

Спектр CNV у пациентов с нарушениями психомоторного развития

Федотов Д.*, Кашеварова А., Скрыбин Н., Лопаткина М., Васильева О.,
Толмачева Е., Беляева Е., Минайчева Л., Петрова В., Равжаева Е.,
Салюкова О., Сивоха В., Фадюшина С., Сеитова Г., Назаренко Л., Лебедев И.

Научно-исследовательский институт медицинской генетики, Томский НИМЦ РАН, Томск, Россия

* dmitry.fedotov@medgenetics.ru

Ключевые слова: вариации числа копий участков ДНК; CNV; нарушения психомоторного развития; интеллектуальные нарушения

Мотивация и цель: Согласно МКБ-11, нарушения психомоторного развития (НПР) – группа поведенческих и когнитивных расстройств, возникающих в период развития нервной системы (интеллектуальные нарушения, расстройство развития языка или речи, расстройство аутистического спектра, расстройство развития учебных навыков, расстройство развития координации движений, синдром дефицита внимания с гиперактивностью, а также синдром стереотипных движений). К НПР могут приводить как средовые, так и генетические факторы. С развитием методов молекулярного кариотипирования был выявлен такой тип хромосомного дисбаланса, вносящий значительный вклад в этиологию НПР, как вариации числа копий участков ДНК (copy number variation, CNV), которые представляют собой фрагмент ДНК размером более одной тысячи пар нуклеотидов, по числу копий отличающийся от референсного генома [1]. Ежегодно в базе данных NCBI PubMed публикуется свыше 100 исследований относительно спектра CNV среди пациентов с НПР. В результате появляется необходимость в систематизации имеющейся информации. Данное исследование представляет собой комплексный анализ спектра CNV, опубликованного в мировой литературе, а также сравнение имеющейся информации с собственными данными.

Методы и алгоритмы: На первом этапе исследования был проведен поиск литературы в базе данных NCBI PubMed с использованием двух поисковых запросов: “neurodevelopmental disorders, copy number variation” ($n = 1410$) и “neurodevelopmental disorders, aCGH” ($n = 208$). В анализ включались работы, в рамках которых были обследованы только пациенты с НПР на предмет наличия CNV в их геномах. При этом публикации должны были содержать полный список обнаруженных aberrаций с указанием версии сборки генома, все CNV должны были быть классифицированы по клинической значимости. Кроме этого, публикация должна была быть на русском или английском языке и находиться в открытом доступе. В конечный анализ включались исследования, которые содержали вышеуказанную информацию ($n = 7$) [2–8]. На втором этапе была сформирована собственная выборка пациентов с НПР. Пробанды были обследованы с помощью матричной сравнительной геномной гибридизации на микрочипах «SurePrint G3 Human CGH Microarray, 8×60K (Agilent Technologies, США). При доступности материала родителей проводилась количественная ПЦР в режиме реального времени с целью верификации варианта и установления его

происхождения. Все CNV классифицировали на пять групп в соответствии с рекомендациями Американской коллегии по медицинской генетике и геномике совместно с проектом «Клинический геномный ресурс» [9]. Индивиды, имеющие вероятно доброкачественные и доброкачественные абберации, а также CNV размером более 10 Мб, исключались из исследования. Для исключения возможного эффекта взаимодействия нескольких вариантов в конечные группы были включены пациенты с одной патогенетически значимой CNV – 345 (литературные данные) и 209 (собственные данные). На последнем этапе проводилось сравнение спектров CNV относительно числа копий, происхождения и размера. Статистическая обработка данных проводилась с помощью критерия хи-квадрат. Результат считался значимым при $p < 0.05$.

Результаты: Частота патогенных CNV в собственной выборке превышает количество данных аббераций, описанных в литературе – 37.3 и 25.2 % соответственно ($p = 0.003$) (табл. 1).

Таблица 1. Характеристика CNV в сравниваемых группах

Характеристика CNV	Литературные данные	Собственные данные
Клиническая значимость		
Патогенные	87/345 (25.2 %)	78/209 (37.3 %)
Вероятно патогенные	54/345 (15.7 %)	17/209 (8.1 %)
Неопределенной клинической значимости	204/345 (59.1 %)	114/209 (54.5 %)
Число копий		
Микроделеции	162/345 (46.9 %)	97/209 (46.4 %)
Микродупликации	182/345 (52.7 %)	109/209 (52.2 %)
Микротрипликации	1/345 (0.3 %)	3/209 (1.4 %)
Происхождение		
<i>De novo</i>	55/150 (36.7 %)	15/64 (23.4 %)
Материнское	55/150 (36.7 %)	25/64 (39.1 %)
Отцовское	40/150 (26.7 %)	24/64 (37.5 %)
Размер		
До 0.5 м.п.о.	175/345 (50.7 %)	145/209 (69.4 %)
0.5–1 м.п.о.	48/345 (13.9 %)	21/209 (10.0 %)
1–5 м.п.о.	106/345 (30.7 %)	36/209 (17.2 %)
5–10 м.п.о.	16/345 (4.6 %)	7/209 (3.3 %)

При этом значительное количество вариантов было классифицировано как CNV с неопределенной клинической значимостью в обеих группах, что связано с трудностями клинической интерпретации не рекуррентных вновь выявляемых аббераций. В выборке CNV как из литературных, так и собственных данных микроделеции и микродупликации встречались с сопоставимой частотой, с незначительным преобладанием микродупликаций – 52.7 и 52.2 %. Микродупликации традиционно считаются менее патогенными, зачастую имея менее выраженный клинический эффект, и могут расцениваться как варианты с неопределенной клинической значимостью. В обеих группах чаще встречаются CNV, унаследованные от родителей, по сравнению с возникшими *de novo* – 63 % (литературные данные) и 76 % (собственные данные) ($p = 0.059$). Унаследованный характер абберации, с одной стороны, уменьшает вероятность ее патогенности; а

с другой стороны, может указывать на ее неполную пенетрантность и наличие дополнительных генетических, эпигенетических и средовых факторов, влияющих на развитие заболевания. В группе литературных данных преимущественно встречаются CNV размером до 0.5 м.п.о. – 50.7 %, при этом среди CNV из собственных данных частота таких aberrаций значительно выше – 69.4 % ($p < 0.001$). Размер выявляемых вариантов обусловлен, прежде всего, разрешающей возможностью метода анализа. Очевидно, что накопление крупного массива данных относительно вероятно патогенных CNV и вариантов с неопределенной клинической значимостью в одних и тех же регионах может способствовать описанию новых синдромов. В литературных и собственных данных чаще встречались CNV в регионах 7q33, 7q36.2, 8p23.3, 11q23.3, 15q11.2, 16p13.11, 16p12.3, Xp22.31, которые были классифицированы как вероятно патогенные или варианты с неопределенной клинической значимостью.

Заключение: Проведенный анализ позволил описать структуру CNV у пациентов с НПР, выделить частые варианты в одних и тех же регионах, что необходимо для более глубокого понимания их возможного патогенного эффекта. Однако высокая частота определенной CNV не может являться основным фактором для интерпретации ее как патогенного варианта. Существует необходимость в экспериментальном подтверждении патологических эффектов часто встречающихся вероятно патогенных вариантов и CNV с неопределенной клинической значимостью на модельных системах.

Финансирование: Исследование поддержано грантом Российского научного фонда № 21-65-00017, <https://rscf.ru/project/21-65-00017/>.

CNV spectrum in patients with neurodevelopmental disorders

Fedotov D.*, Kashevarova A., Skryabin N., Lopatkina M., Vasilyeva O., Tolmacheva E., Belyaeva E., Minaycheva L., Petrova V., Ravzhaeva E., Salyukova O., Sivoha V., Fadyushina S., Seitova G., Nazarenko L., Lebedev I.

Research Institute of Medical Genetics, Tomsk National Research Medical Center, RAS, Tomsk, Russia

* dmitry.fedotov@medgenetics.ru

Key words: copy number variation; CNV; neurodevelopmental disorders; intellectual disabilities

Motivation and Aim: According to ICD-11, neurodevelopmental disorders (NDDs) are a group of behavioral and cognitive disorders that occur during the nervous system development (intellectual disability, language or speech developmental disorder, autism spectrum disorder, learning disabilities, movement coordination disorder, attention deficit hyperactivity disorder, and stereotypic movement disorder). Both environmental and genetic factors can lead to NDDs. Due to the development of molecular karyotyping methods, a type of chromosomal imbalance that contributes significantly to the etiology of NDDs has been identified – a copy number variation (CNV), which is a DNA fragment of more than one thousand nucleotide pairs that differs in copies from the reference genome [1]. Annually, more than 100 studies regarding the spectrum of CNVs among patients with NDDs are published in the NCBI PubMed database. As a result, there is a need to systematisation of available information. This study was aimed as a comprehensive analysis of the CNV spectrum published in the world literature, and comparison of the available information with our own data.

Methods and Algorithms: In the first stage, a literature search was performed in the NCBI PubMed database using two search terms ‘neurodevelopmental disorders, copy number variation’ ($n=1410$) and ‘neurodevelopmental disorders, aCGH’ ($n=208$). Studies that examined only patients with NDDs for the presence of CNVs in their genomes were included in the analyses. In this case, publications had to contain a complete list of detected aberrations with the genome assembly version, and all CNVs had to be interpreted by clinical significance. In addition, the publication had to be in Russian or English and open access. Studies that contained the above information ($n=7$) were included in the final analysis [2–8]. At the second stage, our own sample of patients with NDDs was formed. The probands were screened using array comparative genomic hybridization (aCGH) on SurePrint G3 Human CGH Microarray, 8×60K (Agilent Technologies, USA). In case of parental material availability, quantitative real-time PCR was performed to verify the variant and establish its origin. All CNVs were classified into five groups according to the recommendations of the American College of Medical Genetics and Genomics in conjunction with the Clinical Genomic Resource Project [9]. Individuals with likely benign and benign aberrations and CNVs larger than 10 Mb were excluded from this study. In order to exclude the possible interaction effect of several variants, patients with one pathogenically significant CNV were included in the final groups, 345 (literature data) and 209 (own data). In the last step, set of CNV was compared with respect to the number of copies, origin and size. Statistical analysis of the data was performed using chi-square criteria. The result was considered significant at $p < 0.05$.

Results: The frequency of pathogenic CNVs in our own data exceeds the number of these aberrations found in the literature, 37.3 and 25.2 %, respectively ($p = 0.003$) (Table 1).

Table 1. Characterisation of CNVs in the compared groups

Characteristic	Literature data	Own data
Clinical significance		
Pathogenic	87/345 (25.2 %)	78/209 (37.3 %)
Likely pathogenic	54/345 (15.7 %)	17/209 (8.1 %)
Uncertain clinical significance	204/345 (59.1 %)	114/209 (54.5 %)
Number of copies		
Microdeletions	162/345 (46.9 %)	97/209 (46.4 %)
Microduplications	182/345 (52.7 %)	109/209 (52.2 %)
Microtriplications	1/345 (0.3 %)	3/209 (1.4 %)
Origin		
De novo	55/150 (36.7 %)	15/64 (23.4 %)
Maternal	55/150 (36.7 %)	25/64 (39.1 %)
Paternal	40/150 (26.7 %)	24/64 (37.5 %)
Size		
Up to 0.5 Mb	175/345 (50.7 %)	145/209 (69.4 %)
0.5–1 Mb	48/345 (13.9 %)	21/209 (10.0 %)
1–5 Mb	106/345 (30.7 %)	36/209 (17.2 %)
5–10 Mb	16/345 (4.6 %)	7/209 (3.3 %)

At the same time, a significant number of variants were classified as CNVs with uncertain clinical significance in both groups, which is associated with difficulties in clinical interpretation of non-recurrent new aberrations. In both the CNV sample from the literature and our own data microdeletions and microduplications occurred with

equal frequency with a slight predominance of microduplications, 52.7 and 52.2 %. Microduplications have traditionally been considered less pathogenic, often having less clinical effect, and may be considered as variants of uncertain clinical significance. Parentally inherited CNVs were more common in both groups, 63 % (literature data) and 76 % (own data) ($p = 0.059$). The inherited nature of the aberration, on the one hand, reduces the likelihood of its pathogenicity; on the other hand, it may indicate its incomplete penetrance and the existence of additional genetic, epigenetic and environmental factors influencing the development of the disease. CNVs as small as 0.5 Mb are predominantly found in the literature data group – 50.7 %, while the frequency of such aberrations is significantly higher among CNVs from our own data – 69.4 % ($p < 0.001$). The size of the identified variants is primarily determined by the resolution of the analysis method. Obviously, the accumulation of a large amount of data regarding likely pathogenic CNVs and variants with uncertain clinical significance in the same regions may contribute to the description of new syndromes. In the literature and our own data, CNVs in regions 7q33, 7q36.2, 8p23.3, 11q23.3, 15q11.2, 16p13.11, 16p12.3, Xp22.31 were more frequently observed and classified as likely pathogenic variants or aberrations with uncertain clinical significance.

Conclusion: This analysis made it possible to describe the structure of CNV in patients with NDD, allowed to identify frequent CNVs in the same regions, which is necessary to better understand the possible pathogenic effect of some variants. However, the high frequency of occurrence of a particular CNV cannot be the main factor for interpreting it as a pathogenic aberration. There is a need to experimentally validate the pathological effects of frequently occurring likely pathogenic variants and CNVs of uncertain clinical significance on model systems.

Funding: This study is supported by the Russian Science Foundation grant No. 21-65-00017, <https://rscf.ru/project/21-65-00017/>

Список литературы/References

1. Kearney H.M., Thorland E.C., Brown K.K. et al. Working group of the american college of medical genetics laboratory quality assurance committee. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med*. 2011;13(7):680-685. doi 10.1097/GIM.0b013e3182217a3a
2. Chehbani F., Tomaiuolo P., Picinelli C. et al. Yield of array-CGH analysis in Tunisian children with autism spectrum disorder. *Mol Genet Genomic Med*. 2022;10(8):e1939. doi 10.1002/mgg3.1939
3. Baccarin M., Picinelli C., Tomaiuolo P. et al. Appropriateness of array-CGH in the ADHD clinics: A comparative study. *Genes Brain Behav*. 2020;19(6):e12651. doi 10.1111/gbb.12651
4. Zarrei M., Burton C.L., Engchuan W. et al. A large data resource of genomic copy number variation across neurodevelopmental disorders. *NPJ Genom Med*. 2019;(4):26. doi 10.1038/s41525-019-0098-3
5. Lindstrand A., Eisfeldt J., Pettersson M. et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. *Genome Med*. 2019;11(1):68. doi 10.1186/s13073-019-0675-1
6. Lopes F., Torres F., Soares G. et al. Genomic imbalances defining novel intellectual disability associated loci. *Orphanet J Rare Dis*. 2019;14(1):164. doi 10.1186/s13023-019-1135-0
7. Firouzabadi S.G., Kariminejad R., Vameghi R. et al. Copy number variants in patients with autism and additional clinical features: Report of VIPR2 duplication and a novel microduplication syndrome. *Mol Neurobiol*. 2017;54(9):7019-7027. doi 10.1007/s12035-016-0202-y
8. Fry A.E., Rees E., Thompson R. et al. Pathogenic copy number variants and SCN1A mutations in patients with intellectual disability and childhood-onset epilepsy. *BMC Med Genet*. 2016;17(1):34. doi 10.1186/s12881-016-0294-2
9. Riggs E., Andersen E., Cherry A. et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med*. 2022;22(9):245-257. doi 10.1038/s41436-019-0686-8