



УДК 575.162

Многофакторное снижение размерности (MDR-анализ) в генетических исследованиях

Т. Л. Лебедь¹, Н. В. Жур², Н. В. Шепелевич²

¹ Республиканский научно-практический центр спорта, Минск, Республика Беларусь

² Полесский государственный университет, Пинск, Республика Беларусь

Аннотация

Исходя из предположения, что некоторые варианты генов в комбинации могут проявлять кумулятивный или эпистатический эффект, является актуальным изучение сочетания генотипов независимо действующих и взаимодействующих полиморфных генов. В статье описано применение непараметрического метода снижения мультифакторной размерности (MDR, Multifactor Dimensionality Reduction) на когорте высококвалифицированных спортсменов с целью определения наиболее предпочтительного сочетания генотипов генов нейромедиаторных систем, определяющих функционирование высшей нервной системы. На основе полученных данных о моделях межгенных взаимодействий проведен кластерный анализ, показывающий характер взаимодействия между полиморфными локусами генов нейротрансмиттерных систем. Метод MDR в наших исследованиях был также применён для выявления неблагоприятного сочетания генотипов в развитии таких патологий как нефропатия и нарушение ремоделирования костной ткани у юных спортсменов.

Ключевые слова: генетика, генетические исследования, анализ данных, биоинформатика, статистика, межгенные взаимодействия

Multifactorial dimensionality reduction (MDR analysis) in genetic research

T. L. Lebed¹, N. V. Zhur², N. V. Shapialevich²

¹ Republican Scientific and Practical Sports Center, Minsk, Republic of Belarus

² Polesky State University, Pinsk, Republic of Belarus

Abstract

Based on the assumption that some gene variants in combination may show cumulative or epistatic effects, it is relevant to explore the combination of genotypes of independently acting and interacting polymorphic genes. The article describes the application of nonparametric method of multifactor dimensionality reduction (MDR) on a cohort of highly qualified athletes in order to determine the most preferable combination of genotypes of genes of neurotransmitter systems that determine the functioning of the higher nervous system. Based on the obtained

data on the patterns of intergenic interactions, we performed cluster analysis showing the type of interaction between polymorphic loci of neurotransmitter system genes. In addition, we obtained models of the most favorable combination of genotypes for cyclic and acyclic sports. The MDR method in our researches was also applied to identify the unfavorable combination of genotypes in the development of such pathologies as nephropathy and disorder of bone tissue remodeling in young athletes.

Keywords: genetics, genetic research, data analysis, bioinformatics, statistics, intergenic interactions

Введение

Развитие и доступность современной приборной базы, появление компетентных специалистов позволило существенно расширить за последнее десятилетие распространённость генетических исследований в Республике Беларусь, которые тесно связаны с селекцией, биохимией, криминалистикой, медициной и спортом.

ДНК-тесты широко применяются для диагностики наследственных и инфекционных заболеваний, выявления предрасположенности к различным болезням, идентификации личности, верификации генетически модифицированных организмов и многих других задач. В растениеводстве, сельском и лесном хозяйстве – коммерческая паспортизация растений, анализ генетического родства и происхождения особей, сортов, форм, насаждений; исследование генетической структуры популяций и ее динамики; изучение уровня генетического разнообразия видов [1]. Изучение роли нуклеиновых кислот в хранении и передаче генетической информации раскрывает широкие возможности для создания различных ДНК-тестов, в частности для персонализированной медицины – онкогенетических тестов в том числе. Таким образом, современные подходы генетических исследований стали гораздо доступнее в различных отраслях биологической науки.

Информация, полученная в ходе таких исследований, специфична. При этом подходы и методология проведения исследований существенно разнятся в зависимости от цели исследования, имеющейся приборной базы и доступности реагентов. Учитывая выше сказанное, можно лишь констатировать, что отсутствует единый подход обработки и анализа генетической информации. Наиболее регламентированы медицинские исследования, что связано скорее всего с автоматической обработкой лабораторных исследований, для удобства врача и стандартизации процесса. В то время как в научных исследованиях существует множество биоинформатических приемов, позволяющих конвертировать данные из одного формата данных в другой, обрабатывать их с помощью различных статистических методов. Зачастую исследователь самостоятельно должен сделать выбор способов и методов окончательной обработки результатов и их интерпретации.

Цель работы – поиск межгенных взаимодействий спортсменов, обеспечивающих спортивную успешность, а также сочетания генотипов, способствующих развитию нефропатии почек и нарушению ремоделирования костной ткани.

Методы

Одной из целей спортивной генетики является выявление полиморфизмов, связанных с развитием основных спортивных качеств. Успех в достижении этой цели будет зависеть

от стратегии исследования, которая выявляет и учитывает важность взаимодействия множества генетических факторов и факторов окружающей среды в развитии успешной физической работоспособности. Аналогично можно отметить поиск генетических причин развития мультифакторных патологий. Одним из традиционных подходов к моделированию взаимосвязи между дискретными предикторами, такими как генотипы, и дискретными фенотипическими исходами является логистическая регрессия. Это параметрический статистический подход для соотнесения одной или нескольких независимых или объясняющих переменных (полиморфизмов) с зависимой переменной или переменной результата (статуса спортсмена, результативность, спортивная успешность, формирование патологии), которая следует биномиальному распределению. Наличие слишком большого количества независимых переменных по отношению к количеству наблюдаемых исходов является хорошо известной проблемой и является примером «проклятия» размерности. Один из часто используемых методов, который позволяет делать статистически обоснованные выводы о влиянии межгенных взаимодействий на фенотипы, это метод снижения мультифакторной размерности (MDR, Multifactor Dimensionality Reduction). MDR представляет собой непараметрический метод, не требующий строгой модели, и успешно применяемый для обнаружения эпистатических взаимодействий как в теоретических, так и в эмпирических исследованиях. Метод оперирует такими категориальными данными, как однонуклеотидные полиморфизмы (SNP), и сокращает размерность данных с множеством локусов, объединяя генотипы в группы высокого или низкого риска или ответа, что позволяет более эффективно выявлять генетические комбинации, связанные с клиническими исходами [2, 3, 4].

Выявление характера эффекта и моделей межгенных взаимодействий с помощью MDR-анализа проводилось в 3-х различных исследованиях: 1. Оценка частоты встречаемости сочетаний генотипов полиморфных локусов (L/S 5HTT, T102C 5HT2A, G2319A DAT1, G472A COMT и I/D ACE) среди спортсменов циклических и ациклических видов спорта (350 высококвалифицированных спортсменов, специализирующихся в циклических (биатлон, конькобежный спорт, академическая гребля, плавание, гребля на байдарке и каноэ) и ациклических (волейбол, тхэквондо, синхронное плавание, хоккей с шайбой, пулевая стрельба, самбо) видах спорта) [5].

2. Выявление сочетания диагностируемых маркеров ACE, PPARG2, VEGFA, APOB, NOS3, KCNJ11, TCF7L2, SLC30A8 и IGF2 генетически гетерогенного и многофакторного заболевания ХБП в белорусской популяции. Группа наблюдения была представлена пациентами с сахарным диабетом 2 типа (36 человек без патологии почек, 24 человека с отягощением – диабетическая нефропатия) [6].

3. Оценка генетических причин снижения минеральной плотности костной ткани (МПК) у юных спортсменов по генетическим маркерам – Arg577Ter ACTN3 (rs1815739), A/G EPAS1 (rs1867785), +294T/C PPARA (rs4253778), Gly482Ser PPARGC1A (rs2016520), C/T VDR (rs731236), G/T COL1A (rs1800012). Общее число обследованных профессиональных спортсменов составило 60 человек, 30 из которых представляли Брестское областное училище олимпийского резерва (БрУОР) и 30 спортсменов – специализированного по спорту класса лица ПолесГУ. Также спортсмены были распределены на 2 подгруппы: виды спорта с преимущественным тренировочным и соревновательным процессом в спортивном зале (22 человека) и виды спорта с преимущественным тренировочным и соревновательным процессом на открытом воздухе (38 человек) [7].

Результаты и обсуждение

Анализ межгенных взаимодействий выявил специфическую комбинацию генотипов пяти генов нейромедиаторных систем, наиболее предпочтительных для спортсменов высокой квалификации циклических и ациклических видов спорта.

Продемонстрирован наибольший предсказательный потенциал для полиморфизма T102C 5HT2A (2,77 %). Для всех моделей наилучшим прогностическим признаком предпочтительно гетерозиготное носительство этого гена (генотип СТ). Установление взаимосвязей между генами и фенотипическими проявлениями заболевания ХБП является важнейшим условием для выделения ключевых звеньев патогенеза, затрагивающих определённые физиологические функции, посредством которых и формируется клиническая картина заболевания. Построенная MDR модель показывает наличие сильно выраженного синергического эффекта в отношении одновременного вклада полиморфизмов следующих генов: rs5219 KCNJ11 и rs7903146 TCF7L2 – общий вклад составил 16,36%.

В результате проведенного анализа была выявлена единственная статистически достоверная повторяющаяся модель комбинаций генотипа, влияющая на развитие нарушений МПК, что продемонстрировало гипотетическое распределение случаев встречаемости комбинации генов. Важным выводом является следующее: носительство ТТ гена PPAR α и СС гена ACTN3 в совокупности с нарушениями МПК может приводить к повышению частоты возникновения стресс-травм даже в рамках тренировочного процесса. Также был смоделирован комплекс генов, находящихся в антагонистических взаимодействиях, в рамках рассчитанной статистически достоверной модели VDR, ACTN3 и PPARGC1A.

Выводы

Исходя из выше приведенных результатов наблюдается очевидная тенденция к способности обнаруживать межгенные взаимодействия и возможности смоделировать данные.

Установление взаимосвязей между генами и фенотипическими проявлениями заболевания является важнейшим условием для выделения ключевых звеньев патогенеза, затрагивающих определённые физиологические функции, посредством которых и формируется клиническая картина заболевания. Таким образом полученные с помощью MDR результаты по комплексу генов и генотипов позволяют дополнить и расширить известные представления по оценке индивидуальной предрасположенности к развитию потенциальных нарушений в организме человека и выделить группы риска развития заболевания, определить спортивную успешность.

Однако не стоит обнадеживаться тем, что результаты, полученные на моделях с большим количеством генов, позволят проводить прогнозирование. Незначительные изменения в классификации исходов (фенотипов) приводят к существенным изменениям статистической модели прогнозирования, создаваемой методом MDR. Поэтому важно включать в исследование лишь те генетические маркеры, непосредственно с которыми связаны изучаемые молекулярные изменения [8].

Список литературы

1. Ковалевич А., Падутов В., Баранов О. Полногеномное секвенирование-новый этап генетических исследований // Наука и инновации. – 2015. – №. 5(147). EDN: <https://elibrary.ru/tzfqqx>

2. *Motsinger A.* Multifactor dimensionality reduction: an analysis strategy for modelling and detecting gene-gene interactions in human genetics and pharmacogenomics studies. *Human Genomics*. 2007. Vol.2. No.5. P. 318-328. DOI: <https://doi.org/10.1186/1479-7364-2-5-318>
 3. *Пономаренко И.В.* Использование метода Multifactor Dimensionality Reduction (MDR) и его модификаций для анализа ген-генных и генно-средовых взаимодействий при генетико-эпидемиологических исследованиях (обзор) // *Научные результаты биомедицинских исследований*. 2019. Т.5, № 1. С. 4-21. EDN: <https://elibrary.ru/umbgkq>
 4. *Multifactor Dimensionality Reduction* [Электронный ресурс] 24.12.2014. – Режим доступа: <https://sourceforge.net/projects/mdr/> – Дата доступа: 05.11.2024.
 5. *Шепелевич Н.В., Маринич В.В., Мельнов С.Б.* Генетические маркеры нейромедиаторов и психофизиологические особенности квалифицированных спортсменов различных видов спорта // *Прикладная спортивная наука*. – 2024. – Т. 1. – №. 19. – С. 47-54. URL: <https://rep.polessu.by/handle/123456789/31678>
 6. *Жур Н. В., Лебедь Т. Л., Кручинский Н. Г.* Применение анализа генетических исследований для диагностики и профилактики нефрологических осложнений у пациентов с сахарным диабетом второго типа // *Здоровье для всех: научно-практический журнал*. – 2023. – №. 1. – С. 28-36. URL: <https://rep.polessu.by/handle/123456789/29037>
 7. *Жур Н. В., Евдюков С.В., Лебедь Т.Л., Маринич Т.В., Кручинский, Н.Г.* Молекулярно-генетические методы в диагностике и прогнозировании индивидуальной предрасположенности к нарушению минеральной плотности костной ткани // *Вестник Полесского государственного университета. Серия природоведческих наук*. – 2024. – №. 1. – С. 66-72. URL: <https://rep.polessu.by/handle/123456789/31760>
 8. *Иванюкович В.А., Николаенко Е.А., Мельнов С.Б., Жур Н.В., Лебедь Т.Л.* Влияние классификации исходов на результат моделирования и прогнозирования методом MDR // *Сахаровские чтения 2023 года: экологические проблемы XXI века : материалы 23-й Международной научной конференции, Минск, Республика Беларусь, 18-19 мая 2023 г / Международного государственного экологического ин-т им. А.Д. Сахарова Белорусского государственного ун-та - Минск : МГЭИ им. А.Д. Сахарова БГУ, 2023. – Ч. 2. – С. 301-305. URL: <https://rep.polessu.by/handle/123456789/28924>*
-

Сведения об авторах

Лебедь Татьяна Леонидовна – лаборатория биохимии ГУ «Республиканский научно-практический центр спорта», Минск, Республика Беларусь.

Е-mail: hlebus@mail.ru

Жур Владимир Николаевич – кафедра биохимии и биоинформатики биотехнологического факультета УО «Полесский государственный университет», Пинск, Республика Беларусь.

Е-mail: unizhur@gmail.com

Шепелевич Наталья Васильевна – кафедра биохимии и биоинформатики биотехнологического факультета УО «Полесский государственный университет», Пинск, Республика Беларусь.

Е-mail: shnv1977@mail.ru

Для цитирования:

Лебедь Т.Л., Жур Н.В., Шепелевич Н.В. Многофакторное снижение размерности (MDR-анализ) в генетических исследованиях // Российский журнал информационных технологий в спорте. – 2024. – Т. 1, № 3. – С. 16–21. DOI: <https://doi.org/10.62105/2949-6349-2024-1-3-16-21> EDN: CGKDYH

Cite as:

Lebed T.L., Zhur N.V., Shapialevich N.V. Multifactorial dimensionality reduction (MDR analysis) in genetic research. *Russian Journal of Information Technology in Sports*, 2024, 1 (3), pp. 16–21 (in Russian). DOI: <https://doi.org/10.62105/2949-6349-2024-1-3-16-21> EDN: CGKDYH

Статья поступила в редакцию: 20.11.2023

Статья принята в печать: 08.12.2024

Статья опубликована: 16.12.2024