

ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ

УДК 575:636

ГЕНОМНАЯ СЕЛЕКЦИЯ. I. ПОСЛЕДНИЕ ТЕНДЕНЦИИ
И ВОЗМОЖНЫЕ ПУТИ РАЗВИТИЯ

© 2020 г. Ю. А. Столповский^{1, *}, А. К. Пискунов^{1, **}, Г. Р. Свищева^{1, 2, ***}

¹Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

²Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения
Российской академии наук, Новосибирск, 630090 Россия

*e-mail: stolpovsky@mail.ru

**e-mail: aleksei.piskunov@gmail.com

***e-mail: gulsvi@mail.ru

Поступила в редакцию 19.11.2019 г.

После доработки 03.02.2020 г.

Принята к публикации 06.02.2020 г.

Существует прогноз, что мировой спрос на пищевые продукты животного и растительного происхождения увеличится на 74% уже к 2050 г. (Продовольственная и Сельскохозяйственная Организация ООН). Удовлетворение этого спроса без разрушительного воздействия на окружающую среду возможно лишь при условии сохранения принципов органического сельского хозяйства, а также внедрения новых технологий в животноводстве и растениеводстве. Ключевую роль в этом процессе может сыграть геномная селекция как один из наиболее перспективных и безопасных методов улучшения генетических качеств сельскохозяйственных животных и растений. В данном обзоре обобщены сведения о геномной селекции, обозначены возможные точки роста данного направления, показано как строится геномная оценка племенной ценности, каковы ключевые условия необходимые для ее реализации, обсуждаются преимущества и ограничения геномной и маркерной селекции.

Ключевые слова: геномная селекция, однонуклеотидный полиморфизм, племенная ценность, доместичированные виды животных, референсная популяция.

DOI: 10.31857/S0016675820090143

В течение длительного времени ученые стремились найти значимые взаимосвязи между хозяйственно полезными признаками и генетическими маркерами, чтобы целенаправленно вести селекцию для выявления и закрепления в популяциях ценных аллелей. Применение генетических маркеров для решения селекционных задач получило название маркерная или маркер-направленная селекция. Впервые идея использования маркеров в селекции была теоретически сформулирована А.С. Серебровским и его коллегами еще в 20-х годах XX в. [1]. В широком понимании генетическим маркером является любая наследуемая модификация структурных генов (аллелей), анонимных нуклеотидных последовательностей или хромосом, с которыми сцеплена группа аллелей интереса. Однако теоретические результаты, полученные в ходе огромного количества научных работ по изучению генетических маркеров — от групп крови, белков и ферментов до различных типов ДНК-маркеров — так и не нашли масштабного применения при разведении одомашненных животных, а предлагаемые гене-

тические методы не стали востребованным инструментом на практике. Во многом значимость полученных результатов нивелировалась незначительным эффектом от их проявления. Наибольшие достижения были получены при изучении такого типа маркеров как генные мутации, которые могут приводить к тому или иному заболеванию, влиять на продуктивность, а также отражать происхождение отдельных животных и филогенез пород в целом. Подробная эволюция генетических маркеров и методов представлена в обзорах [2, 3].

В течение последнего десятилетия в генетических исследованиях доместичированных видов животных произошла если не революция, то кардинальное или фундаментальное изменение подхода, связанного с оценкой генетического потенциала животных, их продуктивности, жизнеспособности и здоровья. Почти во всех ведущих мировых селекционных и научных центрах — в первую очередь, США, Германии, Франции, Австралии, Китая, Норвегии и Нидерландов — сформированы основные подходы к новейшей технологии селек-

ции – геномной селекции, которые активно применяются в животноводстве. Геномная селекция базируется на использовании информации о полных геномах сельскохозяйственных животных и растений и методах функциональной геномики и биоинформатики для выявления взаимосвязи между вариантами геномных локусов и степенью проявления хозяйственно значимых признаков. С развитием подходов геномной селекции появилась новая возможность оценивать генетический потенциал животных и растений, соответственно вести селекцию и создавать качественно иные линии, породы животных и сорта растений. В дальнейшем геномная селекция вместе с развитием геномных биобанков и статистико-математического аппарата, служащего для обработки генетических данных, может значительно увеличивать селекционную точность и надежность оценок племенной ценности животных.

В данной работе проанализированы основные направления и проблемы, связанные с развитием геномной селекции в животноводстве – стратегической отрасли сельского хозяйства для большинства государств.

ГЕНЕТИЧЕСКАЯ АРХИТЕКТУРА СЛОЖНОГО ПРИЗНАКА И ПЕРЕХОД К ГЕНОМНЫМ ДАННЫМ

Генетическая изменчивость, свойственная большинству фенотипических признаков животных, может быть как моногенной, так и полигенной. Признаки с менделевской моделью наследования, т.е. признаки, фенотипическое проявление которых полностью определяется малым числом или одним геном, являются скорее исключением, чем правилом. У животных список менделевских признаков включает такие характеристики как масть, сверхплодовитость, удвоение мускулатуры и т.п. В целом, за исключением ряда характерных для породы признаков, наследственных дефектов и летальных мутаций, практически все экономически важные признаки в животноводстве являются сложными полигенными признаками. Например молочная, мясная и яичная продуктивности, белковый и жировой состав молока, устойчивость к большинству заболеваний. Фенотипическая выраженность таких признаков находится под влиянием условий развития и содержания животного, а также обусловлена малыми генетическими эффектами большого числа локусов, которые могут быть рассеяны по всему геному. У крупного рогатого скота (КРС), например, наследуемость полигенных признаков, определяемая как доля вариации, объясняемая генетическими эффектами, колеблется преимущественно в интервале от 5 до 60% [4]. С помощью методов картирования (анализа сцепления и анализа ассоциаций) для многих признаков удалось выявить

большое количество локусов с относительно большими эффектами. Однако их суммарный эффект был значительно ниже общей наследуемости признаков [5].

Становление геномики как дисциплины в 80-х годах XX в. привело к развитию маркерной селекции, основанной на выявлении и направленном отборе генов (аллелей), оказывающих влияние на важные в сельском хозяйстве количественные признаки. Благодаря развитию и применению технологий молекулярного маркирования и статистических методов анализа данных был осуществлен масштабный поиск локусов количественных признаков (quantitative trait loci, QTL) – областей ДНК, содержащих гены, либо сцепленные с ними участки, оказывающие значимый эффект на выраженность количественного признака. Такие локусы были картированы у всех сельскохозяйственных животных для большого числа количественных признаков, характеризующих продуктивность, плодовитость и развитие разных генетических заболеваний [6–8]. В частности у КРС были картированы QTL, оказывающие существенное влияние на развитие синдрома дефицита лейкоцитарной адгезии, приводящего к гибели молодняка в результате иммунного дефицита, лейкоза, скрепи, губкоподобной энцефалопатии и многих других [9].

Наибольшее количество молекулярно-генетических маркеров, ассоциированных с хозяйственно полезными признаками, идентифицировано у крупного рогатого скота. С помощью генотипирования однонуклеотидных полиморфизмов (single nucleotide polymorphism, SNP) было исследовано огромное число генов у различных пород КРС. Например было показано, что гены каппа-казеина (*CSN3*), пролактин (*PRL*), соматотропин (*GH*), бета-лактоглобулин (*BLG*), диацетил-глицерин О-ацилтрансфераза (*DGATI*), рилизинг-фактора (гипоталамического фактора транскрипции *PIT-1*) и др. [9] ассоциируют с показателями молочной продуктивности. С помощью генотипирования SNP были изучены гены *LEP*, *RORC*, *SCD*, *GH* и *TG* на наличие взаимосвязи с показателями мясной продуктивности и репродуктивными показателями [10–12]. У КРС с помощью SNP маркеров проводились исследования распространенного заболевания – комплексного порока позвоночника (CVM), проявление которого связано с миссенс-мутацией в гене *SLC35A3*, и исследования BLAD-синдрома, проявляющегося при наличии точечной мутации в кодирующей части аутосомного гена *CD18* [13, 14].

У различных пород свиней с помощью генотипирования SNP был изучен полиморфизм большого числа генов, в частности генов гамма-субъединицы протеинкиназы А (*PRKAG3*), гипофизарного транскрипционного фактора (*POU1F1*), рецептора меланокортина 4 (*MC4R*) и инсулино-

подобного фактора роста 2 (*IGF2*) для выявления эффектов на показатели мясной продуктивности; также был изучен полиморфизм генов *RYR1* и *KPL2* на взаимосвязь с некоторыми наследственными заболеваниями [15–17]. Генотипированием по гену *ESR* стало возможным оценивать репродуктивные качества сельскохозяйственных животных, весьма значимые при их разведении [18].

Несмотря на достигнутые успехи в области картирования генетических маркеров, ассоциированных со многими сложными болезнями, показателями качества продукции и продуктивности, селекционеры не в полной мере используют результаты научных исследований на практике, а селекция до сих пор во многом строится на принципах, заложенных в XVIII–XX вв., которые впрочем остаются актуальными и сегодня. Главный селекционный принцип – производить отбор “лучших” животных, выбраковывать “худших” и скрещивать “лучших с лучшими”, основываясь на показателях собственной продуктивности, а также на показателях продуктивности их потомков и предков. Примером наиболее известного метода в молочном скотоводстве может служить оценка быка-производителя по продуктивности собственных дочерей. Эффективность метода была доказана, однако, чтобы выбрать быка-улучшателя, необходимо затратить много времени и средств.

Маркерная селекция так и не получила широкого распространения в животноводстве. По мнению ряда исследователей она не может служить достаточно эффективным инструментом для оценки племенной ценности в селекционных программах, поскольку суммарный эффект всех QTL, обнаруженных с помощью маркерной селекции, значительно меньше общей наследуемости изучаемого признака [19]. Вопрос о пользе внедрения маркерной селекции – дискуссионный, так как с нашей точки зрения к настоящему времени исследовано небольшое количество генов, кроме того до сих пор не изучены механизмы их взаимодействия. Напомним, что на сегодняшний день известно более 200 генов у КРС и свиней, связанных с хозяйственно полезными признаками и болезнями (<https://omia.org/home/>).

Ситуация отчасти изменилась с появлением и развитием геномной селекции, которая, несмотря на частое противопоставление маркерной селекции, по сути является продуктом ее развития, инструментом, конечная цель которого – найти функционально значимые мутации в геноме. Геномная оценка племенной ценности или геномная селекция – это результат одновременного тестирования генома по всему доступному объему данных, как правило генотипических данных SNP-маркеров, равномерно покрывающих весь геном. Геномная селекция строится на том же предположении, что и стандартная система оцен-

ки племенной ценности: генетическая вариация количественного признака является результатом аддитивного влияния большого количества (сотни миллионов) маркеров – участков ДНК, равномерно распределенных по всему геному, каждый из которых имеет крайне малый эффект. Таким образом, точное функциональное картирование маркеров, в отличие от QTL в маркерной селекции, не является обязательным условием для осуществления геномной селекции [20]. Такой подход стал возможен прежде всего благодаря коммерческой доступности SNP-чипов высокого разрешения, а также благодаря развитию эффективных статистических методов.

Стоит напомнить, что в 1918 г. Р. Фишером была предложена модель для описания изменчивости количественных признаков, называемая моделью бесконечно малых эффектов (Fisher's infinitesimal model). Эта модель предполагает, что выраженность всех признаков определяется бесконечным числом несвязанных друг с другом локусов (мутаций), каждый из которых имеет крайне малый эффект [21]. Несмотря на то, что данная модель далека от реально существующих биологических механизмов (число генов конечно и они взаимодействуют друг с другом), она оказалась чрезвычайно полезной в селекции животных. Модель Фишера служила основой для разработки селекционных программ до тех пор, пока не появились геномные технологии, позволяющие определить конечное число ДНК-маркеров, ассоциированных с признаком.

Согласно модели Фишера величина генетического прироста (ΔG , genetic gain) определяется тем, насколько отбираемые в качестве родителей особи лучше, чем популяция в целом. На этот параметр влияют три основных фактора. Во-первых, это генетическая вариация (σ), называемая также генетическим стандартным отклонением: если в популяции существует большая вариация фенотипических признаков, обусловленная генетически, то легче найти особей с лучшими показателями, чем у остальных. Во-вторых, это интенсивность селекции (i), которая определяется используемой для разведения долей популяции. Если в качестве родителей будут выбраны все особи, то интенсивность селекции и соответственно генетический прирост будут стремиться к нулю. Таким образом, для увеличения генетического прогресса выгодно поддерживать как можно меньшую пропорцию отбираемых особей. Однако при этом также уменьшается абсолютное число животных-родителей и, соответственно, увеличивается риск инбридинга. С этой позиции наиболее эффективной стратегией для оптимизации генетического прогресса без негативных последствий инбридинга является увеличение численности так называемой “ядерной” популяции. В-третьих, генетический прогресс прямо пропорционален параметру (r) – точности

отбора, который зависит от наследуемости признака и применяемых методов его оценки. Кроме того, генетический прогресс (в расчете на единицу времени) тем выше, чем короче цикл отбора — интервал между поколениями (L). Таким образом, генетический прирост в год может быть оценен как:

$$G = \frac{ir}{L}. \quad (1)$$

Из формулы (1) следует, что чем выше будет генетическое превосходство отбираемых особей, чем точнее оценка их племенной ценности, чем меньше интервал между поколениями, тем большего генетического прироста можно достигнуть. Примером эффективности модели Фишера может служить показатель сокращения времени для набора веса рыночной реализации у куриц с 16 до 5 недель в США за период в 30 лет [22]. Однако несмотря на все успехи модель Фишера плохо отражает реально существующие биологические механизмы, поскольку геном всех животных состоит пусть из очень большого, но все же конечного числа элементов (например около 10^9 пар оснований для человека и КРС). Очевидно также, что одни элементы генома влияют на выраженность фенотипических черт в большей степени, чем другие.

Методы геномной селекции позволяют изучать сотни тысяч генетических маркеров, распределенных по всему геному, получать результаты и соответственно принимать решение о дальнейшей судьбе животных по меркам зоотехнической науки практически мгновенно. В связи с этим интерес к развитию методов геномной селекции огромен.

Как правило селекционная оценка сельскохозяйственных животных состоит из двух основных составляющих: контроля продуктивности и оценки по качеству потомства методом BLUP (Best Linear Unbiased Prediction). В случае геномной селекции оценка сельскохозяйственных животных происходит на основе геномной информации о животном с использованием данных референсной популяции, что позволяет более точно оценить влияние генетических маркеров на тот или иной признак и произвести отбор наиболее лучших особей для селекции.

СОВРЕМЕННАЯ ИСТОРИЯ ГЕНОМНОЙ СЕЛЕКЦИИ

В 2004 г. в США стартовал проект по геномной селекции КРС. С помощью генетического анализатора Illumina был осуществлен ресиквенс геномов 392 животных из 14 пород КРС. В результате было выявлено 444792 SNP маркеров, из которых были отобраны 54000 (~12%) маркеров с высокой степенью детектирования. С 2007 г. началось

практическое использование SNP-чиповой технологии после того, как удалось сконструировать чип, под названием SNP50 BeadChip [23]. Так с использованием вышеуказанного микрочипа был проведен анализ 14 пород лошадей и 18 эволюционно родственных им видов; было выявлено более 54000 полиморфных SNP маркеров [24]. За последние десять лет опубликован обширный материал по геномной селекции [25–27]. Мировой тренд — замена существующих SNP-чипов на чипы, включающие каузальные (причинные) SNP, что, как ожидается, приведет к повышению точности и упрощению методов геномной селекции [28].

Согласно принятому определению [29], SNP — это однонуклеотидные позиции в геномной ДНК, для которых в популяции имеются различные варианты последовательностей (аллели). SNP являются вариантами по одному нуклеотиду, которые не меняют общую длину ДНК последовательности в этом регионе. SNP находятся в избытке по всему геному. Например, в геноме человека они представлены с частотой один SNP на каждые 1000 пар оснований [30]. Большая часть SNP находится в некодирующих областях генома и не имеет прямого влияния на фенотип. Однако некоторые SNP, которые располагаются в экспрессируемых областях, могут влиять на экспрессию генов и индуцировать изменения в структуре белка.

Внедряя в практику животноводства ДНК-маркеры данного типа, можно проводить точную идентификацию генотипов животных, несущих желательные фенотипические особенности, и на их основе вести селекцию. Таким образом, есть возможность более рационально использовать генетический потенциал сельскохозяйственных животных.

ПРИНЦИПЫ ГЕНОМНОЙ СЕЛЕКЦИИ

Одним из факторов, определивших развитие геномной селекции, стала так называемая “геномная революция”. Вслед за расшифровкой последовательностей генома человека были расшифрованы геномы основных видов сельскохозяйственных животных (КРС, свиньи, овцы, козы и курицы) [31]. Качество расшифровки на протяжении довольно длительного времени оставалось относительно невысоким, поскольку используемые методы “считывали” короткие фрагменты ДНК, что приводило к ряду ошибок при генерации итоговой последовательности. В последние годы стали широко доступны новые методы, в частности оптическое картирование, которые позволили в 2017–2018 гг. получить и проанализировать высококачественные последовательности геномов всех видов сельскохозяйственных животных и даже таких сложных объектов как пшеница, генетический материал которой

превышает по объему геном человека более чем в пять раз, и обогащен большим количеством повторов [32]. Все полученные к настоящему времени последовательности могут быть найдены в Ensembl Database (<https://www.ensembl.org/>).

В результате секвенирования геномов животных стало доступным огромное количество информации об имеющейся в них генетической вариации. Такая информация была определена путем сопоставления последовательностей высокого разрешения, полученных от нескольких десятков или сотен особей, представляющих разные породы или популяции. Наиболее распространенной формой генетической вариации является SNP. Разумеется существуют и другие формы, например структурная вариация, которые также играют важную роль в генетической изменчивости. Детальный каталог SNP представлен в крупномасштабном исследовании “1000 Genomes Project” (2015 г.), где были проанализированы последовательности всех трех миллиардов нуклеотидов у каждого из 2405 человек из 26 различных популяций по всему миру. При сопоставлении полученных последовательностей были обнаружены 88 миллионов “точек расхождения”, то есть SNP. У КРС количество SNP приблизительно соответствует таковому у человека [31]. Тем не менее большой объем выборок индивидов, который необходим для геномной селекции в поисках редких генетических вариаций, делает полногеномное секвенирование довольно дорогостоящим, в особенности для видов с “большим” геномом. Более рациональным подходом с экономической и временной точек зрения является использование SNP-чипов, создание и распространение которых также можно рассматривать как одно из достижений “геномной революции”. По мере получения информации об имеющихся SNP, они становились более доступными в открытых базах данных (например, в NCBI dbSNP), что стимулировало коммерческие компании к разработке разнообразных SNP-чипов, которые позволяют одновременно проанализировать от десятков до сотен тысяч полиморфных вариантов (аллелей). Как только SNP-чипы стали коммерчески доступны, геномная селекция быстро получила широкое распространение сначала в молочном скотоводстве, а затем и в других областях сельского хозяйства, как эффективная, быстрая и легко реализуемая альтернатива трудоемкому и дорогостоящему стандартному методу оценки по потомству. В настоящее время SNP-чипы, охватывающие весь геном, доступны практически для всех видов домашнего скота. Стоимость таких исследований в несколько раз ниже, чем стоимость полногеномного секвенирования. SNP-чипы также широко используются в исследованиях полногеномного анализа ассоциаций (genome-wide association study, GWAS). В настоящее время с помощью SNP-чипов генотипировано около трех миллионов быков и одного миллиона свиней [25, 27]. Довольно распространенным подходом также является частичное

секвенирование генома, когда определяются последовательности в относительно небольших, но функционально-значимых участках генома, содержащих большое количество SNP.

Снижение стоимости полногеномного секвенирования в последние годы позволило повторно секвенировать с большим разрешением геномы животных, входящих в референсные популяции и, соответственно, получить высокоточную и надежную оценку племенной ценности. На сегодняшний день ресеквенировано около 2500 быков и сотни голов свиней, овец, коз и домашней птицы [27]. Наиболее известным, но далеко не единственным проектом, развиваемым в этом направлении, является “1000 Bulls Genome Project” [33]. Данный и подобные ему проекты призваны максимально глубоко охарактеризовать внутри- и межпородную генетическую вариацию животных.

Важно отметить, что полногеномное секвенирование сотен животных из референсной популяции позволяет импутировать (восстанавливать) недостающие генотипы миллионов распространенных SNP у гораздо большего числа животных из рабочей выборки, которые уже были генотипированы. Этот подход может быть реализован с использованием иерархической схемы, где верхний уровень – это референсная выборка из нескольких сотен животных, промежуточный уровень представляют животные, генотипированные по SNP высокой плотности, а нижний уровень (наиболее населенный) – животные, генотипированные по SNP низкой плотности. Генотипическая информация о SNP затем проецируется из верхних двух уровней на животных нижнего уровня с использованием двухэтапной стратегии импутации (two-step imputation strategy). В последние годы этот подход становится все более популярным, поскольку позволяет получить точную и надежную геномную оценку племенной ценности за наименьшую стоимость. Эта стратегия также позволила выявить целый ряд летальных рецессивных мутаций, а также идентифицировать множество QTL, связанных с хозяйственно ценными признаками, например, биохимическими параметрами молока [33, 34].

ОСНОВНЫЕ ЭТАПЫ ГЕНОМНОЙ СЕЛЕКЦИИ

Цикл геномной селекции состоит из четырех главных этапов (рис. 1).

1. Создание референсной выборки животных

Первый и наиболее важный этап геномной селекции, проводимый внутри одной породы, – это создание референсной выборки (называемой также референтной, калибровочной и тренировочной). Эта выборка должна состоять из большого числа животных этой породы, для каждого

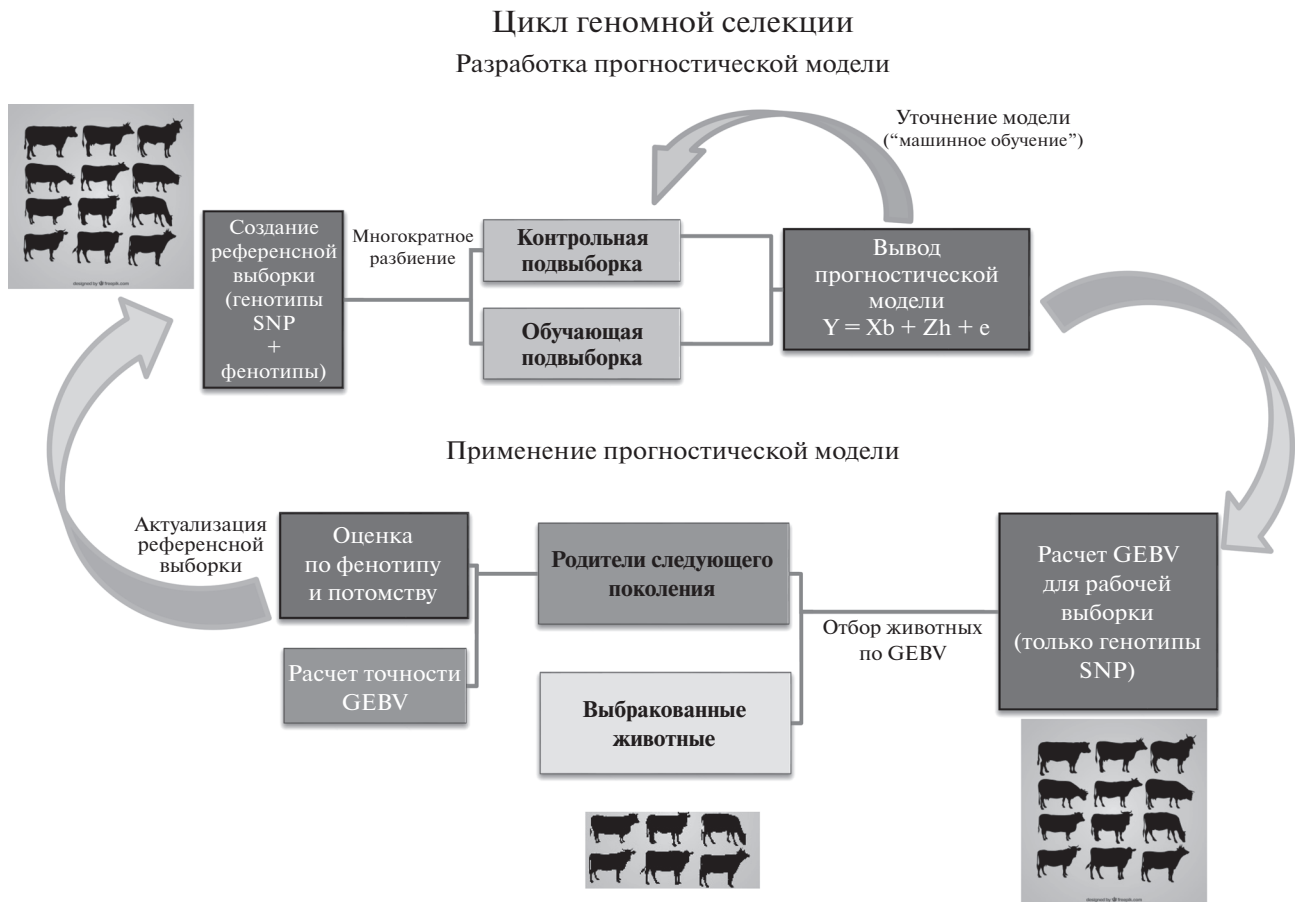


Рис. 1. Цикл геномной селекции в животноводстве.

из которых должны быть получены данные как о геноме, так и о признаках (генотипы и фенотипы). Референсная выборка позволит анализировать связи между результатами и признаками. От качества референсной выборки, в частности от качества оценивания интересующих признаков, критически зависит точность и надежность дальнейшей геномной оценки племенной ценности животного. Отсутствие достоверной оценки хозяйственно-полезных признаков может быть основным препятствием к реализации программ геномной селекции в некоторых странах. В РФ вопрос создания и поддержания референсной популяции также актуален.

2. Выбор модели и оценка модельных параметров по референсной выборке

Вторым этапом геномной селекции является построение математической регрессионной модели, устанавливающей статистическую зависимость между результатами генотипирования и измеренными хозяйственно полезными признаками. Для определения оптимальных модельных параметров в модели рассматривается возможность “машинного обучения” с разбиением референсной

выборки на контрольную и обучающую подвыборки. С использованием построенной модели на референсной выборке оценивается эффект каждого SNP на изучаемый признак [35]. На основе полученных эффектов выводится расчетное модельное уравнение.

3. Применение модели к рабочей выборке

Третьим этапом геномной селекции является применение разработанной модели к рабочей выборке, чтобы вычислить геномные оценки племенной ценности животных (Genomic Estimated Breeding Values, GEBV) на основе лишь геномных данных. Далее можно проводить отбор животных для дальнейшей селекционной работы, не дожидаясь оценки фенотипических признаков и данных о потомстве.

4. Актуализация референсной выборки и оценка точности геномной оценки

Четвертым этапом геномной селекции является использование данных о фенотипах и потомстве, полученных на рабочей выборке, для актуа-

лизации референсной выборки, чтобы увеличить точность геномной оценки.

Таким образом, в основе геномной селекции лежит относительно простая с математической точки зрения схема.

ПРОБЛЕМЫ ГЕНОМНОЙ СЕЛЕКЦИИ

Оценка эффектов SNP на исследуемый признак в референсной выборке основана на модели линейной регрессии, где зависимыми переменными являются значения признака животных, а предикторами – генотипы всех SNP. Одна из основных проблем на данном этапе происходит из-за того, что количество SNP во много раз превосходит число животных в референсной выборке или, другими словами, количество предикторов в регрессионном уравнении превышает число зависимых переменных. Стандартные математические методы, основанные на линейной регрессии, например метод наименьших квадратов, непригодны для оценки эффектов SNP, поскольку регрессионная модель параметрически перегружена. В геномной селекции для преодоления модельной сверхпараметризации обычно используется смешанная линейная модель BLUP (Best Linear Unbiased Prediction) [36, 37]. Эта модель основана на условном предположении о том, что эффекты всех SNP-маркеров на исследуемый признак малы и распределены нормально. Таким образом, с помощью модели BLUP удастся предсказать размер эффекта каждого SNP и оценить его значимость за счет использования дополнительной информации о распределении эффектов всех SNP-маркеров.

Одним из модельных параметров, задаваемых исследователем при BLUP-основанном анализе, является так называемый параметр сжатия (λ). Этот параметр масштабирует дисперсию нормального распределения эффектов SNP. Чем больше λ , тем выше уровень среднего шума относительно эффекта генетических маркеров, другими словами, тем более “сжатыми” будут эффекты SNP (арифметические значения меньше). Чем больше число используемых маркеров, тем выше должен быть параметр сжатия. Для выбора оптимального значения λ обычно используется контрольная подвыборка, являющаяся частью референсной выборки. Параметр сжатия подбирается таким образом, чтобы добиться максимального значения коэффициента корреляции между измеренной и предсказанной величинами признака, а значит и максимальной точности геномной оценки в обучающей подвыборке (рис. 2). При этом разделение референсной выборки на контрольную и обучающую может выполнено несколькими способами, где каждый раз определяется оптимальный параметр сжатия, а для итогового уравнения используется среднее арифметическое из полученных значений [3].

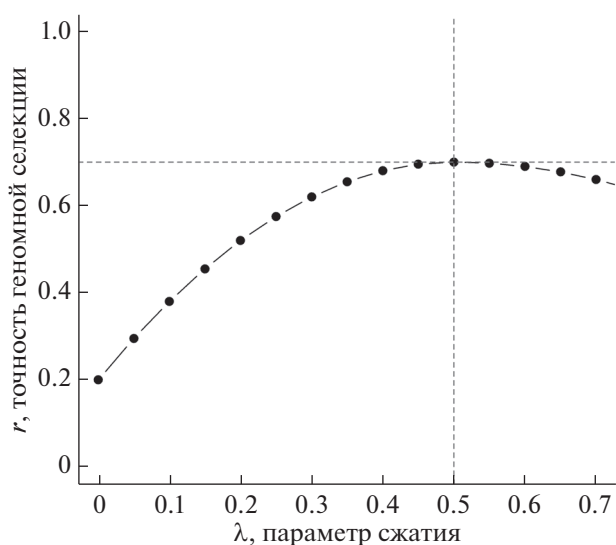


Рис. 2. Зависимость точности геномной оценки племенной ценности животных от параметра сжатия λ (ось абсцисс).

Альтернативным способом получения предсказательного уравнения для величины исследуемого признака является геномная модель BLUP или G-BLUP (Genomic BLUP) [38]. Ее особенность заключается в предварительной оценке близости геномов исследуемых животных и в последующем использовании этой информации при оценке исследуемых признаков. В крайне упрощенном виде геномную оценку животного с применением G-BLUP можно представить как поиск наиболее генетически близкого родственника в референсной популяции и присвоении оцениваемому животному соответствующей величины признака [19].

Одно из преимуществ G-BLUP в сравнении с BLUP состоит в том, что первая модель в меньшей мере опирается на генеалогическую информацию, которая во-первых не всегда доступна, а во-вторых будучи доступна, не всегда точна. Более того, анализ признаков менделевского типа является более точным при использовании генотипов SNP-маркеров, чем при использовании генеалогических данных [27]. Например, на основе только генеалогической информации полные сибсы (родные братья и сестры) имеют одну и ту же оценку племенной ценности, в то время как на основе генотипической информации о SNP возможно проследить передачу аллелей тех локусов, по которым родители гетерозиготны и, соответственно, присваивать сибсам различные более точные оценки племенной ценности.

До появления геномной селекции элитные быки-производители, получившие одинаковую оценку племенной ценности на основе генеалогической информации, должны были проходить дорогостоящее и затратное по времени тестиро-

вание по потомству. Этот процесс включал оценку производительности сотен дочерей каждого быка и, таким образом, занимал не менее пяти лет, поскольку необходимо было дожидаться, как минимум, первой лактации после половозрелости и отела. Стоимость оценки по потомству в расчете на одного быка составляет около 50000 долларов США [25], в то время как содержание одного быка-производителя на Центральной станции искусственного осеменения в Быково (Московская обл.) обходится примерно в 600000 рублей. Оценки, полученные с применением метода G-BLUP, при наличии качественной референсной выборки обладают приблизительно той же точностью, что и оценки по потомству. Однако G-BLUP-основанные оценки могут быть получены уже сразу после рождения быка или даже при SNP-генотипировании бластоцист (см. раздел “Репродуктивные технологии в геномной селекции”, [38a]). Причем для признаков, обладающих высокой наследуемостью (например для показателей молочной продуктивности), точность прогноза племенной ценности с помощью G-BLUP моделей может быть несколько ниже, чем при сочетанном использовании BLUP-моделей и методов оценки по потомству, что впрочем компенсируется (с экономической точки зрения) более ранним получением оценки. Для признаков с низкой наследуемостью (например для фертильности) метод G-BLUP позволяет получить не только более раннюю, но и более точную оценку [25, 27]. Например, недавно были показаны результаты внедрения семилетней программы по геномной селекции у молочного скота в США [39]. Для признаков с высокой наследуемостью (таких как молочная продуктивность) удалось добиться увеличения генетического прироста до размера 50–100%, а по признакам с низкой наследуемостью (таких как число соматических клеток в молоке и фертильность у коров) до размера 300–400%. В этой же работе [39] было показано, что в результате применения геномной селекции продолжительность цикла отбора для элитных быков-производителей сократилась с 7 до 2.5 лет, при этом уровень инбридинга не повышался [39]. Данные результаты, по-видимому, получены при исследовании больших выборок и не могут быть ориентирами для большинства популяций.

За многие десятилетия селекционной работы были накоплены данные о племенной ценности целых поколений животных. По понятным причинам большинство из них не могут быть генотипированы. Однако информацию об этих животных стало возможным включить в систему геномной оценки благодаря все чаще используемому подходу single-step BLUP (ssBLUP, называемому также гибридной оценкой). Этот подход основан на одновременном использовании генеалогической информации и генотипических данных SNP-маркеров, где

негенотипированные животные получают смоделированную SNP-уровневую оценку на основе их сходства с генотипированными особями по фенотипическим признакам [40]. Здесь, как и в случае определения коэффициента сжатия λ в BLUP-модели, выборка может быть многократно разделена на контрольную и обучающую с целью определения оптимальных параметров модели. Таким образом, BLUP и ssBLUP могут рассматриваться как модели машинного обучения в геномной селекции.

Точность геномной оценки племенной ценности особи определяет, насколько GEV соответствует реальной величине признака. Чем выше точность, тем в меньших пределах геномная оценка отличается от истинной величины признака. Высокая (0.6–0.8) точность GEV позволяет отбирать для дальнейшей племенной работы особей с желаемыми параметрами продуктивности, в то время как при низкой точности GEV (0.4 и ниже) возможно лишь отбраковывать животных с крайне нежелательными генетическими характеристиками (приведены ориентировочные значения точности геномной оценки для молочного скота). Например максимальная точность результатов BLUP-основанного анализа в случае использования только информации о родителях составляет 0.71 (квадратный корень из 0.5 – доли генетической вариации между полными сибсами).

В случае G-BLUP точность геномной оценки может быть определена как:

$$r = \sqrt{\frac{Nh^2}{Nh^2 + q}} \quad (2)$$

Здесь N – размер референсной выборки, h^2 – коэффициент наследуемости признака, а q – число независимых сегментов хромосом, оцененное по формуле $q = 2PL$, где P – количество предков в выборке, а L – эффективная длина генома в сантиморганах (сМ). Один сантиморган – участок хромосомы, на котором вероятность рекомбинации составляет 1%. При грубой оценке один миллион пар оснований примерно соответствует 1 сМ. Эффективная длина генома у КРС и человека составляет ~ 3000 сМ. Формула (2) позволяет оценить необходимый размер референсной выборки для получения требуемого уровня точности геномной оценки. Значения точности геномной оценки, полученные путем вычисления по данной формуле, крайне близки к эмпирическим величинам [41]. Из этой формулы также наглядно следует, что для получения одной и той же точности оценки в случае признаков с низкой наследуемостью необходимы очень большие размеры референсной популяции. Причем, как следует из формулы, при использовании максимально больших референсных популяций точность G-BLUP-оценки может приближаться к единице даже в отсутствие

данных о потомстве. На рис. 3 проиллюстрированы эти закономерности.

Создание референсной популяции — наиболее сложный и дорогостоящий этап геномной селекции. Референсная популяция должна быть значительного объема: десятки и сотни тысяч особей. Именно объем референсной популяции имеет решающее значение для достижения достаточного уровня точности геномной оценки племенной ценности животных. Например, в работе [41] при размере референсной выборки в 1000 голов КРС голштинской породы точность оценки составила около 0.35, при увеличении выборки до 5000 голов точность возросла до 0.68. Таким образом, геномная селекция эффективна в основном для широко распространенных пород, так называемых трансграничных или индустриальных. Эффективный размер референсной выборки может быть уменьшен за счет использования животных, близкородственных тем особям, генотип которых надо предсказать. Таким образом, уменьшается число независимых сегментов хромосом, q . В частности, точность геномной оценки может быть увеличена за счет включения в референсную выборку родителей тех особей, чей фенотип надо предсказать. Разумеется при расширении референсной выборки за счет близкородственных особей может повышаться риск того, что полученные прогностические уравнения будут малонадежны для генетически более отдаленных индивидов. Нахождение баланса между двумя параметрами — родственной близостью референсной популяции к оцениваемым особям и в то же время ее генетическим разнообразием, является одним из ключевых условий при создании качественных программ геномной селекции. Приведенные выше теоретические закономерности были наглядно проиллюстрированы эмпирическими данными в недавней работе [42], в которой крайне подробно исследовалось влияние размера референсной популяции и степени родства между гибридами на точность геномной оценки племенной ценности животных.

Отсутствие объемных референсных выборок на протяжении долгого времени являлось основным фактором, ограничивающим эффективность геномной селекции в развивающихся странах, а также для малочисленных пород [28]. Сейчас одной из основных тенденций в геномной селекции является объединение референсных популяций, в том числе на уровне международных программ [43], а также, если говорить о молочном животноводстве, включение в геномную оценку не только быков-производителей, но и коров, с целью повышения точности геномной оценки [44]. В некоторых случаях включение в референсную популяцию наряду с быками и коров привело пятикратному увеличению размера референсной выборки и улучшению точности отбора на 12% [45]. Генотипирование самок позволяет также избежать моногенных забо-

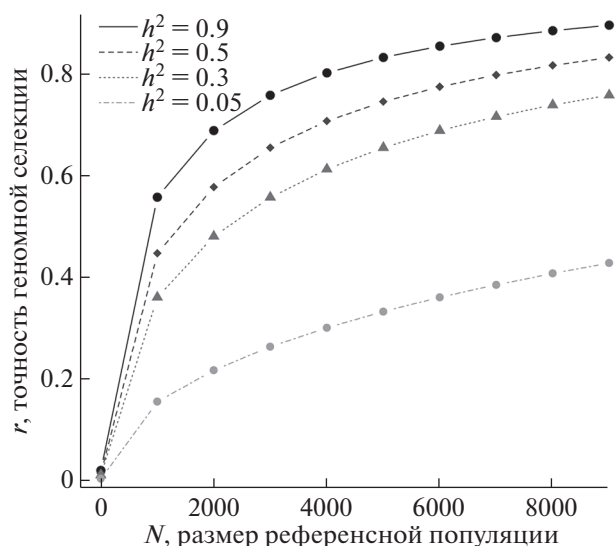


Рис. 3. Зависимость точности геномной оценки племенной ценности животных (r) от объема референсной популяции (N) и наследуемости признака (h^2).

леваний, что может быть крайне актуально в случае сохранения и разведения малочисленных местных пород. Кроме того, точность геномной оценки может быть повышена за счет непрерывной актуализации референсной популяции путем добавления геномных и фенотипических данных вновь рождающихся животных [46].

Как было сказано ранее, прогностические модели, используемые сейчас для вычисления GEBV и картирования QTL, линейны, то есть они предполагают, что однонуклеотидные полиморфизмы пропорционально и аддитивно влияют на выраженность признака. В настоящее время происходит разработка моделей, которые, как ожидается, смогут учитывать нелинейность действия отдельных полиморфизмов и QTL. Их эффективность, однако, в настоящее время не превышает эффективности линейных моделей [37].

Помимо игнорирования множественных эффектов отдельных полиморфизмов, линейные модели не учитывают сложного взаимодействия между SNP. Например, усиления или ослабления влияния на выраженность признака одного гена другим. Учитывать такого рода механизмы возможно лишь при интеграции в систему оценки данных биохимии, цитологии и физиологии, что в настоящее время представляется довольно сложной и комплексной задачей. Кроме того, этот подход требует применения ряда новых и еще не вполне освоенных генетиками математических методов, таких как моделирование структурными уравнениями. Включение в модель дополнительных переменных, например, характеризующих взаимодействие между SNP, потребует также увеличения объема используемых выборок.

Тем не менее, вполне вероятно, что следующий значимый шаг в повышении точности геномной селекции будет связан именно с использованием моделей, учитывающих сложное взаимодействие между различными элементами генома, отражающих реальные биохимические и физиологические процессы.

ЗАКЛЮЧЕНИЕ

Именно сочетание молекулярно-генетических и традиционных методов позволит решить многие проблемы генетики и селекции животных. При этом использование генетических маркеров в селекции в сочетании с методами молекулярно-генетического анализа может значительно с одной стороны упростить, с другой повысить эффективность всего селекционного процесса.

Решение вопросов, связанных с продовольственной безопасностью, повышением эффективности селекции, сохранением и созданием новых отечественных пород сельскохозяйственных животных для индустриального и органического сельского хозяйства требуют определенных биотехнологических подходов. Для животноводства таковым является метод геномной селекции, который позволяет определить более точно и быстро генетический потенциал животного, чем традиционные методы селекции, что в свою очередь несет значительные экономические выгоды. Геномная селекция особенно актуальна в тех случаях, когда оценка признака другими методами является дорогостоящей, либо вовсе невозможна, например, в тех случаях, когда затруднена оценка фенотипов.

В нашей стране и в мире все большее значение приобретает проблема эффективного управления генетическими ресурсами в области сельского хозяйства. Широкое разнообразие животноводческих ресурсов является важнейшим условием для адаптации и развития систем производства сельскохозяйственной продукции. Внедрение методов геномной селекции в животноводство Российской Федерации — это шанс (вероятно на данном этапе один из выходов) оценить реальный потенциал отечественного генофонда животных и на его базе создать конкурентоспособные стада в относительно кратчайшие сроки. Геномная селекция имеет высокий инновационный потенциал, однако ее внедрение в России требует разработки собственных систем оценки племенной и генофондной ценности животных или адаптации международных систем.

Необходимыми условиями для реализации новой селекционной стратегии являются создание объемной референсной выборки, достоверная оценка фенотипических признаков, а также вы-

бор оптимальных математических моделей. В нашей стране эти задачи еще предстоит решить.

Следует также учитывать, что потенциал классических селекционных методов еще не вполне реализован, а точность геномной оценки в настоящее время достигла определенного предела. Вполне вероятно, что следующий значимый шаг в повышении точности геномной селекции будет связан с использованием моделей, учитывающих сложное взаимодействие между различными элементами генома, отражающих реальные биохимические и физиологические процессы. Этот и другие возможные пути развития геномной селекции, в частности, применение постгеномных и репродуктивных технологий, будут рассмотрены во второй части обзора.

Работа выполнена при поддержке гранта РНФ № 19-76-20061.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. *Серебровский А.С.* Генетический анализ. М.: Наука, 1970. 342 с.
2. *Хлесткина Е.К.* Молекулярные маркеры в генетических исследованиях и в селекции // Вавилов. журн. генет. сел. 2013. Т. 17. № 4/2. С. 1044–1054.
3. *Ilska J.J., Meuwissen T.H.E., Kranis A., Woolliams J.A.* Use and optimization of different sources of information for genomic prediction // *Genet. Sel. Evol.* 2017. V. 49(90). <https://doi.org/10.1186/s12711-017-0365-7>
4. *Berry D.P., Bermingham M.L., Good M., More S.J.* Genetics of animal health and disease in cattle // *Irish Veter. J.* 2011. V. 64. № 5. <https://doi.org/10.1186/2046-0481-64-5>
5. *Legarra A., Croiseau P., Sanchez M.P. et al.* A comparison of methods for whole-genome QTL mapping using dense markers in four livestock species // *Genet., Selection, Evol.: GSE.* 2015. V. 47. № 6. <https://doi.org/10.1186/s12711-015-0087-7>
6. *Chamberlain A.J., McPartlan H.C., Goddard M.E.* The number of loci that affect milk production traits in dairy cattle // *Genetics.* 2007. V. 177. № 2. P. 1117–1123. <https://doi.org/10.1534/genetics.107.077784>
7. *Cai Z., Guldbrandtsen B., Lund M.S., Sahana G.* Prioritizing candidate genes for fertility in dairy cows using gene-based analysis, functional annotation and differential gene expression // *BMC Genomics.* 2019. V. 20(255). P. 255–265. <https://doi.org/10.1186/s12864-019-5638-9>
8. *Kiser J.N., White S.N., Johnson K.A. et al.* Identification of loci associated with susceptibility to *Mycobacterium avium* subspecies paratuberculosis (Map) tissue infection in cattle // *J. Anim. Sci.* 2017. V. 95. № 3. P. 1080–

1091.
<https://doi.org/10.2527/jas.2016.1152>
9. *Hu Z.L., Park C.A., Reecy J.M.* Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB // *Nucl. Ac. Res.* 2019. V. 47. № D1. P. D701–710.
<https://doi.org/10.1093/nar/gky1084>
 10. *Lusk J.L.* Association of single nucleotide polymorphisms in the leptin gene with body weight and backfat growth curve parameters for beef cattle // *J. Anim. Sci.* 2007. V. 85. № 8. P. 1865–1872.
<https://doi.org/10.2527/jas.2006-665>
 11. *Barendse W., Bunch R.J., Kijas J.W., Thomas M.B.* The effect of genetic variation of the retinoic acid receptor-related orphan receptor C gene on fatness in cattle // *Genetics.* 2007. V. 175. № 2. P. 843–853.
<https://doi.org/10.1534/genetics.106.064535>
 12. *Matsushashi T., Maruyama S., Uemoto Y. et al.* Effects of bovine fatty acid synthase, stearyl-coenzyme A desaturase, sterol regulatory element-binding protein 1, and growth hormone gene polymorphisms on fatty acid composition and carcass traits in Japanese Black cattle // *J. Anim. Sci.* 2011. V. 89. № 1. P. 12–22.
<https://doi.org/10.2527/jas.2010-3121>
 13. *Марзанов Н.С., Турбина И.С., Ескин Г.В. и др.* Скрининг гена дефицита лейкоцитарной адгезии у черно-пестрого голштинизированного скота // *С.-х. биол.* 2003. Т. 6. С. 23–30.
 14. *Thomsen B., Horn P., Panitz F. et al.* A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation // *Genome Res.* 2006. V. 16. № 1. P. 97–105.
<https://doi.org/10.1101/gr.3690506>
 15. *Ryan M.T., Hamill R.M., O'Halloran A.M. et al.* SNP variation in the promoter of the PRKAG3 gene and association with meat quality traits in pig // *BMC Genetics.* 2012. V. 13. № 66.
<https://doi.org/10.1186/1471-2156-13-66>
 16. *Song Y., Xu L., Chen Y. et al.* Genome-wide association study reveals the PLAG1 gene for knuckle, biceps and shank weight in simmental beef cattle // *PLoS One.* 2016. V. 11. № 12. e016831.
<https://doi.org/10.1371/journal.pone.0168316>
 17. *Wang Z., Chen Q., Liao R. et al.* Genome-wide genetic variation discovery in Chinese Taihu pig breeds using next generation sequencing // *Animal Genet.* 2017. V. 48. № 1. P. 38–47.
<https://doi.org/10.1111/age.12465>
 18. *Cochran S.D., Cole J.B., Null D.J., Hansen P.J.* Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle // *BMC Genet.* 2013. V. 14(49).
<https://doi.org/10.1186/1471-2156-14-49>
 19. *Goddard M.E., Hayes B.J.* Mapping genes for complex traits in domestic animals and their use in breeding programs // *Nat. Rev. Genet.* 2009. V. 10. P. 381–391.
<https://doi.org/10.1038/nrg2575>
 20. *Meuwissen T.H., Hayes B.J., Goddard M.E.* Prediction of total genetic value using genome-wide dense marker maps // *Genetics.* 2001. V. 157. № 4. P. 1819–1829.
 21. *Barton N.H., Etheridge A.M., Véber A.* The infinitesimal model: Definition, derivation, and implications // *Theor. Popul. Biol.* 2017. V. 118. P. 50–73.
<https://doi.org/10.1016/j.tpb.2017.06.001>
 22. *Zuidhof M.J., Schneider B.L., Carney V.L. et al.* Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 2005 // *Poult. Sci.* 2014. V. 93. № 12. P. 2970–2982.
<https://doi.org/10.3382/ps.2014-04291>
 23. *Matukumalli L.K., Lawley C.T., Schnabel R.D. et al.* Development and characterization of a high density SNP genotyping assay for cattle // *PLoS One.* 2009. V. 4(4): e5350.
<https://doi.org/10.1371/journal.pone.0005350>
 24. *McCue M.E., Bannasch D.L., Petersen J.L. et al.* A high density SNP array for the domestic horse and extant *Perissodactyla*: utility for association mapping, genetic diversity, and phylogeny studies // *PLoS Genet.* 2012. V. 8(1): e1002451.
<https://doi.org/10.1371/journal.pgen.1002451>
 25. *Wiggans G.R., Cole J.B., Hubbard S.M., Sonstegard T.S.* Genomic selection in dairy cattle: the USDA experience // *Annu. Rev. Anim. Biosci.* 2017. V. 5. P. 309–327.
<https://doi.org/10.1146/annurev-animal-021815-111422>
 26. *Dekkers J.C.* Application of genomics tools to animal breeding // *Curr. Genomics.* 2012. V. 13. № 3. P. 207–212.
<https://doi.org/10.2174/138920212800543057>
 27. *Georges M., Charlier C., Hayes B.* Harnessing genomic information for livestock improvement // *Nat. Rev. Genet.* 2019. V. 20. № 3. P. 135–156.
<https://doi.org/10.1038/s41576-018-0082-2>
 28. *Mrode R., Ojango J.M.K., Okeyo A.M., Mwacharo J.M.* Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: Current status and future prospects // *Front. Genet. Genomic.* 2019. V. 9(694).
<https://doi.org/10.3389/fgene.2018.00694>
 29. *Brookes A.J.* The essence of SNPs // *Gene.* 1999. V. 234. № 2. P. 177–186.
[https://doi.org/10.1016/s0378-1119\(99\)00219-x](https://doi.org/10.1016/s0378-1119(99)00219-x)
 30. *Sachidanandam R., Weissman D., Schmidt S.C. et al.* International SNP map working group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms // *Nature.* 2001. V. 409. P. 928–933.
 31. *Nicolazzi E.L., Biffani S., Biscarini F. et al.* Software solutions for the livestock genomics SNP array revolution // *Anim. Genet.* 2015. V. 46. № 4. P. 343–353.
<https://doi.org/10.1111/age.12295>
 32. *Keeble-Gagnère G., Rigault P., Tibbits J. et al.* Optical and physical mapping with local finishing enables megabase-scale resolution of agronomically important regions in the wheat genome // *Genome Biol.* 2018. V. 19(112).
<https://doi.org/10.1186/s13059-018-1475-4>
 33. *Daetwyler H.D., Capitan A., Pausch H. et al.* Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle // *Nat. Genet.* 2014. V. 46. № 8. P. 858–865.
<https://doi.org/10.1038/ng.3034>
 34. *Kadri N.K., Harland C., Faux P. et al.* Coding and non-coding variants in HFM1, MLH3, MSH4, MSH5, RNF212 and RNF212B affect recombination rate in cattle // *Genome Res.* 2016. V. 26. № 10. P. 1323–1332.
<https://doi.org/10.1101/gr.204214.116>
 35. *Kaler A.S., Purcell L.C.* Estimation of a significance threshold for genome-wide association studies // *BMC*

- Genomics. 2019. V. 20. № 1(618).
<https://doi.org/10.1186/s12864-019-5992-7>
36. *Van Raden P.M.* Efficient methods to compute genomic predictions // *J. Dairy Sci.* 2008. V. 91. № 11. P. 1414–1423.
<https://doi.org/10.3168/jds.2007-0980>
37. *Kemper K.E., Bowman P.J., Hayes B.J. et al.* A multi-trait Bayesian method for mapping QTL and genomic prediction // *Genet. Sel. Evol.* 2018. V. 50. № 10.
<https://doi.org/10.1186/s12711-018-0377-y>
38. *Clark S.A., van der Werf J.* Genomic best linear unbiased prediction (gBLUP) for the estimation of genomic breeding values // *Methods Mol. Biol.* 2013. V. 1019. P. 321–330.
https://doi.org/10.1007/978-1-62703-447-0_13
- 38а. *Столовский Ю.А., Свищева Г.Р., Пискунов А.К.* Геномная селекция. II. Перспективные направления // *Генетика.* 2020. Т. 56. № 10. С. 00.
39. *García-Ruiz A., Cole J.B., Van Raden P.M. et al.* Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection // *Proc. Natl Acad. Sci. USA.* 2016. V. 113. № 28. P. 3995–4004.
<https://doi.org/10.1073/pnas.1519061113>
40. *Fernando R.L., Cheng H., Golden B.L., Garrick D.J.* Computational strategies for alternative single-step Bayesian regression models with large numbers of genotyped and non-genotyped animals // *Genet. Sel. Evol.* 2016. V. 48(96).
<https://doi.org/10.1186/s12711-016-0273-2>
41. *Hayes B., Goddard M.* Genome-wide association and genomic selection in animal breeding // *Genome.* 2010. V. 53. № 11. P. 876–883.
<https://doi.org/10.1139/G10-076>
42. *Edwards S.M., Buntjer J.B., Jackson R. et al.* The effects of training population design on genomic prediction accuracy in wheat // *Theor. Appl. Genet.* 2019. V. 132. № 7. P. 1943–1952.
<https://doi.org/10.1007/s00122-019-03327-y>
43. *Eynard S.E., Croiseau P., Laloë D. et al.* Which individuals to choose to update the reference population? Minimizing the loss of genetic diversity in animal genomic selection programs // *Genes, Genomes, Genetics.* 2018. V. 8. № 1. P. 113–121.
<https://doi.org/10.1534/g3.117.11172018>
44. *Dehnavi E., Mahyari S.A., Schenkel F.S., Sargolzaei M.* The effect of using cow genomic information on accuracy and bias of genomic breeding values in a simulated Holstein dairy cattle population // *J. Dairy Sci.* 2018. V. 101. № 6. P. 5166–5176.
<https://doi.org/10.3168/jds.2017-12999>
45. *Boison S.A., Utsunomiya A.T.H., Santos D.J.A. et al.* Accuracy of genomic predictions in Gyr (*Bos indicus*) dairy cattle // *J. Dairy Sci.* 2017. V. 100. № 7. P. 5479–5490.
<https://doi.org/10.3168/jds.2016-11811>
46. *Silva R.M.O., Fragomeni B.O., Lourenco D.A.L. et al.* Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population // *J. Anim. Sci.* 2016. V. 94. № 9. P. 3613–3623.
<https://doi.org/10.2527/jas.2016-0401>

Genomic Selection. I. Latest Trends and Future Trajectories

Yu. A. Stolpovsky^{a, *}, A. K. Piskunov^{a, **}, and G. R. Svishcheva^{a, b, ***}

^aVavilov Institute of General Genetics, Russian Academy of Science, Moscow, 119991 Russia

^bFederal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

*e-mail: stolpovsky@mail.ru

**e-mail: aleksei.piskunov@gmail.com

***e-mail: gulsvi@mail.ru

According to the forecast of the Food and Agriculture Organization of the United Nations global demand for food products of animal and plant origin will increase by 74% by 2050 Satisfying this demand without a destructive impact on the environment is possible only by introducing of new technologies in animal husbandry and crop production and by protection and development of organic agriculture. Genomic selection is supposed to play a key role in this process representing one of the most promising and safe methods for improving the genetic qualities of farm animals and plants. This review summarizes information about genomic selection: it shows how the genomic assessment of breeding value is realized and reveals the key conditions necessary for its implementation; advantages and limitations of genomic and marker selection are also discussed. The review also outlines possible development trajectory for genomic selection.

Keywords: genomic selection, single nucleotide polymorphism, breeding value, domesticated animal species, reference population.