

RESEARCH

Aniridia y su relación con otras patologías

Alberto Montero Solera^{??}
, Teresa Vega Martínez
, Alexandra Rodríguez Serantes
and Xian Li Xia

Abstract

Keywords: aniridia; patologías; iris; ocular

1 Introducción

La aniridia es un trastorno congénito del desarrollo ocular caracterizado por la ausencia parcial o total del iris, la estructura que da color al ojo y que regula la cantidad de luz que entra en la pupila. Esta condición, además de afectar la estética del ojo, puede causar una serie de problemas visuales significativos, como fotofobia (sensibilidad a la luz), nistagmo (movimientos involuntarios de los ojos) y disminución de la agudeza visual. Los individuos con aniridia a menudo presentan malformaciones oculares asociadas, como cataratas, glaucoma (enfermedad ocular que daña el nervio óptico, esencial para la visión) y problemas en la córnea, lo que agrava su salud visual y afecta gravemente su calidad de vida [[1].

No obstante, la aniridia no es solo un trastorno ocular, sino que también está vinculada a diversas patologías sistémicas, lo que sugiere que sus mecanismos genéticos y funcionales son complejos y multifactoriales[2]. Por ejemplo, estudios han mostrado que esta enfermedad puede estar asociada con el síndrome de WAGR, una condición genética que incluye anomalías renales, tumores y problemas de desarrollo, además de la aniridia [3]. Otros síndromes relacionados incluyen el síndrome de Gillespie, caracterizado por discapacidad intelectual, y el síndrome de Axenfeld-Rieger, que afecta tanto los ojos como otros órganos [4]. Estas asociaciones con condiciones sistémicas destacan la importancia de investigar en profundidad los mecanismos genéticos que subyacen a la aniridia, así como su relación con otras patologías.

Uno de los avances más significativos en la investigación de la aniridia ha sido la identificación del gen **PAX6** como regulador clave del desarrollo ocular. PAX6 es un factor de transcripción que juega un papel crucial en la diferenciación de diversas estructuras oculares, regulando la expresión de otros genes esenciales para este proceso [5]. Las mutaciones en PAX6 han sido vinculadas no solo con la aniridia, sino también con otras anomalías oculares como la displasia corneal (alteración en el desarrollo o estructura de la córnea) y el glaucoma congénito [6]. Sin embargo, para entender completamente la heterogeneidad clínica de la aniridia, es fundamental analizar cómo interactúa PAX6 con otros genes en redes genéticas más amplias.

Además de **PAX6**, varios genes han demostrado ser importantes para entender la complejidad de la aniridia y sus trastornos asociados, entre ellos **FOXC1**, **WT1**, **FOXE3** y **PITX2**.

Las mutaciones en **FOXC1** [7] están asociadas con el síndrome de Axenfeld-Rieger, un trastorno que incluye glaucoma y otras anomalías oculares [8]. Este gen codifica un factor de transcripción que, al igual que **PAX6**, regula el desarrollo del ojo, y su interacción con otros genes es clave para comprender las manifestaciones clínicas variadas que se presentan en estos pacientes. La integración de **FOXC1** en los análisis de redes génicas permite explorar cómo las vías de desarrollo ocular pueden estar interconectadas con otras rutas reguladoras sistémica.

Otro gen relevante es **WT1**, conocido por su papel en el desarrollo renal y la formación de tumores. Las mutaciones en **WT1** están vinculadas al síndrome de WAGR. La investigación sobre **WT1** desde una perspectiva de biología de sistemas permite investigar cómo los defectos en los programas de desarrollo que afectan tanto a los ojos como a otros órganos pueden estar mediadas por la interacción de redes génicas compartidas [9].

FOXE3 es un gen sin intrones que pertenece a la familia de factores de transcripción "Forkhead", que regulan la expresión de otros genes específicos de tejido cruciales para la diferenciación y desarrollo. La proteína que se codifica desempeña un importante papel en la formación del cristalino en vertebrados, siendo un factor de transcripción específico. Además, las mutaciones de este gen están relacionadas con disgenesia mesenquimal (desarrollo incompleto del tejido mesenquimático) del segmento anterior del ojo y afaquia primaria congénita (ausencia del cristalino).

Finalmente, el gen **PITX2** está involucrado en la morfogénesis craneofacial y cardiovascular. Las mutaciones en este gen están asociadas con el síndrome de Axenfeld-Rieger, afectando no solo a los ojos, sino también al desarrollo de órganos internos, debido a su rol en la simetría del desarrollo[10]. Desde la perspectiva de la biología de sistemas, **PITX2** representa un nodo crítico dentro de redes regulatorias que conectan el desarrollo de múltiples órganos, lo que resalta cómo las disfunciones en un solo gen pueden dar lugar a manifestaciones clínicas en varios sistemas.

El objetivo de este trabajo es examinar las patologías asociadas al fenotipo de la aniridia desde el punto de vista de la biología de sistemas, enfocándonos en los genes que tienen una relación funcional con **PAX6**. A través de la identificación de genes asociados y el análisis de redes génicas y moleculares mediante técnicas como la creación de clusters[11] y el análisis de redes funcionales [12], se tratará de avanzar en la comprensión de cómo las interacciones entre estos genes contribuyen a la presentación clínica de la aniridia y sus trastornos relacionados. Este enfoque no solo amplía nuestro conocimiento sobre la aniridia, sino que también ofrece información valiosa sobre los mecanismos patológicos que interconectan diversas condiciones sistémicas.

2 Materiales y métodos

En esta sección serán expuestos las diferentes herramientas y métodos utilizados para llevar a cabo el proyecto.

2.1 Materiales

2.1.1 Human Phenotype Ontology

La **Human Phenotype Ontology (HPO)** es una base de datos estructurada que organiza y describe las características clínicas (fenotipos) observadas en humanos, utilizando términos estandarizados y jerárquicos [?].

2.1.2 STRING

La **STRING Database** es una base de datos que integra información sobre interacciones entre proteínas, incluyendo interacciones físicas y funcionales, en una amplia variedad de organismos [?].

2.1.3 Python

Python es un lenguaje de programación versátil, fácil de aprender y ampliamente usado en desarrollo web, ciencia de datos, IA y más, gracias a su sintaxis clara y su gran ecosistema de herramientas [?].

2.2 Metodología

2.2.1 Obtención de genes, enfermedades y la red de interacción

El primer paso que se llevó a cabo para comenzar el proyecto fue la obtención de los genes y enfermedades relacionadas con el fenotipo seleccionado. Para ello utilizamos la API de HPO, mediante una función solicitamos estos archivos introduciendo como parámetro el código de la aniridia (HP:000526). Si la respuesta es exitosa, guardamos tanto la lista de enfermedades como la lista de genes en nuestra carpeta de resultados. Más adelante utilizaremos el archivo relacionado con las patologías, ahora nos centraremos en la lista de genes.

Una vez obtenidos los genes, utilizamos la API de STRINGDB para descargar la red de interacción de proteínas. Con otra función hacemos un request a su servidor y solicitamos toda esta información, guardando el archivo `red_inicial.tsv` y la imagen de la red (`red_inicial.png`) en la carpeta de resultados.

2.2.2 Análisis de la red de interacción

Todo lo anterior descrito fue programado en Python, ahora haremos uso de R y su librería iGraph. Importamos la librería y leemos la información de `genes_filtrados.txt`. Con varias funciones incluidas dentro de la librería, averiguamos distintos parámetros: `silos` y `nodes`.

2.2.3 Propagación de red

Para analizar de manera exhaustiva los genes funcionalmente relevantes asociados con nuestro fenotipo de interés, aplicamos el algoritmo **DIAMOnD (Disease Module Detection)** a una red de interacción proteína-proteína (PPI). Este procedimiento nos permitió expandir un módulo génico inicial definido por el conjunto de genes semilla previamente seleccionado mediante la integración de información de interacción genética y funcional. Los genes semilla iniciales, seleccionados por su relevancia biológica y su asociación conocida con el fenotipo en estudio, fueron: **WNT10B, SEM1, WT1, PAX6, y NF1**. Estos genes representan el núcleo inicial desde el cual se propagaron las interacciones para identificar genes adicionales funcionalmente relacionados. Respecto a la red base, utilizamos un archivo que contiene todas las interacciones proteína-proteína humanas disponibles en la base de

datos STRING. El algoritmo DIAMOnD se utilizó para expandir el módulo génico inicial. En este método, en cada iteración, el algoritmo evalúa los nodos vecinos de los genes ya presentes en el módulo y asigna un puntaje basado en el número de conexiones con los genes actuales del módulo. El nodo con el puntaje más alto se añade al módulo. Este proceso se repite hasta alcanzar el número máximo de nodos adicionales (en este caso, 200 nodos) o hasta que no queden candidatos relevantes. Una vez completada la expansión, se construyó un subgrafo que incluye todos los genes del módulo expandido junto con las aristas que los conectan.

2.2.4 Clustering

Para continuar con el análisis se desarrolló un algoritmo de clustering con el objetivo de identificar agrupaciones en la red de genes obtenida en el apartado anterior.

El **clustering** consiste en dividir un conjunto de datos en grupos o comunidades, de forma que los nodos de un mismo grupo tengan mayor relación entre sí que con los de los demás grupos.

Aplicado en el campo de la biología de sistemas, este método permite analizar redes complejas como las redes de interacción proteína-proteínas, rutas metabólicas y redes de regulación génica. A menudo, las comunidades identificadas con este tipo de clustering se corresponden a grupos de proteínas, genes (como es en este caso) o componentes moleculares que trabajan juntos para realizar una función específica. Para este proyecto se emplea un código para realizar el clustering de los genes de la red usando la **biblioteca igraph**. En este lo primero que se hace es cargar los datos de las interacciones entre genes desde el archivo de la red obtenida previamente en formato tabular.

A continuación, se crea un grafo no dirigido con los datos previos. Se estudian algunas propiedades de la red.

Después, se calculan las **métricas de centralidad**, es decir, el **grado de centralidad** (mide el número de conexiones de cada nodo), la **centralidad de cercanía** (mide la distancia media entre un nodo y los demás) y la **conectividad** (mide la densidad de conexiones en el grafo).

Una vez obtenidas las medidas se aplican varios métodos de clustering. Estos son **Edge Betweenness Clustering**, basado en la intermediación de aristas e ir eliminando iterativamente las aristas con mayor betweenness; **Clustering Basado en el Algoritmo Fast Greedy**, este consiste en agrupar nodos que tienen más conexiones entre ellos que con otros, optimizando de manera eficiente la modularidad; y **Clustering Basado en el Algoritmo Louvain**, el cual encuentra comunidades en redes grandes optimizando su estructura jerárquica para maximizar la modularidad. Además, se realiza la identificación de comunidades con **Link Communities**, método que considera las relaciones entre aristas, es decir, entre nodos.

Estos resultados se guardan como imágenes en la carpeta del proyecto "results" y se añadirán en el apartado de resultados de este informe.

Por último, se buscan vecinos interesantes de los genes de interés en el estudio (**WNT10B**, **WT1**, **SEM1**, **PAX6**, **NF1**). Estos se guardan en un archivo plano que se empleará para el enriquecimiento funcional.

2.2.5 Enriquecimiento funcional de los clusters

Para profundizar en la caracterización funcional del cluster seleccionado en el apartado anterior, hemos llevado a cabo un análisis de enriquecimiento funcional utilizando un enfoque basado en la base de datos STRINGdb. En este análisis, evaluamos específicamente dos categorías: procesos biológicos (Process) y ontología del fenotipo humano (HPO, Human Phenotype Ontology). Además de poder evaluar los procesos biológicos en los que están envueltos estos genes, otro de los objetivos era determinar si entre los fenotipos resultantes del análisis de enriquecimiento se encuentran aquellos descritos en HPO como relacionados con el fenotipo de interés. Durante el proceso se ha llevado a cabo la conversión a identificadores UniProt de los genes del cluster mediante el uso de la API de MyGene.info, asegurando compatibilidad con la API de STRINGdb.

Posteriormente, dichos identificadores UniProt fueron sometidos al análisis de enriquecimiento utilizando la base de datos STRINGdb, con un enfoque en las categorías Process y HPO.

En este proceso nosotros consideramos oportuno establecer un umbral de significancia estadística de **FDR < 0.001**, para identificar términos significativamente enriquecidos. Se evaluaron las asociaciones entre los genes del cluster y términos fenotípicos relevantes descritos en HPO.

Como resultado de la ejecución del algoritmo desarrollado se obtienen dos archivos en formato csv. Uno recoge los resultados para la categoría HPO y los de la categoría Process.

2.2.6 Obtencion de fenotipos comunes

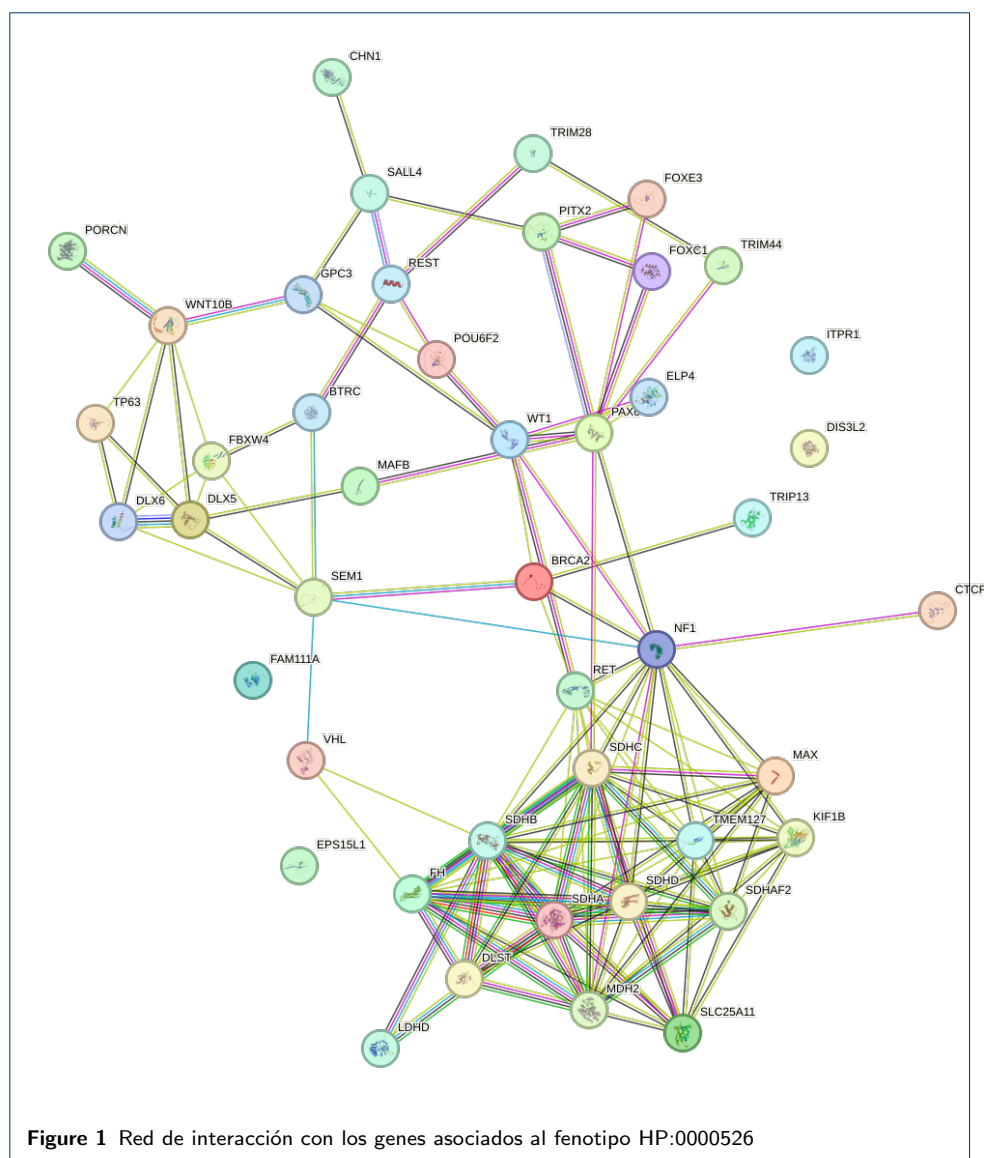
Para finalizar el proyecto es de interés ver si los fenotipos obtenidos en el enriquecimiento anterior están relacionados con las enfermedades relacionadas con nuestro fenotipo obtenidas al comienzo del proyecto. Para ello desarrollamos un script donde se obtuvieron todos los fenotipos asociados a dichas enfermedades. Posteriormente utilizamos otro script para ver los fenotipos comunes entre los obtenidos en el proceso de enriquecimiento y los relacionados con las enfermedades asociadas a la aniridia. Con esto buscamos indagar entre la relación entre estos fenotipos comunes, las enfermedades seleccionadas y nuestro fenotipo.

3 Resultados

3.1 Red de interacciones inicial

3.2 Propagación de la red

A partir de los genes semilla WNT10B, WT1, SEM1, PAX6, y NF1, se generó una red de interacciones proteína-proteína (PPI) utilizando datos de STRINGdb. Se estableció un umbral de puntaje de interacción de 700 para incluir únicamente conexiones de alta confianza. La red inicial centrada en los genes semilla fue posteriormente expandida mediante el algoritmo DIAMOnD, incluyendo 200 nodos adicionales que se encuentran altamente conectados con los genes semilla y entre sí. La red generada señala importantes interacciones entre los genes semilla y otros desconocidos (como la relación WNT10B - TGFB2), lo cual permite continuar el estudio hacia el objetivo del artículo.

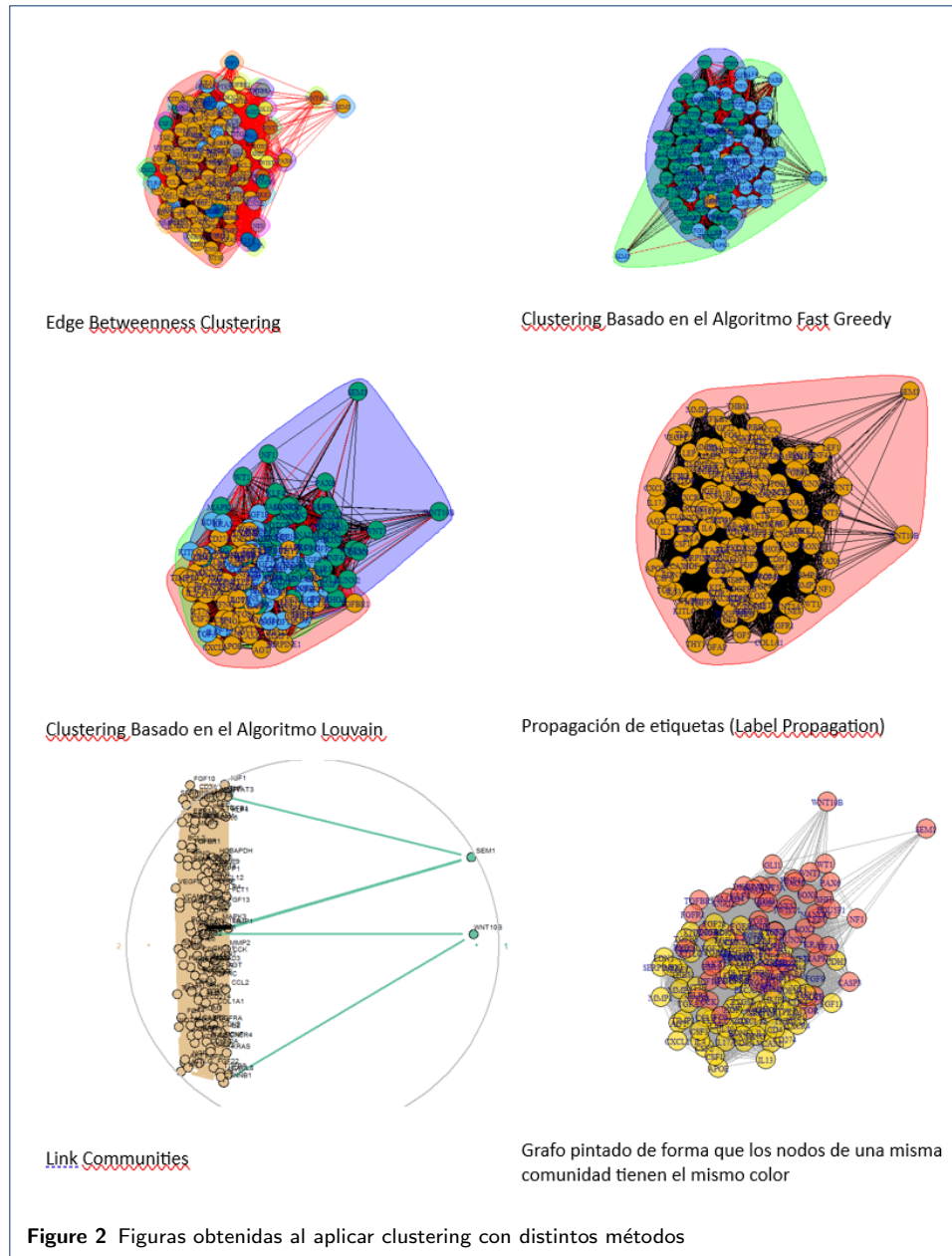


3.3 Clustering

Principalmente el estudio de los resultados se centrará en los genes semilla (WNT10B, WT1, SEM1, PAX6, NF1). En común, se puede observar que SEM1 aparece en varias posiciones destacadas en los grafos; WT1 y PAX6 están en regiones densas (centrales) de las comunidades, lo que indica que tienen muchas conexiones internas; y NF1 se sitúa en los bordes de las comunidades, por lo que podría actuar como puente entre diferentes comunidades; y WNT10B también es central en algunas comunidades.

3.4 Enriquecimiento y combinación de fenotipos

Al realizar un análisis de enriquecimiento considerando la categoría “Process” generamos un archivo que nos permite comprender en qué procesos biológicos clave están involucrados los genes del cluster seleccionado. Esto ayuda a identificar ru-



tas metabólicas, mecanismos celulares y funciones moleculares que pueden estar asociados con el fenotipo de interés.

Por otra parte, realizar el enriquecimiento funcional considerando categoría “HPO” es crucial para identificar cómo los genes analizados están asociados con características fenotípicas específicas, especialmente aquellas relacionadas con enfermedades humanas. Por último, combinar los fenotipos resultantes del enriquecimiento y los asociados a las enfermedades asociadas a nuestro fenotipo permite confirmar si los términos fenotípicos enriquecidos en los clusters tienen una relación directa con las enfermedades previamente identificadas. Si hay coincidencias, esto valida que los genes analizados están realmente relacionados con el fenotipo de interés.

Además, el establecer identificaciones entre las relaciones genotipo-fenotipo ayuda a establecer conexiones concretas entre los genes y las manifestaciones clínicas observadas en pacientes. Esto es especialmente útil para comprender cómo los genes influyen en el desarrollo de ciertas enfermedades o características clínicas.

Finalmente, podemos ver que 284 fenotipos de los 668 obtenidos tras el enriquecimiento están relacionados con enfermedades relacionadas con nuestro fenotipo a estudiar. Este resultado sugiere que el análisis de enriquecimiento funcional está capturando información biológica y clínica relevante. La conexión entre los fenotipos obtenidos y las enfermedades relacionadas con la aniridia valida que los genes analizados están efectivamente implicados en procesos relacionados con el fenotipo de interés. Este hallazgo podría abrir la puerta a la identificación de procesos centrales o vías moleculares críticas en estas enfermedades. Por otra parte, los fenotipos compartidos pueden ayudar a priorizar genes o grupos de genes específicos para estudios adicionales. Estos genes podrían ser considerados como potenciales biomarcadores o dianas terapéuticas para tratar tanto la aniridia como las enfermedades relacionadas. Los fenotipos no relacionados también son interesantes, ya que podrían reflejar aspectos únicos del fenotipo de interés o nuevas áreas de estudio.

4 Discusión

El análisis realizado a lo largo de este artículo ofrece un enfoque integral para el estudio de la **aniridia**, destacando la relevancia de la integración de herramientas de bioinformática y bases de datos especializadas en fenotipos humanos como HPO (Human Phenotype Ontology) para la identificación y caracterización de genes asociados. A partir de un fenotipo inicial relacionado con la aniridia, el estudio utilizó una red proteica generada a través de StringDB para identificar genes clave y extender la red a través de propagación. Este enfoque permitió la construcción de una red de 200 nodos, centrado en los genes semilla WNT10B, WT1, SEM1, PAX6, y NF1, los cuales son biológicamente relevantes para los mecanismos subyacentes de la enfermedad.

La selección de estos genes semilla resultó fundamental para orientar el análisis posterior de la red, ya que su implicación en procesos relacionados con el desarrollo ocular y los fenotipos asociados a la aniridia ya había sido descrita previamente en la literatura. La propagación de la red con base en estos genes permitió capturar interacciones adicionales y posibles moduladores que podrían estar implicados en la etiología del fenotipo estudiado. Este enfoque asegura no solo la identificación de elementos centrales en la red, sino también de genes secundarios o periféricos que podrían influir en los procesos moleculares relacionados.

El clustering de la red, una etapa crítica del análisis, permitió agrupar genes en comunidades funcionales que reflejan su co-implicación en procesos biológicos específicos. El enriquecimiento funcional de los clusters identificados reveló que al menos la mitad de los fenotipos asociados a estos grupos están directamente relacionados con los fenotipos de la aniridia descritos en HPO. Este resultado refuerza la robustez del enfoque metodológico, ya que confirma la coherencia entre los datos derivados de la red y el conocimiento previo. Además, destaca la capacidad de las técnicas de clustering para identificar módulos funcionales con una alta relevancia biológica.

Los genes semilla identificados, en particular PAX6, un gen base en el desarrollo ocular, y WT1, relacionado con procesos de desarrollo y cáncer, resaltan la complejidad de las interacciones genéticas y moleculares que subyacen a la aniridia. La identificación de genes adicionales en la red ampliada sugiere que el fenotipo no solo está condicionado por estos genes principales, sino también por redes más amplias de interacciones genéticas. Estos hallazgos abren nuevas posibilidades para explorar terapias dirigidas o identificar biomarcadores adicionales para diagnóstico y pronóstico.

El enriquecimiento funcional subraya la relación entre los procesos moleculares y los fenotipos clínicos observados en pacientes con aniridia. Esta conexión entre fenotipo y genotipo tiene aplicaciones prácticas en genética clínica, ya que permite refinar los diagnósticos y mejorar las predicciones relacionadas con manifestaciones fenotípicas adicionales. Por ejemplo, la implicación de genes relacionados con el desarrollo del sistema nervioso y el sistema inmunológico podría explicar fenómenos clínicos más allá de los defectos oculares.

Aunque el análisis de la red y el clustering resultaron útiles para identificar genes y procesos asociados, existen limitaciones inherentes al uso de bases de datos como StringDB y HPO. Las redes generadas dependen de los datos disponibles y pueden estar sesgadas hacia interacciones conocidas, subestimando las interacciones novedosas o contextuales. Además, el enriquecimiento funcional se basa en anotaciones existentes, por lo que la interpretación de resultados depende de la cobertura y calidad de estas anotaciones.

Futuros estudios podrían ampliar el análisis utilizando otras metodologías complementarias, como datos transcriptómicos o epigenómicos de tejidos afectados por la aniridia. También sería valioso realizar validaciones experimentales para confirmar las predicciones de la red y explorar los mecanismos moleculares subyacentes en modelos animales o líneas celulares. Finalmente, la integración de datos clínicos específicos de pacientes podría refinar aún más el análisis, permitiendo una personalización del diagnóstico y el tratamiento.

5 Conclusiones

Abreviaciones

Indicar lista de abreviaciones mostrando cada acrónimo a que corresponde

Disponibilidad de datos y materiales

https://github.com/monteero13/project_template.git

Contribución de los autores

Usando las iniciales que habéis definido al comienzo del documento, debeis indicar la contribución al proyecto en el estilo: J.E : Encargado del análisis de coexpresión con R, escritura de resultados; J.R.S : modelado de red con python y automatizado del código, escritura de métodos; ... OJO: que sea realista con los registros que hay en vuestros repositorios de github.

Author details

ETSI Informática, Universidad de Málaga, Málaga, España.

References

1. Landsend, E.C.S., Lagali, N., Utheim, T.P.: Congenital aniridia – A comprehensive review of clinical features and therapeutic approaches (2021). doi:10.1016/j.survophthal.2021.02.011
2. Black, G.C.M., Moosajee, M.: Chapter 17 - aniridia. In: Black, G.C.M., Ashworth, J.L., Sergouniotis, P.I. (eds.) *Clinical Ophthalmic Genetics and Genomics*, pp. 389–391. Academic Press, ??? (2022). doi:10.1016/B978-0-12-813944-8.00017-2. <https://www.sciencedirect.com/science/article/pii/B9780128139448000172>
3. López, M.: Relación entre aniridia y otras patologías no oculares

4. Law, S.K., Sami, M., Piri, N., Coleman, A.L., Caprioli, J.: Asymmetric phenotype of axenfeld-rieger anomaly and aniridia associated with a novel pitx2 mutation. *Molecular Vision* **17**, 1231–1238 (2011)
5. Robles López, K.L., González del Ángel, A.E.: Análisis molecular del Gen PAX6 con alteraciones oculares compatibles con disgenesias de segmento anterior. Trabajo terminal, especialidad. Acceso Abierto; Palabras clave: Medicina y Ciencias de la Salud; Anomalías del ojo (2012). <http://repositorio.pediatrica.gob.mx:8180/handle/20.500.12103/3129>
6. Calvão-Pires, P., Santos-Silva, R., Falcão-Reis, F., Rocha-Sousa, A.: Congenital aniridia: Clinic, genetics, therapeutics, and prognosis. *International Scholarly Research Notices* **2014**, 305350 (2014). doi:10.1155/2014/305350
7. National Center for Biotechnology Information (NCBI): FOXC1 forkhead box C1 [Homo sapiens (human)]. <https://www.ncbi.nlm.nih.gov/gene/2296>. Accessed: 2024-10-12 (2024)
8. Reis, L.M., Maheshwari, M., Capasso, J., Atilla, H., Dudakova, L., Thompson, S., Zitano, L., Lay-Son, G., Lowry, R.B., Black, J., Lee, J., Shue, A., Kremlikova Pourouva, R., Vaneckova, M., Skalicka, P., Jedlickova, J., Trkova, M., Williams, B., Richard, G., Bachman, K., Seeley, A.H., Costakos, D., Glaser, T.M., Levin, A.V., Liskova, P., Murray, J.C., Semina, E.V.: Axenfeld-rieger syndrome: more than meets the eye. *Journal of Medical Genetics* **60**(4), 368–379 (2023). doi:10.1136/jmg-2022-108646
9. Pelletier, J., Bruening, W., Li, F., *et al.*: Wt1 mutations contribute to abnormal genital system development and hereditary wilms' tumour. *Nature* **353**, 431–434 (1991). doi:10.1038/353431a0
10. French, C.R.: Mechanistic insights into axenfeld-rieger syndrome from zebrafish foxc1 and pitx2 mutants. *International Journal of Molecular Sciences* **22**(18), 10001 (2021). doi:10.3390/ijms221810001
11. Ben-Dor, A., Yakhini, Z.: Clustering gene expression patterns. In: *Proceedings of the Third Annual International Conference on Computational Molecular Biology*, pp. 33–42 (1999)
12. Flores Camacho, F.: 'mecanismos funcionales' para explicar los fenómenos: una perspectiva de análisis desde los recursos conceptuales. *Enseñanza de las ciencias (Núm. Extra)*, 1521–1524 (2009). [Consulta: 16 octubre 2024]