

RESEARCH

Análisis de redes génicas asociadas al fenotipo HP:0002344 – Progressive neurologic deterioration mediante biología de sistemas

Patricia Rodríguez Lidueña*, Achraf Ousti El Moussati, Aissa Omar El Hammouti Chachoui and Hugo Salas Calderón

*Correspondence: capi13@uma.es
ETSI Informática, Universidad de Málaga, Málaga, España
Full list of author information is available at the end of the article

Abstract

Keywords: sample; article; author

1 Resumen

Las enfermedades neurodegenerativas tienen en común el fenotipo HP:0002344 – Progressive neurologic deterioration, caracterizado por la pérdida progresiva de funciones neuronales. Su heterogeneidad genética dificulta la identificación de mecanismos comunes, por lo que la biología de sistemas ofrece un marco integrador para su estudio. El presente trabajo analiza cuatro patologías representativas: Spinocerebelar ataxia with epilepsy, Leigh syndrome, Frontotemporal lobar degeneration with TDP-43 inclusions y Rett syndrome, asociadas a los genes MECP2, IARS2 y GRN. Se propone integrar información de bases de datos biomédicas y construir redes de interacción proteína–proteína (PPI) para identificar rutas moleculares compartidas, genes candidatos y posibles dianas terapéuticas implicadas en la neurodegeneración progresiva.

2 Introducción

Las enfermedades neurodegenerativas constituyen uno de los mayores desafíos biomédicos contemporáneos debido a su alta prevalencia y a la ausencia de tratamientos curativos eficaces [1]. Aunque sus manifestaciones clínicas son diversas, comparten un rasgo común: la degeneración neurológica progresiva, entendida como un deterioro gradual e irreversible de las funciones neuronales que afecta al movimiento, la cognición y la función sensorial [2].

Este rasgo se encuentra descrito en la Ontología del Fenotipo Humano (HPO) bajo el identificador HP:0002344 – Progressive neurologic deterioration, que agrupa un conjunto heterogéneo de enfermedades hereditarias y adquiridas, entre ellas ataxias, encefalopatías metabólicas y demencias frontotemporales [3]. Aunque cada una de estas patologías es ultrarrara (frecuencia $\geq 1:1.000.000$), su frecuencia combinada alcanza aproximadamente 1 por cada 100.000 personas, lo que convierte a este fenotipo en un problema relevante dentro de las enfermedades neurológicas raras [4].

La HPO constituye una herramienta clave para vincular fenotipos clínicos con genes causales mediante un vocabulario jerárquico estandarizado [5]. Esta ontología

ha mejorado la interpretación de variantes genéticas y ha facilitado el diagnóstico computacional de enfermedades raras basadas en descripciones fenotípicas precisas. Sin embargo, la heterogeneidad genética y fenotípica de las enfermedades asociadas a HP:0002344 dificulta la identificación de mecanismos comunes de patogénesis. Para acotar el análisis, nos centraremos en un conjunto representativo de cuatro patologías con este fenotipo: Spinocerebellar ataxia with epilepsy (ORPHA:778), Leigh syndrome (ORPHA:506), Frontotemporal lobar degeneration with TDP-43 inclusions (OMIM:607485) y Rett syndrome (ORPHA:778). Estas patologías ejemplifican distintos mecanismos de neurodegeneración —epigenético, metabólico y proteopático—, permitiendo explorar posibles rutas convergentes en el deterioro neurológico progresivo.

En la ataxia espinocerebelosa con epilepsia y el síndrome de Rett, las mutaciones en MECP2 alteran la regulación epigenética y la expresión génica neuronal. MECP2 codifica una proteína que se une al ADN metilado y modula la transcripción de genes esenciales para la plasticidad sináptica y la maduración neuronal; su disfunción produce desequilibrios en la excitabilidad cortical y pérdida progresiva de funciones cognitivas y motoras.

La enfermedad de Leigh, asociada a mutaciones en IARS2, afecta la función de la isoleucil-ARNt sintetasa mitocondrial, lo que compromete la traducción proteica dentro de la mitocondria y conduce a un déficit energético severo, estrés oxidativo y degeneración neuronal en regiones con alta demanda metabólica. Por su parte, la degeneración lobar frontotemporal con inclusiones de TDP-43, causada por mutaciones en GRN, se caracteriza por la pérdida de programulina, una proteína implicada en la homeostasis lisosomal, la inflamación y la supervivencia neuronal. La deficiencia de programulina conduce a la acumulación de inclusiones citoplasmáticas de TDP-43, disfunción sináptica y neuroinflamación crónica.

A partir de esta base, se plantea la hipótesis de que la degeneración neurológica progresiva surge como resultado de alteraciones en redes moleculares interconectadas, compartidas por distintas enfermedades raras con orígenes genéticos diversos.

La integración de información procedente de bases de datos como HPO, OMIM, STRINGdb y GeneCards permitirá construir una visión sistémica del fenotipo HP:0002344, identificando redes funcionales compartidas, genes candidatos no descritos y nodos terapéuticos potenciales. Este enfoque de biología de sistemas busca arrojar luz a los mecanismos moleculares de convergencia que subyacen al deterioro neurológico progresivo y aportar una base para el desarrollo de estrategias terapéuticas dirigidas [6, 7, 8] .

3 Objetivos

El presente trabajo tiene como finalidad estudiar los mecanismos moleculares implicados en el fenotipo **HP:0002344 – Progressive neurologic deterioration** desde un enfoque de biología de sistemas. Para ello, se plantean los siguientes objetivos específicos:

- 1 **Identificar** los genes y las enfermedades asociadas al fenotipo HP:0002344 a partir de bases de datos como *HPO*, *OMIM* y *GeneCards*.

- 2 **Describir** los mecanismos moleculares implicados en las enfermedades representativas—Spinocerebellar ataxia with epilepsy, Leigh syndrome, Frontotemporal lobar degeneration y Rett syndrome— y sus genes causales (*MECP2*, *IARS2* y *GRN*).
- 3 **Construir y analizar** redes de interacción proteína-proteína (*PPI*) a partir de los genes identificados mediante la base de datos *STRINGdb*, evaluando métricas topológicas como el grado, la centralidad y la modularidad.
- 4 **Detectar** grupos funcionales de genes y realizar un análisis de enriquecimiento para identificar procesos biológicos y rutas moleculares comunes.
- 5 **Desarrollar** y documentar el código necesario para reproducir el análisis completo y garantizar la trazabilidad de los resultados en el repositorio de *GitHub*.

4 Materiales y Métodos

4.1 4.1 Materiales

El análisis se basó en información procedente de bases de datos biomédicas de referencia. Se seleccionó el término **HP:0002344 – Progressive neurologic deterioration** de la Human Phenotype Ontology (HPO) [4, 6] como punto de partida, recopilando los genes asociados a este fenotipo. La relación gen-enfermedad se contrastó con OMIM [9] y GeneCards [?], lo que permitió confirmar y ampliar la evidencia disponible sobre cada gen.

Las interacciones entre las proteínas codificadas por dichos genes se obtuvieron mediante STRINGdb (versión 12.0) [?], considerando únicamente interacciones de alta confianza (combined score ≥ 0.7) en *Homo sapiens*.

El procesamiento y análisis de redes se realizaron en R (versión 4.3.2), empleando los paquetes *igraph*, *tidygraph*, *ggraph* y *linkcomm* [?, ?, ?, ?]. Todos los scripts, tablas y visualizaciones se encuentran disponibles en el repositorio de GitHub del proyecto, garantizando la reproducibilidad del estudio.

4.2 4.2 Metodología

4.2.1 4.2.0 Flujo de trabajo general

El procedimiento seguido en el estudio puede resumirse en las siguientes etapas:

- 1 Selección del fenotipo HP:0002344 en HPO.
- 2 Extracción de genes asociados al fenotipo.
- 3 Validación de las asociaciones gen-enfermedad mediante OMIM y GeneCards.
- 4 Obtención de interacciones proteína-proteína (*PPI*) en STRINGdb.
- 5 Construcción y depuración de la red en R.
- 6 Cálculo de métricas estructurales de la red.
- 7 Identificación de comunidades mediante Louvain, Walktrap y linkcomm.
- 8 Análisis detallado de los genes de interés: *MECP2*, *IARS2* y *GRN*.
- 9 Enriquecimiento funcional de los módulos identificados (GO y KEGG).
- 10 Interpretación integrada de resultados.

4.2.2 4.2.1 Recopilación de genes y construcción de la red

A partir del término HP:0002344 se descargó la lista de genes relacionados desde HPO [4]. Tras eliminar duplicados, se verificó la validez de las asociaciones

gen-enfermedad mediante OMIM [9] y GeneCards [?], obteniendo un conjunto depurado de genes candidatos.

Estos genes se introdujeron en STRINGdb (v12.0) [?] para obtener la red de interacciones proteína-proteína (PPI), restringiéndonos a interacciones con evidencia experimental o provenientes de bases de datos curadas. La red se exportó y procesó en R, transformándose en un objeto `igraph` para su análisis estructural.

4.2.3 4.2.2 Análisis estructural de la red

Se caracterizó la topología global mediante métricas estándar de redes. El *grado* permitió identificar genes altamente conectados (posibles hubs), mientras que las centralidades de intermediación y cercanía ayudaron a detectar nodos clave en la comunicación entre módulos.

Se calculó el coeficiente de agrupamiento para evaluar la tendencia de los nodos a formar clústeres, y la modularidad global para estimar el grado de organización de la red en comunidades funcionales.

4.2.4 4.2.3 Identificación de comunidades funcionales

La detección de comunidades se realizó mediante dos algoritmos ampliamente utilizados: Louvain [?] y Walktrap [?], implementados en `igraph`. Además, se aplicó el enfoque de comunidades solapadas mediante el paquete `linkcomm` [?], permitiendo que un mismo gen pudiera participar en distintos módulos cuando intervenía en procesos independientes.

Cada comunidad se evaluó considerando su cohesión, su tamaño y la presencia de genes previamente relacionados con enfermedades neurológicas.

4.2.5 4.2.4 Análisis de genes de interés

Se estudiaron en mayor profundidad los genes **MECP2**, **IARS2** y **GRN**, seleccionados por su relevancia clínica en patologías asociadas al deterioro neurológico progresivo.

El gen *MECP2* está implicado en el síndrome de Rett [7] y en el síndrome por duplicación *MECP2* [10]. El gen *IARS2* se ha asociado a síndrome de Leigh con deficiencia combinada en fosforilación oxidativa [8]. Por último, *GRN* está implicado en la degeneración lobar frontotemporal [11].

Para cada gen se generó una subred local con sus interacciones más próximas, visualizada mediante `ggraph`, destacando su conectividad y su relación con nodos relevantes dentro de la red global.

4.2.6 4.2.5 Enriquecimiento funcional

Finalmente, las comunidades más representativas se sometieron a análisis de enriquecimiento funcional mediante las herramientas integradas en STRINGdb. Se evaluaron las categorías de Gene Ontology (procesos biológicos, funciones moleculares y componentes celulares) y las rutas KEGG, considerando significativos los términos con p ajustado < 0.05 (corrección Benjamini-Hochberg).

Este análisis permitió relacionar los módulos detectados con procesos neuronales y rutas metabólicas asociadas al deterioro neurológico progresivo, en consonancia con los mecanismos descritos en la literatura biomédica [1, 2].

5 Resultados

6 Discusión

7 Conclusiones

Abreviaciones

Indicar lista de abreviaciones mostrando cada acrónimo a que corresponde

Disponibilidad de datos y materiales

Debéis indicar aquí un enlace a vuestro repositorio de github.

Contribución de los autores

Usando las iniciales que habéis definido al comienzo del documento, debéis 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. Kelser, B.M., Teichner, E.M., Subtirelu, R.C., Hoss, K.N.: A review of proposed mechanisms for neurodegenerative disease. *Frontiers in Aging Neuroscience* **16**, 1370580 (2024)
2. Gao, H.-M.: Why neurodegenerative diseases are progressive (2008)
3. Wakap, S.N., et al.: Estimating cumulative point prevalence of rare diseases. *European Journal of Human Genetics* **28**, 165 (2020)
4. Human Phenotype Ontology Consortium: HPO Term HP:0002344. <https://hpo.jax.org/browse/term/HP:0002344>. Accessed: YYYY-MM-DD
5. Wakap, S.N., et al.: Estimating cumulative point prevalence of rare diseases. *Orphanet Journal of Rare Diseases* **14**, 261 (2019)
6. Human Phenotype Ontology Consortium: Human Phenotype Ontology (HPO) database. <https://hpo.jax.org/>. Accessed: YYYY-MM-DD
7. Ehrhart, F., Coort, S.L.M., Cirillo, E., Smeets, E., Evelo, C.T., Curfs, L.M.G.: Rett syndrome – biological pathways leading from mecp2 to disorder phenotypes. *Orphanet Journal of Rare Diseases* **11**(1), 158 (2016). doi:10.1186/s13023-016-0545-5
8. Dong, Q., Yin, X., Fan, S., Zhong, S., Yang, W., Chen, K., Wang, Q., Ma, X., Mahlatsi, R.L., Yang, Y., Lyu, J., Fang, H., Wang, Y.: lars2 mutations lead to leigh syndrome with a combined oxidative phosphorylation deficiency. *Orphanet Journal of Rare Diseases* **19**(1), 305 (2024). doi:10.1186/s13023-024-03310-x
9. Amberger, J., Bocchini, C., Hamosh, A.: A new face and new challenges for Online Mendelian Inheritance in Man (OMIM®). *Human Mutation* (2011). doi:10.1002/humu.21466
10. de Palma, L., Boniver, C., Cassina, M., Toldo, I., Nosadini, M., Clementi, M., Sartori, S.: Eating-induced epileptic spasms in a boy with mecp2 duplication syndrome: insights into pathogenesis of genetic epilepsies. *Epileptic Disorders* **14**(4), 414–417 (2012). doi:10.1684/epd.2012.0546
11. Karamysheva, Z.N., Tikhonova, E.B., Karamyshev, A.L.: Granulin in frontotemporal lobar degeneration: Molecular mechanisms of the disease. *Frontiers in Neuroscience* **13**, 395 (2019). doi:10.3389/fnins.2019.00395