



XI JORNADA CIENTÍFICA PROVINCIAL DEL CAPÍTULO DE LA SOCIEDAD DE GENÉTICA HUMANA Y MÉDICA DE HOLGUÍN, 2023

IDENTIFICACIÓN DE RUTAS MOLECULARES COMUNES ENTRE LA ACTIVACIÓN/REPLICACIÓN RETROVIRAL Y LA ATAXIA ESPINOCEREBELOSA TIPO 2

Ramiro Pupo Ramírez¹
Luis Almaguer Mederos²

¹ Ingeniero en Bioinformática, Universidad de las Ciencias Informáticas, La Habana, Cuba, puporamiro23@icloud.com

² Doctor en Ciencias Biológicas, Profesor Titular, Investigador Titular, Universidad Goethe, Frankfurt am Main, Alemania, lalmaguermederos@gmail.com

Resumen:

Introducción. Varias líneas de evidencia sugieren que rutas moleculares asociadas a la activación/replicación retroviral son componentes de relevancia fisiopatológica para enfermedades neurodegenerativas, incluyendo varias enfermedades poliglutamínicas. Sin embargo, se desconoce cuáles rutas moleculares asociadas a la activación/replicación retroviral son de relevancia fisiopatológica para la Ataxia Espinocerebelosa tipo 2, enfermedad poliglutamínica que alcanza las mayores tasas de incidencia y prevalencia a nivel mundial en la provincia Holguín, Cuba. **Objetivo.** Identificar rutas moleculares comunes y de relevancia fisiopatológica, entre la activación/replicación retroviral y la Ataxia Espinocerebelosa tipo 2. **Métodos.** Se realizaron estudios de redes de interacción de proteínas y análisis de vías biológicas mediante *Cytoscape* y *g:GOST*. **Resultados.** El análisis topológico de la intermediación de la red de interacción arrojó que existe un clúster central conformado por las proteínas HNRNPA1, HNRNPK, SYNCRIP, TP53, DDX5, DDX17, TSG101 y EGFR. El análisis de enriquecimiento de vías mostró que este clúster involucra procesos esencialmente relacionados con el metabolismo de ARNs. **Conclusiones.** Las evidencias obtenidas mediante análisis bioinformático apoyan la concepción relativa a la existencia de rutas moleculares comunes a la activación/replicación retroviral y a la Ataxia Espinocerebelosa tipo 2, que involucran procesos relacionados con el metabolismo de ARNs mediados por proteínas de unión a ARN, incluyendo la regulación del empalme de ARN por medio del empalmasoma, la regulación positiva de la maduración de microARN y el silenciamiento de la expresión génica mediado por ARN, y que son de relevancia fisiopatológica.

Palabras clave: Ataxia Espinocerebelosa tipo 2, bioinformática, neurodegeneración, retrovirus, rutas moleculares.

Introducción

La Ataxia Espinocerebelosa tipo 2 (SCA2) es una enfermedad neurodegenerativa poliglutamínica, progresiva e invalidante, que alcanza las más elevadas tasas de incidencia y prevalencia a nivel mundial en la provincia Holguín, Cuba (1). La mutación causante de la SCA2 consiste en la expansión de una secuencia repetitiva de CAG en el primer exón del gen *ATXN2*, por encima de 31 unidades. El gen *ATXN2* codifica la ataxina-2, una proteína citoplasmática de unión a ARN (2–4). La ataxina-2 con dominio poliglutamínico expandido adquiere una ganancia de función neurotóxica, que conduce a una prominente atrofia del cerebelo y de otras estructuras del sistema nervioso central (5,6).

Notablemente, secuencias repetitivas de CAG de 27 a 34 unidades en el gen *ATXN2* se han asociado a un incremento en el riesgo de padecer de Esclerosis Lateral Amiotrófica (ELA), sugiriendo roles esenciales para este gen en la fisiología del sistema nervioso (7–10). Se ha demostrado que la reducción de los niveles de expresión de la ataxina-2 mejora el fenotipo clínico en un modelo transgénico para la ELA (11), que la proteína TDP43 es secuestrada en agregados citoplasmáticos formados por la ataxina-2 expandida en un modelo *knock-in* de la SCA2 y que varias características de este modelo se superponen con rasgos característicos de la ELA, incluyendo la disminución de la biogénesis de colesterol y neurotoxicidad inducida por ARN (12). Estas evidencias sugieren la existencia de mecanismos fisiopatológicos moleculares comunes a la ELA y la SCA2.

Desde hace varias décadas se han acopiado evidencias observacionales, experimentales e *in silico*, que sugieren que retrovirus exógenos o endógenos tienen roles de importancia fisiopatológica en el contexto de enfermedades neurodegenerativas humanas, incluyendo la ELA y a enfermedades poliglutamínicas (13–16). Sin embargo, se desconoce la relevancia fisiopatológica de rutas moleculares asociadas a la activación y replicación retroviral para las ataxias poliglutamínicas y, en particular, para la Ataxia Espinocerebelosa tipo 2.

En consecuencia, en la presente investigación se persigue el objetivo de identificar rutas moleculares comunes y de relevancia fisiopatológica, entre la activación/replicación retroviral y la Ataxia Espinocerebelosa tipo 2.

Métodos

Se obtuvieron listas de proteínas relacionadas con la activación/replicación retroviral y la Ataxia Espinocerebelosa tipo 2 por medio de minería de datos en la base de datos *Ingenuity Pathway Analysis* (IPA, Qiagen) (<http://www.ingenuity.com>). Las listas de proteínas fueron importadas en la aplicación *Cytoscape* (versión 3.8.1) (17) y fueron generadas redes de interacción de proteínas con el uso de la aplicación STRING (v.11.5) en *Cytoscape* (v.3.9.1) (18).

Para visualizar las redes de interacción de proteínas correspondientes a la activación y replicación retroviral o a la Ataxia Espinocerebelosa tipo 2, se utilizó *Cytoscape* (v.3.9.1). Este software también fue usado para realizar análisis estadísticos y topológicos y para la identificación de conjuntos solapados de proteínas, o sub-redes, a partir de las redes de interacción de proteínas previamente obtenidas. Para mejorar la visualización de las redes obtenidas de

interacción de proteínas, se aplicó un filtro de organización de nodos ("Organic Layout") (19). Por otra parte, se redujo el universo de proteínas en las redes de interacción por medio de un filtro de centralidad de intermediación ("*Betweenness Centrality*") de CentiScaPe (19).

Para la detección de regiones densamente conectadas que pueden representar complejos moleculares, se utilizó MCODE (en inglés: "*Molecular Complex Detection*"), una aplicación de *Cytoscape* (20). Otra aplicación de *Cytoscape*, ClueGo (21), se utilizó para realizar análisis de enriquecimiento de vías biológicas para el conjunto núcleo identificado por MCODE. Adicionalmente, se utilizó *CluePedia* de *Cytoscape*, una herramienta de búsqueda de nuevos marcadores potencialmente asociados a rutas biológicas (22). Alternativamente y siguiendo la misma metodología que con ClueGO, se utilizó la herramienta g:GOST de g:Profiler (<http://biit.cs.ut.ee/gprofiler/>) (23). Para la visualización de los resultados obtenidos mediante g:GOST se utilizó el gráfico de Manhattan (24).

Resultados

Se obtuvieron dos redes de interacción proteína-proteína, una perteneciente a proteínas relacionadas con la activación/replicación retroviral y otra de proteínas que interactúan con la Ataxina-2. La primera cuenta con un total de 245 nodos y 3903 aristas y la segunda con 204 nodos y 1854 aristas. El análisis de las redes ofreció coeficientes de clusterización de 0,579 y 0,461 para la red de activación/replicación retroviral y las de proteínas que interactúan con Ataxina-2, respectivamente, lo que implica un grado de conectividad moderado (25).

La red resultante de la fusión de las redes antes descritas, cuenta con un total de 436 nodos y 5721 aristas, y un coeficiente de clusterización de 0,528. Mediante el filtrado de esta red para las proteínas de intermediación o de mayor importancia detectadas en el análisis topológico, se obtuvo una subred integrada por un total de 13 proteínas: HNRPK, TP53, SYNCRIP, DDX17, DDX6, DDX5, HNRPA1, TSG101, EGFR, MOV10, SRC, ZC3HAV1 y ABCE. Este grupo central de proteínas significativas detectadas en el análisis topológico y coincidentes en la intermediación, presentan un alto grado de centralidad y son capaces de conectar con el resto de nodos de la red de forma eficaz.

Por medio del algoritmo MCODE se logró identificar un clúster central en la subred de la intermediación analizada, siguiendo su configuración predeterminada con una puntuación MCODE de 6,0, lo cual indica una capacidad predictiva del 90% (20). El clúster está conformado por ocho proteínas: HNRNPA1, HNRNPK, SYNCRIP, TP53, DDX5, DDX17, TSG101 y EGFR. La presencia de este clúster central sugiere la existencia de rutas moleculares comunes entre la Ataxia Espinocerebelosa tipo 2 y la activación/replicación retroviral.

El enriquecimiento funcional de vías biológicas y ontología génica de ClueGo, realizado a las ocho proteínas del clúster central en la subred de la intermediación, arrojó que este clúster está significativamente asociado a varios procesos biológicos, entre los que se encuentran la ruta intrínseca de la apoptosis, la regulación del empalme de ARNm por medio del empalmasoma, la regulación positiva de la maduración de microRNA y otros relacionados con el silenciamiento de la expresión génica mediado por ARN (Figura 1).

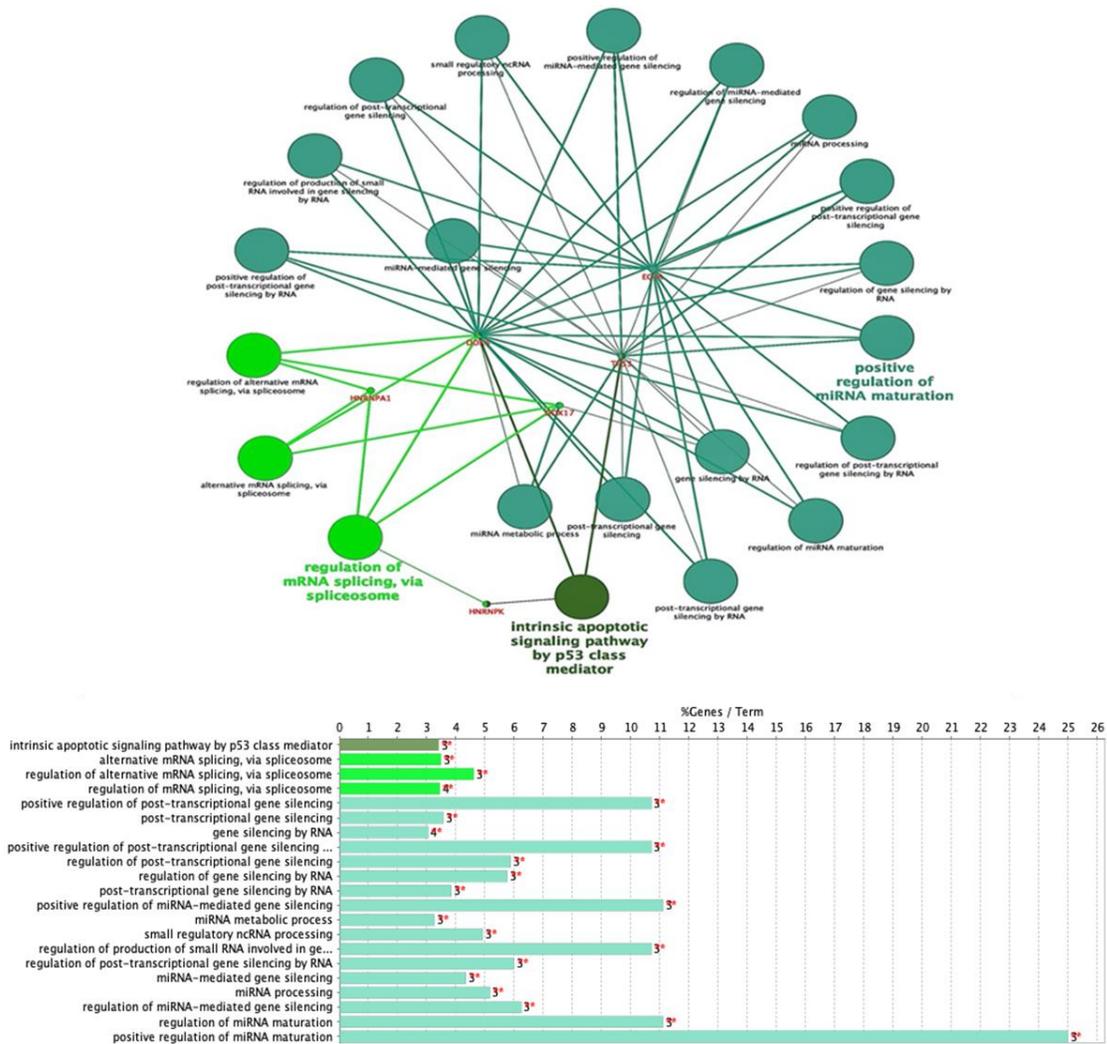


Figura 1. Análisis funcional mediante ontología génica de ClueGo/CluePedia.

Con la herramienta g:GOST de g:Profiler se realizó un análisis del clúster central de proteínas detectadas por MCODE de *Cytoscape*, con el fin de encontrar términos biológicos estadísticamente enriquecidos desde varias fuentes de datos. La herramienta g:GOST permitió realizar un mapeo o perfil funcional de los genes introducidos previamente (Figura 2).

Los resultados obtenidos son coherentes con el conocimiento actual en lo relativo a los procesos de activación/replicación retroviral. Se conoce que para completar su ciclo de infección, los retrovirus secuestran recursos de las células hospederas entre los que se encuentran, de modo particular, proteínas de unión a ARN (26). También se ha demostrado que varias proteínas de unión a ARN son componentes fundamentales de los mecanismos de defensa celular contra la infección por virus (26). Como se mencionó anteriormente, varias de las proteínas identificadas en el clúster central de la subred conformada por las proteínas que interactúan con ataxina-2 y que están vinculadas a la activación/replicación retroviral, son proteínas de unión a ARN.



Figura 2. Gráfico de Manhattan en g:GOSt y procesos biológicos sobrerrepresentados.

Por otro lado, varias proteínas de unión a ARN tienen relevancia fisiopatológica en el contexto de enfermedades neurodegenerativas (2, 13, 16). Notablemente, en el clúster central en la subred de la intermediación analizada, fueron identificadas las proteínas HNRNPA1 y SYNCRIP, proteínas de unión a ARN asociadas a enfermedades neurodegenerativas. De modo significativo, también la ataxina-2 es una proteína de unión a ARN asociada a enfermedades neurodegenerativas (2), y ha sido identificada como un sustrato candidato a degradación por la proteasa 3Cpro de poliovirus en células HeLa (27). Tales observaciones sugieren que la ataxina-2 podría tener algún rol de importancia en los mecanismos de respuesta antiviral a nivel celular, y que tal rol pudiera estar relacionado con su capacidad para unirse y modular el metabolismo de ARNs.

Las evidencias obtenidas en la presente investigación sugieren la existencia de rutas moleculares comunes entre la Ataxia Espinocerebelosa tipo 2 y la activación/replicación retroviral, que esencialmente involucran procesos relacionados con el procesamiento de ARN con la participación de proteínas de unión a ARN. Se requerirá la realización de estudios experimentales para validar los resultados obtenidos y para caracterizar los roles de las proteínas identificadas en el clúster de intermediación, en el contexto de la activación/replicación retroviral y de la Ataxia Espinocerebelosa tipo 2.

Conclusiones

Se identificaron, mediante análisis bioinformático, rutas moleculares de relevancia fisiopatológica comunes a la activación/replicación retroviral y a la Ataxia Espinocerebelosa tipo 2, que involucran el procesamiento y metabolismo de ARNs, la regulación del empalme de ARN por medio del empalmasoma, la regulación

positiva de la maduración de microARN y el silenciamiento de la expresión génica mediado por ARN.

Referencias bibliográficas

1. Scott SS de O, Pedroso JL, Barsottini OGP, França-Junior MC, Braga-Neto P. Natural history and epidemiology of the spinocerebellar ataxias: Insights from the first description to nowadays. *Journal of the Neurological Sciences*. 2020; 417:117082.
2. Auburger G, Sen NE, Meierhofer D, Başak AN, Gitler AD. Efficient Prevention of Neurodegenerative Diseases by Depletion of Starvation Response Factor Ataxin-2. *Trends in Neurosciences*. 2017; 40(8):507-16.
3. Fittschen M, Lastres-Becker I, Halbach M, Damrath E, Gispert S, Azizov M, et al. Genetic ablation of ataxin-2 increases several global translation factors in their transcript abundance but decreases translation rate. *Neurogenetics*. 2015; 16(3):181-92.
4. Lastres-Becker I, Nonis D, Eich F, Klinkenberg M, Gorospe M, Kötter P, et al. Mammalian ataxin-2 modulates translation control at the pre-initiation complex via PI3K/mTOR and is induced by starvation. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 2016; 1862(9):1558-69.
5. Klockgether T, Mariotti C, Paulson HL. Spinocerebellar ataxia. *Nat Rev Dis Primers*. 2019; 5(1):24.
6. Scoles DR, Pulst SM. Spinocerebellar Ataxia Type 2. En: Nóbrega C, Pereira de Almeida L, editores. *Polyglutamine Disorders* [Internet]. Cham: Springer International Publishing; 2018 [citado 18 de septiembre de 2022]. p. 175-95. (Advances in Experimental Medicine and Biology). Disponible en: https://doi.org/10.1007/978-3-319-71779-1_8.
7. Rubino E, Mancini C, Boschi S, Ferrero P, Ferrone M, Bianca S, et al. ATXN2 intermediate repeat expansions influence the clinical phenotype in frontotemporal dementia. *Neurobiology of Aging*. 2019; 73: 231.e7-231.e9.
8. Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. 2010; 466(7310):1069-75.
9. Farhan SMK, Gendron TF, Petrucelli L, Hegele RA, Strong MJ. OPTN p.Met468Arg and ATXN2 intermediate length polyQ extension in families with C9orf72 mediated amyotrophic lateral sclerosis and frontotemporal dementia. *Am J Med Genet B Neuropsychiatr Genet*. 2018; 177(1):75-85.
10. Glass JD, Dewan R, Ding J, Gibbs JR, Dalgard C, Keagle PJ, et al. ATXN2 intermediate expansions in amyotrophic lateral sclerosis. *Brain*. 2022; 145(8):2671-6.
11. Becker LA, Huang B, Bieri G, Ma R, Knowles DA, Jafar-Nejad P, et al. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature*. 2017; 544(7650):367-71.
12. Canet-Pons J, Sen NE, Arsović A, Almaguer-Mederos LE, Halbach MV, Key J, et al. Atxn2-CAG100-KnockIn mouse spinal cord shows progressive TDP43 pathology associated with cholesterol biosynthesis suppression. *Neurobiol Dis*. 2021; 152:105289.

13. Römer C. Viruses and Endogenous Retroviruses as Roots for Neuroinflammation and Neurodegenerative Diseases. *Front Neurosci.* 2021; 15:648629.
14. Schultz JL, Nopoulos PC, Gonzalez-Alegre P. Human Immunodeficiency Virus Infection in Huntington's Disease is Associated with an Earlier Age of Symptom Onset. *J Huntingtons Dis.* 2018; 7(2):163-6.
15. Rex C, Nadeau MJ, Douville R, Schellenberg K. Expression of Human Endogenous Retrovirus-K in Spinal and Bulbar Muscular Atrophy. *Front Neurol.* 2019; 10:968.
16. Li W, Pandya D, Pasternack N, Garcia-Montojo M, Henderson L, Kozak CA, et al. Retroviral Elements in Pathophysiology and as Therapeutic Targets for Amyotrophic Lateral Sclerosis. *Neurotherapeutics.* 2022; 19(4):1085-101.
17. Shannon P, Markiel A, Ozier O, S. Baliga N, T. Wang J, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks [Internet]. 2003 [citado 16 de junio de 2022]. Disponible en: <https://genome.cshlp.org/content/13/11/2498.full>.
18. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res.* 2019; 18(2):623-32.
19. Brandes U, Pich C. Centrality estimation in large networks. *Int J Bifurcation Chaos.* 2007; 17(07):2303-18.
20. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics.* 2003; 4(1):2.
21. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009; 25(8):1091-3.
22. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics.* 2013; 29(5):661-3.
23. Hoffman GB Veronique Voisin, Ruth Isserlin, Robin Haw, Michael. Pathway and Network Analysis of -Omics Data (July 2020) [Internet]. 2020 [citado 13 de enero de 2023]. Disponible en: https://baderlab.github.io/CBW_Pathways_2020/
24. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 2019; 47(W1):W191-8.
25. Li Y, Shang Y, Yang Y. Clustering coefficients of large networks. *Information Sciences.* 2017; 382-383:350-8.
26. Lisy S, Rothamel K, Ascano M. RNA Binding Proteins as Pioneer Determinants of Infection: Protective, Proviral, or Both? *Viruses.* 2021;13(11):2172.
27. Jagdeo JM, Dufour A, Klein T, Solis N, Kleifeld O, Kizhakkedathu J, et al. N-Terminomics TAILS Identifies Host Cell Substrates of Poliovirus and Coxsackievirus B3 3C Proteinases That Modulate Virus Infection. *Journal of Virology.* 2018; 92(8): e02211-17.