

Análise de uma rede de similaridade genética entre a perturbação do espectro do autismo e comorbilidades do foro neurológico e neuropsiquiátrico

Disease similarity network analysis of autism spectrum disorder and comorbid neurological and neuropsychiatric disorders

Joana Vilela^{1,2}, Hugo Martiniano^{1,2}, Ana Rita Marques^{1,2}, João Xavier Santos^{1,2}, Célia Rasga^{1,2}, Guiomar Oliveira^{3,4}, Astrid Moura Vicente^{1,2}

astrid.vicente@insa.min-saude.pt

(1) Departamento de Promoção da Saúde e Doenças Não Transmissíveis, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal

(2) BiolSI-Biosystems & Integrative Sciences Institute. Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

(3) Unidade de Neurodesenvolvimento e Autismo. Serviço do Centro de Desenvolvimento da Criança. Centro de Investigação e Formação Clínica. Hospital Pediátrico, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

(4) Clínica Universitária de Pediatria e Centro de Imagem Biomédica e Investigação Translacional de Coimbra. Faculdade de Medicina, Universidade de Coimbra, Coimbra, Portugal

_Resumo

A Perturbação do Espectro do Autismo (PEA) é uma perturbação do neurodesenvolvimento com apresentação clínica heterogénea, nível de gravidade variável e ocorrência de múltiplas comorbilidades. A PEA tem uma arquitetura genética complexa que se reflete na sua heterogeneidade clínica, existindo evidência de uma sobreposição de genes alterados entre esta condição e diversas comorbilidades do foro neurológico e neuropsiquiátrico. Neste estudo, construímos uma rede de interação entre doenças baseada na similaridade genética, para explorar a componente genética partilhada entre a PEA e comorbilidades neurológicas e neuropsiquiátricas. As doenças analisadas incluem o Défice Intelectual (DI), a Perturbação de Hiperatividade/Défice de Atenção (PHDA) e a Epilepsia, bem como outras doenças neuropsiquiátricas como a Esquizofrenia (SCZ) e a Perturbação Bipolar (PB). Usando a base de dados de doenças da DisGeNET, a similaridade genética entre as doenças analisadas foi calculada a partir do coeficiente de Jaccard entre pares de doenças, e o algoritmo de Leiden foi usado para identificar comunidades de doenças na rede. Identificámos uma comunidade heterogénea de doenças geneticamente mais semelhantes à PEA, que inclui a Epilepsia, a PB, a PHDA com apresentação combinada, e algumas perturbações no espectro da SCZ. Esta abordagem permitiu obter uma maior clarificação acerca da componente genética partilhada entre a PEA e comorbilidades neurológicas e neuropsiquiátricas, com implicações importantes para a nosologia, fisiopatologia e o tratamento o personalizado da doença.

_Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with heterogeneous clinical presentation, variable severity, and multiple comorbidities. A complex underlying genetic architecture matches the clinical heterogeneity, and evidence indicates that several co-occurring brain disorders share a genetic component with ASD. In this study, we established a genetic similarity disease network approach to explore the shared genetics between ASD and frequent comorbid brain diseases, namely Intellectual Disability (ID), Attention-Deficit/Hyperactivity Disorder (ADHD), and Epilepsy, as well as other rarely co-occurring neuropsychia-

tric conditions in the Schizophrenia (SCZ) and Bipolar Disease spectrum (BP). Using sets of disease-associated genes curated by the DisGeNET database, disease genetic similarity was estimated from the Jaccard coefficient between disease pairs, and the Leiden detection algorithm was used to identify network disease communities. We identified a heterogeneous brain disease community that is genetically more similar to ASD, and that includes Epilepsy, BP, ADHD combined type, and some disorders in the SCZ. This approach enabled further clarification of genetic sharing between ASD and brain disorders. Understanding gene sharing across disorders has important implications for disease nosology, pathophysiology, and personalized treatment.

_Introdução

A Perturbação do Espectro do Autismo (PEA) é uma perturbação do neurodesenvolvimento caracterizada por défices na comunicação e na interação social, e por padrões comportamentais repetitivos (1). A PEA apresenta-se como um espectro clínico, dado que os indivíduos afetados apresentam graus variáveis de gravidade clínica, agravados por diversas comorbilidades. A PEA tem uma arquitetura genética complexa, com centenas de genes implicados na sua etiologia (2).

Diversas comorbilidades neurológicas e neuropsiquiátricas podem acompanhar a PEA (3). No entanto, a combinação de diferentes comorbilidades varia entre os indivíduos afetados, o que poderá ser resultado da contribuição de diferentes fatores de risco. Frequentemente, a PEA ocorre simultaneamente com a Perturbação de Hiperatividade/Défice de Atenção (PHDA) e a Epilepsia em crianças com alterações cromossó-

micas raras (4). Existe evidência de que a PEA e a Epilepsia têm uma sobreposição significativa de genes com alterações (5). Em linha com esta constatação, foi demonstrado que as crianças que têm um irmão mais velho diagnosticado com PEA têm 70% maior probabilidade de ter Epilepsia do que crianças sem irmãos mais velhos com PEA (6). A análise de dados genómicos resultantes de estudos com um grande número de participantes com PEA mostram que a PEA e a Esquizofrenia (SCZ) têm uma partilha genética significativa (7), e que a PEA e o Défice Intelectual (DI) compartilham múltiplos genes com alterações (8). Esta sobreposição genética é mais evidente no quadro de síndromes nas quais os indivíduos apresentam tanto PEA como DI, como é o caso da síndrome de Phelan-McDermid. Esta síndrome envolve alterações estruturais numa região cromossómica que inclui o gene *SHANK3*, gene que está associado tanto ao DI como à PEA. Existe evidência de que mutações neste gene ocorrem em cerca de 1,7% dos indivíduos com DI, 0,5% dos indivíduos com PEA e 2% das pessoas com PEA que também apresentam DI moderado a profundo (9).

_Objetivos

Neste estudo, procurámos identificar o nível de similaridade genética entre a PEA e as comorbilidades do foro neurológico e neuropsiquiátrico, incluindo os subtipos de algumas das doenças analisadas, que possa explicar em parte a heterogeneidade clínica da PEA.

_Material e métodos

Acedemos à base de dados da DisGeNET (<https://www.disgenet.org/>), que inclui dados genéticos relativos a todo o espectro de doenças humanas (10) como fonte de informação sobre associações doença-doença e associações gene-doença. Para cada par de doenças, a DisGeNET fornece o Índice de Jaccard de similaridade de doenças, que calcula a fração de genes compartilhados entre as doenças, de acordo com a proporção do número de genes compartilhados pelo número total de genes implicados em ambas as doenças de um determinado par.

Procedemos à seleção dos termos incluídos na DisGeNet relacionados com a PEA e com diversas perturbações neuropsiquiátricas (3) elencadas na **tabela 1**. Selecionámos ainda os termos relacionados com a Epilepsia. A correspondência entre a nomenclatura definida pelo *Unified Medical Language System* (UMLS) usada na DisGeNET e o *Diagnostic and statistical manual of mental disorders, 5th edition* (DSM-5) (1), foi feita usando os códigos da *International Classification of Diseases ICD-10/ICD-10-CM*, que pertencem ao sistema de codificação médica desenvolvido pela Organização Mundial de Saúde (OMS) (11).

A rede de semelhanças doença-doença foi construída utilizando o Índice de Jaccard para calcular a similaridade genética entre cada par de doenças, e usando a similaridade como peso das ligações que conectam cada par de doenças na rede. A rede foi gerada usando o pacote *igraph* do R (<https://igraph.org/>). As comunidades de doenças da rede foram identificadas usando o algoritmo de Leiden de deteção de comunidades (12) implementado no pacote do R *LeidenAlg* (13).

O nível de conectividade de cada doença dentro da rede é medido pelo número de ligações a outras doenças, definido pelo grau do nó correspondente a essa doença. Os tamanhos dos círculos na rede refletem o grau do nó. A espessura da linha que liga cada par de doenças é proporcional ao peso da ligação e representa o grau de similaridade entre pares de doenças, calculado pelo índice de Jaccard.

Tabela 1: ↓ Termos relativos às doenças selecionadas, de acordo com os códigos da *International Classification of Diseases ICD-10/ICD-10-CM*, o sistema de codificação médica desenvolvido pela OMS para catalogar diversas condições de saúde por categorias de doenças semelhantes. Todas as categorias de doenças foram obtidas a partir do DSM-5, exceto a Epilepsia, por ser uma doença neurológica.

Código ICD-10/ ICD-10-CM	Designação do <i>Unified Medical Language System</i> (UMLS)	Categoria do DSM-5
F23; F23.9	Perturbação Psicótica Breve (PPB)	Espetro da Esquizofrenia e outras perturbações psicóticas
F90; F90.9	Perturbação de Hiperatividade/Défice de Atenção (PHDA)	Perturbações do neurodesenvolvimento
F90.2	Perturbação de Hiperatividade/Défice de Atenção (PHDA), apresentação combinada	Perturbações do neurodesenvolvimento
F84.0	Perturbação do Espectro do Autismo (PEA)	Perturbações do neurodesenvolvimento
F31; F31.9	Perturbação Bipolar (PB)	Perturbação bipolar e perturbações relacionadas
F31.81	Perturbação Bipolar (PB) tipo II	Perturbação bipolar e perturbações relacionadas
F06.1	Catatonía	Ocorre no contexto de diversas categorias
F34.0	Perturbação Ciclotímica (PC)	Perturbação bipolar e perturbações relacionadas
F22.0	Perturbação Delirante (PD)	Espetro da Esquizofrenia e outras perturbações psicóticas
G40;G40.9; G40.909	Epilepsia	(não se aplica)
F90.1	Perturbação Hipercinética (PC)	Perturbações do neurodesenvolvimento
F21	Perturbação da Personalidade Esquizotípica (PPE)	Espetro da Esquizofrenia e outras perturbações psicóticas
F79	Défice Intelectual (DI)	Perturbações do neurodesenvolvimento
F70-F79.9	Défice Intelectual (DI)	Perturbações do neurodesenvolvimento
F70	Défice Intelectual (DI), leve	Perturbações do neurodesenvolvimento
F71	Défice Intelectual (DI), moderado	Perturbações do neurodesenvolvimento
F72	Défice Intelectual (DI), grave	Perturbações do neurodesenvolvimento
F73	Défice Intelectual (DI), profundo	Perturbações do neurodesenvolvimento
F29	Psicose não-orgânica (PNO) não especificada	Espetro da Esquizofrenia e outras perturbações psicóticas
F42;F42.9; F42.8	Perturbação obsessivo-compulsiva (POC)	Perturbação obsessivo-compulsiva e perturbações relacionadas
F25	Perturbação esquizoafetiva (PE), tipo bipolar	Espetro da Esquizofrenia e outras perturbações psicóticas
F20.81	Perturbação esquizofreniforme (PEF)	Espetro da Esquizofrenia e outras perturbações psicóticas

_Resultados

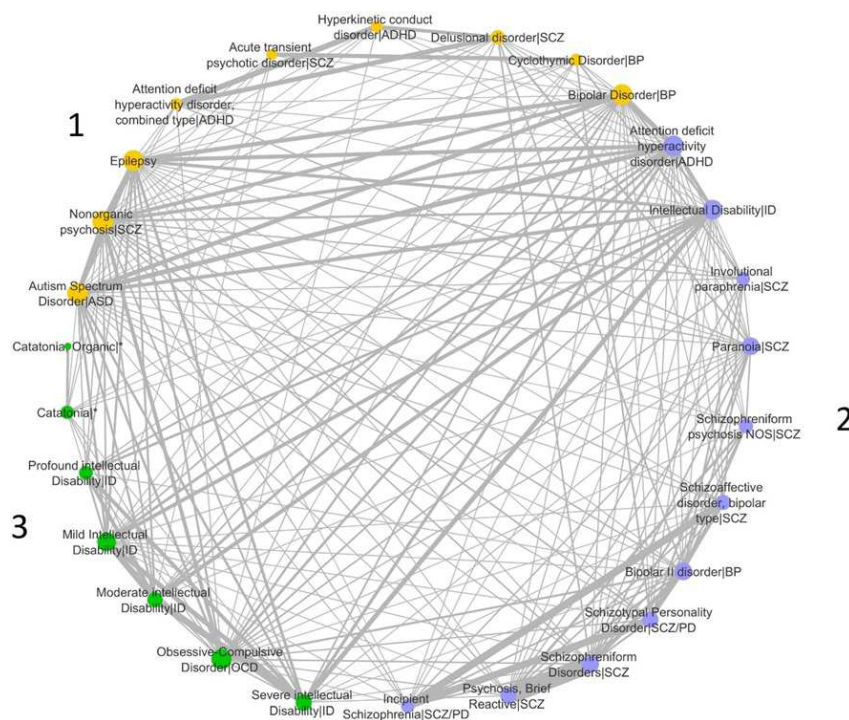
Neste estudo analisámos 309 associações doença-doença com base na similaridade genética entre um conjunto de doenças (**tabela 1**) extraídas das associações doença-doença presentes na DisGeNET. Foram identificadas três comunidades de doenças (**figura 1**). A PEA insere-se na Comunidade 1 juntamente com a Epilepsia, a PHDA (apresentação combinada), a Perturbação Hipercinética (PH), a Perturbação Bipo-

lar (PB), a Perturbação Ciclotímica (PC), e algumas das doenças incluídas no SCZ como a Perturbação Psicótica Breve (PPB). A Comunidade 2 integra a maioria das doenças do espectro da SCZ como a Perturbação da Personalidade Esquizotípica (PPE), a Perturbação esquizoafetiva (PE) tipo bipolar, ou a Perturbação esquizofreniforme (PEF), bem como a PHDA, a PB do tipo II e o DI. Os subtipos de DI estão também presentes na Comunidade 3, juntamente com a POC e a Catatonía (**figura 1**).

A PEA, a PB, a Epilepsia e a Psicose Não Orgânica (PNO) na Comunidade 1, são as doenças que apresentam um maior número de conexões com outras doenças em toda a rede (**figura 1**). A força da conexão entre doenças, definida pelo índice de Jaccard que calcula a similaridade genética entre cada par, é representada pela espessura da linha que as une (**figura 1**), sendo as linhas mais espessas representativas das conexões mais fortes. Observámos que as conexões mais fortes dentro da comunidade da PEA são estabelecidas com a Epilepsia e a PB. A PEA apresenta ainda conexões muito fortes com PHDA e o DI na Comunidade 2 (**figura 1**). PHDA e DI são as doenças

com maior número de ligações a outras doenças (maior grau de nó) na Comunidade 2. Na comunidade 3, a POC e o DI leve (incluindo dificuldades discretas de aprendizagem, comunicação e autonomia), e grave (necessidade de apoio em todas as atividades quotidianas) apresentam o maior número de ligações com outras doenças (**figura 1**). Os subtipos de DI estão agrupados principalmente na Comunidade 3, e podemos observar uma forte conexão entre o DI (na Comunidade 2) e a maioria desses subtipos, independentemente do nível de gravidade. Fortes conexões são também detetadas entre a PHDA e a Perturbação Delirante (PD), e entre a Epilepsia e a PB.

Figura 1: Rede de similaridade de doenças.



As diferentes comunidades estão indicadas com cores diferentes na rede, e as doenças incluídas em cada comunidade são exibidas na figura. Os nós da rede (círculos) representam as doenças. As ligações representam as associações doença-doença. A espessura da linha é proporcional ao peso da ligação e representa o grau de similaridade entre pares de doenças, calculado pelo índice de Jaccard. Os tamanhos dos círculos refletem o número de ligações que uma doença possui com outras doenças da rede, definido pelo grau do nó. ASD, Perturbação do Espectro do Autismo; ADHD, Perturbação de Hiperatividade/Défice de Atenção; SCZ, Doença incluída no Espectro da Esquizofrenia e outras perturbações psicóticas; ID, Doença incluída no Défice Intelectual; BP, Doença incluída na Perturbação Bipolar e perturbações relacionadas; SCZ/PD, Doença incluída na Perturbação da Personalidade e no Espectro da Esquizofrenia e outras perturbações psicóticas; OCD, Doença incluída na Perturbação obsessivo-compulsiva e perturbações relacionadas; *Ocorre no contexto de distúrbios de diferentes categorias de doenças.

_Discussão

Neste estudo investigámos a componente genética comparilhada entre a PEA e comorbilidades do foro neurológico e neuropsiquiátrico. A nossa hipótese era a de que a variabilidade clínica que caracteriza a PEA, nomeadamente a ocorrência variável de comorbilidades nos doentes, possa ser parcialmente explicada por alterações genéticas compartilhadas com as referidas comorbilidades. Através da análise de uma rede de interações doença-doença, demonstrámos que a PEA é uma das cinco doenças com mais conexões na rede de similaridade de doenças produzida. Este resultado indica que a PEA compartilha alterações em genes com várias das doenças analisadas.

A natureza heterogénea da Comunidade 1 sugere que a PEA tem semelhanças genéticas com um número maior de doenças do que a SGZ, a PB e o DI. Esta observação é concordante com a nossa hipótese, dado que a apresentação clínica da PEA é bastante variável, composta por múltiplas comorbilidades, e tendo em consideração que esta variabilidade é frequentemente atribuída ao grande número de genes que têm sido implicados na sua etiologia (2).

O algoritmo de deteção de comunidades usado posicionou o DI e a PEA em comunidades separadas. Embora se saiba que a PEA e o DI partilham genes, estes resultados mostram que a sobreposição genética existente entre PEA e DI não é suficiente para colocar ambas as doenças na mesma comunidade. Isto pode indicar que indivíduos com PEA e DI podem ter mutações num subconjunto de genes que estão mais relacionados com o DI do que os que caracterizam os indivíduos com PEA sem DI.

_Conclusões

Este estudo fornece novas evidências de uma arquitetura genética compartilhada entre a PEA e outras doenças neurológicas e neuropsiquiátricas, incluindo comorbilidades frequentes da PEA.

Estudos adicionais deverão ser desenvolvidos no futuro com o objetivo de expandir esta análise para outras comorbilidades da PEA que não estão no âmbito da doença mental. Pretende-se focar em outros tipos de doenças, como as que afetam o sistema imunitário e que representam um fator de risco para a ocorrência de alterações no desenvolvimento neurológico (14), podendo estar relacionadas com uma constituição geneticamente determinada do microbioma (15).

Referências bibliográficas:

- (1) American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Arlington, VA: American Psychiatric Association, 2013.
- (2) Ramaswami G, Geschwind DH. Genetics of autism spectrum disorder. *Handb Clin Neurol*. 2018;147:321-29. <https://doi.org/10.1016/B978-0-444-63233-3.00021-X>
- (3) White SW, Maddox BB, Mazefsky CA. The Oxford Handbook of Autism and Co-Occurring Psychiatric Conditions. Oxford: Oxford University Press, 2020. doi: 10.1093/oxfordhb/9780190910761.001.0001
- (4) Lo-Castro, A., and Curatolo, P. Epilepsy associated with autism and attention deficit hyperactivity disorder: Is there a genetic link? *Brain Dev*. 36, 185–193. <https://doi.org/10.1016/j.braindev.2013.04.013>
- (5) Novarino G, Baek ST, Gleeson JG. The sacred disease: the puzzling genetics of epileptic disorders. *Neuron*. 2013 Oct 2;80(1):9-11. <https://doi.org/10.1016/j.neuron.2013.09.019>
- (6) Christensen J, Overgaard M, Parner ET, et al. Risk of epilepsy and autism in full and half siblings-A population-based cohort study. *Epilepsia*. 2016 Dec;57(12):2011-2018. <https://doi.org/10.1111/epi.13595>
- (7) Brainstorm Consortium; Anttila V, Bulik-Sullivan B, Finucane HK, et al. Analysis of shared heritability in common disorders of the brain. *Science*. 2018 Jun 22;360(6395):eaap8757. <https://doi.org/10.1126/science.aap8757>
- (8) Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. 2020 Feb 6;180(3):568-584.e23. <https://doi.org/10.1016/j.cell.2019.12.036>
- (9) Soorya L, Leon J, Trelles MP, et al. Framework for assessing individuals with rare genetic disorders associated with profound intellectual and multiple disabilities (PIMD): the example of Phelan McDermid Syndrome. *Clin Neuropsychol*. 2018 Aug-Oct;32(7):1226-1255. <https://doi.org/10.1080/13854046.2017.1413211>
- (10) Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res*. 2020 Jan 8;48(-D1):D845-D855. <https://doi.org/10.1093/nar/gkz1021>
- (11) World Health Organization. ICD-10: International statistical classification of diseases and related health problems: 10th revision, 2nd ed. Geneva: WHO, 2004. (v.3 Alphabetical index). <https://iris.who.int/handle/10665/42980>
- (12) Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*. 2019 Mar 26;9(1):5233. <https://doi.org/10.1038/s41598-019-41695-z>
- (13) Kharchenko P, Petukhov V, Traag VA, et al. leidenAlg: Implements the Leiden Algorithm via an R Interface (1.0.2) [online]. [consult. 01/03/2024]. <https://CRAN.R-project.org/package=leidenAlg>
- (14) Meltzer A, Van de Water J. The Role of the Immune System in Autism Spectrum Disorder. *Neuropsychopharmacology*. 2017 Jan;42(1):284-98. <https://doi.org/10.1038/npp.2016.158>
- (15) Liu Z, Mao X, Dan Z, et al. Gene variations in autism spectrum disorder are associated with alteration of gut microbiota, metabolites and cytokines. *Gut Microbes*. 2021 Jan-Dec;13(1):1-16. <https://doi.org/10.1080/19490976.2020.1854967>