



УДК 577.21/639.371.5

Аквакультура

Генотипирование радужной форели методом полногеномного секвенирования с низким покрытием (low-pass секвенирование, lpWGS) для целей геномной селекции

Н.С. Мюге^{1,2}, Ф.С. Шарко^{1,3}

¹ Всероссийский научно-исследовательский институт рыбного хозяйства и океанографии (ГНЦ РФ ФГБНУ «ВНИРО»), Окружной проезд, 19, Москва, 105187

² Институт биологии развития им. Н.К. Кольцова РАН (ФГБНУ «ИБР РАН»), ул. Вавилова, д.26, Москва, 119334

³ Национальный исследовательский центр «Курчатовский институт» (ФГБУ НИЦ «Курчатовский институт»), площадь Академика Курчатова, д.1, Москва, 123182

E-mail: muge@mail.ru

SPIN-код: Н.С. Мюге – 1916-2289; Ф.С. Шарко – 5357-6805

Цель работы: апробация метода секвенирования с особо низким покрытием (low-pass секвенирование) с дальнейшей импутацией генотипов на радужной форели.

Материалы и методы: выделение ДНК проводили в пласечном формате (4 × 96 образцов), для пробоподготовки и секвенирования использовали набор MGI AgriHigh Low-pass WGS Package, создание референсной базы и импутация выполнены на базе вычислительного кластера Группы Биоинформатики ВНИРО.

Новизна: впервые в России проведено low-pass секвенирование на радужной форели для целей геномной селекции.

Результат: анализ 260 особей радужной форели показал высокую достоверность генотипирования данным методом, в среднем для каждой рыбы достоверно определены около 90% из 12 млн локусов панели геномных маркеров. Показана высокая генетическая дифференциация четырёх изученных селекционных линий радужной форели.

Практическая значимость: апробированный метод low-pass секвенирования может использоваться как более экономичная и значительно более информативная альтернатива традиционным ДНК-микрочипам для генетического анализа. Метод применим как для полногеномного ассоциативного анализа (GWAS), так и для расчёта селекционной ценности при формировании селекционного ядра форели в центре геномной селекции.

Ключевые слова: радужная форель *Oncorhynchus mykiss*, метод секвенирования с низким покрытием, импутация, геномная селекция.

Genotyping of rainbow trout using low-pass whole-genome sequencing for genomic selection

Nikolai S. Muge^{1,2}, Fedor S. Sharko^{1,3}

¹ Russian Federal Research Institute of Fisheries and Oceanography («VNIRO»), 19, Okruzhnoy proezd, Moscow, 105187, Russia

² N.K. Koltzov Institute of Developmental Biology RAS («IDB RAS»), 26, Vavilov Str., Moscow, 119334, Russia

³ National Research Center I.V. Kurchatov Institute (NRC «Kurchatov Institute»), 1, Akademik Kurchatov Square, Moscow, 123182, Russia

Goal: To verify a low-pass sequencing method (lpWGS) with subsequent genotype imputation in rainbow trout.

Materials and Methods: DNA extraction was performed by a customized protocol in plate format (4 × 96 samples), sample preparation and sequencing were performed using the MGI AgriHigh Low-pass WGS Package, reference database creation, and imputation were performed using the VNIRO Bioinformatics Group computing cluster.

Novelty: low-pass sequencing was conducted on rainbow trout for genomic selection purposes for the first time in Russia.

Result: Analysis of 260 rainbow trout individuals using this method demonstrated high genotyping reliability; on average, approximately 90% of the 12 million loci in the genomic marker panel were reliably identified for each fish. High genetic differentiation was demonstrated among the four studied rainbow trout breeding lines.

Practical significance: The verified low-pass sequencing genotyping method can be used as a cheaper and more informative alternative to traditional DNA microarrays. This method enables both GWAS analysis and breeding value estimation for forming a trout breeding nucleus at the genomic selection Center.

Keywords: rainbow trout *Oncorhynchus mykiss*, low-pass sequencing, imputation, genomic selection.

ВВЕДЕНИЕ

Недорогие методы полногеномного генотипирования необходимы для реализации точной геномной селекции в растениях и животных. В мировой селекционной практике до недавнего времени широко использовались ДНК-микрочипы для генотипирования однонуклеотидных полиморфизмов (SNP) при работе с большими выборками в рамках полногеномных ассоциативных исследований (GWAS), геномной селекции (GS) и популяционно-генетических исследований. Для радужной форели *Oncorhynchus mykiss* (Walbaum, 1792) разработан и успешно применяется ДНК-чип Axiom 57K, содержащий 57 тысяч SNP-зондов [Palti et al., 2015]. Данный ДНК-чип используется как для проведения GWAS, так и для рутинного генотипирования рыб из племенного ядра с целью расчёта племенной ценности (breeding value) и проведения геномной селекции. После удаления SNP с низким качеством сигнала и низкой частотой минорного аллеля (minor allele frequency – MAF) для каждой выборки остаются полезными только около 30-35 тыс. SNP из 57 тыс. [Vallejo et al., 2017; Frasin et al., 2020]. Учитывая, что производители ДНК-чипов прекратили поставки высокотехнологичной продукции в Россию, проведение генотипирования сельскохозяйственных животных, включая аквакультурные виды рыб, на ДНК-чипах становится логистически трудновыполнимой задачей.

Полногеномное секвенирование нового поколения (NGS) прочно вошло в практику геномных исследований, обеспечивая возможность расшифровки генетического кода организмов. В отличие от классического подхода, требующего глубокого покрытия (30-50×) для многократного прочтения каждого нуклеотида, в последние годы активно развивается метод полногеномного секвенирования с низким покрытием (low-pass whole genome sequencing, lpWGS или low-coverage WGS, lcWGS). Его суть заключается в существенном снижении средней глубины секвенирования – обычно до 0,1-5× [Lou et al., 2021]. Технически процесс секвенирования включает фрагментацию геномной ДНК на короткие сегменты (200-500 п. н.), лигирование платформенно-специфических адаптеров и параллельное массовое прочтение миллионов таких фрагментов (ридов). Покрытием в данном контексте называют количество ридов, приходящихся на конкретную позицию генома, а среднее покрытие отражает усреднённую глубину прочтения по всем позициям. Снижение этого показателя при lpWGS приводит к значительному удешевлению анализа, сохраняя при этом возможность получения масштабной геномной информации. Однако существует и обратная

сторона данного метода. При секвенировании с высоким покрытием анализ генома достаточно прост и надёжен – все полученные прочтения от данной особи картируются на референсный геном, и каждый нуклеотид особи определяется консенсусом всех прочтений, картированных на участок генома с данным нуклеотидом. В случае с низким покрытием, применяемым при lpWGS, большая доля генома оказывается не покрыта ни одним прочтением, что делает прямой анализ методом консенсуса невозможным. В этом случае генотипирование проводится методом импутации: на основании большого числа (несколько сотен или тысяч) генотипов, полученных путём глубокого секвенирования, создаётся база фазированных гаплотипов – вариантов, переданных от одного из родителей. Таким образом появляется возможность восстановить непокрытые при low-pass полиморфизмы за счёт идентификации гаплотипа по тем полиморфизмам, которые оказались покрыты прочтениями (ридами), и в результате импутации с высокой достоверностью восстанавливаются все полиморфизмы генома. Важным ограничением импутации является её неспособность выявлять *de novo* мутации, отсутствующие в референсной панели. Кроме того, для обеспечения высокой точности восстановления, как правило, учитываются лишь относительно частые варианты с частотой минорного аллеля (MAF) в пределах 1-5%, которые хорошо представлены в референсной выборке [Lou et al., 2021].

В настоящей работе нами проведён эксперимент по генотипированию радужной форели методом low-pass секвенирования с целью верификации достоверности генотипирования и включения этого протокола в программу геномной селекции радужной форели на базе Центра геномной селекции «ВНИРО» (ЦГС ВНИРО).

МАТЕРИАЛ И МЕТОДЫ

Материал для проведения генотипирования

Материалом для секвенирования послужили 4 выборки молоди рыб из одного их аквакультурных хозяйств Северной Осетии, 2 выборки молоди разного возраста, выращенной из икры, завезённой из Испании, и 2 выборки молоди, выращенной из икры, привезённой в разное время из Франции. Данные выборки были отобраны 30 июля 2025 года в рамках проекта по изучению генетических маркеров устойчивости к вирусу инфекционного гемопозитического некроза (IHNV) лососевых. Также в исследование включены генетические образцы от 66 особей форели, содержащихся в Центре геномной селекции ВНИРО и ото-

Таблица 1. Образцы радужной форели, включённые в секвенирование методом lpWGS**Table 1.** Rainbow trout samples included in lpWGS sequencing

Выборка	Дата отбора проб	Происхождение	Кол-во	Идентификатор
SP-S	30.07.2025	Испания, 05.04.2025	47	CO25-1-CO25-47
SP-R	30.07.2025	Испания, 18.01.2025	48	CO25-51-CO25-98
FR_R	30.07.2025	Франция 24.01.2025	46	CO25-101-CO25-139, CO25-141-CO25-147
FR-S	30.07.2025	Франция 01.03.2025	48	CO25-151-CO25-198
VNIRO20	04.08.2025	ЦГС ВНИРО, янв. 2024	22	ONC469-ONC490 ONC511-ONC538
VNIRO8	04.08.2025	ЦГС ВНИРО, янв. 2025	21	ONC491-ONC510, ONC539
ВСЕГО			260	

бренных в ходе проведения бонитировки и чипирования (индивидуального мечения RFID метками) рыб. Общий объём выборок, для которых было проведено полногеномное секвенирование методом low-pass, составил 384 образца, из которых 260 принадлежали радужной форели, а оставшиеся 124 – атлантическому лососю (96 образцов) и другим видам рыб (25 образцов). В данной работе обсуждается генетический анализ только образцов радужной форели (табл. 1).

Выделение ДНК и подготовка библиотек для секвенирования

Производитель набора для Low-Pas секвенирования AgriHigh предлагает два варианта – поставка набора для секвенирования с модулем для выделения ДНК на магнитных частицах MGIEasy Genomic DNA Extraction Prepacked Kit (MGISP-NE384) (384 RXN) и требующего использования автоматизированной системы MGISP-NE384RS или без данного модуля. В набор MGISP-NE384 входят магнитные частицы, расходные материалы и реагенты – лизирующий раствор LS, Протеиназа К, промывочные растворы WB1, WB2 и элюирующий буфер EB.

В связи с отсутствием в лаборатории MGISP-NE384RS, а также с целью отработки более экономичного протокола для дальнейшего массового генотипирования было принято решение провести выделение методом, рутинно применяющимся в Отделе молекулярной генетики ВНИРО для массового выделения ДНК в плашечном формате. Выделение ДНК проводили по стандартной методике с использованием адсорбционных колонок PALL и с использованием лизирующего и промывочных растворов, приготовленных из общелабораторных реагентов (модифицированная методика из Ivanova et al., 2006). Выбор данного метода обусловлен его высокопроизводительностью и дешевизной, не требующей дополнительных дорогостоящих реагентов.

Контроль качества выделенной ДНК проводился путём электрофореза в 1,2% агарозном геле; во всех образцах ДНК присутствовала высокомолекулярная фракция. Выборочное измерение образцов на флуориметре (QUBIT 3) показало, что присутствовавшего в образцах количества ДНК достаточно для приготовления библиотек. Четыре планшета с выделенной ДНК (384 образца) были переданы в Биотек Кампус (Москва) для последующего секвенирования.

Дальнейшие процедуры пробоподготовки – нормализация образцов, подготовка PCR-FREE библиотек, циркуляризация в ДНК-микрошарики (DNA nanoballs – DNB) и загрузка библиотеки в ячейку MGIDL-T7RS проводились с использованием наборов Prep Set for Low-pass WGS (384RXN) и MGIEasy Large-scale PCR-Free FS Library на автоматизированной линии Биотек Кампуса в соответствии с инструкцией к набору AgriHigh Low-pass WGS Package 1536RXN специалистами Биотек Кампуса под контролем команды технического сопровождения российского представительства BGI в рамках первого в России тестового запуска набора AgriHigh Low-pass WGS Package. Секвенирование проводилось на приборе MGI T7. После завершения секвенирования и демультиплексации результаты секвенирования в формате fastq поступили в группу Биоинформатики ФГБНУ «ВНИРО». Работы по выделению ДНК и биоинформационной обработке данных проводились на базе ЦКП «Рыбохозяйственная геномика». Генетические образцы форели депонированы в Биоресурсной коллекции ВБР ВНИРО.

РЕЗУЛЬТАТЫ

Оценка качества секвенирования и картирования

Для 260 образцов радужной форели (*O. mykiss*) был проведён анализ качества данных low-pass

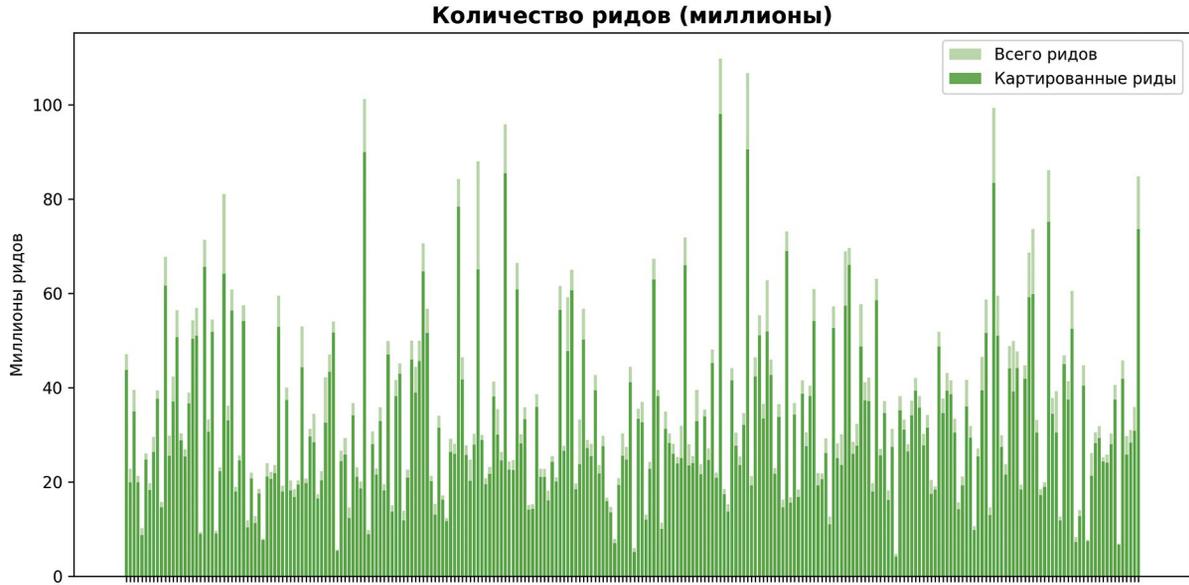


Рис. 1. Распределение картированных/некартированных прочтений по всем образцам

Fig. 1. Distribution of mapped/unmapped reads across all samples

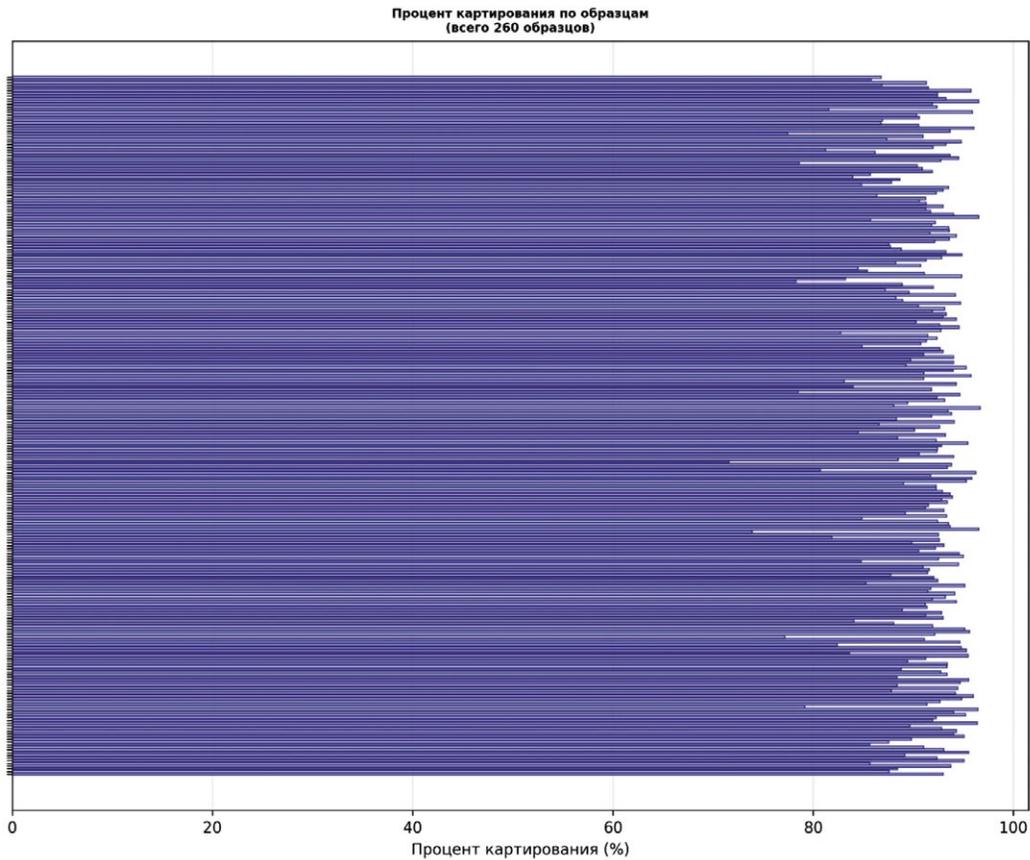


Рис. 2. Распределение процента картирования прочтений (mapped_percent) по всем образцам

Fig. 2. Distribution of the percentage of mapped reads (mapped_percent) across all samples

секвенирования. На образец приходилось в среднем 39 098 611 прочтений, из которых 35 276 784 (90,91%) были успешно картированы на референсный геном

USDA_Отука_1.1 (рис. 1). Средний процент правильно спаренных прочтений составил 81,24%. Дубликаты секвенирования в образцах обнаружены не были. По-

лученные метрики свидетельствуют о высоком общем качестве данных.

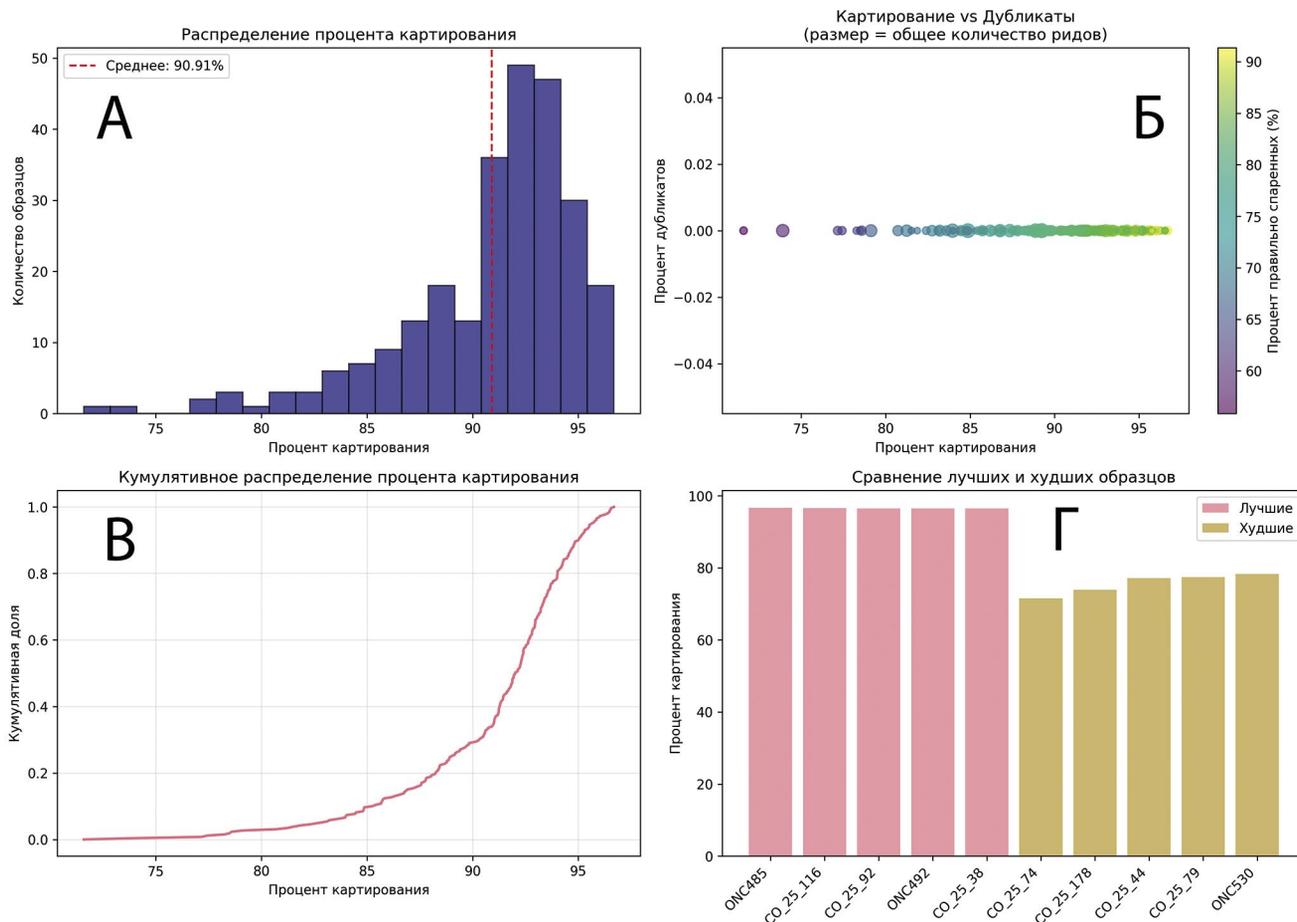
Значение процента картирования (mapped_percent) варьировалось от 71,59% (образец CO_25_74) до 96,69% (образец ONC485) при медиане 92,38% (рис. 2). Более 80% образцов показали уровень картирования выше 90%, что подтверждает высокое качество эксперимента и последующей обработки данных. Основные характеристики картирования представлены на рис. 3.

Показатель правильно спаренных прочтений (properly_paired_percent) варьировался в широких пределах – от 55,85% (CO_25_74) до 91,35% (ONC492), при среднем значении 81,24%. Большинство образцов продемонстрировали уровень парного картирования выше 80%, что приемлемо для дальнейшего анализа. Однако была выявлена группа образцов с аномально низкими показателями, включающая CO_25_74 (mapped_percent=71,59%, properly_paired_percent=55,85%), CO_25_178 (73,91%; 59,34%), CO_25_43 (79,13%; 66,24%) и ONC528 (78,58%; 63,33%). Для этих образ-

цов может потребоваться повторная обработка, дополнительная проверка качества библиотек или специальное исследование причин низкого качества данных.

Процент покрытия генома (coverage %) варьировался от 19,55% (ONC513) до 94,42% (CO_25_57), составляя в среднем 72,65% при медиане 73,46%. Покрытие выше 70%, приемлемое для большинства видов анализа, было отмечено у ~65% образцов. Глубина покрытия (meandepth) колебалась от 0,27× (ONC513) до 6,28× (CO_25_57) со средним значением 2,29× и медианой 1,82×. Большинство образцов (~70%) показали глубину в диапазоне 1,0-3,0×, что соответствует стандартным параметрам low-pass секвенирования.

Среднее качество оснований ДНК (meanbaseq) было стабильно высоким во всех образцах, варьируясь от 35,4 (CO_25_38, CO_25_116) до 36,9 (CO_25_74, CO_25_88) при среднем значении 36,3. Качество картирования (meanmapq) имело больший разброс – от 26,1 (CO_25_74) до 31,4 (CO_25_76, CO_25_92, CO_25_184), со средним 29,9 и медианой 30,1. Около 75% образцов продемонстрировали качество карти-



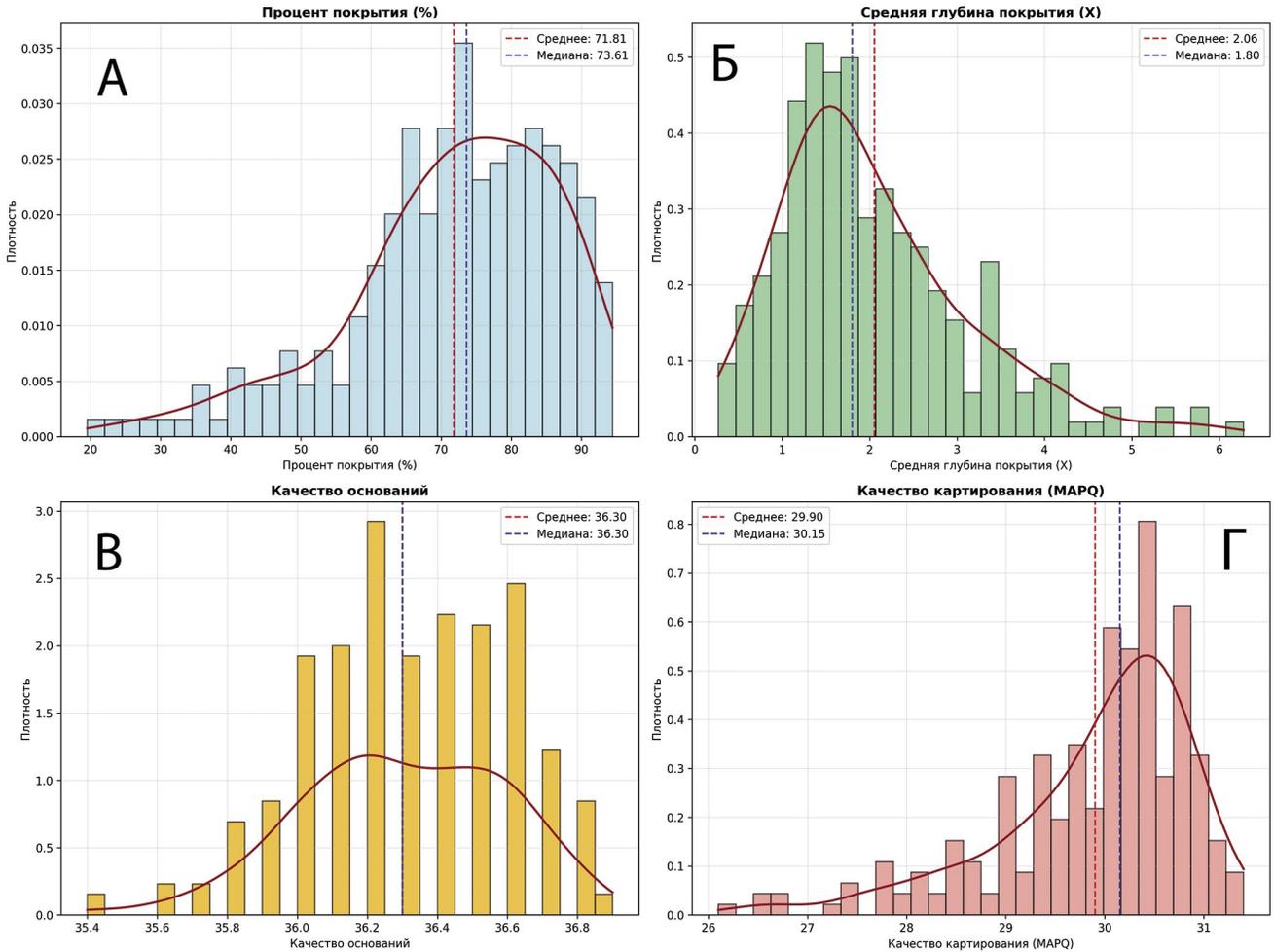


Рис. 4. Распределение образцов по основным метрикам покрытия

Fig. 4. Distribution of samples by main coverage metrics

рования выше 29,0, что считается хорошим показателем. В ходе анализа была выделена группа образцов с экстремально низкими показателями покрытия и глубины:

- ONC513 (coverage 19,55%, depth 0,27×)
- ONC535 (coverage 23,18%, depth 0,34×)
- ONC486 (coverage 26,95%, depth 0,35×)
- ONC493 (coverage 29,17%, depth 0,43×)
- ONC510 (coverage 30,68%, depth 0,46×)

Эти образцы демонстрируют недостаточное для надёжного анализа покрытие и могут требовать повторного секвенирования. Распределение образцов по основным метрикам покрытия представлены на рис. 4.

Создание референсной панели радужной форели для импутации геномов с низким покрытием

Для создания референсной панели было отобрано 108 особей форели, которые ранее были прогеноти-

пированы нами со средним покрытием всего генома 20×. Также были использованы данные из открытых источников: 117 геномов особей дикой американской радужной форели¹ и 410 образцов аквакультурной радужной форели из двух крупнейших племенных центров: Troutlodge Inc. (TLU) и USDA-ARS National Center for Cool and Cold Water Aquaculture (NCCCWA) (табл. 2).

Необработанные данные секвенирования всех 635 особей радужной форели были отфильтрованы по качеству. Для выравнивания прочтений на референсный геном *O. mykiss* (USDA_Омука_1.1) использовалось программное обеспечение Bowtie2 v.2.3.5.1 [Langmead and Salzberg, 2012] с настройками по умолчанию. Затем, для обнаружения полиморфизмов и поиска переменных локусов использовалась программа *bcftools* [Danecsek et al., 2021]. Фильтрация полученного vcf файла с полиморфизмами была выполнена с помощью утилиты *bcftools view*. Итого для 635

¹ <https://www.ncbi.nlm.nih.gov/bioproject/402066>

Таблица 2. Список образцов форели, данные для которых получены из открытых источников
Table 2. List of trout samples for which data were obtained from open sources

Популяция	Месяц нереста	Количество особей
NCCCWA чётная	Февраль	84
NCCCWA нечётная	Февраль	84
TLUM	Май	76
TLUA	Август	19
TLUN	Ноябрь	147
WILD		117

образцов радужной форели в референсную базу для импутации вошло 17037 139 SNP.

Фазирование полученной референсной базы радужной форели было выполнено с помощью программы shareit5 [Hofmeister et al., 2023]. При использовании данных WGS на выборке большого размера была выполнена процедура SHAPEIT5_phase_common в различных крупных регионах хромосом, разделённых на участки по 20 cM. Далее мы выполняли фазирование в фрагментах с перекрывающимися регионами, которые достаточно велики, чтобы иметь достаточное количество гетерозиготных сайтов для этапа лигирования (т. е. сборки всех фрагментов вместе). Фазирование подразумевает разделение унаследованных по материнской и отцовской линии копий каждой хромосомы на гаплотипы для получения полной картины генетической изменчивости.

Для валидации процедуры импутации два образца – MAD51 (Adler) и ONC155 (Norway-IA) – были исключены из референсной базы, после чего их покрытие было искусственно понижено до уровней 0,1x, 0,3x, 0,5x, 0,7x и 0,9x. Полученные наборы данных с пониженным покрытием затем были импутированы и фазированы с использованием программного пакета GLIMPSE2 [Rubinacci et al., 2021]. Процесс включал четыре основных этапа: 1) сегментацию хромосом на регионы для импутации с помощью GLIMPSE2_chunk; 2) подготовку референсной панели, где GLIMPSE2_split_reference разделял её на фрагменты, соответствующие окнам импутации с буферными зонами и данными генетической карты; 3) непосредственное выполнение импутации и фазирования через GLIMPSE2_phase с использованием бинарной референсной панели и списка BAM-файлов с низким

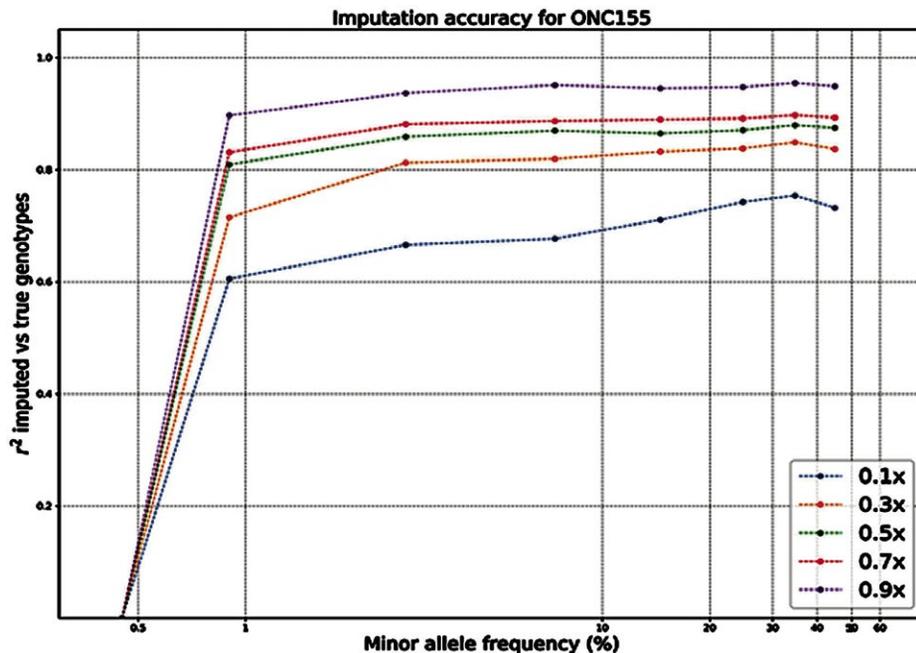


Рис. 5. Соответствие восстановленных и исходных данных для образца ONC155 с понижением покрытия до 0,1x, 0,3x, 0,5x, 0,7x и 0,9x. Чем ближе r^2 к 1, тем точнее импутация

Fig. 5. The measure of agreement between the reconstructed and original data for the ONC155 sample with coverage decreasing to 0.1x, 0.3x, 0.5x, 0.7x, and 0.9x. The closer r^2 is to 1, the more accurate the imputation is

покрытием; 4) лигирование импутированных фрагментов в цельные хромосомные файлы с помощью утилиты GLIMPSE2_ligate.

Для проверки точности импутации, по сравнению с исходным набором данных, мы использовали инструмент GLIMPSE2_concordance, который вычисляет корреляции r^2 между импутированными данными и высоконадёжными генотипами из набора данных с высоким покрытием. Для тестируемых образцов MAD51 и ONC155 с пониженным покрытием была вычислена корреляция r^2 , которая показала, что мера согласованности (concordance) варьирует в зависимости от принадлежности к определённой популяции тестового образца и составляет от 93 до 95% при покрытии генома 0,9× (рис. 4).

Для оценки качества импутации геномных данных форели, полученных методом low-pass секвенирования, был проведён анализ качества статистической импутации генотипов для выборки, состоящей из 260 особей форели. Основной метрикой, используемой для оценки точности этого процесса, является показатель INFO_SCORE (R^2) – представленный в виде вероятности квадрат корреляции между истинным (неизвестным) и импутированным генотипами. Данный показатель также известен как коэффициент детерминации или Imputation Quality Score. В исследованной выборке данный показатель имел следующие значения: минимум – 0,997, среднее – 0,9993, максимум – 1. Данные результаты анализа свидетельствуют о высоком качестве проведённой импутации. Все генетические варианты могут быть использованы для последующих этапов исследования, включая ассоциационные анализы, с высокой степенью достоверности, образцы с недопустимо низким ($\text{INFO_SCORE} < 0,3$) коэффициентом детерминации в анализе отсутствовали.

Проведение кластеризации изученных выборок методом анализа главных компонент

Для проведения сравнительного анализа генетической кластеризации изученных рыб по всему геному для 260 особей были объединены похромосомные VCF-файлы в единый набор данных. В результате импутации было выявлено 12 177 807 полиморфных сайтов. Общее количество неопределённых (N/A) генотипов составило 1 375 937, или 11,3% на образец. Таким образом, для каждой особи в среднем были восстановлены генотипы 88,7% локусов из ~12 млн SNP, представленных в панели фазированных гаплотипов.

Проведение анализа главных компонент выявил наличие четырёх чётких генетических кластеров форели. Выборки форели из одного селекционного центра, хотя и поставленные в разное время, обра-

зуют единые кластеры для форели как французской, так и испанской селекции (рис. 6). Следует отметить включение выборки, маркированной FR-R (6 особей), в кластер испанской форели, что свидетельствует о попадании испанского малька в бассейны с мальком из Франции (бассейны находятся рядом, и такая «контаминация» рыбой из соседних бассейнов – частое явление на рыбозаводных фермах). Выборки форели из Центра Геномной селекции ВНИРО представлены двумя кластерами – один состоит из особей, полученных из икры датского происхождения (выборка VNIRO20), и данный кластер ближе к кластерам французской и испанской форели. Второй кластер представлен полученным в ЦГС ВНИРО вторым поколением форели, имеющим происхождение из селекционного центра Aquagen (Норвегия) и предназначенной для выращивания в условиях морских садковых хозяйств (выборка VNIRO8). Как было ранее нами отмечено по результатам полногеномного секвенирования с высоким (20×) покрытием, форель морской селекции значительно отличается от всех линий форели, предназначенной для выращивания в пресной воде.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В ходе исследования проведена апробация метода low-pass секвенирования для полногеномного генотипирования радужной форели. Полученные данные свидетельствуют о высоком качестве данных, полученных в ходе импутации генотипов по созданной нами референсной базе фазированных гаплотипов, восстановлено более 87% локусов для каждой особи. Немногочисленные генетические образцы, для которых было получено наименьшее покрытие (0,27× – 0,47×), и которые были в ходе предварительного анализа признаны неудовлетворительными, также подверглись импутации, и полученной генетической информации оказалось достаточно чтобы они уверенно кластеризовались со своей популяционной группой.

Технология low-pass секвенирования позиционируется на сегодняшний день как бюджетная и более эффективная замена генотипированию на ДНК-чипах и активно внедряется как рутинная процедура генотипирования в селекции крупного рогатого скота, в свиноводстве и птицеводстве, а также в селекции рыб [Liu et al., 2024]. В нашей работе показана высокая эффективность применения данного вида генетического анализа в геномной селекции радужной форели. Учитывая, что в ценах 2025 года стоимость набора AgriHigh Low-pass WGS Package 384RXN составляет около 800 тыс. рублей (2083 руб. на образец), а также отсутствие на российском рынке ДНК-чипов Axion

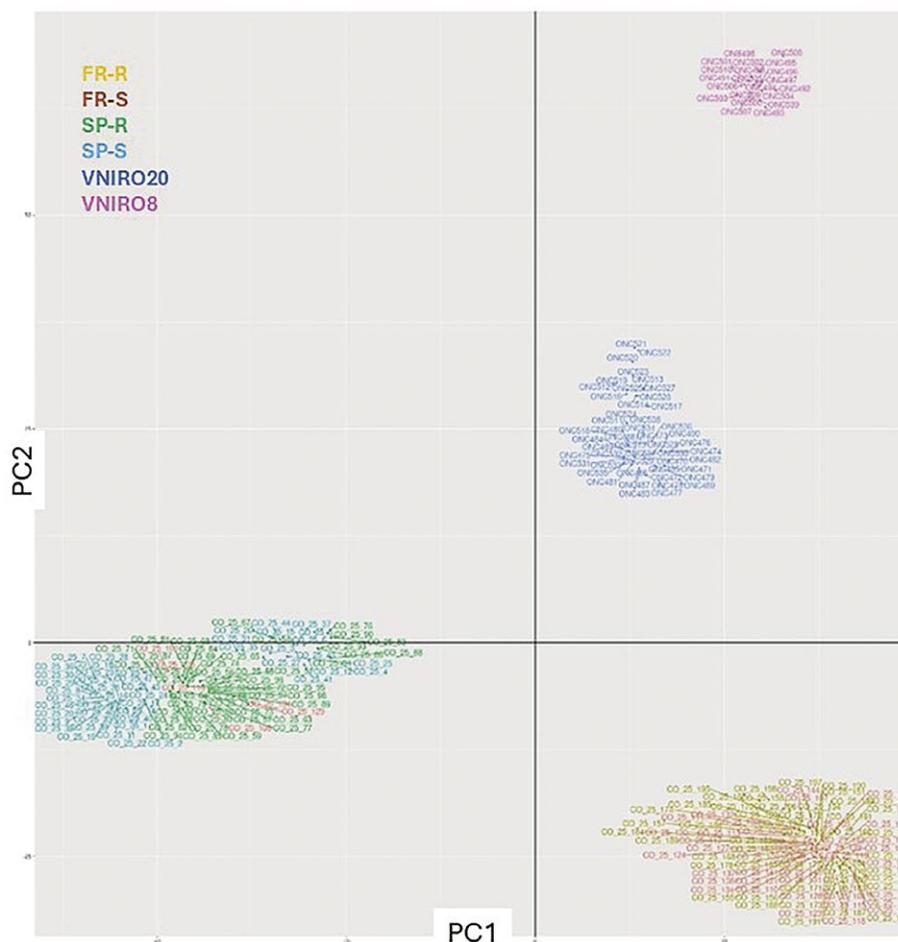


Рис. 6. Кластеризация изученных выборок форели в пространстве первой и второй главных компонент
Fig. 6. Clustering of the studied trout samples in the space of the first and second principal components

57К для форели (стоимость на европейском рынке – около 50 евро за образец), следует признать технологию low-pass секвенирования эффективной заменой генотипированию на ДНК-чипах для целей генотипирования рыб, проведения GWAS анализа и оценки селекционной ценности производителей при формировании селекционного ядра в Центре геномной селекции ФГБНУ «ВНИРО».

Благодарности

Авторы выражают искреннюю благодарность за содействие в проделанной работе коллективу Биотехнологического кампуса за проведение секвенирования, а также компании Хеликон и Московскому представительству компании BGI за предоставление демонстрационного набора Agrihigh.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм

Все применимые этические нормы соблюдены.

Финансирование

Работа выполнена при поддержке гранта Министерства науки и высшего образования Российской Федерации «Геномная селекция как инструмент интенсификации создания новых отечественных пород и линий лососевых рыб для товарной аквакультуры», соглашение № 075-15-2025-479 от 30 мая 2025 г.

ЛИТЕРАТУРА/ REFERENCES

- Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O. et al. 2021. Twelve years of SAMtools and BCFtools // GigaScience. 10 (2). giab008. DOI: 10.1093/gigascience/giab008.
- Fraslin C., Phocas F., Bestin A., Charles M., Bernard M., Krieg F. et al. 2020. Genetic determinism of spontaneous masculinisation in XX female rainbow trout: new insights using medium throughput genotyping and whole-genome sequencing // Sci Rep. 10(1): 17693. DOI: 10.1038/s41598-020-74757-8.

- Hofmeister R.J., Ribeiro D.M., Rubinacci S., Delaneau O. 2023. Accurate rare variant phasing of whole-genome and whole-exome sequencing data in the UK Biobank // *Nature Genetics*. 55(7). DOI: 10.1038/s41588-023-01415-w
- Ivanova N.V., Dewaard J.R., Hebert P.D.N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA // *Molecular Ecology Notes* 6(4): 998-1002. DOI: 10.1111/j.1471-8286.2006.01428.x
- Langmead B., Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2 // *Nature Methods*. 9(4): 357-359. DOI: 10.1038/nmeth.1923
- Liu S., Martin K.E., Snelling W.M., Long R., Leeds T.D., Vallejo R.L., Wiens G.D., Palti Y. 2024. Accurate genotype imputation from low-coverage whole-genome sequencing data of rainbow trout // *G3 Genes|Genomes|Genetics* 14(9).
- Lou R.N., Jacobs A., Wilder A.P., Therkildsen N.O. 2021. A beginner's guide to low-coverage whole genome sequencing for population genomics // *Molecular Ecology*. 30(23): 5966-5993. DOI: 10.1111/mec.16077.
- Palti Y., Gao G., Liu S., Kent M.P., Lien S., Miller M.R., Rexroad C.E., Moen T. 2015. The development and characterization of a 57k single nucleotide polymorphism array for rainbow trout // *Mol Ecol Resour*. 15(3):662-672. DOI: 10.1111/1755-0998.12337.
- Rubinacci S., Ribeiro D.M., Hofmeister R.J., Delaneau O. 2021. Efficient phasing and imputation of low-coverage sequencing data using large reference panels // *Nature Genetics*. 53(1):120-126. DOI: 10.1038/s41588-020-00756-0.
- Vallejo R.L., Leeds T.D., Gao G., Parsons J.E., Martin K.E., Evenhuis J.P., Fragomeni B.O., Wiens G.D., Palti Y. 2017. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture // *Genet. Sel. Evol.* 49(1):17. DOI: 10.1186/s12711-017-0293-6.

Поступила в редакцию 07.12.2025 г.
Принята после рецензий 26.12.2025 г.