange in the alveolus, thereby destroying the pulmonary function. microRNAs (miRNAs) have been suggested to play a critical role in the pathogenesis of multiple respiratory diseases, including the IPF. Our objective is to determine the microRNAs and their target genes through the exhaustive analysis of high-throughput experiments obtained from public databases. The expression profiles of IPF biopsies and control samples were obtained from GEO database. Six datasets (GSE35145, GSE72073, GSE48149, GSE53845, GSE110147, GSE24206) of microarrays and two datasets (GSE150910 and GSE52463) of RNA-seq technology were included in our study. Analysis was performed using R, the quality analysis and normalization were carried out individually by technology and platform. Differential expression analysis was performed using limma, and differentially expressed genes (DEGs) with FDR 0.05 and |logFC| 0.8 were considered significant. The list of overlapping DEGs in 4 or more experiments was obtained using a congruence table. Then, the results were uploaded to miRNet to establish the potential interactions between miRNAs-mRNA. Finally, the network was extracted and visualized using Cytoscape. A total of 150 DEGS overlapped in 4 or more experiments, 90 genes were overexpressed and 60 underexpressed. From this, the miRNAs-mRNA interaction network was obtained, containing 56 predicted miRNAs, 8 of them had a higher degree of connection. The miRNAs miR-126-3p, miR-155-5p, miR-381-3p, miR-335-5p, miR-429, miR-138-5p, miR-200b-3p, miR-24-3p could play a role in extracellular matrix organization and myogenesis in IPF.

PRESENTACIÓN ORAL

ABSTRACT #13

Parallel Algorithm for discovering and comparing three-dimensional proteins patterns

Alejandro Valdés Jiménez, Daniel Jimenez-Gonzalez, Miguel Reyes-Parada and Gabriel Núñez-Vivanco

Identifying conserved (similar) three-dimensional patterns among a set of proteins can be helpful for the rational design of polypharmacological drugs. Some available tools allow this identification from a limited perspective, only considering the available information, such as known binding sites or previously annotated structural motifs. Thus, these approaches do not look for similarities among all putative orthosteric and or allosteric bindings sites between protein structures. To overcome this tech-weakness Geomfinder was developed, an algorithm for the estimation of similarities between all pairs of three-dimensional amino acids patterns detected in any two given protein structures, which works without information about their known patterns. Even though Geomfinder is a functional alternative to compare small structural proteins, it is computationally unfeasible for the case of large protein processing and the algorithm needs to improve its performance. This work presents several parallel versions of the Geomfinder to exploit SMPs, distributed memory systems, hybrid version of SMP and distributed memory systems, and GPU based systems. Results show significant improvements in performance as compared to the original version and achieve up to 24.5x speedup when analyzing proteins of average size and up to 95.4x in larger proteins.

SCB

ABSTRACT #14

The enhancer transcription in Huntington's disease: differential expression and transcriptional regulation of enhancer RNA in mouse

Sebastián Urquiza, Sebastián Contreras, Sandra Arancibia, Alberto J. Martin and Mauricio Saez

In recent years, several studies aimed to identify various classes of non-coding RNAs (ncRNAs). ncRNAs perform various functions at the cellular level, such as post-transcriptional regulation, stability of messenger RNA, and epigenetic control at the chromatin level. In addition to the widely known miRNAs, sRNAs, and lncRNAs, there is yet another type of ncRNA derived from enhancers (eRNAs) whose transcription can be mainly bidirectional and carried out by RNA polymerase II. Enhancers are a type of cis regulatory element of the genomes whose activation has been linked to the production of eRNAs under different stimuli. Importantly, the presence of eRNAs, which is also accompanied by epigenetic marks such as the acetylation of lysine 27 of histone 3 (H3K27ac), serves to determine whether an enhancer is active or not. In addition, it has been described that certain eRNAs could be related to certain pathologies in which transcriptional regulation is one of the key factors. In this study, we characterize eRNAs that are differentially expressed in Huntington's disease, a progressive neurodegenerative disease that produces degeneration at the level of striatal neurons in the brain. It is already known that there are epigenetic and transcriptional alterations in this disease, but it is not yet clear how the alteration of epigenetic regulation can lead to the transcriptional dysregulation of the disease. We identified eRNAs using different enhancer databases (FANTOM5, EnhancerAtlas and our own dataset with epigenetic marks) and RNA-seq and ChIP-seq transcriptomic data of a WT mouse and a transgenic mouse presenting with Huntington's disease (Htt) at different stages of development (four and six weeks). We were able to determine putative eRNAs that are differentially expressed in both conditions, explaining uncharacterized aspects of this disease. Finally, we will find their putative transcription factors associated to active enhancers that could regulate transcription of genes related to the time curse of Hungtington's disease.

PRESENTACIÓN ORAL

ABSTRACT #15

Comparative genomics of Clostridium baratii reveals strain-level diversity in toxin abundance

Claudia Silva Andrade, Alberto J. Martin and Daniel Garrido

Clostridium baratii strains are rare opportunistic pathogens associated with botulism intoxication. They have been isolated from foods, soil and be carried asymptomatically or cause botulism outbreaks. Is not taxonomically related to Clostridium botulinum, but some strains are equipped with BoNT/F7 cluster. Despite their relationship with diseases, our knowledge regarding the genomic features and phylogenetic characteristics is limited. We analyzed the pangenome of C. baratii to understand the diversity and genomic features of this species. We compared existing genomes in public databases, metagenomes, and one newly sequenced strain isolated from an asymptomatic subject. The pangenome was open, indicating it comprises genetically diverse organisms. The core genome contained 28.49% of the total genes of the pangenome. Profiling virulence factors confirmed the presence of phospholipase C in some strains, a toxin capable of disrupting eukaryotic cell membranes. Furthermore, the genomic analysis indicated significant horizontal gene transfer (HGT) events as defined by the presence of plasmids in genomes. Seven strains were equipped with BoNT/F7 cluster. The active site was conserved in all strains, identifying a missing 7-aa region upstream of the active site in C. baratii genomes. This analysis could be important to advance our knowledge regarding opportunistic clostridia and better understand their contribution to disease.

ABSTRACT #16

Identificación de transcritos quiméricos en modelos de Síndrome de RETT

Guillermo Albornoz and Mauricio Saez

Rett syndrome (RTT) is a neuronal disorder, which is characterized by loss of speech and purposeful hand movements, the main cause of this syndrome is the gene methyl-CpG binding protein 2 (MECP2), which main function is related to epigenetics and chromatin architecture, in addition to having activity as both repressor and activator of genes. One of the main genes that have a MECP2 relationship is Long interspersed nuclear elements-1 (LINE-1 or L1), which is an autonomous retrotransposon, and is known for its high expression and mobility in brain tissue and its "cut and paste" movement, which can generate translocations, deletions, and duplications in the genome. This can lead to an increase in gene fusion or chimeric transcripts within existing cases of this syndrome.

In this work, open access databases, mainly European Nucleotide Archive, were used to obtain a total of 100 RNA-seq samples from brain tissue, of which 50 were Mus musculus MECP2-KO models and others 50 control samples. These were analyzed using bioinformatics tools for the detection of chimeric transcript candidates, which used the reports obtained from the tools to then filter the data to see if there were differences in the number of transcripts from the MECP2-KO models and control models.

The results obtained through the analysis show that there is an increase of chimeric transcripts in MECP2-KO models as opposed to control models, which could mean that due to the increase of L1 expression due to MECP2 deficiency, this retrotransposon causes an increase of jumps generating an increase of chimeric transcripts in RETT syndrome models.

ABSTRACT #17

Deciphering the relationship between epigenetic marks and Transcription Factor Binding

J. Sebastian Contreras-Riquelme and Alberto J. Martin

Background:
Transcription Factors (TFs) bind to specific patterns in the DNA to influence cell fate. In eukaryotic cells, the DNA needs to be organized to allow or to impede the binding of the transcriptional machinery. Chromatin structure depends on the action of several proteins, among them, there are proteins that introduce chemical modifications (i.e. acetylation or methylation) in histone tails. Additionally, other proteins can modify the actual DNA, adding methyl groups usually to cytosines. A third regulatory level relays in the 3D chromatin structure, being dynamically regulated by loop-forming proteins such as CTCF, present in the base of ~90% human/mouse chromatin loops. Chromatin loops block or get closer in space regulatory elements, delimiting domains of homogeneous histone modifications.
How the combinations of these factors are related to the binding of TFs remains poorly understood. To unravel this, we applied a Random Forest (RF) algorithm that aims to predict the activation state of TF binding sites (TFBSs) and by doing so, reports the relevance of each mechanism.
Results:
In this work we have used the epigenome of five human cell lines, testing several ways to develop the dataset such as the bin size for chromatin fragmentation and the extension of neighbors of the promoter. Additionally, we experiment by changing hyperparameters of the RF such as the number of trees and max depth that each tree could be extended. Our best model uses majority vote of the tree reaching a True Positive Rate (TPR) of 0.71, a False Positive Rate (FPR) of 0.05, and a Presicion (P) of 0.829. Whilst using Youden's distance threshold we got a TPR of 0.764, a FPR of 0.09 and a P of 0.906. These models also show that H3K4m2/3 and H3K9ac at 1 kilobase up/downstream of the TFBS are the most important histone makrs to define the status of the promoters, while the DNA accessibility increases its importance in a proximal region of the TFBS.
Conclusions:
The results have shown that the combinations of epigenetic modifications near to the promoter site are effective predictors of TFBS activity, whilst the effect of more distant elements needs to be studied in depth.

ABSTRACT #18

Machine Learning approach to determine relationships between epigenetics marks

Leandro Murgas Saavedra, Mauricio Sáez and Alberto J. Martín

Different complex genetic diseases are related to variations in epigenetics that lead to a change in chromatin states, being reflected in highly de-regulated transcription. In some types of cancer, such as colorectal cancer, altered patterns of histone modifications and DNA methylation have been identified and are deemed to be one of the causes behind the altered transcriptional landscape of tumors. Importantly, transcriptional activity of chromatin can be classified into different states depending on patterns of epigenetic marks, but to carry out accurate state assignment many different marks are needed. For this reason, knowing what state each chromatin region is in and observing the changes in these states is a promising source of information to better understand complex pathologies. Thus, the purpose of this work is to reduce the number of epigenetic marks by means of machine learning tools to find non-obvious redundancy relationships between different epigenetic marks, allowing to reduce the number of experiments required to determine chromatin states, increasing the feasibility of this type of analysis.

In this work, various models of the Random Forest type were trained for different epigenetic marks, using data from ChIP-seq experiments of histone modifications, varying parameters of the algorithm and the way in which the training data were provided, to find the optimal way to represent the information. In this way, it was possible to determine various relationships between epigenetic marks, being able to identify which of them have a greater relevance to be able to predict others.

The results obtained to date since this work is currently in progress. Our findings indicate that the prediction of several types of epigenetic marks using information from ChIP-seq experiments of other histone modifications, is possible. Our preliminary results have also proven accurate enough to allow robust chromatin state assignment by combining predicted marks with those from experimental results, increasing in this way the number of experiments required to study alterations on chromatin states without performing many costly experiments.

SCB

ABSTRACT #19

In silico characterization of MAGs from Planctomycetes from Salar de Llamara, Tarapacá, Chile.

Maria Campos, Roberto Veliz, Shelsy Cuellar, Mauricio Acosta and Cecilia Demergasso

Introduction: The High Andean ecosystems of northern Chile host diverse wetlands. However, the Salar de Llamara is the only one in the Atacama Desert. The periodic and multi-site metagenomic analysis allows the reconstruction of genomes and contributes to the knowledge of the eco-functionality present in this type of environment. However, the metagenomic analyses performed so far of the microbiota associated with the Puquios of Salar de Llamara have only considered isolated sites and/or a single station. Among the inhabitants are microorganisms of the phylum Planctomycetes, known for their reproduction by budding, the absence of cell walls and its essential role in the nitrogen cycle. This research aimed to deepen the understanding of the representatives of the phylum and establish their similarities and differences with known Planctomycetes. Methodology: DNA was extracted for metagenome sequencing from gypsum deposition structures, subsequent analysis using SqueezeMeta. The results were then analyzed by taxonomic, functional, abundance, phylogenetic and ANI. Results: We found representatives of the phylum Planctomycetes in Puquios with lower salinity. They are grouped into three clades by their ANI similarity percentage. Two clades have a similarity between 72 and 77% with the available reference genomes of the class Planctomycetia, and the other group, between 74 and 75% with the family Phycisphaeraceae. Conclusions: We recovered six MAGs belonging to Planctomycetes from half of four Puquios sampled on the 2018 and 2019 campaigns. It was possible to observe differences in the content of functional genes associated with different metabolisms and processes such as the cell cycle.

SCB

ABSTRACT #20

New insights into taxonomic and genomic diversity of the Blautia genus, found by a phylogenomic approach

Jose L. Maturana and Juan P. Cardenas

Blautia is a relevant and abundant genus present in the microbiome of human and other mammalian gastrointestinal (GI) tracts. Seventeen accepted Blautia species are currently available, in addition to nine other non-accepted ones. Despite the increasing level of knowledge about Blautia and its role in the GI tract, its genetic and taxonomic diversity is still poorly understood. The growing availability of Blautia genomic sequences in public databases opens the possibility to study this genus from a genomic perspective.

Here, we report the pan-genome analysis and the phylogenomic study of 224 Blautia genomes available in RefSeq. We found 33 different potential species groups at the genomic level, sixteen of them previously unknown. Our comparative analysis showed that the Blautia pan-genome is open, with a relatively small core genome (722 gene families). Additionally, utilizing a set of representative genomes, we applied a gene family gain/loss model, showing that some Blautia lineages experienced massive gene gains (up to ~900 genes), suggesting a primary role of these events during Blautia evolution. Using the HGTector tool to predict horizontal gene transfer (HGT) events, we found that approximately 22% of the genes from the representative genomes set were potentially acquired by HGT, mostly from Clostridiales or other Firmicutes. Additionally, a comparison of the profiles of carbohydrate-active enzymes among Blautia genomes showed a functional differentiation among different taxonomic clades, suggesting a taxonomy-dependent specialization in their carbohydrate degradation abilities.

SCB

PRESENTACIÓN ORAL

ABSTRACT #21

Mechano-Phisiology of titin

Ivana Orellana, Pablo Berrios and Jaime Andres Rivas Pardo

Titin—the largest protein in the human body —is a single polypeptide chain that spans half of the muscle sarcomere. With more than 1 um in length, titin´s function has been traditionally related to muscle elasticity. Titin can be understood as a spring that delimited the maximum extension of the sarcomere. Nevertheless, a few single mutations were recently described suggesting that the function of titin could be more than a simple structural spring. Familial dilated cardiomyopathy, arrhythmogenic cardiomyopathy, and myocardial infarction are only a few examples where point mutations on titin have been found. We have designed a progressive research proposal that combines protein engineering, steered molecular dynamics (SMD), and AFM-based force spectroscopy, in an effort to reproduce the physiological environment that titin experiences within the muscle. We aim to relate if the thiol redox unbalances found in myocardial infarction can be related to changes in the elasticity of titin domains. Through SMD and AFM, we assayed the elastic properties of I10 and I91 titin domains in the presence of glutathione, the molecular signature of thiol oxidation. Our results suggest that the oxidation of the thiols after mechanical unfolding prevents the refolding of I10 and I91 domains significantly, providing a possible mechanism that explains the lack of elasticity in the infarcted tissue. Our findings provide evidence that could bring the attention to review function of titin to a current role in muscle contraction.

ABSTRACT #22

Identificando la naturaleza de las interacciones moleculares entre un microorganismo beneficioso de la microbiota gastrointestinal humana y receptores tipo toll (TLRs)

Nicolás Peña-Vilches and Melissa Alegría-Arcos

La bacteria Akkermansia muciniphila es un microorganismo simbionte de la microbiota intestinal, encontrándose mayoritariamente en las capas de mucosa del tracto gastrointestinal. Diversos estudios han relacionado la abundancia de A. muciniphila con efectos beneficiosos en la salud del huésped incluyendo una menor predisposición al desarrollo de obesidad, diabetes tipo 2, además de influenciar procesos de homeostasis de glucosa , reducir niveles de triglicéridos en la sangre y contribuir a la salud intestinal general. Específicamente, Amuc_1100, una proteína de membrana externa de A. muciniphila involucrada en la formación del pili juega un rol importante en el mecanismo de interacción con el huésped y activan el Receptor de Tipo Toll 2 (Toll-like receptor 2 - TLR2) regulando la respuesta inmune en el intestino, sin embargo, el mecanismo molecular de este proceso no está totalmente determinado. La reciente resolución de la estructura cristalográfica de Amuc_1100 ofrece más oportunidades de examinar a nivel molecular la interacción de esta proteína con receptores TLR, asociación que posee un potencial terapéutico relevante, sobre todo en un contexto farmacológico para el desarrollo de nuevos tratamientos asociados a patologías relacionadas a la salud intestinal. El presente trabajo tiene por objetivo realizar un estudio computacional del complejo proteína-proteína entre Amuc_1100 y receptores TLR humano, determinando las bases estructurales que modulan la interacción tomando ventaja de las estructuras disponibles. Para esto, se utilizaron técnicas de modelado molecular, acoplamiento molecular (Docking) específicos para complejos proteína-proteína y protocolos de dinámica molecular que permiten determinar a nivel in silico el posible modo de unión y los contactos relevantes que gobiernan la asociación entre Amuc_1100 y los receptores TLR2.

ABSTRACT #23

Identification of xyloglucan endotransglycosylase/hydrolase (XTH) genes differentially expressed in raspberry fruit ripening and analysis of their promoter regions

Claudia Rivera-Mora, Aníbal Gomez, Aníbal Ayala, Paz E. Zúñiga, Karla Jara-Cornejo, Carlos R. Figueroa and Lida Fuentes-Viveros

Background

Raspberry is a fruit of economic interest with a short shelf life due it softens quickly. Changes in fruit firmness have been associated with the cell-wall modifications due to different enzymes such as xyloglucan endotransglycosylase/hydrolases (XTHs). In strawberry, the transcriptomic and the promoter regions analysis relate the expression of these genes to phytohormones and stress regulation.

We previously identified 19 XTHs genes in raspberry genome, but its expression is unknown during raspberry fruit ripening. Therefore, the main objective of this work is to identify the XTHs genes differentially expressed in flower and fruit transcriptome database and to analyze their promoter regions.

Methods

The BLAST analysis was performed on the differential expressed genes library of raspberry cv. Heritage available in Rosaceae Genome Database, and the Fragments per kilo base of transcript per million mapped fragments values were determined. Analysis of the 2000 bp promoter region of the XTHs genes was performed using the PlantPan and Plantcare servers.

Results

Fifteen XTHs were identified as differentially expressed genes. Nine of these genes are expressed more in flowers than in fruits, four have a higher expression in green fruit than in pink fruit, and eight increase their expression at the pink stage. The analysis of promoter regions indicates that they have cis-elements associated with phytohormones such as auxins, abscisic acid and gibberellin.

Conclusions

This study provides valuable information for future investigations about the role of XTHs enzymes in raspberry fruit firmness and their transcriptional regulation during ripening.

SCB

ABSTRACT #24

A tool for RNA structure prediction and refinement

Simón Poblete

The understanding of the biological function of RNA molecules requires detailed knowledge of their three-dimensional structure. For this purpose, many computational approaches have been developed during recent years, which may incorporate experimental NMR, X-ray or cryo-EM data as well. Among these tools, coarse-grained models are particularly useful due to their versatility: they allow to bridge different levels of resolution, explore the conformational space or constitute a first-line of attack for all-atom Molecular Dynamics simulations. These features become especially relevant when dealing with large RNA structures.

In this contribution, we describe a systematic approach for designing all-atom models of an icosahedral RNA virus genome and for refining structures of long noncoding RNAs generated by a coarser description using SAXS data. The simulations employ the SPQR model, a nucleobase-centred RNA representation with a set of tools written in Python. In both cases, the experimental information is fundamental for guiding the design and refinement process, which requires the disentanglement of RNA loops and the proper geometrical arrangement of secondary structure elements through specific and malleable energy restraints. The procedure described can also be extended to Molecular Dynamics and potentially, to other coarse-grained descriptions.

PRESENTACIÓN ORAL

ABSTRACT #25

Discovering the potential role of lncRNA in expression dynamics of POMC neurons during obesity

Valentina Opazo and Estefania Tarifeño

The hypothalamic POMC neurons play an important role in feeding behavior because they have the capability of decrease food intake and increase energy expenditure in response to metabolic need. Nevertheless, multiple evidence shows that these cells are dysregulated in organism with obesity and doesn't exist a clear molecular mechanism that explain this effect. In this context, our aim is search potential lncRNA that can modulate the expression of neighbor's genes in POMC neurons of obesity mice. We use RNA-seq data associated to Lyu et. al. 2020 research, obtained from POMC neurons of mice feed with standard diet or High Fat Diet (HDF) for 8 weeks. From this data and based on cis regulation performed by lncRNA, we computed a genomic colocalization using the software BedTools-Intersect between coding genes and lncRNA differentially expressed. Subsequently using FPKM values, we realized a Spearman coexpression between pairs of lncRNA-mRNA that had been previously identify in colocalization analysis. The results were 41 pairs of lncRNA-mRNA colocalized and coexpressed, representing lncRNAs with common expression patterns to their neighboring coding genes and therefore potential regulatory elements of expression changes in POMC neurons. Finally, we execute a GO enrichment to identify biological pathways modulated by lncRNA, highlighting lipid metabolism, phosphatidylcholine biosynthesis and negative apoptosis regulation, three important pathways involving in effects of obesity induced for diet on POMC neurons. Altogether, the results suggest that lncRNA contribute to gene dysregulation affecting POMC neurons during obesity and our future perspective is evaluate the interaction between lncRNA-mRNA using different predictors tools.

ABSTRACT #26

In silico identification of methylases genes and expression analysis under metal exposition in Enterococcus faecalis

Victor Aliaga-Tobar, Gabriel Gálvez and Mauricio Latorre

Transcriptional regulation is often seen as the role of transcription factors or the action of non-coding RNAs. However, a third layer is the DNA methylation, a process in which DNA bases are chemically modified in order to reduce/increase the affinity of transcriptional machinery. Recent studies have revealed the importance of bacterial methylations in several physiological processes such as virulence, motility and metals homeostasis. The latter is of main importance in pathogenic species that must deal with the increase and absence of metal ions, principally iron and copper, in an efficient way to avoid death from toxicity or starvation. Therefore, in this work, we focus methylases genes in Enterococcus faecalis, a commensal and nosocomial pathogen involved in persistent infections. First, we downloaded ~20,000 protein sequences from five bacterial methylases types described in the literature in order to identify methylases genes in E. faecalis. The analysis resulted in the identification of two methylases genes for E. faecalis, one related to 5mC and the other to 4mA methylations. Subsequently, we use transcriptomic data from metal exposition to analyze their expression. The data revealed that all genes are expressed in E. faecalis, although any of them showed an increase or decrease in either iron or copper exposition, suggesting a constitutive expression at least in these conditions. Finally, we are planning to extend the transcriptomic analysis to antibiotic data, recently available in public database. To our knowledge, this is the first work related to methylases genes and their expression under metals exposition.

ABSTRACT #27

Comparative analysis of the bacterial community found in soils near and far from the rhizosphere of plant growth in a copper mine tailing

Jaime Ortega, Gabriel Gálvez and Mauricio Latorre

Copper mine tailings are an extreme environment derived from wastes from mining activities. These tailings present a high concentration of metals and extremely low pH, which are unfavourable growth conditions for most of organisms, whether they are prokaryotes or complex eukaryotes, such as plants. Interestingly, there are sectors in Cauquenes tailings where pines and willows can live and thrive. An interesting factor that can greatly influence plant growth is the interaction between microbes and plant roots, and rhizosphere bacteria could reduce the environmental stress in these trees. In this work, samples from four sectors from the Cauquenes tailing were studied. Specifically, samples were taken from soils close to the rhizosphere of willow (zone 1) and pine (zone 2) growth, and from soils far from this rhizosphere, belonging to the Muro (zone 3) and Torre (zone 4) sectors of the tailing. In each sample, the pH was determined. Results showed that the pH of the zones 1 and 2 had mostly neutral pH, ranging from 6 to 7, whereas the pH from zones 3 and 4 were mildly acidic, ranging from 4 to 5. To research the composition of the bacterial community in this tailing, total DNA extraction was carried and sequenced, and OTUs were assigned. Interestingly, these results showed that Proteobacteria phyla was highly represented in zones 1 and 2 but had low abundance in zones 3 and 4. A similar phenomenon was observed regarding Firmicutes phyla. In this case, it had the most abundance in zones 3 and 4, while the abundance in zones 1 and 2 was low. Actinobacteria phyla had a high abundance in zone 1, but relatively low in the other zones, while Acidobacteria was found in zone 2 but not in the other zones. Interestingly, we found bacterial species associated with potential applications in biotechnology, such as: Alkane degradation (Smithella spp.), alleviation of drought stress in crops (Variovorax paradoxus), root development in beans (Bradyrhizobium elkanii), among others. Using the OTUs information as a template, a co-occurrence network model representing the putative interactions between species existing in these samples was constructed. By comparing these networks, we found unique interactions between the OTUs belonging to the rhizosphere-associated samples. Of note, these results showed that most of the negative interactions in this network were from OTUs belonging to the Corynobacterium genus. These negative interactions may be related to the ability of these OTUs to generate antibiotic-like metabolites that may be harmful to the other members of the community. On the other hand, one of the OTUs with high positive interactions was from the genus Singulisphaera. These OTUs can produce secondary metabolites which can stabilize the microbial community. Also, this genus was found to be linked with high shoot biomass in plants, highlighting the importance of this OTUs in the rhizosphere environment. Overall, our results provide valuable information regarding the microbiome in zones with tree growth in Cauquenes tailing.

ABSTRACT #28

Modelo de homeostasis de hierro en Enterococcus faecalis. Versión 2.0

Sebastián Gomez, Jorge Torres, Victor Aliaga and Mauricio Latorre

Enterococcus faecalis es una bacteria oportunista que debe pasar una variedad de obstáculos para llegar a colonizar en el cuerpo humano. Uno de estos es la deprivación de hierro, micronutriente esencial utilizado como cofactor en diversos procesos. La adaptación al déficit de este metal se basa en la capacidad de E. faecalis de reconfigurar la red transcripcional relacionada a elementos de homeostasis de hierro. Dada la importancia del hierro para el metabolismo bacteriano, los mecanismos de homeostasis de hierro están parcialmente descritos siendo la última actualización el año 2012. Por esto, el siguiente trabajo tiene como objetivo actualizar el modelo de homeostasis de hierro en E. faecalis. Para ello se realizó una exhaustiva búsqueda en la literatura en cuanto a mecanismos de homeostasis de Fe en bacterias desde el año 2012 utilizando las palabras “Bacterial”, “Iron” y “Homeostasis”. Posteriormente, mediante la herramienta BlastP se buscaron estos elementos en el proteoma de E. faecalis V583. Como resultado de esta búsqueda se generó un repositorio de 15 nuevos elementos de homeostasis en bacterias, de los cuales 6 están codificados en el genoma de E. faecalis. Entre los elementos encontrados están; DbhA y IceT, relacionados a síntesis e ingreso de sideróforos, respectivamente; hemolisina III (hlyIII) asociado a la ruptura de glóbulos rojos para obtener hierro hemínico; finalmente, ATPasas de salida de hierro como CtpD, PfeT y FetA. Es interesante notar que en el presente trabajo se identificaron 3 mecanismos de homeostasis asociados a la expulsión de Fe del citoplasma, mientras que en el modelo antiguo solo existía uno. Finalmente, se realizó un análisis de conservación de estos mecanismos dentro del orden Lactobacillales. Los resultados indicaron un alto nivel de conservación dentro de las especies estudiadas, lo que sostiene la importancia de las proteínas involucradas en la homeostasis de Fe en bacterias.

ABSTRACT #29

Identificación de ARN no codificantes involucrados en la respuesta a cobre en Enterococcus faecalis.

Jorge Torres, Sebastián Gomez, Victor Aliaga-Tobar and Mauricio Latorre

El cobre es un micronutriente y potencial tóxico bacteriano. En Enterococcus faecalis, los niveles de cobre intracelular están regulados estrictamente por el factor de transcripción CopY. Sin embargo, la acción de este regulador no explica la respuesta total transcripcional de la bacteria cuando se expone a una concentración elevada de cobre. Como posible mecanismo complementario, el presente trabajo buscó identificar ARN no codificantes (ARNnc) los cuales podrían controlar la expresión de genes de respuesta a cobre en E. faecalis. En primer lugar, utilizando una selección de 585 posibles ARNnc codificados en el genoma de E. faecalis se identificaron mediante el software intaRNA 1804 posibles genes blancos. Como resultado, se encontró un total de 8 ARNnc (3 cis y 5 trans) posiblemente controlando la expresión de 8 blancos transcripcionales involucrados en la respuesta a cobre. Estos ARNnc poseen una alta complementariedad de secuencia en cis con transcritos que codifican para una superóxido dismutasa y proteínas de estrés térmico (RBP RNA-binding protein). Interesantemente, los ARNnc con acción en trans, podrían regular transcritos directamente involucrados en la homeostasis de cobre, tales como el regular CopY (EF0297) y la ATPasas de eflujo del metal CopA (EF0298). Estos resultados permiten profundizan en la compresión de la homeostasis de Cu y evidenciar posibles mecanismos complementarios relacionados a la respuesta transcripcional de E. faecalis frente a un exceso de cobre.

ABSTRACT #30

ANILLO: Systems Biology Center for the study of extremophile communities from mining tailings.

Mauricio Latorre, Lorena Pizarro, Angélica Reyes, Alex Di Genova, Vinicius Maracaja-Coutinho, Valentina Parra and Emilio Vilches

With the opening of the "El Teniente" mine in 1905, Chile launched the large-scale exploitation of copper called the "Gran Minería del Cobre". Since then, 757 mining tailings have been registered, of which 85% have been abandoned or remained inactive. In particular, the Cauquenes tailing located in the O'Higgins Region is, to date, the oldest and largest copper tailing reservoir of the material deposited by El Teniente. In this context, identifying and characterizing communities of extremophile microorganisms that inhabit the Cauquenes tailing will provide valuable information about the structure of these communities and how they have been maintained or changed over time. In addition, studying their metabolic capacities and how they interact inside the community is essential for understanding the adaptation of organisms to an extreme environment. The latter has not been practically addressed to date in any mining tailings, and exploring this part of the Chilean Heritage might be the source of attractive biotechnological products.

For these reasons, and through the integration of the various capacities of national and international researchers, this project seeks to establish the foundations for creating a Systems Biology Center to study the communities that inhabit mining tailings. At the research level, we seek for: i) Characterization of the structure of the extremophile communities; ii) Identification and validation of the metabolic potentials of the communities and their members; iii) To catalog and classify obtained information through the development of–a genomic database for collected strains and; iv) Biotechnology applications. Overall, we aim to uncover the functional potential of species that inhabit extreme mining environments.

With a strong regional commitment and using a multidisciplinary and comprehensive perspective, our project will generate valuable molecular, genomic, and phenotypic information about microorganisms from extreme environments, data will be fully available for the Chilean bioinformatic community to promote new bridges of collaboration.

SCB

ABSTRACT #31

Construction of a transcriptional regulatory network activated by heart failure

Juan-Francisco Silva Agüero, Victor Aliaga Tobar, Ximena Calle Chalco, Valentina Parra Ortiz and Mauricio Latorre Mora

Background: Heart failure (HF) is a complex condition where the heart cannot pump enough blood to the body. HF onset is preceded by cardiac hypertrophy (CH), an adaptative response to the increased workload that is characterized by an increase in cardiomyocyte size. Cardiac stressors such as Norepinephrine (NE) activate a global transcriptional response, controlled by different signaling cascades of master transcription factors (TFs, e.g., NFAT, GATA4, MEF2). However, considering the global scale of this transcriptional reprogramming we hypothesized that additional TFs could be involved in this complex response. Methods: Using a systems biology approach, we build a human genome-scale transcriptional regulatory network (TRN) to identify novel TF involved in CH/HF. First, we selected different public articles and databases to consolidate a human TRN template that comprises TF-TF and TF-gene interactions. Second, we analyze public human HF transcriptomic datasets (RNAseq) to obtain differentially expressed genes. By integrating HF global gene expression into the human TRN, we generate a model of HF-response TRN. Interestingly, the network contains several TF involved in CH/HF along with novel uncharacterized factors, selecting BCL6 and SREBF1 as putative new regulators. To study BCL6 and SREBF1 transcriptional response under a CH condition, we performed RT-qPCR in neonatal rat cardiomyocytes treated with NE (20 μM, 48 h). Results and conclusions: We observed an increase in BCL6 and SREBF1 mRNA levels, correlated with the overexpression of classical hypertrophy marker genes (ANP, BNP and β-MHC1), and HF-response TFs (ATF4, NFATc4 and XBP1sp). This is the first report showing a human TRN for HF, a model which contains interesting new TF activated during CH, such as BCL6 and SREBF1. Our next goal is to evaluate the effects of BCL6 knockdown on cardiomyocyte hypertrophic growth and mitochondrial fragmentation.

SCB

ABSTRACT #32

Gene regulatory network of periodontitis: identification of transcription factors involved in bone resorption

Josefa Nuñez-Belmar, Emilio A Cafferata, Luis González-Osuna, Alberto J. M. Martin, Emiliano Vicencio, Rolando Vernal and Cristian Cortez

Periodontitis is the most prevalent osteolytic disease worldwide. It is a chronic inflammatory disease in which the dysbiotic subgingival microbiota triggers host immune response, responsible for alveolar bone resorption and finally, teeth loosening. Indeed, pathological tooth-supporting alveolar bone resorption is the critical hallmark of periodontitis. Although this process has been widely studied, there is little evidence about gene regulatory dynamics and recognition of hub genes involved in bone loss. This study aimed to build a gene regulatory network (GRN) with experimental transcriptomic data and identify relevant transcript factors (TF) in disease states related to periodontitis-related alveolar bone resorption. From mice affected with ligature-induced periodontitis, total RNA was extracted from periodontal tissue samples and subsequently sequenced for transcriptomic studies. Non-ligated mice were used as controls. Transcript quantification per-gen of 3 periodontitis and control samples was performed. The GRN was built with the referential GRN DoRothEA and algorithm LoTo (graphlet comparison and determination of topological differences). Downstream analysis: a subnetwork was built with most representatives nodes (>1 log2 fold.change and <0.05 p-value) and its indirect edges from GRS, and gene ontology enrichment analysis was performed. The GRN comprised 4.957 nodes, 246 TF, and 11.828 edges, and the subnetwork, 239 nodes, 81 TF, and 949 edges. The most represented KEGG pathway in the subnetwork was osteoclast differentiation. The TFs with higher outdegree related to osteoclast differentiation were Myc, Sp1, E2f1, Trp53, E2f4, Stat1, and Nfkb1. There were 5 regulatory interactions absent in the treatment network: Esr2 (TF) interaction with Fos, Pgr, Jun, Ptprc, and Serpine1. The GRN analysis was effective in identifying hub TFs and TF targets in the pathogenesis of periodontitis; in particular, osteoclast differentiation. For these reasons, and through the integration of the various capacities of national and international researchers, this project seeks to establish the foundations for creating a Systems Biology Center to study the communities that inhabit mining tailings. At the research level, we seek for: i) Characterization of the structure of the extremophile communities; ii) Identification and validation of the metabolic potentials of the communities and their members; iii) To catalog and classify obtained information through the development of–a genomic database for collected strains and; iv) Biotechnology applications. Overall, we aim to uncover the functional potential of species that inhabit extreme mining environments.

With a strong regional commitment and using a multidisciplinary and comprehensive perspective, our project will generate valuable molecular, genomic, and phenotypic information about microorganisms from extreme environments, data will be fully available for the Chilean bioinformatic community to promote new bridges of collaboration.

SCB

PRESENTACIÓN ORAL

ABSTRACT #33

Ensamblaje y anotación del genoma de Cistanthe Longiscapa, una mirada genómica al desierto florido

Patricio Tapia-Reyes, Claudio Meneses and Anibal Riveros

El fenómeno de desierto florido es un evento de floración masiva que ocurre tras las escasas precipitaciones del Desierto de Atacama. Se desconocen las características que permiten a estos organismos vegetales o a sus semillas sobrevivir en este entorno poco favorable. Cistanthe longiscapa es una planta endémica del Desierto de Atacama predominante del desierto florido en sectores desérticos y costeros. Esta especie pertenece al orden de los caryophyllales y sus semillas pueden germinar con bajas cantidades de agua. En este trabajo se generó un ensamblaje híbrido (usando lecturas de illumina y PacBio) del genoma de C. longiscapa, a partir del cual se realizó una caracterización exploratoria del genoma en conjunto al análisis de genómica comparativa utilizando sus genes predichos y anotados. El ensamble generado tiene un tamaño total de 797 Mb, distribuidos en 739 scaffolds, con un N50 de 2,73 Mb siendo de 10,61 Mb en el scaffold más largo. De este genoma un 36,75% corresponde a 42.760 genes anotados y un 41,17% a elementos repetitivos como transposones, retrotransposones y repeticiones simples. La genómica comparativa se basó en la búsqueda de zonas de sintenia entre los genomas de C. longiscapa, Amaranthus hypochondriacus y Chenopodium quinoa, destacando la conservación de un segmento de aproximadamente 20 Mb del cromosoma 1 de A. hipocondriacus con C. longiscapa, el cual es abundante en genes de adaptabilidad al estrés abiótico. Además, se generó un árbol filogenético multimétrico con expansiones y contracciones de grupos de genes ortólogos de C. longiscapa con un conjunto de otros 21 proteomas de organismos dicotiledóneos, en el cual se destacan posibles eventos de duplicación genética ancestrales de C. longiscapa, además de una diversificación evolutiva mas antigua que otros organismos caryophyllales incluidos en el estudio. De este estudio exploratorio, se generó un genoma de alta calidad con métricas competitivas con otros ensambles de novo dentro del orden de las caryophyllales. La cantidad de genes anotados y el árbol multimétrico indican eventos de duplicación del genoma que expandieron en gran medida la cantidad de copias disponible para determinados genes, lo cual puede asociarse con la plasticidad tanto en ambientes desérticos como costeros. Este trabajo es el primer paso para caracterizar desde un punto de vista molecular los mecanismos de sobrevivencia de la espcie predominante del fenómeno del desierto florido.

SCB

ABSTRACT #34

Coevolution of inter- and intradomain contacts in RfaH give rise to its novel fold in the NusG family of transcription factors

Pablo Galaz-Davison, Ernesto A. Román and César A. Ramírez-Sarmiento

BACKGROUND: The NusG protein family is structurally and functionally conserved through all domains of life, as they directly bind to RNA polymerases to regulate transcription. The RfaH subfamily is known for displaying distinct features than NusG, as they regulate the expression of virulence factors in Enterobacteria in a sequence-dependent manner. This feature is achieved through a structural interconversion between an active fold, the canonical NusG architecture, and an autoinhibited and novel fold. This structural change is called fold-switching, and proteins displaying this feature are classified as metamorphic, but how the autoinhibited fold emerged throughout evolution is unknown.

METHODS: In this work, we used deposited genomic and metagenomic sequences through Hidden Markov Models to construct a complete RfaH multiple sequence alignment (MSA). We further filtered these sequences by predicting their metamorphic features at the secondary structure level using JPred. Coevolutionary signals were calculated from this MSA using DCA and Gremlin, and these signals were further used to generate structure-based models for molecular dynamics (MD) to predict the structures of RfaH via simulated annealing.

RESULTS: Using this approach, we show that a metagenomic enrichment and metamorphic filtering protocol improves the chances of correctly identifying physical contacts of the novel fold of RfaH from sequence information. MD simulations using these contacts led to structures of autoinhibited and active RfaH with RMSD values around 5Å, and simulations of subsets of these residue pairs show that non-native contacts play a key role in achieving the autoinhibited RfaH fold.

CONCLUSIONS: Emergent coevolutionary signals found in RfaH sequences encode both the autoinhibited and active folds of this protein, shedding light on key interactions in this metamorphic protein.

FUNDING: FONDECYT 1201684, ANID Millennium Science Initiative Program ICN17_022

SCB

ABSTRACT #35

STUDY OF THE VARIATION IN THE ACETYLATION LEVELS OF H3K27 AND TRANSCRIPTIONAL CHANGES ASSOCIATED WITH CRE ELEMENTS IN HUNTINGTON DISEASE MODELS

Sandra Arancibia-Opazo, Mario Sánchez, J. Sebastián Contreras-Riquelme, Alberto J.M. Martin and Mauricio Sáez-Venegas

Huntington's disease (HD) is a neurodegenerative disorder caused by an abnormal expansion in the number of CAG trinucleotide repeats within the HTT gene. In HD models, the mutant huntingtin protein (mHtt) is capable of sequestering different proteins as a nuclear aggregate such as CREB-binding protein (CBP), causing altered acetylation of neuronal histones, cellular toxicity, and causing deregulation of the cAMP response element binding protein (CREB). Given this, CREB could decrease its transcriptional activity and therefore decrease the transcriptional activity of the CRE element. In this work, using the information of HD murine models identify the variation in the acetylation levels of H3K27 and transcriptional changes associated with CRE elements.

Our findings suggest that the decrease observed in the levels of acetylation of H3K27 (H3K27ac) in the genes close to CRE elements could be due to lower levels of CBP available when it is sequestered by mHtt. These findings explain the decrease of the CREB-CRE function and the observed lower activation levels of CRE related genes. Subsequently, we explain these effects by using context-specific Gene Regulatory Networks that model how transcriptional regulation is carried out in normal and illness conditions. We compared these networks to gather information about key regulatory changes in the early stages of HD considering the variation in H3K27ac levels produced by the decrease in CBP activity and the subsequent misregulation in CRE cascade.

SCB

PRESENTACIÓN ORAL

ABSTRACT #36

The fold-switch energy profile of the metamorphic protein KaiB

Maira Rivera, Pablo Galaz-Davison, Ignacio Retamal-Farfan and Cesar Ramirez-Sarmiento

I. Background

The metamorphic protein KaiB is a key regulator of the cyanobacterial clock, whose fold-switch leads to the cyanobacterial night physiology and defines its periodicity. The fold-switch of KaiB includes not only a change in secondary and tertiary structure of its C-terminal half, but also in oligomerization state. Although the structures of both native states are well-studied, little is known about its fold-switch mechanism.

II. Methods

Here, we used confinement molecular dynamics (CCR-MD) and folding simulations using structure-based models (SBMs) to unveil the fold-switch energy profile of KaiB. Additionally, FoldX and the Frustratometer web server were employed to assess mutational changes in localized energetic frustration.

III. Results

The thermodynamics of the fold-switch of KaiB, as ascertained by CCR-MD, describe that the transformation energy of KaiB from a ground-state tetramer (gsKaiB) to a fold-switch monomer (fsKaiB) is 33 kcal/mol, being the tetramer and dimer the most stable states. The energy barriers for this process were then explored by dual-basin SBM simulations, showing that the limiting step for the transformation is the dimer dissociation.

The energy determined by CCR was decomposed at the per-residue level and demonstrated that only residues 51-100 exhibit a differential stability pattern with preference for either the gsKaiB or the fsKaiB fold. We then computationally mutated six residues which could be important for the transformation kinetics, with some mutations being compatible with minimized energetic frustration in both folds.

IV. Conclusions

By using CCR-MD and dual basin SBM-MD we were able to diagram a full energy profile for the transformation of KaiB. This combination allowed us to predict specific sites to study the transformation kinetics on this metamorphic protein.

Funding: ANID FONDECYT 1201684 and 3190731

SCB

PRESENTACIÓN ORAL

ABSTRACT #37

Seasonal protein-coding gene expression and identification of non-coding RNAs in the Chilean Altiplano fish Orestias ascotanensis

Rodrigo Maldonado, Elvis Acevedo, Vinicius Maracaja-Coutinho and Martin Montecino

Orestias ascotanensis (Cyprinodontidae) is a teleost pupfish endemic to springs that feed into the Ascotan saltpan in the Chilean Altiplano (3,700 m.a.s.l.), and therefore, represents an opportunity to study adaptations to high-altitude aquatic environments. We have sequenced 42 RNA and 42 small RNA libraries prepared from gills, skin and muscle tissues from 14 individuals collected in summer and winter. We characterized the gene expression of protein-coding genes (PCG), long non-coding RNAs and microRNAs. Analysis of PCG expression shows that all samples grouped according to the tissue they come from in one of three clearly separated clusters of the principal component analysis (PCA) performed using DEseq2. On the other hand, gene expression of samples from tissues collected in different seasons are not clearly separated into clusters in the PCA. Despite this, we were able to identify 2,531, 934 and 2,870 differentially expressed genes (FC>0, P>0.05) between season in gills, muscle and skin respectively. Interestingly, seasonal gene expression of muscle tissue shows a better separation between summer and winter samples in the PCA compared to the other tissues, and consequently present the highest number of differentially expressed genes. Genes that increase their expression (FC>2, P>0.05) during summer in muscle tissue includes gene ontology (GO) categories such as GPCR signaling, blood circulation and transport processes. Downregulated genes(FC<0.75, P>0.05) in this tissue during summer includes genes associated to fatty acid biosynthesis and regulation of cell cycle and regulation of cell cycle.

Using several tools (including CPAT, Transdecoder, RNAmining and FEELnc), we were able to identify 3,762 long non-coding RNAs (lncRNAs) expressed in O. ascotanensis. These lncRNAs include novel and previously described lncRNAs that are conserved in other vertebrates (i.e., Yam1, Msx1-as and Megamind/TUNA and Cyrano (OIP5-as)). Furthermore, we performed a genome-based prediction of the miRNome encoded in O. ascotanensis genome using Bowtie1, BLAT and StructRNAfinder and then we looked for evidence of expression of these predicted microRNAs (miRNAs) loci with our small RNA seq data. Using miRDeep2 we identified 222 miRNA candidates with seed regions previously described in teleost and 139 miRNA candidates with non-homologous seed sequence in teleost, and therefore potentially novel miRNAs present in O. ascotanensis.

Future work will focus on establishing if RNA-RNA interaction networks might regulate the changes in gene expression observed during seasonal acclimatization in this fish. Due to the conservation of lncRNAs in Orestias, we believe that the characterization of the mRNA-miRNA-lncRNA networks during seasonal acclimatization in this fish, will help us to increase our understanding of the non-coding mediated regulatory networks in other species.

SCB

ABSTRACT #38

A comparative study of the norepinephrine transporter (NET) and acetylcholinesterase as targets of dual molecules: An in silico approach.

Luis Dinamarca Villarroel, Angélica Fierro and Rocío Gutierrez

Recent studies in the field of Alzheimer's disease (AD) research show that norepinephrine (NE) has different functions modulating memory processing in the hippocampus, amygdala, and prefrontal cortex, among other brain areas. The brain's principal nucleus that provides NE is the Locus coeruleus, which is affected in AD early stages, producing significant neuronal dead. New evidence proves that selective norepinephrine reuptake inhibitors improve memory and reduce the accumulation of Aβ protein in the brain. Combining this new target with the available therapies like acetylcholinesterase (Ache) inhibitors may lead to new molecules that modulate the noradrenergic and cholinergic systems, both impaired in AD. The computational comparative study of these two macromolecules binding sites provides new and relevant information at a structural level to design new active compounds. This study was carried out using molecular modeling to determine the tridimensional structure of NET, docking studies for both proteins, molecular dynamics simulations at an atomistic level, combined with statistical algorithm PocketMatch to evaluate the binding similarities between NET and AChe. Both binding sites have a similar electrostatic potential distribution, a similar type of residues, and a significant presence of aromatic and negatively charged residues, making both proteins binding sites rich in electrostatic density. Finally, we proposed a common site to evaluate an available NET inhibitor in Ache, proving that both sites are similar.

ABSTRACT #39

Identifying key residues for substrate specify and transport mechanism on GTRs thought Fast Dynamic Docking Guided by Adaptive Electrostatic Bias

Carlos Peña, Christa Kanstrup, Osman Mirza, Ingo Dreyen, Hussam Nour and David Ramírez

Nitrate/peptide transporter family (NPF) is one of the largest transporter families in the plant kingdom, with members capable to transport nitrate, peptides, phytohormones and glucosinolate defense compounds. However only a few NPF transporters have been studied so far, and little is known about their interaction with substrates and physiological functions. For instance, the glucosinolate transporters (GTRs) have shown to be essential for glucosinolate uptake in Arabinosis thaliana. However, the structural determinants for substrate specify and transport mechanisms remain unknown.

Aiming to understand the transport mechanism of GTRs, we selected GTR3 member of NPF because while GTR1 and GTR3 shares a 60% of identity, GTR1 transport 4MTB and I3M glucosinolates, meanwhile GTR3 only transports I3M. GTR3 was modeled in inward-facing conformation, and subject to electrostatics-inspired bias method based in MD simulations, where 4MTB and I3M are driven to the binding site. For each system, 50 replicas were performed, and we took the MDs with the lowest energies as the most favorable route of entry in which we identify the most interacting residues on the entry route (out of equilibrium) and inside the binding site (in equilibrium). This method provides good accuracy predictions at an affordable computational cost, allowing the identification of key molecular determinants in the transport mechanisms and specificity of GTR transporters.

SCB

ABSTRACT #40

IMPORTANCE OF MOBILE GENETIC ELEMENTS ON RESISTANCE GENES DISSEMINATION IN ANTARCTIC SOILS

Felipe Andrés Burgos, Andrés Santos, Leticia Barrientos and Jaime Martinez-Urtaza

Antibiotic resistance genes (ARGs) are undergoing a particularly rapid geographic expansion in various ecosystems, including pristine environments such as Antarctica. The study of ARGs mechanisms could provide a better understanding on their origin, evolution, and dissemination in these pristine environments. Thus, the aim of this study was to describe the importance of mobile genetic elements as a possible dissemination mechanism for ARGs in Antarctica. For this purpose, five soil metagenomes belonging to Deception Island in Antarctica were analyzed. Metagenomes were sequenced on the illumina NovaSeq 6000 platform and subsequently assembled with metaSpades. The search for antibiotic resistance genes was performed with Deeparg v2.0, using the ARBD and CARD databases, while the BacMet v2.0 database was used for the environmental resistance genes identification. Afterwards, the presence of plasmids and phages associated with the detected resistance genes was assessed using Plasflow tool v1.1 and DeepVirFinder v1.0, respectively. Results showed that the main detected antibiotic resistance genes detected in the five metagenomes correspond to efflux proteins, antibiotic efflux regulators (ARLR, GOLS and CRR) and antibiotic deactivators (KSGA and ARR). Regarding to environmental stress resistance genes, the dominant types of resistance were associated to nickel, silver, zinc, sodium dodecyl sulfate, copper and deoxycholic acid. Afterwards, it was determined that 80% of antibiotic resistance genes and 25% of environmental resistance genes were associated with plasmids and phages sequences. Aforementioned suggest a relevant role of these mobile genetic elements as reservoirs and potential drivers for resistance genes dissemination in this environment. Finally, these results are valuable as they provide us with information on mobile elements and their potential contribution to the spread of clinically important antibiotic resistance genes and environmental resistance genes in ecosystems.

ABSTRACT #41

Mapping Targets and Drugs by Indication

Jonathan David Hurtado Pachón, Carlos Peña Varas and David Ramírez

Over the past few years, a rational design of the Multi-Target targeted Directed Ligand (MTDL) therapeutic approach has been promoted as a powerful strategy in the development of potential therapies for neurological disorders. However, this approach (MTDL) requires an exhaustive and rigorous methodology to track and discover all targets, along with drugs and compounds that have been shown to have pharmacological potential, to find common features between targets and successful drugs and use those features to design new pharmaceutical entities that can act on multiple targets that are part of the complex pathophysiology of a particular disease state, thus increasing the likelihood of success in a clinical trial.

Network pharmacology is a powerful tool in drug discovery targeting neurodegenerative diseases (NDs), due to the complex process of selecting key target.

We present a flexible, automated, fast, and efficient workflow using the open-source KNIME software. This workflow was designed to work with neurodegenerative diseases. However, its use can be extended to several pathologies. The main goal of this workflow is to capture relevant information about an indication of interest by integrating different databases such as ChEMBL, TTD, DRUGBANK, STRING, and OPEN TARGETS, to obtain a complete list of targets and drugs, (with its phase of the study), as well as the most important protein-protein interactions in the modulation of the biological process associated with the disease.

ABSTRACT #42

Saliva proposes a bacterial-immune conversation in recurrent aphthous stomatitis

César Rivera, Andrea Camargo, Romina Hernández-Olivos and Estefanía Nova-Lamperti

There are currently no preventative options for recurrent aphthous stomatitis, and the only available treatments are palliative. This is partly due to a poor understanding of its etiopathogenesis. The objective of our research was to characterize the proteome present in the saliva of people with recurrent aphthous stomatitis.

We conducted a case-control study. The saliva of persons with recurrent aphthous stomatitis (n = 36) was examined using mass spectrometry-based proteomics in the presence and absence of lesions. We also compared these profiles with healthy people with no history of oral ulcers (n = 31). The identified sequences were aligned against the human proteome available in Uniprot and the oral microbiome available in the Human Oral Microbiome Database.

We discovered that the existence of lesions is linked to the biological processes of antigen presentation and T cell activation. The MAIT cells (mucosal-associated invariant T-cells) are the protagonists of this activity, according to the interaction networks, which could be activated against bacteria in the oral microbiota that create vitamin B2 and B9 metabolites. Our findings show that salivary fluid biopsy can be used to learn more about the interactions between the oral microbiota and the defensive immunological response, which can lead to epithelium damage. From a medical perspective, we believe that recurrent aphthous stomatitis is caused by specific populations of bacteria with vitamin synthesizing activity, which would cause a population of T cells to become activated. More trials are needed to confirm this conceptual framework.

We appreciate funding from FONDECYT (ANID-Chile) N°11180170 and Red Estatal de Odontología (Chile) N°REO19-012.

SCB

PRESENTACIÓN ORAL

ABSTRACT #43

Integrative genomics shed light on evolutionary forces shaping acidophilic lifestyle

Carolina González, Eva Vergara, Mark Dopson, Jorge Valdés and David S. Holmes

Extreme acidophiles thrive in environments rich in protons (pH values < 3) and often high levels of dissolved heavy metals. They are distributed across the three domains of the Tree of Life including members of the Proteobacteria. The Acidithiobacillia class is formed by the neutrophilic genus Thermithiobacillus along with the extremely acidophilic genera Fervidacidithiobacillus, Igneacidithiobacillus, Ambacidithiobacillus, and Acidithiobacillus. Phylogenomic reconstruction revealed a division in the Acidithiobacillia class correlating with the different pH optima that suggested the acidophilic genera evolved from an ancestral neutrophile within the Acidithiobacillia. Genes and mechanisms denominated as “first line of defense” were key to explaining the Acidithiobacillia acidophilic lifestyle including preventing protons influx that allows the cell to maintain a near neutral cytoplasmic pH and differs from the neutrophilic Acidithiobacillia ancestors that lacked these systems. Additional differences between the neutrophilic and acidophilic Acidithiobacillia included the higher number of genes copies in the acidophilic genera coding for “second line of defense” systems that neutralize and/or expel protons from cell. Gain of genes such as hopanoid biosynthesis involved in membrane stabilization at low pH and the functional redundancy for generating an internal positive membrane potential revealed the transition from neutrophilic properties to a new acidophilic lifestyle by shaping the Acidithiobacillaceae genomic structure. The presence of a pool of accessory genes with functional redundancy provides the opportunity to “hedge bet” in rapidly changing acidic environments. Although a core of mechanisms for acid resistance was inherited vertically from an inferred neutrophilic ancestor, the majority of mechanisms, especially those potentially involved in resistance to extremely low pH were obtained from other extreme acidophiles by horizontal gene transfer (HGT) events.

SCB

ABSTRACT #44

Physical exercise as a possible non-pharmacological preventive treatment of Alzheimer's disease based on protein-protein interaction.

Alexis Sepulveda, Tabata Barbosa and David Ramírez

BACKGROUND
Alzheimer's disease (AD) is a neurodegenerative disease characterised by the progressive loss of neuronal connections. There are 3 hypotheses about its physiopathologic mechanism: cholinergic, β-amyloid, and the neurofibrillary tangles hypothesis. On the other hand, physical activity (PA) could be a possible response to prevent AD, based on its physiological effects, for example, it is hypothesised that PPARGC1A, signalled by PA, inhibits the creation of β-amyloid, responding to the β-amyloid hypothesis that suggests that abnormal amyloid precursor protein cleavage gives rise to insoluble β-amyloid peptides that bind forming senile plaques. Based on this premise, a protein-protein network (PPI) between AD and PA was characterized and analyzed using bioinformatics tools.
METHOD
1) The PPI network correlating AD and AP was characterized using KNIME software. 2)The most relevant proteins of the PPI network were determined by topological analysis using R-studio software and visualizing the PPI network with the most relevant proteins labelled using Cytoscape software. 3) Experimental approaches will be designed to verify the results obtained above.
RESULTS
1) 927 interactions between 606 proteins grouped into 8 components were identified for the PPI network between AE and PA. 2) 22 relevant proteins were determined, being 17 proteins directly related to PA and 5 indirectly related to PA. 3) Work is currently underway on this specific objective.
CONCLUSIONS
The achievement of these objectives will allow the identification of the proteins related to PA with greater implication on AD, which will subsequently allow the use of PA as a possible non-pharmacological preventive treatment for AD to be determined.

ABSTRACT #45

Identification of a peptide binding motif of zonula occludens toxin of Vibrio parahaemolyticus PMC53.7 to proteinase-activated receptor 2

Cristian Iribarren Rojas, Katherine Garcia and Melissa Alegria-Arcos

i) Background: Zonula occludens toxin (Zot) is an enterotoxin elaborated by Vibrio cholerae which increases intestinal permeability. Zot disassembles the intercellular tight junctions through redistribution of the actin cytoskeleton. Zot mimics the effect of zonulin by binding to the FCIGIRL motif, which activates the PKC-α signaling pathway to proteinase-activated receptor 2 (PAR2). V. parahaemolyticus clinical non-toxigenic strain PMC53.7, possesses a Zot protein that induces redistribution of actin cytoskeleton in Caco-2 cells in absence of the motif FCIGRL. ii) Methods. Different bioinformatics tools such as Toxin pred*, MEME** were used to predict which of the peptide motifs of Zot-PMC53.7 that would bind PAR2. Additionally, we built independent peptides from the Zot-PMC53.7 sequence and used molecular docking to predict which motif could be binding to PAR2. iii) Results. Using different tools and predictive methods in silico, we selected one peptide candidate (TASASSLP) in the carboxyl terminal-end of PMC53.7 that would bind to PAR2. iv) Conclusions: Our results showed an approximation in silico of a possible peptide motif to bind directly to PAR2. This peptide motif could explain how Zot-PMC53.7 proteins induce the disassembly of the intercellular tight junctions by the redistribution of actin cytoskeleton. Finally, we will propose this candidate peptide for testing in vitro using Caco-2 cell culture.

ABSTRACT #46

Common pharmacophores between BuChE, BACE-1 and CB-2 open new avenues in the treatment of Alzheimer's disease by multitarget directed ligands

Jordan Alegria Espinoza, Lily Arrué Ayala, José Carlos Estanislao Márquez Montesinos and David Ramírez Sánchez

Neurodegenerative diseases associated with dementia are highly challenging, among them Alzheimer's disease (AD) is the one that most frequently affects the elderly. Regarding this pathophysiology, three hypotheses have been proposed: cholinergic, amyloid, and neurofibrillary tangles, which all coincide in an interruption of synaptic transmission. There are very few drugs approved for the treatment of AD, including donepezil, tacrine and rivastigmine, which are cholinesterase inhibitors, and galantamine and memantine, which are inhibitors of NMDA receptors. Although several therapeutic strategies have been developed, the therapeutic effects are not targeting to improve cognitive functions. Thus, due to the multifactorial nature of this disease and the lack of effective drugs important challenges arise, for example: What considerations should be considered when designing a safe and effective multitarget drug against AD? Faced with this challenge, the development of computational polypharmacology seems promising. The main goal of this project is to find the basis of the molecular architecture of MTDLs against Alzheimer's disease, through the exploration and improvement of existing drugs for the different molecular targets where Butyrylcholinesterase (BuChE), Beta Secretase 1 (BACE-1) and Cannainoid Receptor 2 (CB2) are of interest for this study, among other targets related to the Alzhimer's physiopatological hypotheses. Which by means of virtual screening will serve to identify new candidates for multi-target drugs as a new alternative for the treatment of AD. Our methodology will let us to find common pharmacophoric between them, that are key to design novel MTDLs. The emerging principles can potentially be applied to other pathology/targets, expanding the manner MTDLs are designed.

Agencia Nacional de Investigación y Desarrollo, Programa de Cooperación Internacional. Proyecto Nº REDES190074 - Multi-target drug design against neurodegenerative diseases.

SCB

ABSTRACT #47

A receptophore model for binding sites of local anesthetics in relevant atrial ion channels

José Carlos Estanislao Márquez Montesinos, David Ramírez, Gabriel Núñez-Vivanco and Wendy González Díaz

Introduction: An alteration of ion channels behavior generates abnormal cardiac phenomena such as atrial fibrillation (AF), the most common arrhythmic condition worldwide. In this context, potassium channels hKv1.5 and hTASK-1 as well as sodium channel hNav1.5 seem to be promising drug targets for AF treatment because of their relevance in this illness. It is well known that ropivacaine, bupivacaine, and lidocaine (local anesthetics-LAs-) exhibit a multi-target behavior, because each LA can inhibit these three channels simultaneously. Interestingly, these types of drugs have been shown better efficacy and safety parameters than those highly selective ones. For that reason, the presence of a common nature among their binding sites (BSs), defined as receptophore, is a plausible hypothesis for the discovery or rational design of novel drugs with polypharmacological profiles. Methods: Molecular Dynamics (MD) were performed for each ion channel. Reported residues that have a significant role on channel-LAs interaction were used to perform a MD clustering and retrieve the centroid of the most populated cluster. Then, a comparison of geometric patterns of these centroid-BSs was performed by pairs, using Geomfinder. Finally, an atom coordinates K-means clustering where performed in R-studio to generate residue groups for each BS and its geometric center calculation as well as distance measurements were performed in VMD, to propose the receptophore model. Discussion and conclusion: A preliminary receptophore model for LAs BS is presented, enlightening the common nature of hKv1.5, hTASK-1, and hNav1.5 binding sites. This has a tetrahedral shape where the vertices are the geometric center of residue groups which share similar physicochemical properties, i.e. (1 and 2) aliphatic, (3) aromatic and (4) sulfur residue groups.

PRESENTACIÓN ORAL

ABSTRACT #48

Identification of experimental and clinical characteristics for severity and pulmonary sequelae in Chilean COVID-19 patients using Artificial intelligence

Kevin Aguilar-Valdés

Background

COVID-19 is an infectious disease caused by SARS-CoV-2. In the clinic, the severity of COVID-19 is classified into three levels as mild, moderate and severe, and patients who survive exhibit diverse and persistent sequelae. Currently, the relationship between Chilean COVID-19 patient characteristics and the degree of severity and sequelae post COVID-19 remains unknown. Thus, the aim of this study was to analyze different clinical, experimental, and demographic parameters with artificial intelligence to determine their association with severity and sequelae in COVID-19 patients.

Methods

Data from 60 Chilean patients with a confirmed diagnosis of mild, moderate, and severe SARS-CoV-2 infection was obtained. Acute respiratory distress syndrome defined severity and pulmonary tests 6-months after infection defined lung sequalae. From the full data set (327-characteristics) including parameters from blood test, clinical evaluation, questionaries, circadian cycle and pulmonary function, relevant parameters were identified using three algorithms for feature selection and two models of machine learning to classify severity and sequelae.

Results

Our results revealed that the most important features in the modeling of the degree of severity and pulmonary sequelae were respiratory distress during acute infection, abnormal computerized axial tomography scan and spirometry, increased levels of anti-SARS-CoV-2 antibodies and cytokines, age, and high insulin levels. In addition, classification models with a performance of up to 100% was generated using the Random Forest and XGBoost methods, contributing to the analysis of information generated by the pandemic.

Conclusions

Artificial intelligence methods identified patterns and the most relevant characteristics associated with severity and the level of pulmonary sequelae in patients with COVID-19. Our results suggest that in Chilean population, sustained inflammation and glucidic metabolic alteration during and after infection could retard fully recuperation in COVID-19 patients.

SCB

ABSTRACT #49

Antifungal drug discovery by chemical similarity-guided network-based inference

Carlos Vigil Vásquez, Maria Jimenez Socha, Patricia Ortiz Bermudez and Andreas Schüller

Target identification is a key step in the discovery and development of novel drugs. Computational methodologies have been developed to be able to predict possible interactions between a drug candidate and a biological target of interest, generally using information obtained from either the drug's chemical structure or the target three-dimensional structure. Network-based methods have shown great potential thanks to the easy integration of information from different sources, the possibility of extracting novel information from graphical representation, and the fast and accessible application of methods based on networks from other fields outside of natural sciences. Here we present a novel method that combines network-based-inference with chemical similarity to predict the possible targets a compound could have. This method employs a tripartite drug-drug-target network constructed from the protein-ligand interaction annotations and the drug-drug chemical similarity on which the resource-spreading algorithm derived from traditional social network method predicts the potential biological targets for both known drugs and new chemical entities or failed drugs in clinical trials. Through multiple cross-validation and time-split validation procedures over a series of datasets we showcase that our proposed method exceeds the state-of-the-art performance obtained by other predictive methods based around this formalism and, through an extensive bioinformatic and topological analysis, is flexible enough to be adapted to different procedures commonly used in the development of novel drugs. As a proof-of-principle, we showcase the use of the proposed method as a tool for the discovery of novel antifungals for clinically relevant organisms. Therapies for invasive fungal infections of immunosuppressed patients have not changed in more than 50 years, presenting problems of spectrum, tolerability, and toxicity. Our results indicate that antifungal activity of some of the strongly predicted associations obtained in our drug repurposing procedure were confirmed both experimentally and bibliographically.

SCB

ABSTRACT #50

Comparative and system analysis of Leishmania spp. non-coding RNAs through developmental stages

J. Eduardo Martinez, Victor Aliaga-Tobar, Carolina Gonzalez, Rubens Monte-Neto, Alberto J. M. Martin and Vinicius Maracaja-Coutinho

Leishmania spp. is the causal agent of several illnesses called leishmaniases; neglected diseases that seek to be eradicated in the coming years. The life cycle of these parasites is very complex and involves different host and stress environments. Recently, many studies have shown several protein coding genes that are involved directly with the Leishmania lifecycle and its interaction with the human host. However, little is known about the role of non-coding RNAs (ncRNAs) in lifecycle progression. Here, we aimed to identify and characterize genomic structure, expression patterns and function of ncRNAs from 16 Leishmania species in a comparative perspective.

We studied 26 strains corresponding to 16 different species of Leishmania genus, using a comparative genomics approach to identify novel ncRNAs. We combined this approach with co-expression network analysis that relies on publicly available RNAseq data to identify likely target genes involved in its post-transcriptional regulation.

Our RNAome analysis revealed the presence of several ncRNAs that are shared through different species and others that allow to differentiate between subgenus; as well as species that are canonically related specifically to visceral leishmaniasis. We found coexpression relationships within coding genes and ncRNAs, thus suggesting possible functional relationships within them. Co-expression network analysis in the amastigote developmental stage for Leishmania braziliensis and Leishmania donovani reveals the presence of miRNA-like co-expressed RNAs with several coding genes involved in starvation and survival in macrophage and histone modification.

This work constitutes the first effort to characterize the Leishmania RNAome, supporting further approaches to better understand the role of ncRNAs in gene regulation, infective process and host-parasite interaction of kinetoplastids.

SCB

PRESENTACIÓN ORAL

ABSTRACT #51

A large-scale genome-based survey of acidophilic Bacteria suggests that genome streamlining is an adaption for life at low pH

Diego Cortez, Gonzalo Neira, Carolina Gonzalez, Eva Vergara and David Holmes

Genome streamlining theory suggests that reduction of microbial genome size optimizes energy utilization in stressful environments. Although this hypothesis has been explored in several cases of low nutrient (oligotrophic) and high temperature environments, little work has been carried out on microorganisms from low pH environments and what has been reported is inconclusive. In this study, we performed a large-scale comparative genomics investigation of more than 260 bacterial high quality genome sequences of acidophiles, together with genomes of their closest phylogenetic relatives that live at circum-neutral pH. A statistically supported correlation is reported between reduction of genome size and decreasing pH that we demonstrate is due to gene loss and reduced gene sizes. This trend is independent from other genome size constraints such as temperature and G+C content. Genome streamlining in the evolution of acidophilic Bacteria is thus supported by our results. Analyses of predicted COG categories and subcellular location predictions indicate that acidophiles have a lower representation of genes encoding extra-cellular proteins, signal transduction mechanisms and proteins with unknown function, but are enriched in inner membrane proteins, chaperones, basic metabolism, and core cellular functions. Contrary to other reports for genome streamlining, there was no significant change in paralog frequencies across pH. However, a detailed analysis of COG categories revealed a higher proportion of genes in acidophiles in the following categories: “Replication and repair”, “Amino acid transport” and “Intracellular trafficking”. This study brings increasing clarity regarding genomic adaptations of acidophiles to life at low pH while putting elements such as the reduction of average gene size under the spotlight of streamlining theory.

ABSTRACT #52

Long non-coding RNAs and RNA modifying enzymes differentially expressed in different types of heart failure

Sonia Vidal Vidal Vilches, Sergio Alejandro Lavandero González and Vinicius Ramos Henriques Maracajá Coutinho

Background: At the transcriptional level there are few studies that assess the association oflong non-coding RNAs (lncRNAs) and RNA modifying enzymes (RMEs) in heart failure. Here, we performed a comparative differential expression analysis of the set of lncRNAs and RMEs that are differentially expressed in the heart of patients with heart failure with preserved ejection fraction (HFpEF), heart failure with reduced ejection fraction (HFrEF) and healthy controls.

Methods: Firstly, we integrated the data obtained from Zenodo (10.5281/zenodo.4114617), the Modomics database for RMEs and the sets of lncRNAs available in GENCODE to identify the sets of differentially expressed lncRNAs and RMEs associated with the different types of heart failure. For that, we usedDESeq2 package, considering ap-adjusted < 0.05 and log2FoldChange> 0.5.

Results: A total of 30 RMEs were identified as differentially expressed in the different compared conditions. Of these, TPRKB presented the greatest change and the most downregulated RME. This is a protein that modifies tRNA and that has been associated with the TP53 gene. Furthermore, a total of 1038 lncRNAs were identified as differentially expressed. The following lncRNAs presented the greatest change: UMODL1-AS1, LINC01736, LINC01565, FHAD1-AS1, LINC02137 and LINC02183. All are known to be associated with various types of cancer and UNQ6494, and also presented as downregulated in both heart failure types.

These results obtained here show some insights into the potential role of both RMEs and lncRNAs in the different types of heart failure. We are currently performing additional integrative and systems biology analysis in order to obtain further insights on how these lncRNAs could be associated with these RNA modifying enzymes in heart failure.

ABSTRACT #53

circRNA transcriptional signature of chronic stress in male and female rat ventral hippocampus

Juan Silva, Wladimir Corrales, Felipe Aguayo, Felipe Olave, Luciano Román-Albasini, Vinicius Maracaja-Coutinho and Jenny Fiedler

Background: Sex differences seems to influence hippocampal plasticity and cognition in several disorders that alters the hippocampal integrity. The ventral hippocampus, which is linked to anxiety behavior, is a highly stress-sensitive structure evidenced by a broad transcriptomic and proteomic change under chronic stress in rodents. The present study describes for the first time the sex differences on circRNAs specific abundance and their putative pathological roles in the rat ventral hippocampus.

Methods: RNA-seq was performed to identify circRNAs using CIRCexplorer3 and CIRIquant. DESeq2 was used to evaluate the circRNAs differential expression (DE) from rat ventral hippocampus under chronic restrain stress (2.5 h/day for 14 days), by considering sex specific effects (n=20). Parental gene enrichment and co-expression analysis using CEMitool were performed to reveal circRNA function. circRNAs were validated by RT-qPCR, and then miRNA-circRNA interaction was predicted with IntaRNA, RNA22, and RNAhybrid. Finally, ViennaRNA was implemented for visualization.

Results: We identified 7.386 high confidence circRNAs, with 963 of them never reported before. Notably, 206 circRNAs displayed basal sex-specific expression, and chronic stress modify the levels of a higher number of circRNAs in females (194) compared to males (145). Female DE circRNAs induced by chronic stress are related to axon guidance and glutamatergic synapses, while male circRNAs are related to morphology and ion transport. These changes in neuronal plasticity can be mediated by miRNA-circRNAs interactions, which were detected in silico that circGabrg3_0001 interacts with miR-485 both in stress and sex-specific context.

Conclusions: circRNAs levels are chronic stress-responsive in the rat ventral hippocampus in a sex specific manner. The circRNAs levels may affect hippocampal neuroplasticity by a competing endogenous mechanism that could have broad impact on hippocampal function.

Funding: FONDECYT 1190899 and 1211731; FONDAP 15130011.

SCB

PRESENTACIÓN ORAL

ABSTRACT #54

Sex-biased impact of chronic stress on m6A modification machinery and neuroplastic pathways in the rat dorsal hippocampus

Wladimir Corrales, Leslye Venegas-Zamora, Matías Alarcón-Mardones, Juan Silva, Felipe Olave, Felipe Aguayo, Luciano Román-Albasini, Vinicius Maracaja-Coutinho and Jenny Fiedler

i) Background: The dorsal hippocampus, a key structure in learning and memory, is highly sensitive to stress. Stress can not only regulate gene expression through epigenetic mechanisms but also through the RNA modification landscape: the epitranscriptome. Recently, the epitranscriptome has drawn much interest, given its involvement in RNA metabolism. Some authors have reported that neurogenesis is regulated by m6A dynamics, a highly enriched RNA modification in the brain, where writer, eraser, and reader proteins play a crucial role. Notably, the m6A landscape has the potential to explain the sex-biased stress behavioral outcomes.

ii) Methods: We performed RNA-seq from female/male rat dorsal hippocampus to evaluate the transcriptomic profile induced by chronic restraint stress (2.5h/day for 14 days, n=20). Functional enrichment analysis was performed to identify relevant elements/pathways involved in m6A dynamics and dorsal hippocampus function. RT-qPCR was used to validate changes in the RNA modification machinery levels. Finally, we used expectation maximization to unveil a putative regulation of glucocorticoids on the RNA modification machinery gene expression.

iii) Results: We identified 1,352 stress-sensitive genes in the female dorsal hippocampus compared to 479 in males. Based on enrichment analysis, we identified several transcripts involved in RNA binding, RNA metabolism, and glutamate signaling in females; and axon guidance, GABAergic, and serotonergic synapse in males. Additionally, among sex-biased expressed genes, we found some components of the m6A machinery: METTL3, ALKBH5, YTHDF1, and YTHDF2. Finally, several m6A machinery genes contain putative response elements for glucocorticoids.

iv) Conclusions: Chronic stress can induce a sex-biased gene expression in the dorsal rat hippocampus, including genes belonging to the m6A machinery. These results suggest that sex-dependent modifications induced by stress could explain sex-biased behaviors.

SCB

ABSTRACT #55

Long non-coding RNAs involved in cardiac differentiation are differentially expressed in iPSCs from Down Syndrome patients

Francisco Sigcho, Wladimir Corrales, Leslye Venegas, Valentina Parra and Vinicius Maracaja-Coutinho

Background: Down syndrome (DS) is a condition characterized by the presence of an extra chromosome 21, which leads to an increased risk of pathologies including congenital heart diseases that affect ~50% of infants with DS. It causes a genetic imbalance that disturbs the transcriptome and various cellular processes, including cardiogenesis. Long non-coding RNAs (lncRNAs) are usually tissue specific, acting as important regulators of cardiac development. Also, cardiogenesis in cells derived from DS patients (3S) is altered compared to non-trisomic individuals (2S). Here, we hypothesize that "lncRNAs associated with cardiogenesis are differentially expressed (DE) in induced pluripotent stem cells (iPSC) from DS patients". Methods: We performed a meta-analysis of public transcriptome datasets focusing specifically on lncRNAs. First, a database search was performed in GEO, using the keywords “iPSCs”, “trisomy” and “Homo sapiens”. Then, three microarray and one RNA-seq datasets that compared the expression profile of iPSCs from 3S patients with 2S individuals were selected. The microarray platforms were reannotated using BioMart, allowing the identification of probed lncRNAs. Next, the sets of 2S/3S DE lncRNAs were determined for each study using GEO2R (p-value <0.05, FC >1.5). For the RNA-seq assay, we used FastQC for the quality control, then with Trim Galore we trimmed the reads; and with Subread, we aligned and quantified the transcripts expression. Following, with DeSeq2 we obtained the DE lncRNAs (p-value <0.05, FC >1.5). Results and Conclusions: We found a total of eight lncRNAs DE in least three studies (shared DE lncRNAs), with only four of them which have been previously associated with cardiogenesis, suggesting that they could affect 3S iPSC cardiomyocyte differentiation. Our next goal is to functionally characterize those lncRNAs and carry out experimental validations using cellular and molecular biology approaches. Methods: RNA-seq was performed to identify circRNAs using CIRCexplorer3 and CIRIquant. DESeq2 was used to evaluate the circRNAs differential expression (DE) from rat ventral hippocampus under chronic restrain stress (2.5 h/day for 14 days), by considering sex specific effects (n=20). Parental gene enrichment and co-expression analysis using CEMitool were performed to reveal circRNA function. circRNAs were validated by RT-qPCR, and then miRNA-circRNA interaction was predicted with IntaRNA, RNA22, and RNAhybrid. Finally, ViennaRNA was implemented for visualization. Results: We identified 7.386 high confidence circRNAs, with 963 of them never reported before. Notably, 206 circRNAs displayed basal sex-specific expression, and chronic stress modify the levels of a higher number of circRNAs in females (194) compared to males (145). Female DE circRNAs induced by chronic stress are related to axon guidance and glutamatergic synapses, while male circRNAs are related to morphology and ion transport. These changes in neuronal plasticity can be mediated by miRNA-circRNAs interactions, which were detected in silico that circGabrg3_0001 interacts with miR-485 both in stress and sex-specific context. Conclusions: circRNAs levels are chronic stress-responsive in the rat ventral hippocampus in a sex specific manner. The circRNAs levels may affect hippocampal neuroplasticity by a competing endogenous mechanism that could have broad impact on hippocampal function.

Funding: FONDECYT 1190899 and 1211731; FONDAP 15130011.

ABSTRACT #56

3D Enzyme Classification.

Fabio Durán-Verdugo, Alejandro Valdés-Jiménez and Gabriel Nuñez-Vivanco

Even when experimental methods to solve the tridimensional structures of proteins have been improved significantly in recent years, the functions assign and their classification remains a challenge. Indeed, most enzymes' function remains unknown.

Machine Learning and Deep learning allow the creation of computational models composed of multiple methods and techniques to data processing, learn and represent data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art of pattern recognition to research domains such as drug discovery and genomics.

This work in progress proposes some results and how to approach problems classification of structures in large datasets created from enzymatic characterization using the methods, techniques, data mining applying machine learning and deep learning. All training data will be acquired from structured databases such as Cath, Scope, Swiss Model, and Protein Data Bank.

ABSTRACT #57

RNAS CIRCULARES ASOCIADOS AL PERFIL DE EXPRESIÓN DE MICRORNAS DESREGULADAS EN INSUFICIENCIA CARDÍACA: EN ANÁLISIS DE SILICO

Camilo Rebolledo, Juan Pablo Silva, Vinicius Maracaja-Coutinho and Nicolas Saavedra

Background: Las enfermedades cardiovasculares (ECVs) son problemáticas de salud que encara el panorama actual debido a su alta tasa de mortalidad a nivel mundial. Nuestro país tiene una situación similar, siendo las ECVs la primera causa de muerte con 27% de las defunciones. Dentro de estas enfermedades, en nuestra población predominan las enfermedades hipertensivas, isquémicas y cerebrovasculares, que predisponen al desarrollo de una insuficiencia cardiaca. Diversos estudios han reportado el rol de miRNAs y circRNAs en diversas enfermedades, y su potencial regulatorio sobre otras especies de ARN. Su forma de acción ha sido descrita a través de tres mecanismos principales, funcionando como un modulador de la acción de miRNAs y mecanismos proteicos, como también modificador de la expresión génica. Uno de los mecanismos descritos hace referencia a la capacidad de los circRNA de funcionar como una esponja de miRNAs, capturando a través de sitios MRE y reduciendo su disponibilidad modulando e impactando en la expresión de los mRNA blancos. Métodos: Se analizaron datos de RNAseq obtenidos desde repositorios públicos utilizando find_circ, CIRI2 y CIRIquant para pesquisar circRNAs, y shortCat para pesquisar miRNAs. Expresión diferencial fue analizada utilizando DESeq2. Pares circRNA-miRNA fueron analizados por complementariedad utilizando RNAhybrid y generando redes de interacción con CytoScape. Finalmente, los datos fueron sometidos a un análisis de enriquecimiento utilizando enrichR. Resultados: Se obtuvo un total de 9 circRNAs y 12 miRNAs diferencialmente expresados, cuyos genes parentales o blancos se encuentran involucrados en vías moleculares vinculadas a insuficiencia cardiaca y procesos asociados, como hipertrofia o remodelación cardiaca. Conclusión: Existe teóricamente relación entre los perfiles de expresión de circRNAs y miRNAs, viéndose en acción dicho mecanismo de esponja descrito, que estaría causando un efecto sobre la expresión génica de tejidos cardiacos en sujetos con insuficiencia cardiaca.

Funding: FONDECYT 1211731; FONDAP 15130011; BECA ANID 21192016.

SCB

ABSTRACT #58

Characterization of competing endogenous RNAs in essential hypertension

Allan Javier Peñaloza Otarola, Thais Ramos, Victor Aliaga Tobar, José Gómez Ponce, Elmer Fernandez and Vinicius Maracaja Coutinho

Hypertension is a pathology that consists of an increase in arterial blood pressure. Primary arterial hypertension is the one which arises without a specific identifiable cause. This type of hypertension is estimated to affect 40% of the world's adult population. In Chile it is estimated that around 27.6% of the population suffers from arterial hypertension and around 27.53% of the affected population dies from disorders derived from this condition. Recently, many studies have focused on investigating the role of long non-coding RNAs (lncRNAs) in the mechanism of endogenous Competing RNAs (ceRNAs) in human diseases. To delve into this genetic regulation mechanism, we compiled information from databases of interaction of coding and non-coding RNAs with miRNAs, building an interaction catalog of ceRNAs with a total of 105,411,466 interactions. On the other hand, to identify the putative deregulated networks in the condition of primary arterial hypertension, we compiled four studies from the GEO and SRA databases evaluated the differential expression of coding genes, lncRNAs and miRNAs using the R package MetaVolcano. This last approach allowed us to identify a total of 7 coding genes (ACSL4, LMBR1L, CCN2, APLP2, ANXA1, HIPK3 and SURF4), that have interactions with the miRNA hsa-miR-92a-3p, which in turn they showed interaction with other 2 lncRNAs (LINC01128 and SNHG5). Interestingly, all target genes identified participate in metabolic pathways associated with renin expression and cardioprotective mechanisms, furthermore, the lncRNAs are known to be associated with the protective response in brain tissue.

ABSTRACT #59

Characterization of structural variants between clonal selections of Vitis vinifera cv. Carmenère

Daniela Araya-Ortega, Felipe Gainza-Cortés and Gonzalo Riadi

The vegetative propagation of cultivars of Vitis vinifera has originated a variety of clonal selections, which are cultivated for wine production in Chile. This propagation is part of the domestication process of Vitis vinifera for years.

This propagation produces mutations that accrue in the genome of somatic cells. Although the majority of them are deleterious, they are also recessive variants of heterozygous alleles, not affecting the fitness of the plant.

In order to do a genomics characterization and compare the Carmenère clonal selections, we sequenced 6 samples (3 biological replicates by clonal selection). We made a Reference assembly of these clonal selections, and we discovered the Structural Variants (SVs), Single Nucleotide Polymorphism (SNPs), and short InDels in them, and performed a preliminary analysis of the effects of these mutations. Then, we associated the mutated genes with metabolic pathways.

Using one clonal selection as the reference genome and the other as the sample, we compared and found 219,283 SNPs and 316,774 short InDels; 99.3% of the effects produced by these mutations are modifiers, affecting non-coding regions; 0.4% are high impact, affecting coding regions, mainly being frameshift variants; 1,697 mRNAs have at least one high impact variation. On the other hand, 165 SVs were found. 84.8% of the effects predicted are modifiers impact and 14.2% are high impact, impacting 29 mRNAs. These effects are mainly deletions of sequences that include a transcript segment.

Using KEGG Pathway Database, we associated 120 mRNAs with 91 pathways. The top pathways are Starch and sucrose metabolism, Fructose and mannose metabolism, Amino sugar and nucleotide sugar metabolism, and Purine metabolism participating respectively 7, 7, 7, and 6 mRNAs.

In conclusion, the order of mutations between clones is in the thousands, lower than the millions of mutations between cultivars reported in the literature. This low number is achieved by using the assemblies by reference and directly comparing the clones.

Altogether, these results suggest that clonal selections have few deleterious mutations but the majority of which are related to metabolic pathways relevant in grapes. Despite this, the characteristics of the plant, apparently, are not affected.

As future work, we propose to study the zygosity of the mutations to see if occur mainly in heterozygous alleles, and for this reason, they do not affect the plant's characteristics

SCB

ABSTRACT #60

Differential expression of transposable elements during aging in female and male mice.

Bairon Hernández and Gonzalo Riadi

Aging is characterized by progressive deterioration. It is associated with increased mortality, increased suffering from pathologies such as cancer, heart and neurodegenerative disease. One of the signs of ageing is the loss of epigenetic marks in the genome. As consequence, the state of heterochromatin (lax state), which allows the expression of elements that were repressed by euchromatin, is increased together with the genomic instability of the DNA. Transposable Elements (TEs), which are among the expressed elements, might be mobilized, potentially contributing to the genomic instability, and they also might be involved in the generation of transcript variants containing their TEs sequence (TEs exonizations).

Given the loss of heterochromatin and the increase in the activity of TEs as an organism age, we propose that there is an increased activity of TEs in the genome over time and there could be differential expression between the youngest age point versus the others age points.

In order to explore the relationship between TE expression and ageing, we used liver tissues RNA-seq data of female (age stages: 22 days, 6, 12, 18 months, 3 biological replicas) and male (2, 6, 14, 18, 24 months, 5 biological replicas). We used SQUiRE pipeline to calculate the differential expression of TEs, taking the youngest versus the others age point. After parsing the results, we used R to graph a volcano plot of the differential expression (DE) of TEs.

In the DE analysis, 1171, 4730 and 4962 TEs were differentially expressed in female (6mm, 12mm and 18mm vs 22dd). The most abundant TEs were LINEs throughout aging, then LTRs, SINEs and DNA (34%, 32%, 29% and 3% on average respectively). Whereas in males, 34, 92, 171, and 225 (6mm, 14mm, 18mm and 24mm vs 2mm) were identified differentially expressed TE. Being the LTRs the most abundant ones, followed by LINEs, DNAs, and SINEs (55%, 24%, 3% and 17% on average respectively).

In conclusion, it is observed that in females the most abundant DE TE are LINEs, whereas in males it was the LTR through aging. On the other hand, the number of DE TEs identified is different between female and male. Maybe this difference is due that female have more biological changes in the first age stage (juvenile and puberty period) in comparison with the male. The next step will be to identify and characterize the genes that are close to the TEs found and see if there is a regulatory relationship between them.

ABSTRACT #61

Ancestral Sequence Reconstruction as means to improve Transposable Elements annotation.

Camilo Fuentes-Beals, Juan Opazo and Gonzalo Riadi

Transposable elements (TEs) are ubiquitous in the genomes of a plethora of species. They have the means to transpose within the genome by activating mechanisms that can copy or cut them to a new position. The general classification of these elements groups them by their method of transposition and their structural composition, providing several categories that have been revised as more information is obtained. Annotations of TEs in genomes are based on sequence similarity of a genome sliding window to the consensus TE sequences that represent a category of said TE.

As the annotation of TEs are bound to this consensus, a question arises. How can we improve the annotations of these elements? The consensus sequences are approximations made with the available information. Improving this information would improve the annotations. Ancestral sequence reconstruction (ASR) might provide the means to do this. Through Bayesian inference, it is possible to determine ancestral states for each position of a given MSA from a set of TEs sequences, together with a phylogenetic tree. Thus, obtaining a common ancestor for said set of TEs.

By means of sequence simulation we have tested the ASR methodology in a TE family from the known subclass Endogenous RetroVirus, ERV, obtained from the human genome. We have managed to obtain over 80% of sequence identity comparing the reconstructed sequences to the simulated one, using the most likely phylogeny from the TE family. Ultimately, we seek to know if the ASR methodology provides an improvement over the actual consensus sequences, which are the ones used for TE annotation.

www.scbioinfo.cl

scbioinfo