

RESEARCH

Estudio del fenotipo: Discalculia

Pascual Mellado Alejandro^{*}
and Ramírez Arco David

^{*}Correspondence:

ale.pas.mel@uma.es

ETSI Informática, Universidad de
Málaga, Málaga, España

Full list of author information is
available at the end of the article

Abstract

La discalculia es una condición que dificulta el desarrollo de ciertas habilidades matemáticas, así como la realización de algunas tareas en las que intervienen operaciones aritméticas. Es similar a la [dislexia](#), fenotipo mas conocido, pero a diferencia de esta, refleja dificultad en eventos lógico-matemáticos, sin que con ello se refleje un déficit en las capacidades cognitivas, ni tampoco se presente dificultades con el tiempo, la medición y el razonamiento espacial.

En el presente proyecto de investigación procederemos a recolectar información, obteniendo asociaciones de genes y patologías en las que estas mismas proteínas están asociadas.

Keywords: Discalculia; Discapacidad matemática; Genes; Interacciones

1 Introducción

(Versión Preliminar)

La Discalculia es una discapacidad de aprendizaje específica que afecta la adquisición de habilidades aritméticas. Aunque la falta de enseñanza, recursos y la baja inteligencia se han relacionado con la etiología de la discalculia, se sabe actualmente que esta discapacidad del aprendizaje es un trastorno cerebral con una predisposición genética familiar [1].

Por otra parte, hay estudios [1, 2] que muestran que una forma de este fenotipo está asociado genéticamente con anomalías tanto funcionales como estructurales del surco intraparietal derecho, llevando esta región un papel fundamental en el desarrollo de las habilidades aritméticas.

Enfermedades como Alzheimer u otras similares tienen que ver con demencias fronto-temporales [3] que están también fuertemente relacionadas con un número significativo de fenotipos relacionados con los genes implicados en nuestro fenotipo a estudiar

1.1 Información sobre los genes a estudiar

A continuación se dará una breve información sobre los genes que de mayor grado de interconexión que hemos encontrado al establecer la red del grafo obtenido con String (véase la figura 3), un recurso que mas tarde usaremos para los experimentos pero que ahora hemos utilizado a modo informativo para encontrar recursos bibliográficos sobre los genes implicados en nuestro fenotipo

SQSTM1[4]: Sequestosoma-1; Receptor de autofagia que interactúa directamente tanto con la carga a degradar como con un modificador de autofagia de la familia MAP1 LC3. Puede regular la activación de NFkB1 por TNF-alfa, factor de crecimiento nervioso (NGF) e interleucina-1. Puede desempeñar un papel en la señalización posterior de titina/TTN en las células musculares.

FUS[5]:Proteína de unión a ARN FUS; Se une tanto al ADN monocatenario como al bicatenario y promueve la hibridación independiente de ATP de los ADN monocatenarios complementarios y la formación del bucle D en el ADN superhelicoidal de doble cadena. Puede desempeñar un papel en el mantenimiento de la integridad genómica; Pertenece a la familia RRM TET.

VCP[6]: ATPasa del retículo endoplásmico de transición; Necesario para la fragmentación de las pilas de Golgi durante la mitosis y para su reensamblaje después de la mitosis. Participa en la formación del retículo endoplásmico de transición (tER). La transferencia de membranas desde el retículo endoplásmico al aparato de Golgi se produce a través de vesículas de transición de 50-70 nm que se derivan de elementos de transición parcialmente rugosos y parcialmente lisos del retículo endoplásmico (tER). La formación de vesículas en el tER es un proceso dependiente de ATP. El complejo ternario que contiene UFD1, VCP y NPLOC4 se une a proteínas ubiquitinadas.

CHMP2B[7]:Proteína corporal multivesicular cargada 2b; Probable componente central de la clasificación endosomal requerida para el complejo de transporte III (ESCRT-III) que está involucrado en la formación de cuerpos multivesiculares (MVB) y la clasificación de proteínas de carga endosomal en MVB. Los MVB contienen vesículas intraluminales (ILV) que se generan por invaginación y escisión de la membrana limitante del endosoma y, en su mayoría, se envían a los lisosomas, lo que permite la degradación de las proteínas de la membrana, como los receptores del factor de crecimiento estimulado, las enzimas lisosomales y los lípidos.

HNRNPA2B1[8]:Ribonucleoproteínas nucleares heterogéneas A2/B1; Ribonucleoproteína nuclear heterogénea (hnRNP) que se asocia con pre-ARNm nacientes y los empaqueta en partículas de hnRNP. La disposición de las partículas de hnRNP en el hnRNA naciente no es aleatoria y depende de la secuencia, y sirve para condensar y estabilizar las transcripciones y minimizar los enredos y los nudos. El empaque juega un papel en varios procesos, como la transcripción, el procesamiento de pre-ARNm, la exportación nuclear de ARN, la ubicación subcelular, la traducción de ARNm y la estabilidad de los ARNm maduros. Forma partículas hnRNP con al menos otras 20 hnRNP diferentes.

2 Materiales y métodos

Se proceden a explicar los distintos recursos que se van a usar para obtener resultados a base de experimentos y procedimientos:

2.1 Métodos

En esta parte exponemos los distintos métodos con los que hemos llevado a cabo los experimentos.

2.1.1 *Análisis por comunidades*

En los análisis de comunidades [9] se busca la detección y el análisis de grupos. Esto es de enorme aplicabilidad en diferentes dominios: La detección de comunidades es un problema que no tiene un objetivos o una solución global esperada, ya que la naturaleza de las comunidades no se conoce de antemano. El problema se vuelve aún más complicado debido al hecho de que las comunidades emergen en la red en varias formas, como disjuntas, superpuestas y jerárquicas [10].

A lo largo de la historia se han propuesto varias heurísticas [11] para abordar estos desafíos, dependiendo de la aplicación en cuestión. Todas estas heurísticas se han materializado en forma de nuevas métricas, que en la mayoría de los casos se utilizan como funciones de optimización para detectar la estructura de la comunidad, o proporcionan una indicación de la bondad de las comunidades detectadas durante la evaluación.

2.1.2 *Enriquecimiento funcional*

Se han desarrollado muchos métodos para inferir y razonar acerca de las redes de interacción molecular. Uno de ellos es el análisis funcional de genes [12]; en éste se anota y se analiza estadísticamente una de listas de genes utilizando métodos estadísticos para identificar anotaciones funcionales sobre los genes analizados para ver si están significativamente relacionados.

Muchas enfermedades humanas están relacionadas [13] entre sí por causas comunes o incluso por patologías parecidas. El conocimiento de estas relaciones se ha explotado durante mucho tiempo para tratar enfermedades similares con los mismos tratamientos. Sin embargo, la mayoría de los enfoques tradicionales para descubrir estas relaciones han dependido de medidas subjetivas, como la similitud en los síntomas, sin obtener el conocimiento completo disponible a través del enriquecimiento.

2.2 Materiales

En esta sección serán explicados los recursos web y las librerías de R utilizadas para los experimentos

2.2.1 *Human Phenotype Ontology*

HPO (Human Phenotype Ontology)[14] es un vocabulario estandarizado de anomalías fenotípicas en enfermedades humanas que utiliza un fenotipado detallado/preciso para poder ser usado a nivel computacional.

2.2.2 *String*

STRING [15] es una base de datos de interacciones proteína-proteína conocidas y previstas. Las interacciones incluyen asociaciones directas (físicas) e indirectas (funcionales); se derivan de la predicción computacional, de la transferencia de conocimiento entre organismos y de interacciones agregadas de otras bases de datos (primarias).

2.2.3 *igraph*

Los objetivos de igraph son proporcionar un conjunto de tipos de datos y funciones para la implementación sencilla de algoritmos gráficos, el manejo rápido de gráficos grandes y permitir la creación rápida de prototipos a través de alta lenguajes de nivel como R

2.2.4 *Linkcomm*

Las comunidades de enlaces revelan la estructura anidada y superpuesta en las redes y descubren los nodos clave que forman conexiones con múltiples comunidades. linkcomm [16] proporciona herramientas para generar, visualizar y analizar comunidades de enlaces en redes de tamaño y tipo arbitrario.

2.2.5 *ClusterProfiler*

clusterProfiler [17] implementa métodos para analizar y visualizar perfiles funcionales de coordenadas genómicas, genes y grupos de genes.

2.2.6 *org.Hs.eg.db*

Anotación de todo el genoma para humanos, basada principalmente en el mapeo utilizando identificadores de genes Entrez

2.2.7 *Gene Ontology (GO)*

La base de conocimientos Gene Ontology (GO) [18] es la fuente de información más grande del mundo sobre las funciones de los genes. Este conocimiento es tanto legible por humanos como por máquinas, y es la base para el análisis funcional de experimentos genéticos y de biología molecular a gran escala en la investigación biomédica.

3 Resultados

Se explicaran ahora los resultados que hemos obtenido usando todos los recursos explicados anteriormente:

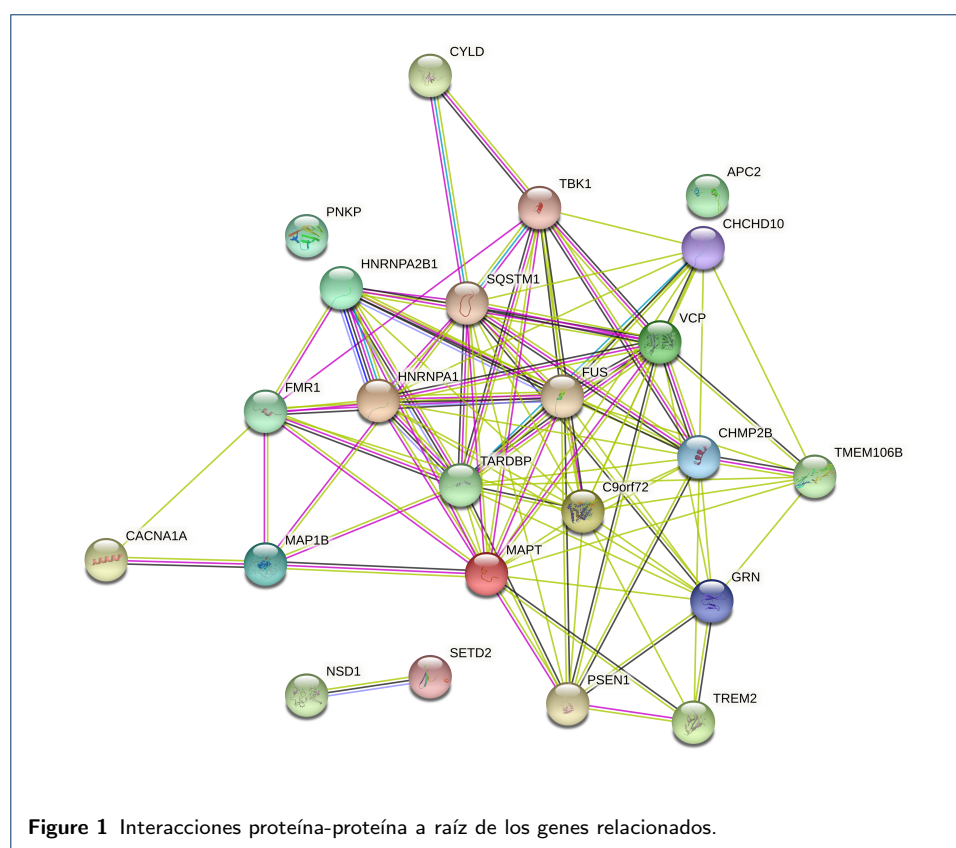
3.1 Human Phenotype Ontology

Tras realizar una búsqueda del fenotipo a investigar, *Dyscalculia*, vemos que se encuentra clasificado como [HP:0002442](#) con un total de 13 fenotipos de otras enfermedades asociadas y un total de 24 genes asociados a esta en HPO.

Descargamos un .csv para cargar el listado de genes con sus links en String

3.2 String

Gracias a las anotaciones en HPO obtenidas del .csv mencionado anteriormente, podemos obtener los genes asociados del fenotipo, y con ello procedemos a la búsqueda de más información acerca de estos utilizando [STRING](#), una base de datos biológica y un recurso web de interacciones entre proteínas.



Una vez cargados los genes y sus relaciones en string, procedemos a descargar los ficheros "string-node-degree.tsv" e "string-interactions.tsv".

Se creara también un objeto network de la librería de R stringdb, con las siguientes características:

Versión=11 Specie=9606 Score threshold=400

Se usara esta network de genes humanos, para buscar genes vecinos de los 24 que teníamos en un principio y realizar una propagación de red con el fin de aumentar el total de genes a estudiar que guarden directa o indirectamente una relación con nuestro fenotipo a estudiar.

3.3 igraph

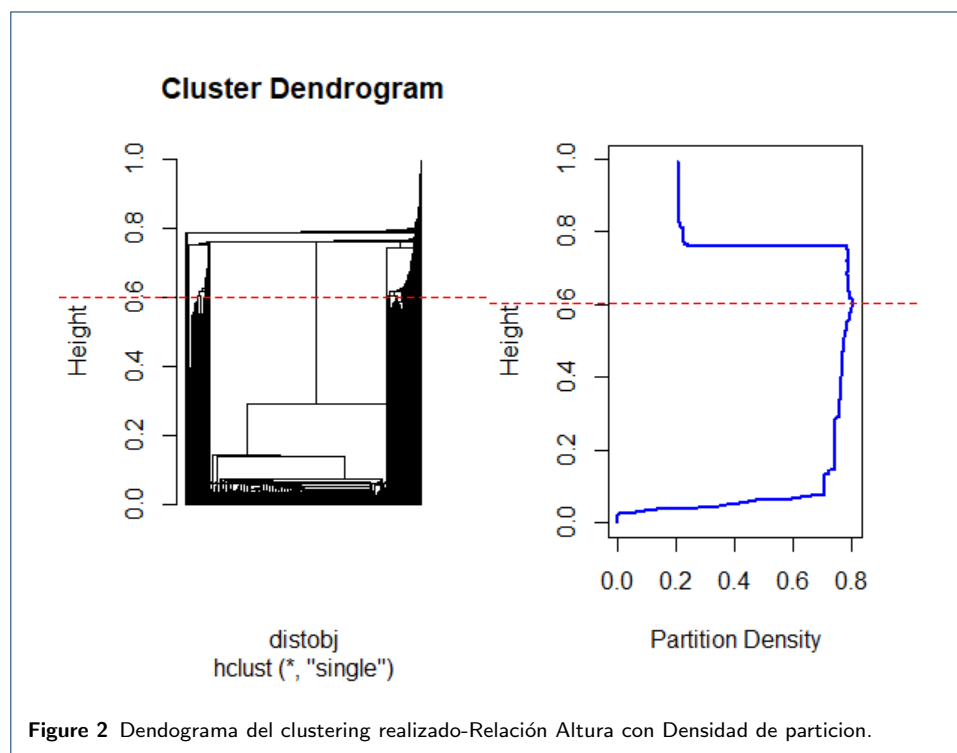
Usaremos igraph para convertir nuestro objeto network (genes originalmente relacionados con el fenotipo mas los genes vecinos de la red general anteriormente citada) en un dataframe que sea procesable por las funciones de clustering para la división en comunidades

3.4 Análisis por comunidades

Se realizan a continuación los correspondientes pasos para clusterizar nuestro conjunto en distintas comunidades y realizar el correspondiente análisis de enriquecimiento funcional a las adecuadas:

3.4.1 LinkComm

Tras una serie de pruebas con distintas funciones de igraph para hacer comunidades, se ha decidido usar la anteriormente citada LinkComm, que nos deja esta imagen como muestra de su trabajo sobre nuestra red de genes:



3.4.2 ClusterProfiler

A continuación se realizara un enriquecimiento funcional con GO (Gene Ontology) para observar las funciones implicadas en la formación de las distintas componentes celulares de las comunidades con un número de genes similar a 20, que supone aproximadamente un 10/100 del total de nuestra red propagada.

Se muestra ahora las principales funciones obtenidas del enriquecimiento:

| ID | Description | GeneRatio | BgRatio | pvalue | p.adjust | qvalue | geneID | Count |
|----------------------------|---------------------------|-----------|-----------|-----------|-----------|--------|---------|-------|
| GO:0035578 | azurophil granule lumen | 2/2 | 91/19869 | 0.0000207 | 0.0001951 | NA | MPO/AGA | 2 |
| GO:0005766 | primary lysosome | 2/2 | 155/19869 | 0.0000605 | 0.0001951 | NA | MPO/AGA | 2 |
| GO:0042582 | azurophil granule | 2/2 | 155/19869 | 0.0000605 | 0.0001951 | NA | MPO/AGA | 2 |
| GO:0005775 | vacuolar lumen | 2/2 | 176/19869 | 0.0000780 | 0.0001951 | NA | MPO/AGA | 2 |
| GO:0034774 | secretory granule lumen | 2/2 | 322/19869 | 0.0002618 | 0.0003858 | NA | MPO/AGA | 2 |
| GO:0060205 | cytoplasmic vesicle lumen | 2/2 | 325/19869 | 0.0002667 | 0.0003858 | NA | MPO/AGA | 2 |

Figure 3 Cabezera de la tabla resultante del enriquecimiento con GO.

4 Discusión

(Versión preliminar de lo que anteriormente fue la introducción)

A raíz de los artículos recomendados por la base de datos STRING [3, 19], suponemos una relación de estos efectos negativos con una mala conexión entre los hemisferios cerebrales, recordemos que el cerebro delega algunas funciones clave como en este caso puede ser la realización de operaciones matemáticas o el reconocimiento numérico y las interconecta a través del Cuerpo Calloso [20]. Por tanto, un mal funcionamiento de este elemento del sistema puede llevar a una incapacidad de conexión de las funciones de los distintos hemisferios.

Tras investigar las distintas interacciones proteína-proteína (véase la figura 3) y observar las enfermedades relacionadas con la discalculia en HPO y los distintos artículos científicos citados anteriormente,suponemos la estrecha relación de la discapacidad con un mal funcionamiento del sistema nervioso. En concreto se podría teorizar que existe una relación con las enfermedades de demencia del complejo frontotemporal del cerebro y la esclerosis lateral amiotrófica, pues estas palabras claves aparecen en la mitad de las enfermedades relacionadas con el fenotipo.

5 Conclusiones

Abreviaciones

HPO: Human Phenotype Ontology
IPP: Interacciones proteína-proteína

Disponibilidad de datos y materiales

[Proyecto en GitHub](#)

Contribución de los autores

A.P.M.: Búsqueda de información extra sobre los genes a estudiar;
D.R.A.: Ampliación y cambios en la introducción;
A.P.M.: Búsqueda de métodos de extracción de información de STRING;
D.R.A.: Encargado de la búsqueda bibliográfica científica sobre la discalculia y posibles genes relacionados

Author details

ETSI Informática, Universidad de Málaga, Málaga, España.

References

- Molko, N., Cachia, A., Rivière, D., Mangin, J.-F., Bruandet, M., Bihan, D.L., Cohen, L., Dehaene, S.: Functional and structural alterations of the intraparietal sulcus in a developmental dyscalculia of genetic origin. *Neuron* **40**, 847–858 (2003). doi:[10.1016/S0896-6273\(03\)00670-6](#)
- Shalev, R.S., Gross-Tsur, V.: Developmental dyscalculia. *Pediatric Neurology* **24**, 337–342 (2001). doi:[10.1016/S0887-8994\(00\)00258-7](#)
- Walterfang, M., Luders, E., Looi, J.C.L., Rajagopalan, P., Velakoulis, D., Thompson, P.M., Lindberg, O., Östberg, P., Nordin, L.E., Svensson, L., Wahlund, L.-O.: Shape analysis of the corpus callosum in alzheimer's disease and frontotemporal lobar degeneration subtypes. *Journal of Alzheimer's Disease* **40**, 897–906 (2014). doi:[10.3233/JAD-131853](#)
- Komatsu, M., Kageyama, S., Ichimura, Y.: p62/sqstm1/a170: Physiology and pathology. *Pharmacological Research* **66**, 457–462 (2012). doi:[10.1016/j.phrs.2012.07.004](#)
- Schwechheimer, C., Deng, X.-W.: The cop/det/fus proteins—regulators of eukaryotic growth and development. *Seminars in Cell & Developmental Biology* **11**, 495–503 (2000). doi:[10.1006/scdb.2000.0203](#)
- van den Boom, J., Meyer, H.: Vcp/p97-mediated unfolding as a principle in protein homeostasis and signaling. *Molecular Cell* **69**, 182–194 (2018). doi:[10.1016/j.molcel.2017.10.028](#)
- Parkinson, N., Ince, P.G., Smith, M.O., Highley, R., Skibinski, G., Andersen, P.M., Morrison, K.E., Pall, H.S., Hardiman, O., Collinge, J., Shaw, P.J., Fisher, E.C.: Als phenotypes with mutations in chmp2b (charged multivesicular body protein 2b). *Neurology* **67**, 1074–1077 (2006). doi:[10.1212/01.wnl.0000231510.89311.8b](#)
- Alarcon, C.R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., Tavazoie, S.F.: Hnrnpa2b1 is a mediator of m6a-dependent nuclear rna processing events. *Cell* **162**, 1299–1308 (2015). doi:[10.1016/j.cell.2015.08.011](#)
- Chakraborty, T., Dalmia, A., Mukherjee, A., Ganguly, N.: Metrics for community analysis. *ACM Computing Surveys* **50**, 1–37 (2018). doi:[10.1145/3091106](#)
- Shen, H.-W.: Detecting the Overlapping and Hierarchical Community Structure in Networks, pp. 19–44 (2013). doi:[10.1007/978-3-642-31821-4_2](#)
- Kennedy, S.R., Prost, S., Overcast, I., Rominger, A.J., Gillespie, R.G., Krehenwinkel, H.: High-throughput sequencing for community analysis: the promise of dna barcoding to uncover diversity, relatedness, abundances and interactions in spider communities. *Development Genes and Evolution* **230**, 185–201 (2020). doi:[10.1007/s00427-020-00652-x](#)
- Poirel, C.L., Owens, C.C., Murali, T.M.: Network-based functional enrichment. *BMC Bioinformatics* **12**, 14 (2011). doi:[10.1186/1471-2105-12-S13-S14](#)
- Suthram, S., Dudley, J.T., Chiang, A.P., Chen, R., Hastie, T.J., Butte, A.J.: Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Computational Biology* **6**, 1000662 (2010). doi:[10.1371/journal.pcbi.1000662](#)
- Robinson, P.N., Köhler, S., Bauer, S., Seelow, D., Horn, D., Mundlos, S.: The human phenotype ontology: A tool for annotating and analyzing human hereditary disease. *The American Journal of Human Genetics* **83**, 610–615 (2008). doi:[10.1016/j.ajhg.2008.09.017](#)
- Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., Jensen, L.J., von Mering, C.: The string database in 2021: customizable protein protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research* **49**, 605–612 (2021). doi:[10.1093/nar/gkaa1074](#)
- Ahn, Y.-Y., Bagrow, J.P., Lehmann, S.: Link communities reveal multiscale complexity in networks. *Nature* **466**, 761–764 (2010). doi:[10.1038/nature09182](#)
- Yu, G., Wang, L.-G., Han, Y., He, Q.-Y.: clusterprofiler: an r package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology* **16**, 284–287 (2012). doi:[10.1089/omi.2011.0118](#)
- Zheng, Q., Wang, X.-J.: Goeast: a web-based software toolkit for gene ontology enrichment analysis. *Nucleic Acids Research* **36**, 358–363 (2008). doi:[10.1093/nar/gkn276](#)
- Ber, I.L.: Genetics of frontotemporal lobar degeneration: An up-date and diagnosis algorithm. *Revue Neurologique* **169**, 811–819 (2013). doi:[10.1016/j.neurol.2013.07.014](#)
- Ewing-Cobbs, L., Hasan, K.M., Prasad, M.R., Kramer, L., Bachevalier, J.: Corpus callosum diffusion anisotropy correlates with neuropsychological outcomes in twins discordant for traumatic brain injury. *AJNR. American journal of neuroradiology* **27**, 879–881 (2006)