

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, así como los resultados obtenidos con cada una de ellas.

### 2.1 Materiales

#### 2.1.1 *Human Phenotype Ontology*

La **Human Phenotype Ontology (HPO)** es un proyecto que desarrolla un vocabulario estandarizado y jerárquico para describir fenotipos humanos, con el objetivo de apoyar el diagnóstico y la investigación de enfermedades genéticas. Además de definir y organizar características fenotípicas específicas, la HPO permite identificar los genes y variantes genéticas asociadas a cada fenotipo, de esta manera facilita el análisis de correlaciones genotipo-fenotipo a los usuarios.

#### 2.1.2 *STRING*

**STRING** es una base de datos que integra distintas fuentes de información y que además permite analizar y visualizar interacciones entre proteínas y genes. Pueden ser de dos tipos: interacciones físicas o interacciones funcionales. Las interacciones físicas permiten identificar compatibilidades y relaciones, mientras que las funcionales determinan si dos o más proteínas son compatibles aún sin tener interacción directa. Así pues, se pueden realizar perfiles filogenéticos y conocer genes con patrones similares, además de aportar información sobre patrones moleculares, mapas bioquímicos y mejorar la comprensión de los sistemas biológicos.

#### 2.1.3 *Python*

**Python** es uno de los lenguajes de programación más usados del mundo. Se trata de un lenguaje de alto nivel con una sintaxis sencilla y fácil de entender que dispone de multitud de paquetes de código abierto. Se caracteriza por ser interpretado, es decir, no se requiere de compilación previa a instrucciones de lenguaje máquina ya que se emplea un entorno o intérprete para ejecutar el código. La versión utilizada en este proyecto es ?.

#### 2.1.4 *R*

**R** es otro de los lenguajes de programación más importantes en el campo de la estadística. Es un entorno de software libre y también es un lenguaje interpretado. A diferencia de otros softwares de análisis de datos, R tiene un entorno completamente integrado y coherente. Proporciona una amplia gama de técnicas estadísticas y gráficas. La versión utilizada en este proyecto es ?.

#### 2.1.5 *iGraph*

**iGraph** es una biblioteca de código abierto utilizada para el análisis y visualización de redes complejas. Es una herramienta muy útil para el estudio de los sistemas biológicos ya que estos están compuestos por genes, moléculas, proteínas, etc, que interactúan formando grafos complejos de analizar a simple vista. Está implementado en C, aunque también puede ser utilizado en otros lenguajes como R y Python. Esta herramienta presenta muchas ventajas, entre las que destacamos: la integración de todo tipo de datos ómicos, la visualización de datos de forma interactiva, mejora

la comprensión de las relaciones entre las distintas biomoléculas del sistema, permite la identificación de nodos clave en la red biológica, facilita el descubrimiento de grupos funcionales de genes o proteínas y predice cómo ciertas modificaciones podrían perturbar los sistemas.

#### 2.1.6 *Pandas*

**Pandas** se trata de una popular librería de Python de código abierto muy necesaria para el ámbito de Data Science y Machine Learning. Proporciona unas estructuras poderosas y flexibles implementando todas las herramientas necesarias para el análisis de datos, ya que permite la carga, modelado, análisis, manipulación y preparación de datos.

#### 2.1.7 *Algoritmos de clusterización*

En el presente proyecto, uno de los objetivos es la identificación de patrones entre los distintos genes asociados al fenotipo de la aniridia. Para ello, es necesario el agrupamiento de aquellas biomoléculas con mayor similitud dentro nuestra red biológica, así que se aplicarán algoritmos nos permitan detectar estructuras sin necesidad de tener conocimiento previo. Vamos a hablar de los algoritmos de Givan-Newman, de optimización voraz, de propagación de etiquetas y de Louvain.

Primero, tenemos el **algoritmo de Givan-Newman**, que es uno de los métodos más utilizados para la detección de comunidades dentro de sistemas complejos de datos. Su funcionamiento se basa en eliminar progresivamente aquellos enlaces que conectan grupos de nodos con mayor densidad dentro de la red, para visualizar finalmente subgrafos desconectados que representan las comunidades.

Por otra parte, el **algoritmo de optimización voraz** es un algoritmo que dado un problema, elige aquellas decisiones localmente óptimas con el propósito de encontrar la solución más óptima globalmente. A medida que va tomando unas decisiones, el algoritmo no reconsidera ningún paso ya realizado.

Además, el **algoritmo de propagación de etiquetas** es una técnica que nos permite encontrar comunidades dentro del sistema velozmente. El funcionamiento se basa en la asignación de etiquetas a nodos, que en cada iteración se van propagando por la red y converge finalmente cuando cada nodo posee la etiqueta perteneciente a su vecino más cercano. Una de las ventajas de este algoritmo es que no es necesaria la predefinición de un número de clústers.

Por último, el **algoritmo de Louvain** se basa en el concepto de modularidad, es decir, tiene como objetivo maximizar el número de aristas dentro de una comunidad y minimizar el número de relaciones entre distintas comunidades. Este algoritmo es muy recomendable para sistemas biológicos muy amplios, ya que se obtienen comunidades compactas y bien definidas.

#### 2.1.8 *Otras librerías empleadas*

#### 2.1.9 *Software empleado para la generación de diagramas*

#### 2.1.10 *GitHub?*

### 2.2 Métodos

BORRADOR: \*OPCIÓN 1\* Para realizar este estudio, el primer paso que se llevó a cabo fue la obtención de los genes relacionados con el fenotipo. Los descargamos de

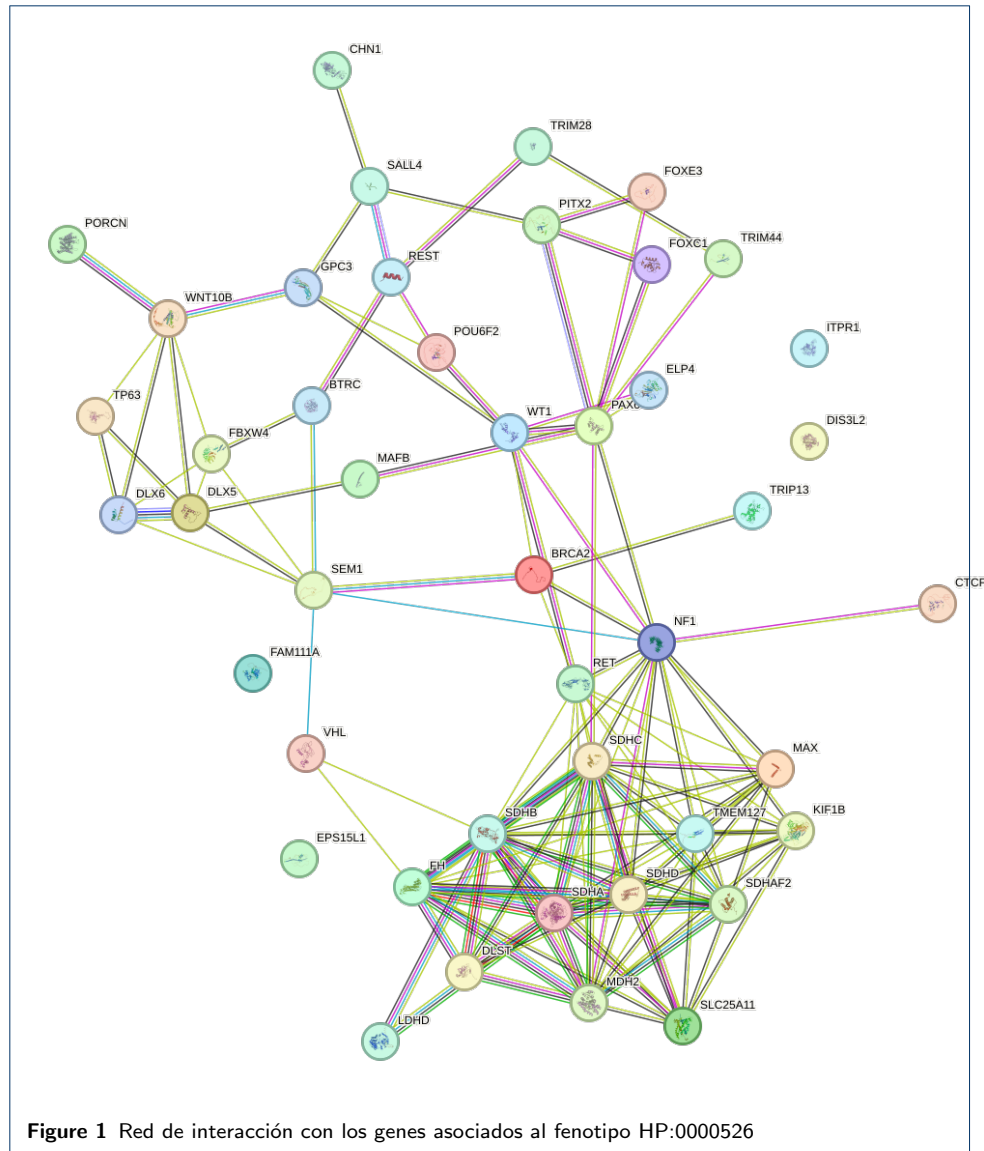
la página Human Phenotype Ontology mediante la introducción del ID del fenotipo, en nuestro caso "HP:0000526". Una vez obtenidos los genes, realizamos una red de interacción de los genes asociados a nuestro fenotipo mediante la herramienta STRINGDB.

A continuación, se muestran los pasos llevados a cabo para obtener una red de interacción de los genes de un fenotipo concreto:

- Visitamos la página de STRING
- Seleccionamos la opción para introducir múltiples proteínas
- Introducimos solo los nombres de los genes descargados de Human Phenotype Ontology.
- Especificamos la especie, en este caso, *Homo sapiens*.
- STRING generará una red de interacciones, mostrando conexiones basadas en interacciones experimentales, predicciones y bases de datos de coexpresión.
- Se guarda como imagen la red obtenida.

Entre los objetivos del uso de STRING están visualizar las interacciones proteicas conocidas y predichas, identificar posibles proteínas clave o "hubs" dentro de la red y evaluar la conectividad y la funcionalidad de los genes en el contexto del fenotipo elegido.

Tras introducir el nombre de todos los genes asociados a nuestro fenotipo obtuvimos la siguiente red de interacciones:



abras clave y patrones relevantes, y los resultados se guardan en archivos CSV para facilitar su análisis posterior.

Finalmente, descargamos la red de interacciones en formato TSV usando otra solicitud GET a la API de StringDB. Si la descarga es exitosa, el archivo se guarda con el nombre de archivo especificado o, en caso contrario, con el nombre predeterminado `reddescargada.tsv`. Ante cualquier error en la descarga, se imprime un mensaje de advertencia.

Al cargar el archivo TSV descargado, seleccionamos las columnas de nombres preferidos y eliminamos las filas duplicadas. Estos datos se guardan en un archivo de texto llamado `genesigraph.txt`, que contiene únicamente los nombres de los genes relevantes para un análisis adicional de la red biológica.

### *2.2.1 Flujo de trabajo*

### 3 Tipo de metodología

En esta sección se describe la metodología detallada para el análisis de redes génicas en el contexto de la aniridia, empleando el algoritmo DIAMOnD (Disease Module Detection) y técnicas de clustering para identificar genes asociados funcionalmente con PAX6. El objetivo de este análisis es comprender cómo las interacciones entre estos genes pueden influir en la manifestación clínica de la aniridia y en su relación con otros trastornos sistémicos.

#### 3.1 Construcción de la red de interacciones génicas

##### 3.1.1 Obtención de Datos de Interacción Génica

Se recopilieron datos de interacción génica a partir de STRINGdb (v11.5) para construir una red basada en las interacciones conocidas de PAX6 y otros genes asociados a la aniridia. La base de datos STRINGdb proporciona un puntaje de confiabilidad para cada interacción basado en evidencia experimental, coexpresión, bases de datos de conocimiento, y otros criterios. Para garantizar la relevancia biológica en el contexto de la aniridia, solo se seleccionaron interacciones con un puntaje de confianza superior a 0.7 (umbral que indica alta probabilidad de interacción funcional).

##### 3.1.2 Construcción de la Red Inicial

A partir de los datos obtenidos, se construyó una red de interacciones en Python utilizando la librería NetworkX. En esta red, los nodos representan genes y las aristas indican interacciones funcionales entre ellos. Los genes de interés fueron seleccionados como puntos de partida o "semillas" para el análisis de módulos funcionales. Estos incluyeron no solo a PAX6, sino también genes que, de acuerdo con estudios previos, tienen relevancia funcional en la aniridia, como FOXC1, WT1, COL4A1 y PITX2. Estos genes semilla fueron elegidos en base a su relación funcional conocida con PAX6 y su implicación en el desarrollo ocular y otras funciones biológicas relevantes para la patología.

#### 3.2 Aplicación del Algoritmo DIAMOnD para la Identificación de Módulos de Enfermedad

##### 3.2.1 Implementación del Algoritmo DIAMOnD

El análisis del módulo de enfermedad se llevó a cabo mediante la implementación del algoritmo DIAMOnD (Disease Module Detection) en la red génica construida. DIAMOnD fue diseñado para detectar módulos de enfermedad mediante la expansión progresiva de un conjunto de genes basado en la conectividad. Este algoritmo prioriza aquellos genes que muestran mayor número de conexiones hacia el módulo de interés, permitiendo identificar aquellos que, aunque no están directamente implicados en la patología, poseen relaciones funcionales clave.

##### 3.2.2 Parámetros de Configuración en DIAMOnD

El algoritmo se configuró para comenzar con los genes semilla seleccionados (PAX6, FOXC1, WT1, COL4A1, y PITX2) y expandir el módulo añadiendo genes en iteraciones. En cada iteración, se evalúan los genes adyacentes a los genes del módulo existente, y se priorizan aquellos con el mayor número de conexiones (o "vecinos") hacia los genes en el módulo. Este proceso se repite hasta alcanzar un tamaño predefinido para el módulo o hasta que se agoten los genes candidatos que cumplen con el criterio de conectividad.



### 3.2.3 Resultados de la Expansión del Módulo

Al finalizar las iteraciones, se obtiene un módulo expandido de genes que presentan una posible relación funcional con la aniridia. Este módulo incluye no solo los genes semilla, sino también genes adicionales con alta conectividad hacia estos y que podrían estar implicados en vías y funciones relevantes para la patología. Los genes identificados en este módulo se consideran candidatos para un análisis posterior de funciones y vías.

## 3.3 Análisis de Clustering para la Identificación de Submódulos Funcionales

### 3.3.1 Aplicación de Clustering en el Módulo Expandido

Para analizar la estructura interna del módulo expandido, se aplicaron técnicas de clustering en la red de genes seleccionados. Se utilizó el método de clustering jerárquico (enfoque de Ben-Dor et al., 1999) para identificar subgrupos dentro del módulo. Este enfoque permite agrupar genes en función de su similitud de conectividad y su co-ocurrencia en rutas biológicas. El clustering jerárquico es especialmente útil para identificar genes que, aunque puedan estar en el mismo módulo, desempeñan roles en funciones o rutas específicas.

### 3.3.2 Interpretación de los Clusters

Cada cluster resultante fue interpretado en función de las funciones y vías asociadas a los genes incluidos. Este análisis permite observar si ciertos grupos de genes comparten funciones específicas o están involucrados en vías biológicas que podrían contribuir a la patología de la aniridia. Los clusters identificados se utilizan como base para el análisis de enriquecimiento funcional descrito a continuación.

## 3.4 Análisis de Enriquecimiento Funcional de los Clusters

### 3.4.1 Herramientas y Procedimiento para el Enriquecimiento Funcional

El análisis de enriquecimiento funcional se llevó a cabo utilizando GOATOOLS y Enrichr, herramientas que permiten realizar análisis de sobre-representación de términos del Gene Ontology (GO) y otras bases de datos de rutas y procesos biológicos. En este análisis, se seleccionaron las categorías de procesos biológicos, funciones moleculares y componentes celulares, permitiendo observar qué términos están significativamente sobre-representados en cada cluster.

### 3.4.2 Identificación de Funciones y Vías Clave

Se calcularon p-valores ajustados mediante el método de FDR (False Discovery Rate) para los términos enriquecidos, considerando significativos aquellos con valores inferiores a 0.05. Este análisis permite identificar procesos y vías específicas en las que los genes del módulo están implicados, lo cual es esencial para entender cómo la disfunción en estas interacciones contribuye a la presentación clínica de la aniridia. En particular, el análisis se enfocó en vías de desarrollo ocular, rutas de señalización y procesos morfogénéticos.

## 3.5 Validación de los Resultados y Control de Calidad

### 3.5.1 Validación con Estudios Previos y Datos Experimentales

Los resultados obtenidos se compararon con la literatura y datos previos sobre PAX6 y genes asociados. Los genes y las funciones identificadas en el análisis de DIAMOnD

y de enriquecimiento funcional fueron cotejados con estudios experimentales previos para validar su relevancia en el contexto de la aniridia. Esta etapa de validación es crucial para confirmar que los genes incluidos en el módulo son relevantes y para reducir falsos positivos en el análisis.

### 3.5.2 Control de Calidad y Ajuste de Parámetros

Se realizaron análisis de sensibilidad ajustando parámetros clave, como el puntaje de interacción de STRINGdb y el tamaño del módulo, para evaluar la robustez de los resultados. Se examinó el efecto de estos ajustes en la composición del módulo y en los términos de enriquecimiento funcional, asegurando así que los resultados sean consistentes y representativos.

## 4 Resultados

## 5 Discusión

## 6 Conclusiones

### Abreviaciones

Indicar lista de abreviaciones mostrando cada acrónimo a que corresponde

### Disponibilidad de datos y materiales

[https://github.com/monteero13/project\\_template.git](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]