

Universidad de Málaga

Escuela Técnica Superior de Ingeniería Informática



UNIVERSIDAD DE MÁLAGA

Biología de Sistemas

Tarea Investigación

Comorbidades del SARS-CoV2

Nombre

DNI/NIE

Marina Pacheco Rojas	77667027Y
Antonio Vargas Almodóvar	26831919G
Talal Awija	Y4775465N
Bogdan Catalin Miheșan	X7349245D

Profesor:

Aurelio Moya

Doy mi palabra de que he realizado esta actividad con integridad académica.

1. Introducción

1.1. Explicacion:

El proyecto que expondremos en el siguiente invorme trata de la relacion que existen entre las proteinas virales del SARS-CoV2 y otras enfermedades.

1.2. Motivacion:

Las proteínas humanas con las que interaccionan las proteínas virales de SARS-CoV2 pueden estar asociadas a otras enfermedades. De esta forma, introducimos el término de **Cormobilidades**. Este concepto es utilizado para describir dos o más trastornos o enfermedades que ocurren en la misma persona. Pueden ocurrir al mismo tiempo o uno después del otro. La comorbilidad también implica que hay una interacción entre las dos enfermedades que puede empeorar la evolución de ambas. Pero este término vamos a llevarlo a nivel genético. Si analizamos la importancia que tienen las proteínas virales del Sars-CoV2 en otras enfermedades, se pueden predecir los síntomas o enfermedades que produce el virus. En este trabajo vamos a tratar de identificar los targets de Sars-Cov2 mediante el modelado del interactoma funcional **Sars-Humano**, y su mapeo en un interactoma de enfermedades.

2. Materiales y métodos

2.1. Genes Humanos:

Como material, necesitamos los genes humanos que pueden estar relacionados con los genes del Sars-Cov2. Esos datos podemos obtenerlos mediante el estudio conjunto de la PPI del Sars-Cov2 y del Sars-Humano. Pero en este caso, hemos utilizado una tabla ya creada que aparece en **GO (Gene Ontology)** , que contiene diferente información sobre los genes humanos que pueden relacionarse con los genes del Sars-Cov2.

2.2. Manipulación:

Una vez tengamos los materiales, podemos proceder con la manipulación de los mismos. Sustancialmente, lo que vamos a realizar es un mapeado nuestros genes, para posteriormente obtener el gráfico de la red PPI de los mismos. Una vez tengamos la red, la simplificaremos eliminando los nodos que no se encuentran conectados. Posteriormente, vamos a realizar un análisis de la **centralidad**. La centralidad es una medida que determina la importancia de un nodo en la red, y se puede estudiar desde diferentes puntos de vista. Por un lado, está la **centralidad de intermediación (betweenness centrality)**, que se basa en los caminos cortos. Por otro lado, tenemos la **centralidad de proximidad (closeness centrality)** de un nodo, que se basa en la distancia de este con el resto de nodos. En nuestro caso, vamos a usar ambos análisis para determinar una selección de los 50 genes más relevantes.

2.3. GeneCards:

Una vez llegados a este punto, volvemos a la tabla de datos inicial. Podemos observar que se incluye una columna con información acerca de enfermedades relacionadas con los genes. De manera que fijándonos en esos datos y mediante la ayuda de GeneCards, vamos a identificar las distintas enfermedades que están asociadas a nuestros genes. Cabe destacar, que la información ya dada en la tabla está

incompleta, y para nuestros 50 genes, solo aparecen 19 de ellos con datos relativos a las enfermedades, de manera que vamos a obtener qué enfermedades están relacionadas con esos 19 genes. Una vez tengamos esta información, mediante la herramienta GeneCards ya mencionada, podemos adquirir los genes correspondientes a cada enfermedad. Para ver la información más clara, hemos creado una tabla Excel.

	A	B	C	D	E	F	G	H	I
1	Gene	Disease	Associated genes						
2	BRD4	Nut Midline Carcinoma	NUTM1, BRD4, NSD3, ZNF532, MYC, EP300, KMT2C, SOX2, SF3B1, H2AC18						
3	BRD4	Exophthalmic Ophthalmoplegia	KCNJ18, POMC, BRD4						
4	PMPCB	Multiple Mitochondrial Dysfunctions Syndrome 6	PMPCB						
5	ETFA	Multiple Acyl-CoA Dehydrogenase Deficiency	ETFDH, ETFB, FLAD1, ACADS, SLC25A32, SLC22A5, SLC52A1, ACADVL, PRODH						
6	POLA1	Van Esch-O'driscoll Syndrome	POLA1						
7	POLA1	Pigmentary Disorder, Reticulate, With Systemic Manifestations, X-Linked	POLA1						

Para el objetivo de este proyecto, debemos seleccionar algunas enfermedades anteriores. Como se puede ver en la imagen de la tabla, hay enfermedades en las que solo está implicado el gen en cuestión, por lo que escogeremos 3 enfermedades que impliquen varios genes para poder hacer el análisis más eficaz. En nuestro caso, hemos seleccionado las siguientes enfermedades:

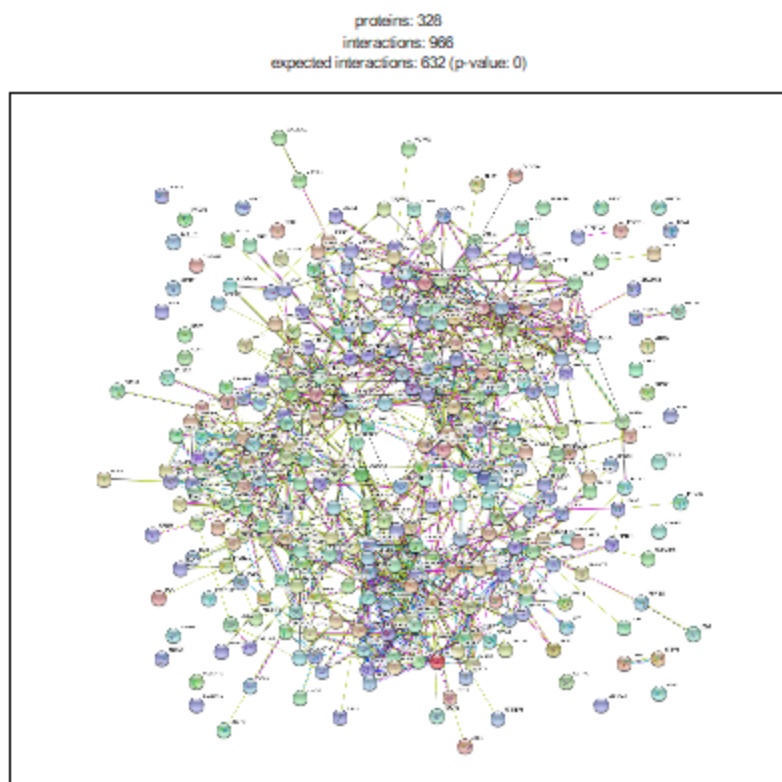
- **Enanismo primordial osteodisplásico microcefálico, tipo II**
- **Esquizofrenia**
- **Microcefalia 3 primaria autosómica recesiva**

Una vez escogidas las enfermedades, continuamos con el siguiente paso. Para cada una de ellas, vamos a obtener la red PPI, para después ver la interacción de todos los genes que las forman en un único interactoma. Finalmente, mapearemos los 50 genes seleccionados al principio del proyecto en el interactoma de las enfermedades.

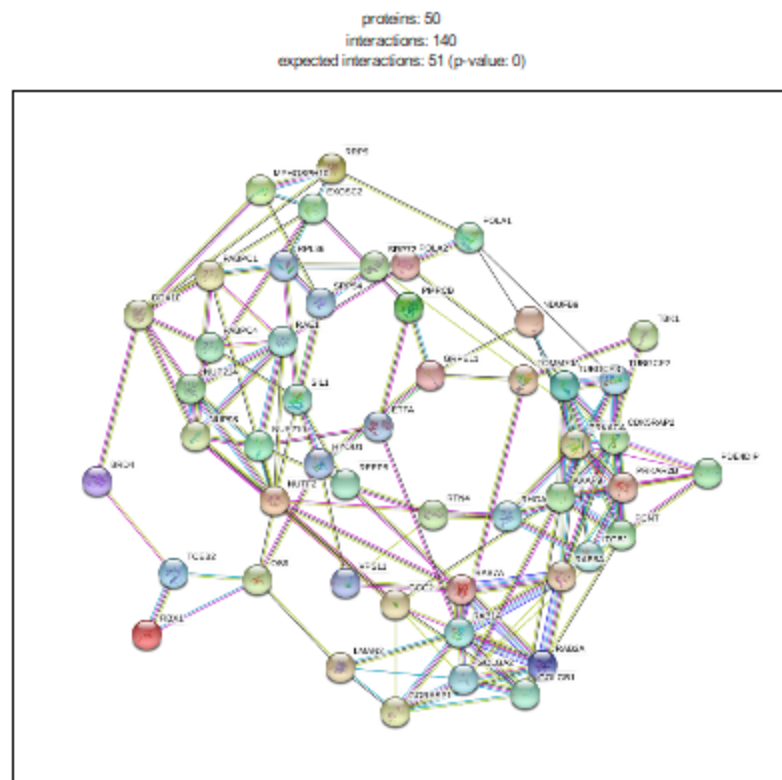
3. Resultados

3.1. Genes Sars-Humano

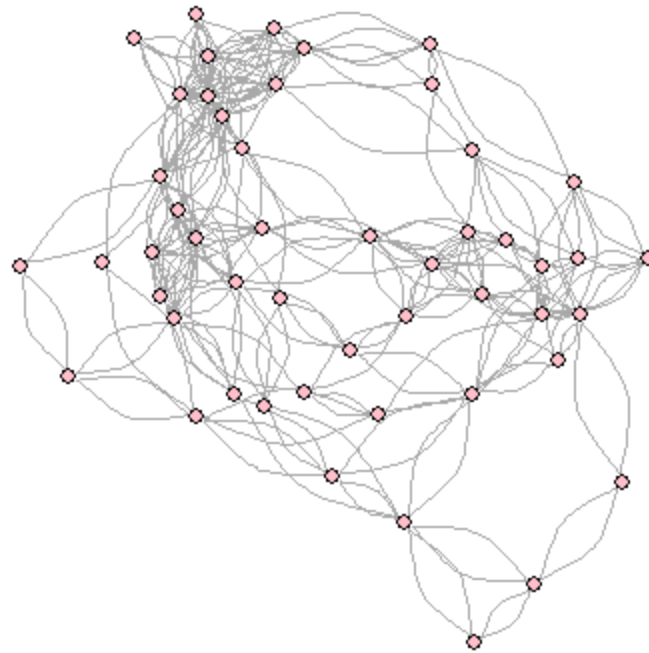
En este apartado vamos a mostrar las redes que hemos ido formando. En primer lugar, observamos el interactoma de todos los genes relevantes para el SARS-CoV2. Esta red consta de 328 nodos (genes) y 1932 interacciones.



Como vemos, hay algunos genes que no están conectados con la componente conexa, por lo que prescindiremos de ellos. A continuación, procedemos al análisis de centralidad ya mencionado, y a la selección de los 50 genes mas relevantes según el estudio de esta medida. Este interactoma constará de 50 nodos y 140 interacciones.



También vamos a observar el gráfico creado mediante el paquete iGraph, tal y como se ha explicado en clase.



3.2.

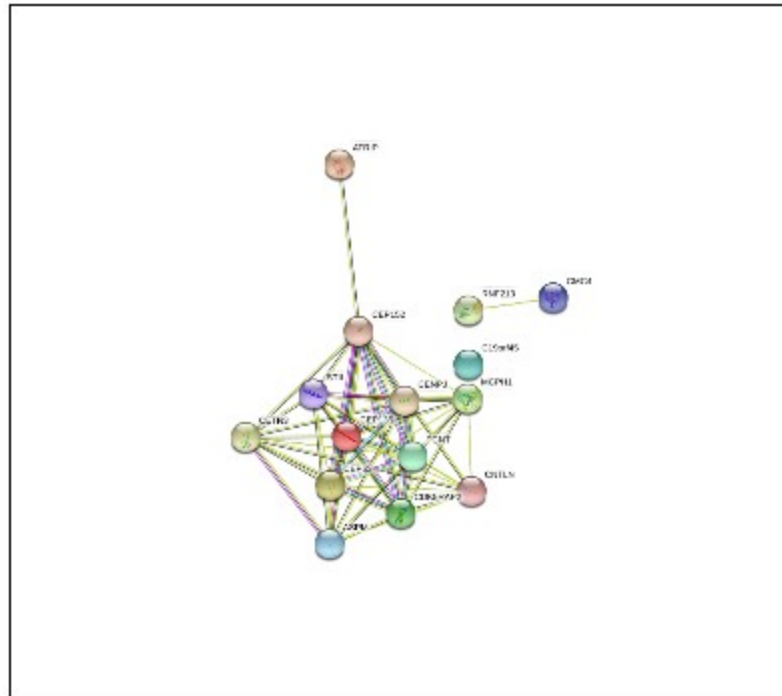
A continuación, vamos a observar los interactomas de las distintas enfermedades:

- **Enanismo primordial osteodisplásico microcefálico, tipo II:**

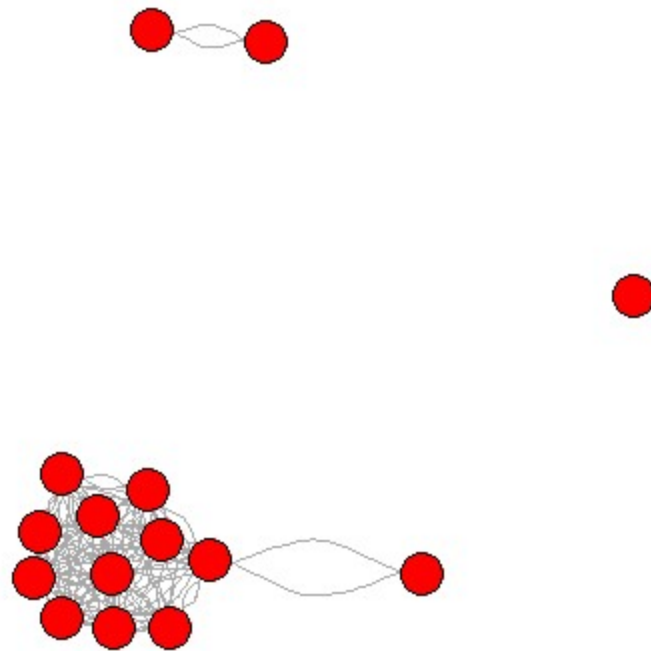
El interactoma consta de 15 nodos y 56 interacciones.

Red string enanismo:

proteins: 15
interactions: 56
expected interactions: 1 (p-value: 0)



Red enanismo:

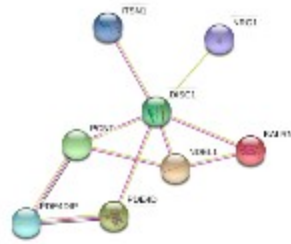


■ Esquizofrenia

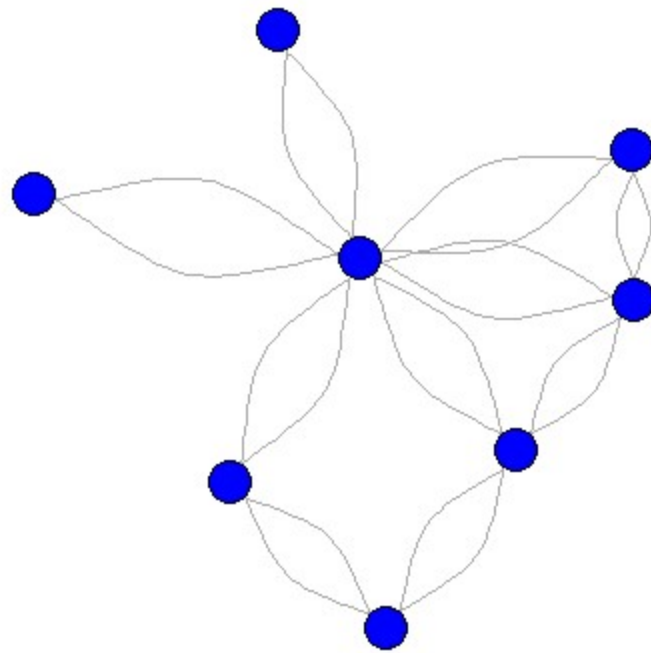
En cuanto a esta red, observamos 8 nodos y 10 interacciones.

Red string esquizofrenia:

proteins: 8
interactions: 10
expected interactions: 0 (p-value: 4.55e-11)



Red esquizofrenia:

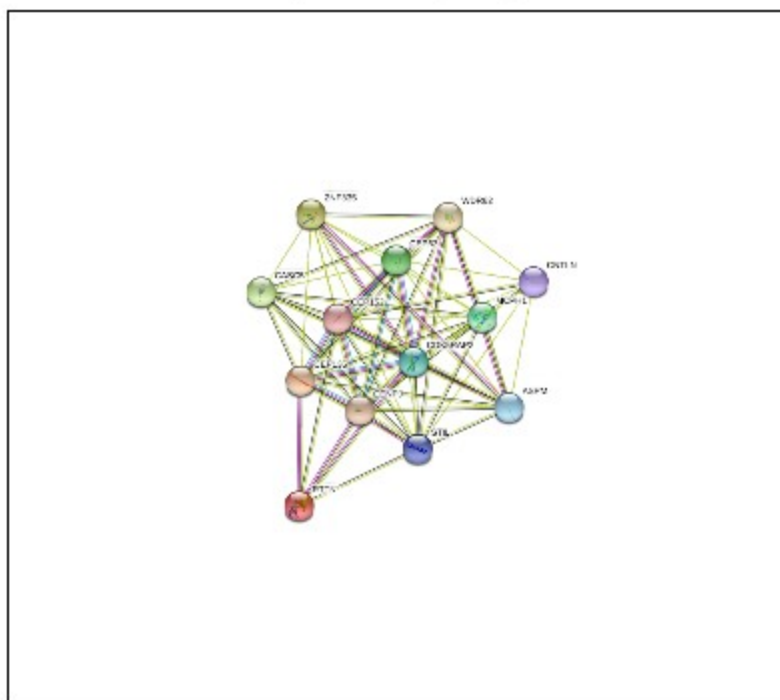


■ **Microcefalia 3 primaria autosómica recesiva**

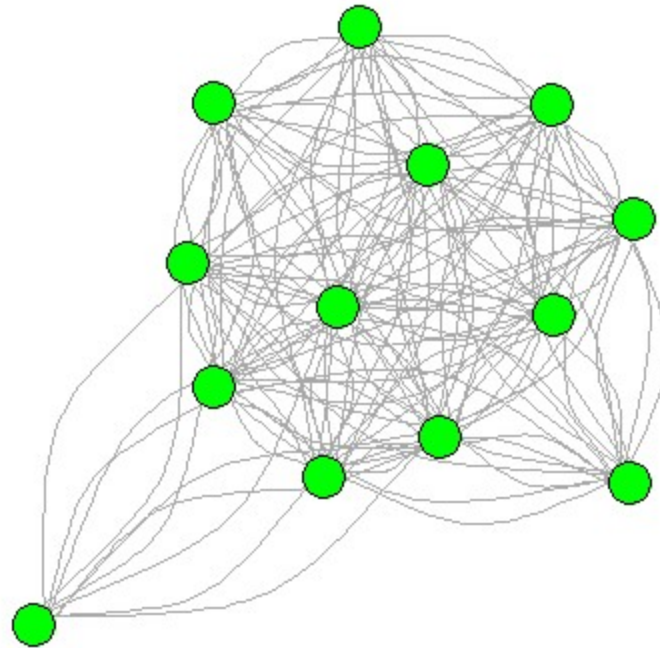
El interactoma correspondiente a esta enfermedad está formado por 13 nodos y 69 interacciones.

Red string microcefalia:

proteins: 13
interactions: 69
expected interactions: 1 (p-value: 0)

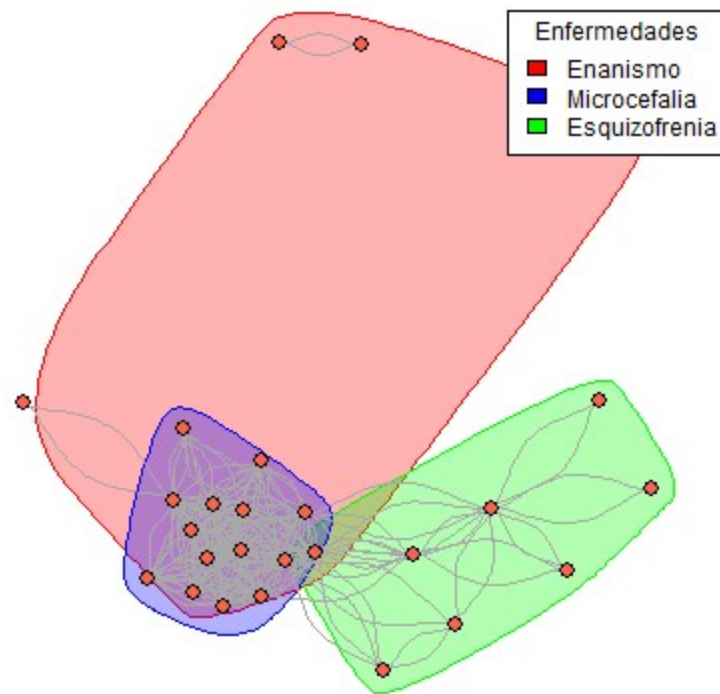


Red microcephalia:

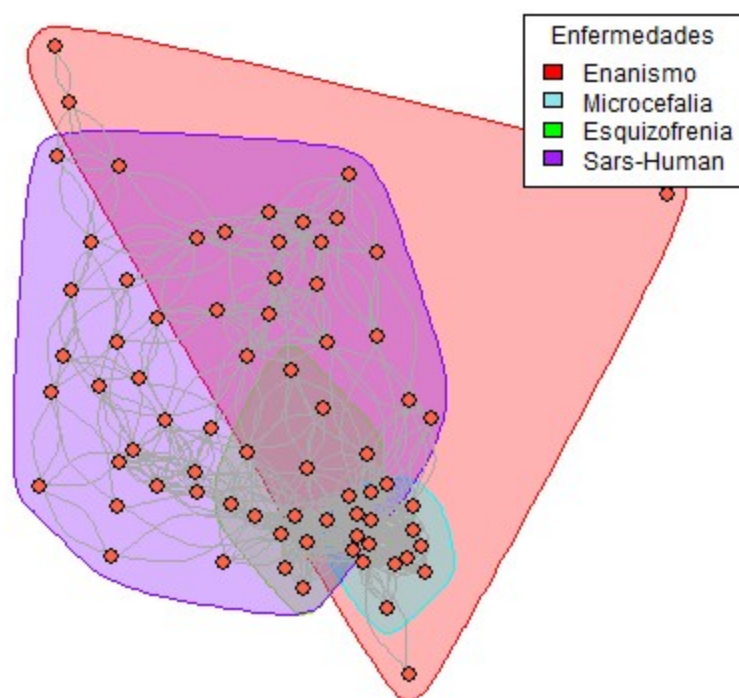


3.3. Redes PPI enlazadas:

La primera red que vamos a observar contiene todos los genes de las 3 enfermedades. Podemos ver con distintas áreas de colores a que enfermedad pertenecen los nodos. En la leyenda aparecen los respectivos colores para cada enfermedad, siendo el rojo para el enanismo, el azul para la microcefalia y el verde para la esquizofrenia.



A continuación, hemos incluido al gráfico la selección de 50 genes que recopilamos al principio del proyecto (recordemos que estos genes son los de SARS-Humano relacionados con el SARS-CoV2). En este caso, El área de color rojo se relaciona con el enanismo, el área celeste con la microcefalia, el área verde con la esquizofrenia y finalmente el área morada, que incluye los 50 genes mencionados.



4. Discusión

Para introducir este apartado, primero haremos una breve explicación de cada enfermedad escogida.

El **enanismo primordial osteodisplásico microcefálico, tipo II** es una enfermedad ósea poco frecuente y una forma de enanismo primordial microcefálico caracterizada por un grave retraso del crecimiento tanto pre como posnatal, con una marcada microcefalia en proporción al tamaño corporal, displasia esquelética y dentición anómala.

Por otro lado, la **esquizofrenia** es un trastorno psiquiátrico calificado por episodios continuos o recurrentes de psicosis. Se caracteriza por la aparición de alucinaciones, delirios y pensamiento desorganizado. Finalmente, la microcefalia 3 primaria autosómica recesiva: es un trastorno poco frecuente, genéticamente heterogéneo, del desarrollo neurogénico cerebral caracterizado por una reducción del perímetro cefálico (PC) al nacer, sin anomalías macroscópicas de la arquitectura cerebral y con grados variables de déficit cognitivo.

Existe una razón por la cual hemos decidido escoger estas enfermedades. Este proyecto consiste en analizar la importancia de las proteínas humanas con las que interaccionan las proteínas virales de SARS-CoV2, para posteriormente examinar la importancia de estas proteínas en otras enfermedades. Es por eso por lo que hemos tratado de mostrar enfermedades en las que coincidan algunos de sus genes.

Si observamos la red de conexión de los genes de las 3 enfermedades solo, podemos ver como en efecto comparten nodos, especialmente la red del enanismo y de la microcefalia. Esto se debe a que este tipo de enanismo (enanismo primordial osteodisplásico microcefálico), tal y como indica su nombre, es microcefálico, por lo que es lógico que tenga muchos genes en común con la microcefalia.

5. Conclusiones

Podemos ver como la biología de sistemas y la medicina de sistemas es un gran factor a favor del análisis y el estudio de enfermedades. Hemos comprobado, como mediante el estudio y la manipulación de las distintas redes, podemos relacionar enfermedades mediante los genes asociados a estas, lo que facilita el avance en la investigación de las distintas afecciones y sus respectivas curas.

En este proyecto hemos visto cómo se pueden predecir los síntomas o enfermedades derivadas de otras, pero es cierto que no se ha relacionado exclusivamente con el SARS-CoV2. Esto es debido a las distintas elecciones de datos que hemos ido realizando a lo largo del proyecto, tanto de genes, como de las enfermedades. Pero nos ha quedado claro (y podemos verlo reflejado en este trabajo) que esta forma de analizar las redes y sus interacciones, basada en la tecnología más reciente de biología de sistemas y medicina de sistemas, pueden hacer un claro avance en la investigación del COVID-19.

Author details

Referencias

1. (???) Orphanet: Enanismo primordial osteodisplásico microcefálico tipo II. URL https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=ES&Expert=2637
2. (???) Orphanet: Microcefalia primaria autosómica recesiva. URL https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=ES&Expert=2512
3. (???) Schizophrenia - Wikipedia. URL <https://en.wikipedia.org/wiki/Schizophrenia>
4. (???) STRING: functional protein association networks. URL https://string-db.org/cgi/input?sessionId=bngNaRiC3k3S&input_page_show_search=on
5. Genecards (???) GeneCards - Human Genes — Gene Database — Gene Search. URL <https://www.genecards.org/>
6. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Von Mering C (2019) STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research 47(D1):D607–D613,