

Идентификация и описание ключевых генов, связанных с патогенезом расстройств аутистического спектра посредством реконструкции и анализа генных сетей

Колос А.В.^{1*}, Можина В.В.², Куцин И.Ю.³, Минасазова А.Р.¹, Ефимочкина С.М.¹, Савина Е.А.¹

¹Первый Московский государственный медицинский университет им. И.М. Сеченова, Москва, Россия

²ФГБОУ ВО МЗ РФ, Волгоградский государственный медицинский университет, Волгоград, Россия

³ФГБОУ ВО МЗ РФ, Омский государственный медицинский университет, Омск, Россия

* kolos_a_v@student.sechenov.ru

Ключевые слова: расстройства аутистического спектра; генные сети; анализ генных онтологий; глутаматергические синапсы

Мотивация и цель: На сегодняшний день точные молекулярные механизмы, лежащие за биологией расстройств аутистического спектра (РАС), не выяснены. Однако количество «надежных» генов, выявленных разными исследовательскими группами, продолжает увеличиваться [1]. Несмотря на гетерогенную природу РАС, определенные гены встречаются одновременно в различных списках, вызывая интерес к пониманию их роли и взаимодействий друг с другом. В этом контексте представляется важным проанализировать эти гены и их взаимосвязи с использованием системного подхода, такого как реконструкция генных сетей.

Методы и алгоритмы: Для получения списка ключевых генов использовалась база данных ассоциации генов с нарушениями нервного развития GeneTrek [2], внутри которой были выбраны списки, содержащие гены, ассоциированные с расстройствами аутистического спектра, а именно: SFARI [3] категории 1 – high-confidence, SPARK [4], SysNDD [5] и DDG2P [6] категории definitive genes, DBD с мутациями, приводящими к потере функции [7]. Реконструкция генной сети была выполнена в STRINGdb, затем экспортирована в Cystoscape для дальнейшего анализа и визуализации генных онтологий. Анализ генных онтологий для списка генов был выполнен с помощью PANTHER Overrepresentation Test [8]. Использовался тест Фишера с поправкой Бонферрони на множественное тестирование. Получившийся список из 71 гена аннотировали с помощью SynGo v1.2 – аннотации синаптических генных онтологий [9].

Результаты: Генная сеть связанная и образует функциональные кластеры, на что указывают значения PPI enrichment p -value генной сети $< 1.0e-16$ и средний коэффициент локальной кластеризации, равный 0.493. Средняя степень узла равна 9.44, а количество ребер – 335, из которых 240 получены путем текстового майнинга научных публикаций, 26 – из результатов экспериментов, 28 – из баз данных и 9 из них – данные коэкспрессии. Большинство коэкспрессируемых пар участвуют в регуляции апоптоза, клеточной пролиферации и нейрогенезе. Согласно SynGO, среди 71 гена 21 ген имеет синаптическую локализацию и/или функцию. В то время как анализ генных онтологий молекулярной функции

(GO molecular function) в PANTHER определил как значимые онтологии: активность протеин-цистеинметилтрансферазы с $p\text{-value} = 3.70\text{E-}02$, среди которых S-аденозилметионин-зависимая активность метилтрансферазы, занимающаяся переносом одноуглеродных групп с $p\text{-value} = 3.95\text{E-}02$, и активность протеинметилтрансферазы с $p\text{-value} = 1.45\text{E-}04$, активность ацетилтрансферазы гистона H3K27 с $p\text{-value} = 3.70\text{E-}02$, неметилованное связывание CpG посредством связывания органических циклических соединений с $p\text{-value} = 4.81\text{E-}03$ и АТФ-зависимая активность ремоделирования хроматина $5.56\text{E-}04$. А анализ оверрепрезентации генных онтологий клеточной локализации выявил как наиболее значимые: пресинаптическую мембрану с $p\text{-value} = 3.13\text{E-}03$, хромосомную и центромерную область с $p\text{-value} = 3.47\text{E-}03$, глутаматергические синапсы $p\text{-value} = 2.73\text{E-}02$ и хроматин с $p\text{-value} = 1.91\text{E-}08$. Что позволяет предполагать, что глутаматергическая гипотеза патофизиологии РАС [10] может стоять за значительным количеством фенотипов РАС.

Выводы: Реконструкция генной сети генов, обнаруживаемых независимыми исследовательскими группами, позволила определить наличие функциональных кластеров РАС. Учитывая высокую гетерогенность, вероятно, что на настоящий момент гены, экспрессирующиеся в пресинаптической мембране, хромосомной и центромерной области, а также глутаматергических синапсах, преобладают в большом количестве фенотипов аутизма. Это представляется важным для развития возможных путей рациональной терапии, которые не представлялись возможными до эпохи успешного открытия генов.

Финансирование: Исследование поддержано РНФ (грант 24-24-00563, «Разработка цифровых образовательных программ в биомедицине»).

Identification and description of key genes associated with the pathogenesis of autism spectrum disorders through reconstruction and analysis of gene networks

Kolos A.V.^{1*}, Mozhina V.V.², Kutsin I.Yu.³, Minasazova A.R.¹, Efimochkina S.M.¹, Savina E.A.¹

¹ I.M. Sechenov First Moscow State Medical University, Moscow, Russia

² Volgograd State Medical University, Volgograd, Russia

³ Omsk State Medical University, Omsk, Russia

* kolos_a_v@student.sechenov.ru

Key words: autism spectrum disorder; gene networks; gene ontology analysis; glutamatergic synapses

Motivation and Aim: To date, the precise molecular mechanisms underlying the biology of autism spectrum disorders (ASD) remain unclear. However, the number of “high-confidence” genes identified by various research groups continues to grow [1]. Despite the heterogeneous nature of ASD, certain genes appear concurrently in different lists, sparking interest in understanding their roles and interactions with each other. In this context, it is important to analyze these genes and their relationships using a systems approach, such as gene network reconstruction.

Methods and Algorithms: To obtain a list of key genes, the GeneTrek database linking genes to neurodevelopmental disorders was utilized [2], from which lists were selected

containing genes associated with autism spectrum disorders, specifically SFARI [3] category 1 – high-confidence, SPARK [4], SysNDD [5], and DDG2P [6] definitive genes categories, as well as DBD with loss-of-function mutations [7]. Gene network reconstruction was conducted in STRINGdb, then exported to Cytoscape for further analysis and visualization of gene ontologies. Analysis of gene ontologies for the gene list was performed using the PANTHER Overrepresentation Test [8]. The Fisher test with Bonferroni correction for multiple testing was employed. The resulting list of 71 genes was annotated using SynGo v1.2 – annotations of synaptic gene ontologies [9].

Results: The gene network is cohesive and forms functional clusters, as indicated by the PPI enrichment p-value of the gene network $< 1.0e-16$ and an average clustering coefficient of 0.493. The average node degree is 9.44, with 335 edges, of which 240 are derived from text mining of scientific publications, 26 from experimental findings, 28 from databases, and 9 from co-expression data. Most co-expressed gene pairs are involved in the regulation of apoptosis, cell proliferation, and neurogenesis. According to SynGO, among the 71 genes, 21 genes have synaptic localization and/or function. Meanwhile, analysis of gene ontologies in molecular function (GO molecular function) using PANTHER identified significant ontologies: protein-cysteine S-methyltransferase activity with a p-value of $3.70E-02$, including S-adenosylmethionine-dependent methyltransferase activity involved in one-carbon group transfer with a p-value of $3.95E-02$ and protein methyltransferase activity with a p-value of $1.45E-04$, histone H3K27 acetyltransferase activity with a p-value of $3.70E-02$, unmethylated CpG binding through binding of organic cyclic compounds with a p-value of $4.81E-03$, and ATP-dependent chromatin remodeling activity at $5.56E-04$. Overrepresentation analysis of gene ontologies in cellular localization identified the most significant as: presynaptic membrane with a p-value of $3.13E-03$, chromosomal and centromeric regions with a p-value of $3.47E-03$, glutamatergic synapses with a p-value of $2.73E-02$, and chromatin with a p-value of $1.91E-08$. This suggests that the glutamatergic hypothesis of ASD pathophysiology [10] may underlie a significant number of ASD phenotypes.

Conclusion: Reconstruction of a gene network of genes identified by independent research groups has allowed for the identification of functional clusters related to autism spectrum disorders (ASD). Given the high heterogeneity, it is likely that currently, genes expressed in the presynaptic membrane, chromosomal and centromeric regions, as well as glutamatergic synapses, prevail in a large number of autism phenotypes. This is crucial for the development of potential pathways for rational therapy that were not feasible before the era of successful gene discovery.

Funding: This work was supported by Russian Science Foundation (grant project 24-24-00563, “Development of digital education programs in biomedicine”).

Список литературы/References

1. Manoli D.S., State M.W. Autism Spectrum Disorder Genetics and the Search for Pathological Mechanisms. *Am J Psychiatry*. 2021;178(1):30-38. doi 10.1176/appi.ajp.2020.20111608
2. Leblond C.S., Le T.L., Malesys S., Cliquet F., Tabet A.C., Delorme R., Rolland T., Bourgeron T. Operative list of genes associated with autism and neurodevelopmental disorders based on database review. *Mol Cell Neurosci*. 2021;113:103623. doi 10.1016/j.mcn.2021.103623
3. Abrahams B.S., Arking D.E., Campbell D.B. et al. SFARI Gene 2.0: a community-driven knowledge-base for the autism spectrum disorders (ASDs). *Mol Autism*. 2013;4:36. doi 10.1186/2040-2392-4-36
4. SPARK Consortium. SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research. *Neuron*. 2018;97(3):488-493. doi 10.1016/j.neuron.2018.01.015
5. Kochinke K., Zweier C., Nijhof B. et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet*. 2016;98(1):149-164. doi 10.1016/j.ajhg.2015.11.024

6. Wright C.F., Fitzgerald T.W., Jones W.D., Clayton S., McRae J.F. et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet*. 2015;385(9975):1305-1314. doi 10.1016/S0140-6736(14)61705-0
7. Gonzalez-Mantilla A.J., Moreno-De-Luca A., Ledbetter D.H., Martin C.L. A Cross-Disorder Method to Identify Novel Candidate Genes for Developmental Brain Disorders. *JAMA Psychiatry*. 2016;73(3):275-283. doi 10.1001/jamapsychiatry.2015.2692
8. Mi H., Muruganujan A., Ebert D., Huang X., Thomas P.D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419-D426. doi 10.1093/nar/gky1038
9. Koopmans F., van Nierop P., Andres-Alonso M., Byrnes A., Cijssouw T., Coba M.P., Cornelisse L.N. et al. SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron*. 2019;103(2):217-234.e4. doi 10.1016/j.neuron.2019.05.002
10. Galineau L., Arlicot N., Dupont A.C., Briand F., Houy-Durand E. et al. Glutamatergic synapse in autism: a complex story for a complex disorder. *Mol Psychiatry*. 2023;28(2):801-809. doi 10.1038/s41380-022-01860-9