

зованием алгоритма VEGAS2, позволило оценить значимость 17 322 генов. Достигнутое минимальное значение $p=7 \cdot 10^{-6}$ характерно для rs7762549 в хромосомной позиции 6p21.1, рядом с которым находится три гена: (*ENPP4* (ectonucleotide pyrophosphatase/phosphodiesterase 4, NCBI Gene ID – 22875), *ENPP5* (ectonucleotide pyrophosphatase/phosphodiesterase 5, NCBI Gene ID – 59084) и *CLIC5* (chloride intracellular channel 5, NCBI Gene ID – 53405)). Так же авторами было отмечено, что ранее ни один из этих генов не отмечался как фактор повышенной вероятности развития РАС.

В качестве маркеров, которые могут быть связаны с РАС, можно выделить следующие: rs16922945 (*STYK1*), rs10888329 (*OR2M4*), rs6672981 (*OR2M4*), rs4397683 (*OR2M4*), rs2447097 (*SGSM2*), rs10966205, rs7933990, rs7762549 (*ENPP4*). С учетом того факта, что GWAS, как правило, проводятся с использованием микрочиповых технологий, а выявленные генетические ассоциации на уровне генома ($p \leq 1,00E-06$) должны в дальнейших популяционных исследованиях проверяться дополнительно, выявленные закономерности не должны приниматься как категорически доказанные. Результаты GWAS, в первую очередь, направлены на то, чтобы дать исследователю область поиска в геноме для выявления локусов (генов), являющихся потенциальными мишенями при лечении пациентов.

ЛИТЕРАТУРА

1. Lai, M. C., Lombardo M. V., Baron-Cohen S. Autism // Lancet. – 2014.
2. Veenstra-Vanderweele, J. Autism as a paradigmatic complex genetic disorder. / J. Veenstra-Vanderweele [et al.] // Annu Rev Genomics Hum Genet. – 2004. – Vol. 5. – P. 379–405.
3. Po-Hsiu, K. Genome-Wide Association Study for Autism Spectrum Disorder in Taiwanese Han Population / K. Po-Hsiu [et al.] // PLoS One. – 2015. – Vol. 10.
4. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia / The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium // Mol. Autism. – 2017. – Vol. 8(21).

МОДЕЛИРОВАНИЕ ПАНЕЛИ SNP-МАРКЕРОВ С ВЫСОКИМ ДИФФЕРЕНЦИРУЮЩИМ ПОТЕНЦИАЛОМ ДЛЯ ОПРЕДЕЛЕНИЯ ЧИСТОПородНОСТИ БИРМАНСКОЙ КОШКИ

PANEL OF SNP MARKERS WITH HIGH DIFFERENTIATION POTENTIAL TO DETERMINE THE PUREBRED OF BIRMAN CATS

Е. В. Снытков

E. Snytkov

*Белорусский государственный университет, МГЭИ им. А. Д. Сахарова БГУ,
г. Минск, Республика Беларусь
evsnytkov@gmail.com*

Belarusian State University, ISEI BSU, Minsk, Republic of Belarus

С использованием биоинформатического метода многомерного сокращения размерности (MDR-анализ) произведен расчет вероятности корректной дифференциации животных бирманской породы кошек от 10 других пород. Полиморфизм rs44083224 (Chr.B1:16760839) обладает высоким дифференцирующим потенциалом.

Using the bioinformatics method Of Multidimensional Reduction of Dimension (MDR-analysis), we calculated the probability of correct differentiation of animals of Burmese cat breed from 10 other breeds. Polymorphism rs44083224 (Chr.B1: 16760839) has a high differentiating potential.

Ключевые слова: бирманская кошка, однонуклеотидный полиморфизм, порода, чистопородность, MDR-анализ, *Felis silvestris catus*.

Keywords: Birman cat, single nucleotide polymorphism, breed, purebred, MDR-analysis, *Felis silvestris catus*.

Бирманская кошка – порода полудлинношерстных кошек колор-пойнтового окраса. Произошла от скрещивания сиамских и персидских кошек. Оценка генетического разнообразия по данным STR-маркеров для данной породы была показана в исследовании Lipinski M.J. et al. 2008.

Ранее нами была показана возможность с использованием данных полногеномных секвенсных проектов коммерческих пород свиней определить наличие породоспецифичных SNP-маркеров (или SNP-маркеров с высоким дифференцирующим потенциалом) для животных ряда пород свиней и дикого кабана [1-3]. Использование MDR-анализа (Multifactor Dimensionality Reduction) для отбора из совокупности потенциально информативных

генетических маркеров именно тех, вклад которых наиболее существенен при решении задачи дифференциации нескольких групп, показан в наших ранних исследованиях [4,5].

Цель исследования – смоделировать с использованием MDR-анализа панель генетических маркеров, пригодную для дифференциации животных вида *Felis silvestris catus* бирманской породы (Birman cat, BIR) от представителей пород Asian domestic cat (ASI), Oriental Shorthair (ORI), Maine Coon Cross (MCC), Siamese (SIA), Burmese (BUR), Napoleon (NAP), Selkirk Rex (SR), British Shorthair (BRI), Devon Rex (DEV), Bengal (BEN), а также охарактеризовать ее с позиций чувствительности, специфичности и общей точности.

Определение генотипа по SNP-маркерам был выполнен с использованием алгоритма SRA Nucleotide BLAST (Sequence Read Archive Nucleotide BLAST) и программы Unipro UGENE v.1.29. Количество включенных в анализ SNP – 49 [Brooks A. 2016].

Были использованы SRA-данные по полногеномному секвенированию (NGS), размещенные в открытом доступе на облачном сервисе DNAnexus (<http://sra.dnexus.com/>), а также в SRA-NCBI – high-throughput DNA and RNA sequence read archive (www.ncbi.nlm.nih.gov/sra). Число полногеномных прочтений для животных вида *Felis silvestris catus* – 78 (Asian domestic cat – 13; Bengal – 3; Birman – 4; British Shorthair – 4; Burmese – 6; Devon Rex – 4; Maine Coon Cross – 10; Napoleon – 6; Oriental Shorthair – 13; Selkirk Rex – 5; Siamese – 10). Общее количество проанализированных сиквенсов – 25 585 652 390.

Построение модели взаимодействий SNP (определение минимального и достаточного количества генетических маркеров для решения поставленной задачи) проводилось с использованием биоинформатического метода многомерного сокращения размерности (MDR v.3.0.2).

В результате ранее проведенного исследования (ROC-анализ) нами было выявлено 9 SNP-маркеров с высоким дифференцирующим потенциалом для бирманской кошки: rs43863896 (Chr.D1:18963036), rs43864283 (Chr.D1:17016430), rs43867932 (Chr.D1:12324748), rs43774667 (Chr.A1:67542780), rs43879120 (Chr.D2:6453747), rs44083224 (Chr.B1:16760839), rs43979058 (Chr.B4:28216789), rs43982208 (Chr.B4:91236832), rs44078625 (Chr.B1:69430915).

В процессе моделирования нами были использованы высоко консервативные настройки поиска конфигурации модели, которые позволили однозначно дифференцировать наличие/отсутствие статистически значимых эффектов: количество атрибутов (attribute count range) – от 1 до n (где n – количество переменных в модели); воспроизводимость модели (cross-validation count) – 10; анализ топ-моделей (track top models) – 1000; поиск конфигурации модели (search method configuration) – exhaustive; классификация ячеек (ambiguous cell assignment) – unclassified.

В результате проведенного моделирования была определена модель, отражающая такое сочетание SNP-маркеров для бирманской кошки, которое позволило наилучшим образом отличить животных этой породы от других пород в рамках данной работы. В частности, модель включала в себя один SNP (rs44083224 (Chr. B1:16760839)) и имела следующие характеристики: adj. Balanced accuracy – 0,8571, Sensitivity – 1,0, Specificity – 0,7143, Cross Validation Consistency – 10/10.

Таким образом, нами предложена и охарактеризована модель, включающая один SNP-маркер, с помощью которой имеется возможность с высокой вероятностью корректно отличить чистопородных домашних кошек бирманской породы (BIR) от особей других 10 пород кошек (Oriental Shorthair (ORI), Maine Coon Cross (MCC), Siamese (SIA), Burmese (BUR), Napoleon (NAP), Selkirk Rex (SR), Asian domestic cat (ASI), British Shorthair (BRI), Devon Rex (DEV), Bengal (BEN)). Однако для увеличения специфичности предложенной модели необходим дальнейший поиск SNP, обладающих высоким потенциалом в дифференциации данной породы кошек.

Таким образом, полученные нами результаты могут лечь в основу создания панели SNP-маркеров для определения чистопородности особей бирманской породы вида *Felis silvestris catus*.

ЛИТЕРАТУРА

1. Кипень, В. Н. Использование полногеномных данных проектов NGS для поиска решения криминалистической задачи по дифференциации диких кабанов и домашних свиней на основе анализа SNP / В. Н. Кипень // Молекулярная диагностика-2017: сб. тр. IX Всероссийской науч.-практ. конф. с междунар. участием. – М., 2017. – С. 305–306.
2. Кипень, В. Н. Моделирование панели породоспецифичных SNP-маркеров для определения чистопородности домашних свиней крупной белой породы / В. Н. Кипень // Молекулярная диагностика-2017: сб. тр. IX Всероссийской науч.-практ. конф. с междунар. участием. – М., 2017. – С. 396–397.
3. Кипень, В. Н. Определение «новых» SNP, обладающих дифференцирующей способностью для различения особей *Sus scrofa domesticus* и *Sus scrofa scrofa* / В. Н. Кипень, Е. В. Снытков // Молекулярная диагностика-2017: сб. тр. IX Всероссийской науч.-практ. конф. с междунар. участием. – М., 2017. – С. 420–421.
4. Кипень, В. Н. Роль полиморфных вариантов в генах семейства P450 в модификации риска развития рака молочной железы – исследование по типу случай-контроль с использованием MDR-анализа / В. Н. Кипень, Е. В. Снытков, Н. С. Смольник, С. Б. Мельнов // Сахаровские чтения 2017 года: экологические проблемы XXI века: материалы 17-й междунар. науч. конф. В 2-х ч. / под общ. ред. С. А. Маскевича, С. С. Позняка. – Минск, 2017. – С. 165–166.
5. Кипень, В. Н. Сочетанное наличие патогенетически значимых полиморфных вариантов в генах, ответственных за контроль фолатного цикла и метилирование ДНК *de novo*, в увеличении риска развития рака молочной железы / В. Н. Кипень, Е. В. Снытков, С. Б. Мельнов // Сахаровские чтения 2017 года: экологические проблемы XXI века: материалы 17-й междунар. науч. конф. В 2-х ч. / под общ. ред. С. А. Маскевича, С. С. Позняка. – Минск, 2017. С. 167–168.