Documentation of an effective method for undertaking molecular genomics in a mollusc: DNA extraction, microsatellite analysis, RAD-seq, and whole genome sequencing in the Horse mussel (*Modiolus modiolus*)

Van Wyngaarden, M. [1], McBride, M.C. [1], Cronmiller, E. [1], Jeffery, N.W. [1], Kess, T. [2], Bradbury, I.R. [2], Sameoto, J.A. [1], and Wringe, B.F [1].

[1]Bedford Institute of Oceanography
1 Challenger Drive
Dartmouth, Nova Scotia B2Y 4A2

[2]Northwest Atlantic Fisheries Centre
80 East White Hills Rd
St. John's NL A1C 5X1

2025

Canadian Technical Report of
Fisheries and Aquatic Sciences 3694

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report
of Fisheries and Aquatic Sciences 3694

2025

Documentation of an effective method for undertaking molecular genomics in a mollusc: DNA extraction, microsatellite analysis, RAD-seq, and whole genome sequencing in the Horse mussel (*Modiolus modiolus*)

by

Mallory Van Wyngaarden[1], Meghan McBride[1], Evan Cronmiller[1], Nick W. Jeffery[1], Tony Kess[2], Jessica Sameoto[1], and Brendan F. Wringe[1]

[1]Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive
Dartmouth, NS, B2Y 4A2

[2]Fisheries and Oceans Canada
Northwest Atlantic Fisheries Centre
80 East White Hills Road
St. John's, NL A1C 5X1

Cat. No. Fs 97-6/3694E ISBN: 978-0-660-77093-2 ISSN 1488-5379

https://doi.org/10.60825/y31a-zy54

Correct citation for this publication:

Van Wyngaarden, M., McBride, M.C., Cronmiller, E., Jeffery, N.W., Kess, T., Bradbury, I.R., Sameoto, J.A., and Wringe, B.F. 2025. Documentation of an effective method for undertaking molecular genomics in a mollusc: DNA extraction, microsatellite analysis, RAD-seq, and whole genome sequencing in the Horse mussel (*Modiolus modiolus*). Can. Tech. Rep. Fish. Aquat. Sci. 3694: viii + 97 p. https://doi.org/10.60825/y31a-zy54

TABLE OF CONTENTS

LIST OF TABLES V
LIST OF FIGURES VI
ABSTRACT VII
RÉSUMÉ VIII
INTRODUCTION 1
MATERIALS AND METHODS 3
SAMPLE COLLECTION 3
DNA EXTRACTION 3
DNA QUANTIFICATION 4
DNA QUALITY SCREENING 4
RAD LIBRARY PREPARATION 4
Adapter Preparation 4
DNA Dilution 5
Restriction Digest 5
Adapter Ligation 5
Sample Pooling and Purification 5
DNA Shearing 6
DNA Fragment Size Verification 6
Restriction Enzyme Fragmentation and Isolation of DNA Fragments 6
Purification of RAD-tag pooled DNA samples 7
Quantification of RAD-tag pooled DNA samples 7
Final Illumina Library Prep 7
WHOLE GENOME ASSEMBLY USING LONG AND SHORT READS 7
Oxford Nanopore Long-Read Sequencing 7
Illumina Short-Read Sequencing 9
Genome Assembly 9
MICROSATELLITE ANALYSIS 9
RAD-SEQ BIOINFORMATIC PIPELINE 10
RESULTS 10
DNA EXTRACTION 10
RAD LIBRARY PREPARATION 10
GENOME ASSEMBLY, RAD LIBRARIES, AND SNP CALLING 10
MICROSATELLITE ANALYSIS AND ALLELE CALLING 11
DISCUSSION 12
ACKNOWLEDGEMENTS 15
REFERENCES 17

TABLES 22
FIGURES 26
APPENDICES 34
APPENDIX 1 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR DNA EXTRACTION AND RAD-SEQ (RADCAP), WITH MODIFICATIONS FOR HORSE MUSSEL 34
Safety Considerations 34
Equipment 34
Supplies 35
Reagents and Solutions 36
Procedures 38
Appendix 1 References 60
Appendix 1 - Figures 61
APPENDIX 2 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR MINION NANOPORE LONG-READ SEQUENCING, WITH MODIFICATIONS FOR HORSE MUSSEL 67
Safety Considerations 67
Equipment 67
Supplies 67
Reagents 67
Procedures 68
Lambda Control Experiment 82
Appendix 2 References 88
APPENDIX 3 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR MICROSATELLITES, WITH MODIFICATIONS FOR HORSE MUSSEL 89
Safety Considerations 89
Equipment 89
Supplies 90
Reagents 91
Solutions 91
Software 91
Procedures 91
Appendix 3 References 97

LIST OF TABLES

Table 1 – Summary of horse mussel samples collected, extracted, and screened for DNA quality and concentration. 22

Table 2 – Distribution of samples between libraries. 23

Table 3 – RAD library quality metrics and labelling indexes. We used the standard dual index primer set from New England Biolabs, with Universal i5 primer and unique i7 primer.....23

Table 4 – Overview of the microsatellite PCR primers used, their performance (resultant product concentration) and the duplex panel in which they were run. Refer to Methods and Appendix 3 for PCR conditions. Note: size ranges may differ if additional populations are genotyped. 24

Table 5 – Locus-specific number of individuals (N), number of alleles (N_a), allelic richness (Ar), observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), and number of populations will null alleles for nine microsatellite loci amplified in 95 individual horse mussels.....25

LIST OF FIGURES

Figure 1 – Map of eight horse mussel collection locations in the Northwest Atlantic Ocean 26

Figure 2 – Proportion of different transposable element types found in the genome of the horse mussel. 27

Figure 3 - Repeat landscape plot for the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs. 28

Figure 4 – Repeat landscape plot for different types of transposable elements identified in the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs. 29

Figure 5 – Detailed repeat landscape plot for different types and sub-types of transposable elements identified in the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs. 30

Figure 6 - Frequency of missing RAD loci per sample (a) before and (b) after filtering for RAD loci present in seven populations and 75% of individuals per population 31

Figure 7 - Frequency of missing SNPs per sample after retaining one random SNP per RAD locus and filtering for a MAF of 0.01 32

Figure 8 Principal components analysis of genotypes for 472 horse mussels using (a) 8414 genome-wide SNPs from sequencing reads aligned to the horse mussel genome and (b) 3500 SNPs detected de novo from sequenced RAD tags, with one random SNP retained per RAD tag 33

ABSTRACT

Van Wyngaarden, M., McBride, M.C., Cronmiller, E., Jeffery, N.W., Kess, T., Bradbury, I.R., Sameoto, J.A., and Wringe, B.F. 2025. Documentation of an Effective Method for Undertaking Molecular Genomics in a Mollusc: DNA Extraction, Microsatellite Analysis, RAD-Seq, and Whole Genome Sequencing in the Horse mussel (*Modiolus modiolus*) Can. Tech. Rep. Fish. Aquat. Sci. 3694 viii + 97 p. https://doi.org/10.60825/y31a-zy54

Molluscs, one of the most diverse and ancient phyla, are globally significant as sources of protein from fisheries and aquaculture. Many species play crucial roles in maintaining ecosystem functions and act as keystone species by engineering biogenic habitats. Despite their ecological and economic importance, molecular genomic studies on molluscs have been limited due to challenges in obtaining quality DNA samples, primarily caused by the co-purification of polyphenolic proteins and mucopolysaccharides.

In this report, we document a series of effective methodologies for conducting molecular genomics in the Horse mussel (*Modiolus modiolus*), a bivalve mollusc known for forming complex benthic structures that enhance biodiversity and provide essential ecosystem services. We provide comprehensive laboratory protocols detailing tissue collection, DNA extraction, purification, and quantification, as well as microsatellite PCR and analysis, and the development and sequencing of restriction site-associated DNA (RAD-seq).

Furthermore, we present methodologies and results for *de novo* whole genome sequencing using both Illumina short-read and benchtop Nanopore long-read technologies. Unmodified manufacturer instructions or standard laboratory procedures are included in full to ensure methodological completeness and facilitate reproducibility We include relevant quality metrics and results for each molecular method. Although these methods may not be universally applicable to all mollusc species, they offer a validated starting point for similar genomic studies.

Our findings contribute valuable techniques and insights for advancing molluscan molecular genomic research, thereby facilitating further ecological and evolutionary studies in this critical phylum.

RÉSUMÉ

Van Wyngaarden, M., McBride, M.C., Cronmiller, E., Jeffery, N.W., Kess, T., Bradbury, I.R., Sameoto, J.A., and Wringe, B.F. 2025. Documentation of an Effective Method for Undertaking Molecular Genomics in a Mollusc: DNA Extraction, Microsatellite Analysis, RAD-Seq, and Whole Genome Sequencing in the Horse mussel (*Modiolus modiolus*) Can. Tech. Rep. Fish. Aquat. Sci. 3694 viii + 97 p. https://doi.org/10.60825/y31a-zy54

Les mollusques comptent parmi les phylums les plus anciens et les plus diversifiés, avec une importance mondiale en tant que sources de protéines issues de la pêche et de l'aquaculture. Plusieurs espèces remplissent également des fonctions écologiques clés, notamment en tant qu'ingénieurs d'habitats biogènes. Toutefois, la recherche en génomique moléculaire chez les mollusques demeure limitée, en grande partie en raison des défis liés à l'obtention d'ADN de haute qualité, notamment à cause de la copurification de composés comme les protéines polyphénoliques et les mucopolysaccharides.

Dans cette étude, nous présentons une série de méthodes efficaces pour mener des analyses génomiques sur *Modiolus modiolus*, un bivalve formant des structures benthiques complexes favorisant la biodiversité et fournissant des services écosystémiques essentiels. Nous détaillons des protocoles complets, de la collecte de tissus à l'extraction, la purification et la quantification de l'ADN, en passant par la PCR, l'analyse des microsatellites et le séquençage de l'ADN associé aux sites de restriction.

Nous incluons également les approches utilisées pour le séquençage *de novo* du génome entier à l'aide des technologies Illumina (courte lecture) et Nanopore (longue lecture), en intégrant les protocoles du fabricant et les procédures standard pour en assurer la reproductibilité. Bien que ces méthodes ne soient pas universellement applicables à toutes les espèces, elles constituent une base solide pour faire progresser les études en génomique moléculaire des mollusques.

INTRODUCTION

Molluscs are globally distributed and occupy nearly every habitat on Earth, including the deepest parts of the oceans to intertidal areas to land. They are also among the oldest (middle of the Cambrian, approx. 500-550 million years ago [Ma]) and most speciose (estimated > 70,000 species) phyla (Ponder and Lindberg 2008; Rosenberg 2014). In aquatic environments, molluscs are used as pollution monitoring tools and to study climate change, ocean acidification, and coastal ecosystem resilience to climate change. While some mollusc species are highly invasive and cause ecosystem-level impacts, others play vital roles in sustaining ecosystem function and even some are keystone species that undertake ecosystem engineering via the production of biogenic habitats. Moreover, they are of global importance as sources of protein from fisheries and aquaculture. In 2023, over 119,000 tonnes of molluscs were landed by commercial fisheries, while an additional 38,000 tonnes were produced by aquaculture (DFO 2025; Statistics Canada 2024).

Despite their importance, only recently there have there been a few large-scale, molecular genomic studies on molluscs, although, as with most other phyla, those numbers appear to be growing. In a review of molluscan-optimized molecular methods, Adema (2021) notes that GenBank contained whole genome sequences for only 84 species (see also Gomes-dos-Santos et al. 2020; Sigwart et al. 2021), yet less than five years later, that number has increased to 370. It is also worth noting that because of publication bias, efforts to undertake whole genome sequencing and/or molecular genomic studies in many more species may have been unsuccessfully attempted. Among the issues contributing to the apparent paucity of molecular genomic studies in molluscs in relation to their ubiquity and importance to humans is phylum-wide difficulties in obtaining suitable DNA samples (Sigwart et al. 2021). Common DNA extraction methods – such as those used for mammals and other vertebrates – can result in the co-purification of polyphenolic proteins and mucopolysaccharides (Adema 2021; Arseneau et al. 2017). Mucopolysaccharides are known to complex with DNA, whereas the phenolic proteins can inhibit DNA-interactive enzymes (e.g., those used during PCR).

Given the abundance and biodiversity of these phyla, it may be unfair to characterise molluscs as being recalcitrant to DNA extraction, when the argument could be made that extraction from vertebrates is facile. It may seem peculiar that the same two types of compounds are responsible for causing issues with molecular biology across an entire phylum. However, despite their biodiversity, molluscs share a number of conserved biological features, including biomineralization and production of copious amounts of mucous. Polyphenolic proteins are suspected to be involved in molluscan biomineralization, while mucopolysaccharides are key constituents of various adhesive structures, such as adhesive gels, byssal threads and mucous, which are common throughout the phylum (Adema 2021). Polyphenols and polysaccharides are also present in other invertebrates, plants (Do and Adams 1991) and fungi (e.g., Huang et al. 2018), and their co-purification in these taxa causes the same issues as observed in molluscs (Adema 2021; Arseneau et al. 2017; Demeke and Adams 1992; Huang et al. 2018). As a result of the difficulties with mollusc DNA extraction, in-house methods and commercial extraction kits (e.g., Omega Bio-tek E.Z.N.A.® Mollusc & Insect DNA Kit) – many of which contain the detergent hexadecyltrimethylammonium bromide (CTAB), which forms complexes with polysaccharides – have been developed for taxa that contain mucopolysaccharides, such as plants and fungi, as well as kits specific to mollusca.

Even though methods exist to reduce the issues with mollusc DNA extraction, the removal of polysaccharides also causes DNA fragmentation, adversely affecting genome sequencing and downstream genome assembly which rely on high-molecular-weight DNA (Sigwart et al. 2021). These challenges are further exacerbated by high repeat content (average 40% from Table 2 of Gomes-dos-Santos et al. 2020), high heterozygosity, and the prevalence of highly divergent haplotypes in mollusc genomes (Gerdol et al. 2020; Gundappa et al. 2022; Peñaloza et al. 2021; Takeuchi et al. 2012; Wang et al. 2019). Each of these factors independently contributes to increased assembly discontinuity and misassembly. For these reasons, a sequencing method that creates long-read data, including long-range genomic information (i.e., proximity location; Hi-C or Dovetail), is strongly suggested as part of a mollusc genome sequencing initiative (Sigwart et al. 2021). Methods that do not rely on high-molecular-weight DNA, such as microsatellites or genotyping by sequencing (e.g. RAD-seq), are less likely to be affected by this fragmentation, but they may be affected by the presence of inhibitors.

While CTAB and other methods are effective in removing polyphenolic proteins and mucopolysaccharides, some evidence suggests other “sticky proteins and biochemical compounds” co-purify with molluscan DNA. While it is unclear if these other compounds inhibit molecular techniques, they may interact with DNA in other, potentially enzymatic, ways. For instance, many researchers dilute DNA extracts 10- to 50-fold prior to PCR, and efficacy of PCR may decline with time from the initial DNA extraction in mollusc samples suggesting that some type of inhibitor(s) or enzymatic reaction(s) remain after purification (Adema 2021). Moreover, molluscan pigments also tend to co-purify with DNA (e.g., Schultzhaus et al. 2019); these are often visible in the extracted DNA pellet (Adema 2021). These pigments can interfere with quantification of DNA concentration and have been shown to lead to up to three-fold overestimation (Adema 2021). The effect appears to be greater for spectrophotometric quantification rather than fluorometry quantification, which relies on fluorescence signal of specific marker dyes that bind to DNA (Adema 2021).

Horse mussel biology

The horse mussel (*Modiolus modiolus*), a broadcast-spawning marine bivalve, forms complex benthic structures that enhance biodiversity and provide crucial ecosystem services (Holt et al. 1998; Kent et al. 2017; Ragnarsson and Burgos 2012; Rees et al. 2008; Sanderson et al. 2008; Shaw et al. 2014; Wildish et al. 1998) That are found across the North Atlantic and Pacific Oceans, where their aggregations, including in the Bay of Fundy, support increased species diversity (Cooper et al. 2019; Gormley et al. 2015; Mackenzie et al. 2022; Wilson et al. 2021). The Bay is part of the Scotian Shelf–Bay of Fundy bioregion and is under consideration for conservation prioritization by the Canadian government, with horse mussel aggregations identified as Ecologically and Biologically Significant Areas (EBSAs) (Buzeta 2014). These EBSAs are threatened by fishing activities, particularly bottom-contact gear (Cook et al. 2013; Kenchington et al. 2007; Kenchington et al. 2006). Understanding the genetic connectivity and structure of these populations could inform conservation strategies for this species and the bioregion.

Despite the use of specifically designed DNA extraction methods to prevent the co-purification of polyphenolic proteins and polysaccharides, issues like DNA fragmentation, may not always occur during molecular genomic methods in molluscs (Daniels et al. 2023; Hartung 2022; Law et al. 2025; Nam et al. 2021; Razkin et al. 2017; Razkin et al. 2016). In fact, commercial kits without CTAB have previously been shown to be effective for the extraction of horse mussel DNA for microsatellite PCR (Gormley et al. 2015; Mackenzie et al. 2018; Mackenzie et al. 2022) and for mitochondrial CO1 gene sequencing (Halanych et al. 2013). For whole genome sequencing and sequencing-based methods, DNA yield, purity and fragment size/size

distribution should be prioritized (Schultzhaus et al. 2019). However, as described above, this is often difficult for molluscs.

Here we describe and provide laboratory procedures using commercially available reagents and kits that we have found to work well in the horse mussel, a species of mollusc, to undertake commonly used molecular genomic methods. The methods include tissue collection for the purposes of DNA extraction and quantification, microsatellite PCR and analysis, and the development and sequencing of restriction site-associated DNA (RAD) libraries. Unmodified manufacturer instructions or standard laboratory procedures are also included in full to facilitate reproducibility and ensure all methods are readily available in one place. For each of these methods, we document and show relevant quality metrics and results. We also provide methodology and describe the results of a *de novo* Illumina short-read and Nanopore long-read-based whole genome sequencing effort that resulted in a genome approximately 60% of the expected size and with a BUSCO (Benchmarking sets of Universal Single-Copy Orthologs) score of 20%. For the RAD-seq and *de novo* whole genome sequencing, we describe and provide links to relevant bioinformatic pipelines.

MATERIALS AND METHODS

SAMPLE COLLECTION

We collected horse mussel samples from seven locations separated by a total of about 1500 km in Atlantic Canada (Table 1; Figure 1). Sampling sites ranged from the Bay of Fundy to the eastern coast of the Avalon Peninsula in Newfoundland (NL). Samples within (sites SI, Bio, GM, BI) and proximate to the Bay of Fundy off Southwest Nova Scotia (NS; 29W) were collected in 2022 using a modified scallop dredge that is 0.61 m wide. This miracle dredge consists of a single gang of scallop miracle gear with a steel frame and a bag made of 82 mm diameter rings, lined with a 27 mm polypropylene mesh. For more information on this sampling method and gear, please refer to Sameoto et al. (2021). Samples from the two remaining locations were collected by scuba diving in 2023 (ESI, coastal NS; IC, Avalon Peninsula, NL).

We transported live mussels to the lab in coolers on ice packs. We used a scallop knife to separate the abductor muscle from the shell walls, and the mussels were opened to expose the viscera. We removed tissue samples (approximately 3 g) from the adductor muscle using dissecting scissors, which were rinsed and flame-sterilized between samples. We placed the extracted tissue in 2 mL centrifuge tubes containing 1.6 mL of 95% ethanol, which were then held at -20 °C until DNA extraction.

DNA EXTRACTION

We extracted DNA in a 96-well plate format using the DNeasy 96 Blood & Tissue kit (Qiagen) and the QIAcube HT instrument (Qiagen). Briefly, we dissected approximately 20 mg of adductor tissue for each sample into labelled Qiagen collection microtubes filled with 200 µL ATL buffer. We incubated these samples for 30 minutes at room temperature to rehydrate the tissue. Once the incubation was complete, we removed the ATL buffer and added 200 µL of a 9:1 ATL/Proteinase K mixture, then added one 3 mm tungsten bead to each tube. We disrupted the tissue using the Qiagen TissueLyser II (10 s on each side of the machine), then incubated the samples for one hour at 56 °C. After one hour, we removed the samples from the incubator, shook them vigorously to break up the tissue, centrifuged the tubes, and then placed the samples back in the incubator at 56 °C to incubate overnight.

The following day, we used a multichannel pipette to transfer the digested tissue samples from the tubes into a Qiagen S-block. We loaded the S-block and all required DNeasy kit reagents and plasticware onto the QIAcube HT deck, then ran the automatic DNeasy Blood and Tissue Extraction program.

A total of 602 horse mussel samples were extracted, with a goal of producing 10 libraries of 48 individuals each for a total of 480. The full laboratory procedures are available in Appendix 1.

DNA QUANTIFICATION

After extraction, we determined the DNA concentration of each sample using the fluorescence-based Quant-IT PicoGreen dsDNA Assay Kit (Invitrogen). First, we prepared a λ DNA standard curve ranging from 0.05 to 2 ng/μL plus a blank. We added 50 μL of each of the DNA standards (2, 1.5, 1.0, 0.5 and 0.05 ng/μL) to five separate wells in a black 96-well half area plate and 50 μL of 1X TE buffer to another well for the blank. We then added 50 μL of a 1:100 PicoGreen working solution to the standards and blank. We mixed the plate via MixMate for 30 seconds at 900 rpm, then incubated the plate in the dark for 5 minutes. We quantified the standard curve and applied it to all genomic DNA (gDNA) samples per the manufacturer's protocol. We mixed the extracted DNA sample plates via MixMate for 30 seconds at 1200 rpm, then centrifuged the plates for 1 minutes at 1500 rpm. We prepared the DNA quantification plates by adding 49 μL of 1X TE buffer and 1 μL of gDNA to each well of a black 96-well half-area plate, using one well per sample. We then added 50 μL of the 1:100 PicoGreen working solution to each well, and mixed and incubated the plate following the methods used for the standards plate. Once incubation was complete, we ran the plate on the microplate reader following the manufacturer's protocol.

DNA QUALITY SCREENING

We screened gDNA samples with concentrations ≥8 ng/μL for DNA quality using Genomic DNA ScreenTapes and Reagents on the TapeStation 4200 (Agilent), following the manufacturer's protocol. We combined 1 μL of each gDNA sample with 10 μL of the Genomic DNA Sample Buffer in a 96-well plate, sealed the plate with foil, vortexed the plate at 2000 rpm for 1 minute, centrifuged the plate 1500 rpm for 1 minute, and then loaded the plate into the TapeStation. Once the run was completed, DNA samples with a clearly identified tight band with minimal smearing were considered suitable (e.g., 'good') for inclusion into RAD-Seq libraries. The TapeStation provides a DNA Integrity Number (DIN) for each sample, and a DIN ≥ 6.5 usually correlates with the visual identification of good quality DNA samples.

RAD LIBRARY PREPARATION

We prepared RAD libraries following the method of Ali et al. (2016), with some modifications (see Appendix 1 for more information).

A total of 545 DNA samples met the initial concentration criteria for inclusion in RAD libraries. Out of these, we selected 480 samples based on DNA quality (often DIN ≥ 6.9), ensuring balanced representation from each sample site. To avoid biases per library, we randomly assigned these samples to one of 10 RAD libraries, each containing 48 samples.

Adapter Preparation

From Ali et al. 2016 appendices, we selected the first 48 (B1-B48) complementary adapters (i.e., the matching forward (F) and reverse (R) RAD-Cap primer barcodes) and then diluted each adapter to 50 μM from the 100 μM stock using EB Buffer (Qiagen). We prepared 0.1 μM

dilutions of the complementary adapters in 1.5 mL Lobind tubes using the following recipe: 100 µL 10X Buffer AB, 1 µL of each RAD-Cap primer barcode (F and R), and 898 µL Ultrapure water (ThermoFisher). Next, we boiled 1 L of reverse osmosis (RO) water in a 2 L glass beaker using a hot plate, then turned off the hot plate. Using a floating tube holder, we floated each of the tubes containing the 0.1 µM combined RAD-CAP primer barcode dilutions in the beaker of hot water for approximately 6 h, until the water had cooled to room temperature. We dried off the tubes, mixed them for 15 s on a vortexer, then centrifuged them briefly to ensure all liquid was at the bottom of the tube. We made a working plate of the 48 prepared adapters, adding 40 µL of each adapter to a 96-well Lobind PCR plate. We sealed the working plate with PCR adhesive film and stored both the working plate and the prepared 0.1 µM primer dilution tubes at -20 °C.

DNA Dilution

Using an epMotion 5075 liquid handling instrument (Eppendorf), we normalized DNA concentration to 10 ng/µL (libraries 1 through 9) or 8 ng/µL (libraries 9 and 10). DNA concentration was lower in libraries 9 and 10 because we did not have enough samples with a concentration ≥10 ng/µL to create all 10 libraries.

Restriction Digest

For each library, we individually digested 48 samples using the *SbfI* enzyme (New England Biolabs; see Appendix 1 for preparation) in 5 mL of NEBuffer 4 (New England Biolabs). In an Eppendorf freezer block, we added 5 µL of *SbfI* enzyme cocktail to each well of a DNA LoBind plate followed by 20 µL DNA. This was then vortexed on the MixMate at 1650 rpm for 30 seconds before being centrifuged, along with the freezer plate, at 1500 rpm for 1 minute.

We placed the plate into a preheated Thermomixer C with a heated lid and 96-well plate block, and ran the following program: 37 °C for 1 hour; 80 °C for 20 minutes to inactivate the enzyme; and then held at 21 °C.

Adapter Ligation

We performed ligation immediately after the plate was removed from the Thermomixer C. In the freezer block, we added 2 µL of ligation master mix (see Appendix 1) to the required wells of an Eppendorf LoBind 96-well PCR plate, then added 12 µL of the *Sbf*I digested DNA samples followed by 2 µL of annealed *Sbf*I biotinylated RAD barcode adapters (0.1 µM), ensuring that each sample received the proper barcode. We vortexed the plate on the MixMate at 1650 rpm for 30 seconds before centrifuging it at 1500 rpm for 1 minute. We placed the plate into a preheated Thermomixer C with a heated lid and 96-well plate block, and ran the following program: 20 °C for 1 hour; 65 °C for 20 minutes to inactivate the enzyme; and then hold at 21 °C. Following ligation, we vortexed and centrifuged the plate again. For each library, we pooled 5 µL of each sample into separately labelled LoBind 1.5 mL tubes, and stored the remainder in the 96-well plate at -20 °C.

Sample Pooling and Purification

To purify the libraries, we calculated the volume of the pooled samples (5 µL × 48 samples = 240 µL) and added an equal volume of Agencourt AMPure XP beads (1:1 ratio) to the tube with the pooled samples. We mixed the samples and beads via pipette and incubated them for 5 minutes at room temperature to bind the DNA to the beads. After incubation, to separate the beads with the bound DNA from the solution, we transferred the tubes to a magnetic rack for 2-5 minutes until the solution cleared. Leaving the tube on the magnet, we carefully removed the

supernatant by pipette, making sure not to disturb the pellet. Continuing to leave the tube on the magnet, we washed the pellet by adding 250 µL of fresh 80% absolute ethanol, incubating for 30 seconds, and removing the ethanol via pipette. We repeated this step for a total of two ethanol washes.

We then removed the tube from the magnet, added 205 µL of Low TE buffer to release the DNA from the beads, and mixed by pipetting 10 times. We incubated the tube at room temperature for two minutes, then placed the tube on the magnetic rack. When the solution cleared (after approximately one minute), we transferred the supernatant to a new, labelled 0.5 mL microcentrifuge tube.

DNA Shearing

We sheared the DNA using a BioRuptor, which had been pre-cooled and filled with a water-ice slurry. From each library, we transferred 100 µL aliquots to two 0.5 mL BioRuptor tubes (A & B). To reduce variability in fragment size ranges within libraries, one tube (A) was sonicated for 10 minutes at 30 seconds ON/15 seconds OFF and the other tube (B) was sonicated for 10 minutes at 30 seconds ON/30 seconds OFF. After sonication, we combined both tubes (A & B) containing the sonicated products for each library in a labelled 1.5 mL LoBind tube.

DNA Fragment Size Verification

We verified the DNA fragment sizes in each library were concentrated between 300 and 500 bp using the TapeStation 4200 with the HS D1000 ScreenTape and Reagents, following the manufacturer's protocol.

Restriction Enzyme Fragmentation and Isolation of DNA Fragments

We added 30 µL of Dynabeads M-280 Streptavidin to a 1.5 mL LoBind tube and placed the tube on a magnet rack. Once the solution was clear, we removed the supernatant using a pipette. We washed the Dynabeads by removing the tube from the magnet, adding 100 µL of the 2X binding and wash buffer, mixing by pipette, returning the tube to the magnet until the solution was clear, and then removing the supernatant. We repeated this process for a total of two washes. Next, we removed the tube from the magnet, resuspended the washed beads in 200 µL of 2X binding and wash buffer, and added 200 µL of the sheared DNA library. We incubated the bead/DNA mixture for 20 minutes at room temperature, occasionally mixing gently.

After incubation, we placed the tube on the magnet until the solution was clear, then removed the supernatant using a pipette. We washed the bead/DNA mixture by removing the tube from the magnet, adding 150 µL room temperature 1X binding and wash buffer, mixing by pipette, returning the tube to the magnet until the solution was clear, and then removing the supernatant. We repeated this process for a total of two washes with the room temperature 1X binding and wash buffer. We then repeated the same wash steps using the 56 °C 1X binding and wash buffer, for a total of two washes with the 56 °C 1X binding and wash buffer. Finally, we repeated the same was steps using 100 µL 1X NEBuffer 4, again for a total of two washes with the 1X NEBuffer 4.

With all eight wash steps completed (two for the beads, six for the bead/DNA mixture) and the supernatant removed from the tube, we resuspended the bead/DNA mixture in 40 µL of *Sbf*I solution and incubated the tube at 37 °C for 1 hour to cut the DNA from the beads at the *Sbf*I cut site. Once incubation was complete, we placed the tube on the magnet rack for 1 minute until the solution was clear and transferred the supernatant to a new, labelled 1.5 mL LoBind tube.

Purification of RAD-tag pooled DNA samples

To each 1.5 mL LoBind tube containing a library of RAD-tag pooled samples, we added 40 µL AMPure XP beads (1:1 ratio), mixed via pipette, and incubated the tube at room temperature for 5 minutes. We placed the tube on the magnet rack for about 2 minutes until the solution cleared and removed the supernatant using a pipette.

Keeping the tubes on the magnet rack, we added 150 µL of fresh 80% ethanol, incubated for 30 seconds at room temperature, and removed the supernatant. We repeated this for a total of two ethanol washes, keeping the tube on the rack and ensuring the bead pellet was not disturbed.

Finally, we removed the tubes from the magnet rack, added 60 µL of low TE to release the purified library of RAD-tag pooled samples from the beads, and mixed via pipette. We incubated the tube for 2 minutes at room temperature and returned the tube to the magnet until the solution cleared. Once clear, we transferred the supernatant to a new, labelled 1.5 mL LoBind tube.

Quantification of RAD-tag pooled DNA samples

We quantified 5 µL from each library using the Qubit dsDNA HS Assay, following the kit manufacturer's instructions.

Final Illumina Library Prep

We completed final preparation of the libraries for sequencing on an Illumina platform following the manufacturer's protocol, including NEBNext end preparation, adaptor ligation, size selection, and cleanup of adaptor-ligated DNA, PCR enrichment of adaptor-ligated DNA, and final cleanup. Final libraries were visualized using TapeStation 4200 using the HS D1000 ScreenTapes and reagents and then quantified via qPCR using the QuantStudio 7 Flex instrument (ThermoFisher) with NEBNext Library Quant Kit for Illumina (New England Biolabs), both following the manufacturer's protocols. Libraries were sequenced at Génome Québec (Montreal, Canada) using 150 bp paired-end sequencing on the Illumina NovaSeq platform.

For all procedures outlined in this section, detailed Aquatic Biotechnology Laboratory standard operating procedures are available in Appendix 1.

WHOLE GENOME ASSEMBLY USING LONG AND SHORT READS

Oxford Nanopore Long-Read Sequencing

To generate a horse mussel reference genome, we performed long-read sequencing using the MinION Mk1B nanopore sequencer (Oxford Nanopore), which requires high-molecular-weight DNA. We selected DNA samples with high extraction concentrations (≥8 ng/µL) and used a TapeStation 4200 with Genomic DNA ScreenTapes and Reagents (Agilent) to determine the length of the DNA fragments in the extraction. We sequenced a sample that balanced high DNA concentration (72.75 ng/µL) with high molecular weight (average peak size around 30 kb) using the Ligation Sequencing Kit V14 (Oxford Nanopore) with the optional DNA Control Sample addition and an R10.4.1 flow cell. Full laboratory procedures are available in Appendix 2.

DNA Repair and End Prep

Mixing by pipette between each addition, we added 47 µL of DNA sample (~1 µg), 1 µL DNA Control Sample, 3.5 µL of the NEBNext FFPE DNA Repair Buffer, 2 µL of the NEBNext FFPE DNA Repair Mix, 3.5 µL of the Ultra II End-prep reaction Buffer, and 3 µL of the Ultra II End-prep Enzyme Mix to a 0.2 mL thin-walled PCR tube, for a total of 60 µL. We incubated the mix

at 20 °C for 5 minutes followed by 65 °C for 5 minutes, then transferred the mix to a 1.5 mL tube. To purify the DNA, we added 60 µL AMPure XP Beads and mixed by flicking the tube. We then incubated the tube on a rotating mixer at room temperature for 5 minutes at 10 rpm. Following incubation, we placed the tube on a magnet stand for up to 5 minutes to separate the beads from the supernatant. Once the solution was clear, we removed and discarded the supernatant. We washed the bead pellet twice while leaving the tube on the magnet, adding 200 µL fresh 80% ethanol, and immediately removing and discarding the ethanol without disturbing the beads. Following the washes, we removed the tube from the magnet, spun it down briefly in a benchtop mini-centrifuge, and placed it back on the magnet. We removed residual ethanol by pipette and left the tube open to air-dry for 30 seconds. Once dry, we removed the tube from the magnet and resuspended the bead pellet in 61 µL of nuclease-free water. We incubated the tube for 2 minutes at room temperature, then placed the tube back on the magnet for up to 5 minutes to separate the beads from the supernatant. Once the solution was clear, we transferred the supernatant to a clean 1.5 mL tube and discarded the beads.

Adapter Ligation and Clean-up

To a new 1.5 mL tube, we added 60 µL of the prepared DNA, 25 µL Ligation Buffer, 10 µL NEBNext Quick T4 DNA Ligase, and 5 µL Ligation Adapter, mixing by pipette between each addition, for a total of 100 µL. We incubated the tube at room temperature for 10 minutes, then added 40 µL resuspended AMPure XP Beads and mixed by flicking the tube. We incubated the tube on a rotating mixer at room temperature for 5 minutes at 10 RPM. Following incubation, the tube was placed on a magnet stand for up to 5 minutes to separate the beads from the supernatant. Once the solution was clear, we removed and discarded the supernatant. We washed the bead pellet twice by adding 250 µL Short Fragment Buffer, flicking the tube to resuspend the beads, briefly spinning the tube down in a benchtop mini-centrifuge, and returning the tube to the magnet for up to 5 minutes to separate the beads from the supernatant, which was then removed and discarded. Following the second wash, the tube was spun down briefly using a bench top mini-centrifuge, then placed back on the magnet. We removed any residual buffer and left the tub open to air-dry for 30 seconds. Once dry, we removed the tube from the magnet and resuspended the pellet in 25 µL of Elution Buffer. We incubated the tube for 10 minutes at 37 °C, then placed the tube on the magnet for 10 minutes. Once the solution was clear, we transferred 25 µL of the supernatant to a new 1.5 mL tube and discarded the beads.

Quantify and Dilute DNA Library

Following the adapter ligation and clean-up, we checked DNA quantity using the fluorescence-based Qubit dsDNA HS Assay, following the manufacturer's protocol. We split the library into three 1.5 mL tubes with 75 ng, 37.5 ng, and 37.5 ng of DNA (Library 1A, 1B, and 1C, respectively). We then added Elution Buffer to bring the total volume to 12 µL per library.

Run Library on MinION

We mixed 1170 µL of Flow Cell Flush and 30 µL of Flow Cell Tether in a 1.5 mL tube, then primed the MinION Flow Cell following manufacturer's directions. Next, we mixed 37.5 µL of Sequencing Buffer, 25.5 µL of Library Beads, and the 12 µL of prepared DNA Library 1A in a 1.5 mL tube. We completed the MinION Flow Cell priming, then loaded the prepared DNA library into the Flow Cell following manufacturer's directions. After approximately 24 hours of sequencing, we washed the flow cell using the Flow Cell Wash Kit, then loaded DNA Library 1B to continue sequencing for 48 hours. Following the sequencing of Library 1B, the available pores in the flow cell dopped below the usable amount. Consequently, so we prepared a new flow cell and loaded Library 1C, which was sequenced for an additional 24 hours.

Illumina Short-Read Sequencing

Genomic DNA from the same individual used for the MinION long-read sequencing was shipped to Genome Quebec (Montreal) for high-depth short-read sequencing on the Illumina NovaSeq platform. A PCR-free shotgun library was developed from the genomic DNA for paired-end 150 bp sequencing, which resulted in 463,332,211 reads with an average quality score of 39. These reads were then demultiplexed, followed by filtering data for quality and trimming Illumina adapters using fastp v0.24.0 (Chen et al. 2018).

Genome Assembly

We used the output from the Nanopore long-read sequencing and the Illumina high-depth short-read sequencing to assemble a horse mussel reference genome. We used Jellyfish (Marçais and Kingsford 2011) and GenomeScope (Vurture et al. 2017) to estimate the horse mussel genome size and then used several genome assembly programs to assemble a hybrid (i.e., a combination of long- and short-reads) genome and a genome based solely on the long-read sequences.

We initially attempted a long-read genome assembly using Canu (Koren et al. 2017). Canu uses long-read sequences in FASTA or FASTQ format from PacBio or Oxford Nanopore sequencing to create a long-read assembly, but requires a minimum of 10X genome coverage (and recommends between 30-60X coverage as the minimum coverage to attempt genome assembly). Our Nanopore data only provided approximately 8X genome coverage, and so Canu would not attempt assembly with this low coverage.

Instead, we used MaSuRCA v4.1.0 (Zimin et al. 2013) with the default parameters as a hybrid genome assembly option, combining the 8X Nanopore coverage and approximately 60X Illumina coverage to first create a long-read genome assembly scaffold, which was ‘polished’ using the short-read sequences. MaSuRCA combines de Bruijn graph assembly using a specific k-mer length, with the overlap-layout-consensus (OLC) method, which computes all pairwise overlaps based on sequence similarity among reads and then creates a consensus sequence from the aligned, overlapping sequences.

As an alternative hybrid assembly option, we then assembled a reference horse mussel genome using WENGAN v0.2 (Di Genova et al. 2021) on the Government of Canada’s General Purpose Science Cluster (GPSC). We compared the ‘A’ and ‘D’ assembly methods, which use Abyss2 (Jackman et al. 2017) and DiscovarDenovo (Weisenfeld et al. 2014) short read assemblers, respectively, and used the –ontraw flag denoting Nanopore reads with an N50 of 15-40 kb. We also specified a genome length, -g, of 2000 to reflect the Jellyfish-estimated genome size of 2 Gb.

Once the genome was assembled, we assessed genome completeness with BUSCO v5.5.0 (Manni et al. 2021) and used Earl Grey v5.0.0 to identify and annotate transposable elements (TE) (Baril et al. 2024).

MICROSATELLITE ANALYSIS

We selected 91 individuals processed in the RAD-seq libraries plus five individuals from a new population (Gulf) that did not meet the DNA quality criteria for RAD-seq to be analysed at 12 microsatellite loci. These loci were developed by Gormley et al. (2015) and Mackenzie et al. (2022) and had previously been analysed successfully in European horse mussels.

Using DNA from the initial extraction (see DNA Extraction above for details), we used an epMotion robot (Eppendorf) to normalize the samples to 10 ng/µL gDNA in 30 µL volume using 10 mM Tris (pH 8.0). Two of the 12 microsatellites failed to produce genotypes during initial

testing, so we ran eight loci within four duplex panels and two as single-plex Polymerase Chain Reactions (PCRs; Table 4). PCRs were conducted We used in 10 µL reactions and consisted of 5 µL Type-it Microsatellite PCR Master Mix (Qiagen), 1 µL RNase-free water, 2 µL primer mix, and 2 µL DNA (10 ng/µL). Cycling conditions consisted of an initial 5 minute denaturing step at 95 °C followed by 28 cycles of: 95 °C for 30 seconds, 60 °C for 90 seconds, 72 °C for 30 seconds and then a final extension at 60 °C for 30 minutes. We prepared the PCR products to be visualized on the fragment analyzer, ABI 3130xl, by adding 9.3 µL HiDi Formamide (ThermoFisher), 0.2 µL GeneScan500 LIZ, and 0.5 µL PCR product. The raw electrophoretograms were analyzed using GeneMapper6 (Applied Biosystems) and all genotypes were verified visually. Full laboratory protocols are available in Appendix 3.

RAD-SEQ BIOINFORMATIC PIPELINE

Full details of the RAD-seq bioinformatic analysis can be found in the GitHub repository ModiolusRADseq (Van Wyngaarden 2025; https://doi.org/10.5281/zenodo.14962721). In brief, we trimmed and demultiplexed sequences using cutadapt v4.7 (Martin 2011) and stacks v2.65 (Catchen et al. 2013; Rivera-Colón and Catchen 2021) with the –bestrad option. We aligned reads to the reference genome using bwa-mem v2.2.1 (Vasimuddin et al. 2019) and samtools v1.20 (Danecek et al. 2021). We used stacks to call and filter SNPs from the aligned reads, then used plink v1.9 (Chang et al. 2015) to filter the data based on Hardy-Weinberg Equilibrium and linkage disequilibrium.

RESULTS

DNA EXTRACTION

Of the 602 DNA extractions from horse mussel tissue, 560 (93%) had a DNA concentration ≥ 8 ng/µL and 466 of the 560 (83%) met or exceeded 10 ng/µL (Table 1).

Lower concentration DNA extracts were not equally spread among sample sites. In fact, Island Cove samples had markedly lower concentrations, with only 20 of 35 (57%) being ≥ 8 ng/µL (Table 1). The size of tissue samples used for DNA extractions were consistent across samples (and thus between sites), so the lower DNA concentration in Island Cove extractions may be related to the collection and/or preservation methods, however we are unaware of differences that may be the cause.

Overall, 592 of the 602 (98%) of the DNA extracted were identified as suitable (‘good’) for RAD-Seq DNA quality based on a tight band and minimal smear which appeared to correspond for the most part with the DNA Integrity Number (DIN) ≥ 6.5 (Table 1). .

RAD LIBRARY PREPARATION

From the 560 DNA extracted samples that met both the quality and concentration RAD-seq library criteria, 480 were selected for RADseq library preparation, balancing samples across collection sites and ensuring high DNA quality (corresponding to samples with a DIN ≥ 6.9) (Table 2). All 10 final libraries met the Génome Québec quality and quantity requirements with an average fragment size range between 510-573 bp and a concentration range of 7.4-25.7 ng/µL.

GENOME ASSEMBLY, RAD LIBRARIES, AND SNP CALLING

A total of 906.89 million Illumina paired-end reads with a GC content of 34.97% and 6.35 million Nanopore long reads were generated from the single individual horse mussel selected for

genome assembly, amounting to approximately 60X short-read and 8X long-read coverage. The horse mussel genome size was estimated to be approximately 2 Gb using the 41 bp k-mer distribution in Jellyfish and GenomeScope.

The horse mussel genome that was assembled using MaSuRCA used both Illumina short-reads and Nanopore long-reads, however, this genome ended up only being a few Mb in length. The potential reason for this small and presumably fragmented assembly may be a result of the strict quality filtering used by MaSuRCA. In the end, we did not use this genome for any analysis.

The WENGAN A assembly method constructed a 1.2 Gb genome, representing approximately 60% of the estimated genome size of 2 Gb. This genome contained 67,119 contigs with an N50 of 24,915 bp.

Earl Grey found 35.2% of the assembled genome to be non-repeat (not TEs), while another 36.0% was designated as unclassified TEs. Within the remaining 28.8% of the assembled genome, 10.2% was Penelope elements, 8.1% was categorized as “Other (Simple Repeat, Microsatellite, RNA)”, 5.7% was LINEs, 3.0% was DNA, 1.5% was LTR, 0.2% was rolling circle elements, and 0.08% was SINEs (Figures 1,2). Of the different classifications, Penelope elements, DNA, rolling circle elements, and the unclassified elements showed evidence of younger TEs based on lower Kimura 2-Parameter (K2P) Distance values, which indicate that TEs in this classification have had less time to diverge from the source. In contrast, SINEs showed mostly older TEs (higher K2P Distance values) and LINEs and LTRs had bimodal distributions of both younger and older TEs (Figures 3,4).

Sequencing the 10 RAD-Seq libraries generated 2.8 billion paired-end reads, ranging from 228 to 364 million reads per library. After filtering and demultiplexing, the number of reads per individual ranged from 622,086 to 21,349,457, with an average of 92.1% of all available reads retained. The BUSCO score generated for the WENGAN genome using the metazoa and mollusca lineage orthologs was approximately 20% complete, yet 72% of the filtered RAD-Seq library reads mapped to this genome assembly with proper pairing.

From the initial SNP catalog of 147 million generated by gstacks, 9,893 SNPs and 472 of 480 individuals were retained after MAF and missing data filters. Plots of the frequency of missing RAD loci per sample before and after filtering show a clear sample site-specific pattern (Figure 6). However, each RAD-Seq library consisted of randomly assigned individuals across sample sites therefore this pattern is likely result of tissue preservation quality rather than a sequencing artefact. Once SNPs were called from loci then filtered for missing data and MAF, using the populations module, this pattern was no longer present (Figure 7). Of the SNPs that passed MAF and missing data filters, 449 were removed by the HWE filter and 1030 were removed due to linkage, thus leading to a final dataset of 8,414 SNPs and 472 individuals.

MICROSATELLITE ANALYSIS AND ALLELE CALLING

Despite all 12 microsatellite loci working on Scottish and Irish horse mussels (Gormley et al. 2015; Mackenzie et al. 2018; Mackenzie et al. 2022), three of them failed to amplify in our samples (Table 4). The remaining nine amplified well in all populations and provided interpretable genotypes, therefore, we suspect the potential reason why three loci failed may be a result of something specific to the interaction (or lack thereof) between these primers and the North American horse mussel genome rather than any co-purified polysaccharides or polyphenolic proteins. This is further reinforced by the fact that the DNA extracts had worked well for RAD-seq which would have been negatively affected by the presence of polysaccharides or polyphenolic proteins. Similarly, successful microsatellite genotyping of five individuals that did not met the RAD-Seq DNA quality criteria (i.e. DNA too fragmented) also lends credence to this conclusion.

All nine loci showed evidence of null alleles in at least one population, with two loci (MM_pp17 and MM_pp24) showing evidence of null alleles in more than four populations. The number of alleles per locus ranged from five (MM13) to 13 (MM_pp17 and MM_pp24) and the allelic richness (adjusted to the lowest sample size of four individuals per population) ranged from 2.66 (MM13) to 4.96 (MM_pp17). The observed heterozygosity (H_o) was lower than expected heterozygosity (H_e) for all loci except MM13. Average H_o across all loci was 0.49 while average H_e across all loci was 0.60. Loci MM_pp17 and MM_pp24 had the greatest difference between observed and expected heterozygosity; these were also the two loci that showed strong evidence of null alleles in more than four populations (Table 5).

DISCUSSION

The effectiveness of molecular genomic techniques depends on the extraction of adequate quantities of high-quality DNA with appropriate fragment size and purity. This can be challenging in molluscs because standard DNA extraction techniques, typically employed for mammals and other vertebrates, often lead to the co-purification of polyphenolic proteins and mucopolysaccharides (Adema 2021; Arseneau et al. 2017). Mucopolysaccharides can form complexes with DNA, while phenolic proteins have the potential to inhibit DNA-interactive enzymes, such as those utilized in PCR.

To surmount this issue, DNA extraction methods that incorporate the detergent hexadecyltrimethylammonium bromide (CTAB), which forms complexes with polysaccharides, are often used (Kess et al. 2015; Sigwart et al. 2021). These can be found within in-house methods or commercial extraction kits specific to Mollusca (Arseneau et al. 2017). This is often combined with additional wash steps as well as DNA extract dilution as further means of reducing the impacts of co-purified inhibitors (Adema 2021). Although removal of polysaccharides is often recommended, doing so causes DNA fragmentation, which affects certain downstream methods, such as genome assembly (Sigwart et al. 2021).

However, the removal of mucopolysaccharides during extraction does not appear to be universally required as there are many existing successful molecular genomic projects that did not remove mucopolysaccharides during extraction (Daniels et al. 2023; Gormley et al. 2015; Halanych et al. 2013; Hartung 2022; Law et al. 2025; Mackenzie et al. 2018; Mackenzie et al. 2022; Nam et al. 2021; Razkin et al. 2017; Razkin et al. 2016). Although, commercial DNA extraction kits specific to mollusc species has been shown to result in DNA of suitable size, quantity, and quality for *de novo* whole genome sequencing (Law et al. 2025; Wang et al. 2019).

Interestingly, while genomic work in molluscs often requires taxa-specific kits and modified laboratory procedures (e.g. Arseneau et al. 2017), we were able to successfully extract DNA, amplify and score microsatellites and undertake reduced recombination sequencing in horse mussel using generally standard laboratory procedures. We used a phylum-agnostic extraction method that uses commercial DNA extraction kits which do not contain CTAB, and a robot to extract high-molecular-weight DNA from horse mussels. DNA extractions were performed on a total of 602 samples with a target DNA concentration of ≥ 8 ng/μL and high DNA quality with a tight band with minimal smearing. Of these 602 samples, 560 met these targets which allowed us to increase the stringency on DNA quality for the final sample selection of 480 for the 10 RAD-Seq libraries.

The success of microsatellite analysis, RAD-seq, and *de novo* whole genome sequencing, as discussed below, demonstrates the efficacy of the DNA extraction methodology. The overall effectiveness of the extraction is evident when considering the combined results, while its suitability for each method is reflected in the degree of its individual successes. This suggests that preliminary testing with less expensive, broadly available genomic kits and standard

procedures may be conducted for other mollusc species to assess their suitability before proceeding with more specialized, mollusc-specific kits and modified protocols.

Whole genome sequencing in molluscs is particularly challenging. Although removal of polysaccharides during extraction is often recommended, doing so causes DNA fragmentation, which prevents optimal operation of sequencing technologies that rely on high-molecular-weight DNA and hinders downstream genome assembly (Sigwart et al. 2021). However, Sigwart et al. (2021) have suggested that, because of the high heterozygosity and repeat content in mollusc genomes, a sequencing method that creates long-read data including long-range genomic information (i.e. proximity location; Hi-C or Dovetail) is required.

We used a combination of 463.33 million Illumina paired-end (~60X) and 6.35 million Oxford Nanopore long-reads from a single individual to create a *de novo* whole genome for the horse mussel. After assembling with WENGEN, the genome comprised 67,119 contigs with an N50 of 24,915 bp, had a GC content of approximately 34.97%, and an assembly length of 1.2 Gb, which represents approximately 60% of the estimated genome size of 2 Gb. Using BUSCO, we found the assembly was approximately 20% complete against the Mollusca lineage orthologs.

Discrepancy between bivalve genome sizes obtained through flow cytometry and sequencing-k-mer-based estimates have been detected in a number of species (e.g. European Flat Oyster, *Ostrea edulis* (Gundappa et al. 2022)*;* Pacific Oyster, *Crassostrea gigas* (Peñaloza et al. 2021) and (Wang et al. 2019); Pearl Mussel, *Pinctada fucata* (Takeuchi et al. 2012); Mediterranean Mussel, *Mytilus galloprovincialis* (Gerdol et al. 2020)). This has been attributed to the affects of extremely large numbers of repeats (Gomes-dos-Santos et al. 2020), high heterozygosity, and a high frequency of highly divergent haplotypes (Gerdol et al. 2020; Gundappa et al. 2022; Peñaloza et al. 2021; Takeuchi et al. 2012; Wang et al. 2019). Each of these factors alone could lead to high fragmentation and misassembly. Additionally, population differences in genome size as a result of genome plasticity has been detected and can add further complication (Gerdol et al. 2020).

A whole genome of the Black-Shelled Pacific Oyster, which is a human-bred variety of Pacific Oyster (*C. gigas*), was developed using a similar method to our horse mussel genome. Wang et al. (2019) extracted DNA using a Qiagen Blood & Cell Culture DNA Mini Kit; then short read sequencing was accomplished with 150 bp paired-end, single-index on a MGISEQ-2000 platform (BGI, Shenzhen, China), which conducts BGI-seq. Long-read sequencing was done through two flow-cells of a GridION DNA sequencer (Oxford Nanopore, Oxford, UK) that produced 61.8 Gb of raw Nanopore reads, which is less than the theoretical 100 Gb. After quality control, 39.9 Gb remained or approximately 67X coverage based on a k-mer-derived genome size estimate of 594 Mb. The short-read genotyping produced 105.6 Gb of raw sequence, which was reduced to 104.9 Gb or approximately 177X coverage after filtering for low-quality reads.

Similarly, our MinION run produced fewer reads than the theoretical 100 Gb, yielding 6.35 Gb across two flow cells, corresponding to approximately 8X long-read coverage based on a k-mer-derived genome size estimate of 2 Gb. Short-read sequencing, using PE150 chemistry on an Illumina NovaSeq6000, generated 906.89 paired-end reads, providing approximately 60X short-read coverage.

The reasons for our assembled genome being markedly shorter than expected are unclear. Invertebrates are notoriously difficult to extract high-molecular weight DNA for a number of reasons, including the aforementioned polysaccharides and other compounds found on and in mollusc tissues that may disrupt DNA elution or amplification. In our case, a lack of adequate sequencing depth is the most likely culprit behind the undersized genome assembly, as we only managed to sequence approximately 8X coverage of the genome with the Nanopore MinION

using extracted DNA that had been shown to be suitable for reduced representation (RAD) and short read (Illumina) sequencing. As noted, many long-read assemblers will not generate usable results with less than 20X coverage, and greater than 30X is typically preferred. Genomes larger than 2 Gb in general also have a high proportion of repetitive DNA, which consists of defunct genes, simple sequence repeats, and transposable elements, all of which make genome assembly difficult without high sequencing depth and computational power. For example, the Hard Clam or Northern Quahog, *Mercenaria mercenaria,* has a comparable genome size of 1.86 Gb, yet only 2% of its genome consists of protein coding genes (Farhat et al. 2022). More than half of the *M. mercenaria* genome consists of repetitive elements, and its completeness was estimated at 76.4% using the Mollusca database with BUSCO (Farhat et al. 2022). This genome used a combination of technologies, including ~85X Illumina HiSeq PE150 coverage, ~20X PacBio Sequel I coverage, and additional PacBio Sequel II and Hi-C (Belton et al. 2012) sequencing, which provided high depth coverage of the genome and identification of chromosomes from the Hi-C contact maps. Furthermore, as previously discussed, discrepancies between estimated genome length and sequenced genome length are commonly observed in molluscs. Although we did not verify metrics for potential mis-assemblies, the majority of RAD-tags successfully mapped to our genome with proper pairing. This indicates that even incomplete genomes can be highly valuable for population genomic analyses.

Numerous examples demonstrate the successful application of RAD-seq and related reduced representation genotyping methods (e.g., Rapture; Ali et al. 2016) in molluscs (Nam et al. 2021; Sturch and D'Aloia 2023). In general, these methods rely on bacterial restriction enzymes to cut DNA at specific sequence-dependent locations within the genome, resulting in a reduced collection of DNA fragments that begin and end with known sequences. The known sequences can be used to ligate DNA barcodes, which provide individual identification, and sequences that allow for the binding of specific DNA primers. By relying on fragmented DNA and primers to amplify target DNA before sequencing, RAD-seq methods are less dependent on high-molecular-weight DNA and can efficiently increase the amount of target DNA relative to non-target DNA using PCR.

Like us, El Ayari et al. (2019) also used the QIAGEN DNeasy Blood & Tissue Kit to extract DNA from Mediterranean Mussel in order to test for fine-scale genetic structure using PCR product length polymorphisms in a small number of ancestry-informative loci identified from previous genome scans. Kess et al. (2018) utilized a CTAB-containing extraction method and *de novo* self-assembly of RAD-tags using STACKS to find SNPs in Littorina (*Littorina saxatilis*), whereas Kess et al. (2021) was able to align the same sequencing data to the *L. saxatilis* genome using the Burrows-Wheeler algorithm in bwa-mem as we did with the horse mussel. Although the filtering criteria differed between Kess et al. (2018) and Kess et al. (2021) (e.g., SNPs present in at least 70% of samples in Kess et al. (2018) versus 80% in Kess et al. (2021), the 2021 study identified more SNPs (11,504 compared to 4,256). Kess et al. (2021) had high genotyping rate with more than 96% of SNPs aligning to mapped locations in the genome. Additionally, the 2021 study leveraged genomic information to infer relationships between phenotypic divergence and putative structural variants, providing insight into the genomic architecture of intertidal adaptation in *Littorina saxatilis.*

We also aligned our RAD-tags to a genome using bwa-mem. While our genome appeared to be only 60% of the estimate size and had only 20% complete BUSCO, we still found 72% of the filtered RAD library reads mapped to this genome assembly with proper pairing and 8,414 SNPs and 472 individuals. We compared this set of SNPs to those we produced through *de novo* self-assembly of RAD-tags using STACKS. We found that aligning RAD-tags to the genome produced approximately three times as many SNPs as the *de novo* SNP calling. However, in spite of differences in the number of SNPs between the aligned and *de novo* datasets, the

results of population genomic analyses using both datasets were qualitatively the same (Figure 8).

Based on our ability to complete SNP-based population genomic analysis, the suitability of the DNA extraction and RAD-seq methods we used for molluscs is demonstrable.

The set of 12 microsatellite makers selected were designed for and found to work in Scottish and Irish horse mussels (Gormley et al. 2015; Mackenzie et al. 2018; Mackenzie et al. 2022). Despite producing PCR product in Scottish and Irish horse mussels, we found that three of the 12 microsatellites failed to amplify. The Scottish and Irish studies had evidence of null alleles across the preponderance of populations. Mackenzie et al. (2018) and Mackenzie et al. (2022) reported the presence of null alleles in all 12 microsatellites for each population. Gormley et al. (2015) detected null alleles in the majority of populations in four of the five microsatellites used. While allelic dropout can be caused by PCR inhibitors, it can also be the result of poor primer design or mutation(s) in the primer binding region(s).

While we do not have access to the Scottish and Irish samples to test, we have a strong suspicion that the null alleles are unrelated to co-purified polysaccharides, polyphenolic proteins, or other types of PCR inhibitors. Indeed, our results suggests primer design or mutation(s) within the primer binding region(s) are most likely the cause. We note that the majority of DNA sample extractions on which we conducted microsatellite analysis had previously worked well for RAD-seq, which has a higher DNA quality criteria than microsatellite methodologies. Additionally, microsatellite genotyping was successful for five individuals that did not meet the DNA quality criteria for RAD-seq also lending further credence to this conclusion.

We have described a method to extract DNA from horse mussel, a species of mollusc, which yielded DNA with sufficient quantity, quality, and purity for successful PCR microsatellite amplification, RAD-seq, and *de novo* whole genome sequencing using both Illumina and Oxford Nanopore technologies. This paper also details the laboratory, analytical and bioinformatic methods used. Comprehensive protocols for molecular biology techniques and the corresponding bioinformatic analysis are provided in the appendices.

We have also presented metrics that demonstrate the successful effectiveness of our DNA extraction, microsatellite and RAD-seq analyses, and whole genome sequencing.

Our success does not imply that this method will work equally well in all mollusc species. As such, we recommend that when working on a new species, researchers use a subset of their samples to compare multiple DNA extraction methods and evaluate the quality of the DNA extract as well as testing for indications of inhibition. DNA quantification, where co-purification of polysaccharides and polyphenolic proteins may occur, is best accomplished using fluorometry (e.g. Qubit) which relies on the specific binding of marker dyes to DNA. The quality of DNA can be checked using gel electrophoresis or a TapeStation (Agilent), which conducts electrophoresis within its proprietary ScreenTapes.

ACKNOWLEDGEMENTS

We thank the staff at Génome Québec for sequencing our RAD libraries and conducting high-depth Illumina sequencing for the horse mussel genome assembly. Past and present employees of the Aquatic Biotechnology Lab contributed significantly to the initial development of the laboratory protocols.

We also thank DFO Research Scientists Drs. Timothy Healy and Kyle Wellband for providing insightful comments during their review of this report that improved its quality and clarity.

Funding was provided by the Fisheries and Oceans Canada Marine Conservation Targets program.

REFERENCES

Adema, C.M. 2021. Sticky problems: extraction of nucleic acids from molluscs. Philos. Trans. R. Soc. London, Ser. B **376**(1825): 20200162. doi:doi:10.1098/rstb.2020.0162.

Ali, O.A., O'Rourke, S.M., Amish, S.J., Meek, M.H., Luikart, G., Jeffres, C., and Miller, M.R. 2016. RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping. Genetics **202**(2): 389-400. doi:10.1534/genetics.115.183665.

Arseneau, J.-R., Steeves, R., and Laflamme, M. 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Mol. Ecol. Res. **17**(4): 686-693. doi:https://doi.org/10.1111/1755-0998.12616.

Baril, T., Galbraith, J., and Hayward, A. 2024. Earl Grey: A Fully Automated User-Friendly Transposable Element Annotation and Analysis Pipeline. Mol. Biol. Evol. **41**(4). doi:10.1093/molbev/msae068.

Belton, J.-M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. 2012. Hi–C: A comprehensive technique to capture the conformation of genomes. Methods **58**(3): 268-276. doi:https://doi.org/10.1016/j.ymeth.2012.05.001.

Buzeta, M.-I. 2014. Identification and Review of Ecologically and Biologically Significant Areas in the Bay of Fundy. DFO. Can. Sci. Advis. Sec. Res. Doc. 2013/065. vi + 59 p.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. **22**(11): 3124-3140. doi:https://doi.org/10.1111/mec.12354.

Centre d'expertise et de services Génome Québec. 2020. *User Guide: Illumina sequencing technologies – DNA-Seq*. Available at: https://genomequebec.com/wp-content/uploads/2023/11/userGuide_CES_MPS_Illumina_DNASeq_7.0_En-1.pdf (Accessed: April 1, 2025).

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience **4**(1). doi:10.1186/s13742-015-0047-8.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics **34**(17): i884-i890. doi:10.1093/bioinformatics/bty560.

Cook, R., Fariñas-Franco, J.M., Gell, F.R., Holt, R.H.F., Holt, T., Lindenbaum, C., Porter, J.S., Seed, R., Skates, L.R., Stringell, T.B., and Sanderson, W.G. 2013. The Substantial First Impact of Bottom Fishing on Rare Biodiversity Hotspots: A Dilemma for Evidence-Based Conservation. PLoS One **8**(8). doi:ARTN e6990410.1371/journal.pone.0069904.

Cooper, J.A., Goodwin, C., Lawton, P., Brydges, T., Hiltz, C., Armsworthy, S., and McCurdy, Q. 2019. Characterisation of the sublittoral habitats of the Brier Island/Digby Neck Ecologically and Biologically Significant Area, Nova Scotia, Canada. Can. Tech. Rep. Fish. Aquat. Sci. 3327: xv +163 p.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. 2021. Twelve years of SAMtools and BCFtools. GigaScience **10**(2). doi:10.1093/gigascience/giab008.

Daniels, B.N., Nurge, J., Sleeper, O., Lee, A.Y., Lopez, C., Christie, M.R., Toonen, R.J., White, C., and Davidson, J.M. 2023. Genomic DNA extraction optimization and validation for genome sequencing using the marine gastropod Kellet's whelk. PeerJ **11**. doi:ARTN e1651010.7717/peerj.16510.

Demeke, T., and Adams, R.P. 1992. The effects of plant polysaccharides and buffer additives on PCR. BioTechniques **12**(3): 332-334.

DFO. 2025. *2023 Atlantic & Pacific coasts commercial landings by province*. Available at: https://www.dfo-mpo.gc.ca/stats/commercial/land-debarq/sea-maritimes/s2023pq-eng.htm (Accessed April 1, 2025).

Di Genova, A., Buena-Atienza, E., Ossowski, S., and Sagot, M.-F. 2021. Efficient hybrid de novo assembly of human genomes with WENGAN. Nat. Biotechnol. **39**(4): 422-430. doi:10.1038/s41587-020-00747-w.

Do, N., and Adams, R. 1991. A simple technique for removing plant polysaccharide contaminants from DNA. BioTechniques **10**(2): 162-166.

El Ayari, T., Trigui El Menif, N., Hamer, B., Cahill, A.E., and Bierne, N. 2019. The hidden side of a major marine biogeographic boundary: a wide mosaic hybrid zone at the Atlantic–Mediterranean divide reveals the complex interaction between natural and genetic barriers in mussels. Heredity **122**(6): 770-784. doi:10.1038/s41437-018-0174-y.

Farhat, S., Bonnivard, E., Pales Espinosa, E., Tanguy, A., Boutet, I., Guiglielmoni, N., Flot, J.-F., and Allam, B. 2022. Comparative analysis of the Mercenaria mercenaria genome provides insights into the diversity of transposable elements and immune molecules in bivalve mollusks. BMC Genomics **23**(1): 192. doi:10.1186/s12864-021-08262-1.

Gerdol, M., Moreira, R., Cruz, F., Gómez-Garrido, J., Vlasova, A., Rosani, U., Venier, P., Naranjo-Ortiz, M.A., Murgarella, M., Greco, S., Balseiro, P., Corvelo, A., Frias, L., Gut, M., Gabaldón, T., Pallavicini, A., Canchaya, C., Novoa, B., Alioto, T.S., Posada, D., and Figueras, A. 2020. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome Biol. **21**(1): 275. doi:10.1186/s13059-020-02180-3.

Gomes-dos-Santos, A., Lopes-Lima, M., Castro, L.F.C., and Froufe, E. 2020. Molluscan genomics: the road so far and the way forward. Hydrobiologia **847**(7): 1705-1726. doi:10.1007/s10750-019-04111-1.

Gormley, K., Mackenzie, C., Robins, P., Coscia, I., Cassidy, A., James, J., Hull, A., Piertney, S., Sanderson, W., and Porter, J. 2015. Connectivity and Dispersal Patterns of Protected Biogenic Reefs: Implications for the Conservation of *Modiolus modiolus* (L.) in the Irish Sea. PLoS One **10**(12): e0143337. doi:10.1371/journal.pone.0143337.

Gundappa, M.K., Peñaloza, C., Regan, T., Boutet, I., Tanguy, A., Houston, R.D., Bean, T.P., and Macqueen, D.J. 2022. Chromosome-level reference genome for European flat oyster (*Ostrea edulis* L.). Evol. Appl. **15**(11): 1713-1729. doi:10.1111/eva.13460.

Halanych, K.M., Vodoti, E.T., Sundberg, P., and Dahlgren, T.G. 2013. Phylogeography of the horse mussel *Modiolus modiolus*. J. Mar. Biol. Assoc. U.K. **93**(7): 1857-1869. doi:10.1017/S0025315413000404.

Hartung, H. 2022. *Development and Application of Genomic Scale Resources for the Analysis of Spisula Solidissima Operational Taxonomic Units*. MSc Thesis. Cornell University.

Holt, T.J., Rees, E.I., Hawkins, S.J., and Seed, R. 1998. Biogenic reefs (volume IX). An overview of dynamic and sensitivity characteristics for conservation management of marine SACs. Scottish Association for Marine Science (UK Marine SACs Project). 170 pages.

Huang, X., Duan, N., Xu, H., Xie, T.N., Xue, Y.R., and Liu, C.H. 2018. CTAB-PEG DNA Extraction from Fungi with High Contents of Polysaccharides. Mol. Biol. **52**(4): 621-628. doi:10.1134/S0026893318040088.

Jackman, S.D., Vandervalk, B.P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S.A., Jahesh, G., Khan, H., Coombe, L., Warren, R.L., and Birol, I. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. Genome Res. **27**(5): 768-777. doi:10.1101/gr.214346.116.

Kenchington, E.L., Kenchington, T.J., Henry, L.A., Fuller, S., and Gonzalez, P. 2007. Multi-decadal changes in the megabenthos of the Bay of Fundy: The effects of fishing. J. Sea Res. **58**(3): 220-240. doi:10.1016/j.seares.2007.04.001.

Kenchington, E.L.R., Gilkinson, K.D., MacIssaac, K.G., Bourbonnais-Boyce, C., Kenchington, T.J., Smith, S.J., and Gordon, D.C. 2006. Effects of experimental otter trawling on benthic assemblages on Western Bank, northwest Atlantic Ocean. J. Sea Res. **56**(3): 249-270. doi:10.1016/j.seares.2006.03.010.

Kent, F.E.A., Last, K.S., Harries, D.B., and Sanderson, W.G. 2017. In situ biodeposition measurements on a *Modiolus modiolus* (horse mussel) reef provide insights into ecosystem services. Estuar. Coast. Shelf Sci. **184**: 151-157. doi:https://doi.org/10.1016/j.ecss.2016.11.014.

Kess, T., Galindo, J., and Boulding, E.G. 2018. Genomic divergence between Spanish *Littorina saxatilis* ecotypes unravels limited admixture and extensive parallelism associated with population history. Ecol. Evol. **8**(16): 8311-8327. doi:https://doi.org/10.1002/ece3.4304.

Kess, T., Brachmann, M., and Boulding, E.G. 2021. Putative chromosomal rearrangements are associated primarily with ecotype divergence rather than geographic separation in an intertidal, poorly dispersing snail. J. Evol. Biol. **34**(1): 193-207. doi:https://doi.org/10.1111/jeb.13724.

Kess, T., Gross, J., Harper, F., and Boulding, E.G. 2015. Low-cost ddRAD method of SNP discovery and genotyping applied to the periwinkle *Littorina saxatilis*. J. Molluscan Stud. **82**(1): 104-109. doi:10.1093/mollus/eyv042.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. **27**(5): 722-736. doi:10.1101/gr.215087.116.

Law, S.T.S., Nong, W., Au, M.F.F., Cheung, L.H.T., Shum, C.W.Y., Lee, S.Y., Cheung, S.G., and Hui, J.H.L. 2025. Genomes of two indigenous clams *Anomalocardia flexuosa* (Linnaeus, 1767) and *Meretrix petechialis* (Lamarck, 1818). Sci. Data **12**(1): 409. doi:10.1038/s41597-025-04748-9.

Mackenzie, C.L., Kent, F.E.A., Baxter, J.M., and Porter, J.S. 2018. *Genetic analysis of horse mussel bed populations in Scotland*. Scottish Natural Heritage Research Report, no. 1000.

Mackenzie, C.L., Kent, F.E.A., Baxter, J.M., Gormley, K.S.G., Cassidy, A.J., Sanderson, W.G., and Porter, J.S. 2022. Genetic Connectivity and Diversity of a Protected, Habitat-Forming Species: Evidence Demonstrating the Need for Wider Environmental Protection and Integration of the Marine Protected Area Network. Front. Mar. Sci. **9**. doi:ARTN 77225910.3389/fmars.2022.772259.

Manni, M., Berkeley, M.R., Seppey, M., and Zdobnov, E.M. 2021. BUSCO: Assessing Genomic Data Quality and Beyond. Curr. Protocol. **1**(12): e323. doi:https://doi.org/10.1002/cpz1.323.

Marçais, G., and Kingsford, C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. Bioinformatics **27**(6): 764-770. doi:10.1093/bioinformatics/btr011.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 **17**(1): 3-10. doi:10.14806/ej.17.1.200.

Nam, B.-H., Kim, H., Seol, D., Kim, H., Noh, E.S., Kim, E.M., Noh, J.K., Kim, Y.-O., Park, J.Y., and Kwak, W. 2021. Genotyping-by-Sequencing of the regional Pacific abalone (*Haliotis discus*) genomes reveals population structures and patterns of gene flow. PLoS One **16**(4): e0247815. doi:10.1371/journal.pone.0247815.

Peñaloza, C., Gutierrez, A.P., Eöry, L., Wang, S., Guo, X., Archibald, A.L., Bean, T.P., and Houston, R.D. 2021. A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. GigaScience **10**(3). doi:10.1093/gigascience/giab020.

Ponder, W., and Lindberg, D.R. 2008. Phylogeny and Evolution of the Mollusca. Univ of California Press.

Ragnarsson, S.A., and Burgos, J.M. 2012. Separating the effects of a habitat modifier, *Modiolus modiolus* and substrate properties on the associated megafauna. J. Sea Res. **72**: 55-63. doi:https://doi.org/10.1016/j.seares.2012.05.011.

Razkin, O., Gómez-Moliner, B.J., Vardinoyannis, K., Martínez-Ortí, A., and Madeira, M.J. 2017. Species delimitation for cryptic species complexes: case study of *Pyramidula* (Gastropoda, Pulmonata). Zool. Scr. **46**(1): 55-72. doi:https://doi.org/10.1111/zsc.12192.

Razkin, O., Sonet, G., Breugelmans, K., Madeira, M.J., Gómez-Moliner, B.J., and Backeljau, T. 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail complex *Pyramidula*: The power of RADseq data. Mol. Phylogen. Evol. **101**: 267-278. doi:https://doi.org/10.1016/j.ympev.2016.05.002.

Rees, E.I.S., Sanderson, W.G., Mackie, A.S.Y., and Holt, R.H.F. 2008. Small-scale variation within a *Modiolus modiolus* (Mollusca: Bivalvia) reef in the Irish Sea. III. Crevice, sediment infauna and epifauna from targeted cores. J. Mar. Biol. Assoc. U.K. **88**(1): 151-156. doi:10.1017/S0025315408000052.

Rivera-Colón, A.G., and Catchen, J. 2021. Population genomics analysis with RAD, reprised: Stacks 2. bioRxiv: 2021.2011.2002.466953. doi:10.1101/2021.11.02.466953.

Rosenberg, G. 2014. A new critical estimate of named species-level diversity of the recent Mollusca. Am. Malacol. Bull. **32**(2): 308-322. doi:Doi 10.4003/006.032.0204.

Sameoto, J.A., Hall, K., Gass, S.E., Keith, D., Kirchhof, S., and Brown, C.J. 2021. Conservation implications of demographic changes in the horse mussel *Modiolus modiolus* population of the inner Bay of Fundy. Mar. Ecol. Prog. Ser. **670**: 93-104. doi:10.3354/meps13741.

Sanderson, W.G., Holt, R.H.F., Kay, L., Ramsay, K., Perrins, J., McMath, A.J., and Rees, E.I.S. 2008. Small-scale variation within a *Modiolus modiolus* (Mollusca: Bivalvia) reef in the Irish Sea. II. Epifauna recorded by divers and cameras. J. Mar. Biol. Assoc. U.K. **88**(1): 143-149. doi:10.1017/S0025315408000040.

Schultzhaus, J.N., Taitt, C.R., Orihuela, B., Smerchansky, M., Schultzhaus, Z.S., Rittschof, D., Wahl, K.J., and Spillmann, C.M. 2019. Comparison of seven methods for DNA extraction from prosomata of the acorn barnacle, *Amphibalanus amphitrite*. Anal. Biochem. **586**: 113441. doi:https://doi.org/10.1016/j.ab.2019.113441.

Shaw, J., Todd, B.J., and Li, M.Z. 2014. Geologic insights from multibeam bathymetry and seascape maps of the Bay of Fundy, Canada. Cont. Shelf Res. **83**: 53-63. doi:https://doi.org/10.1016/j.csr.2013.12.015.

Sigwart, J.D., Lindberg, D.R., Chen, C., and Sun, J. 2021. Molluscan phylogenomics requires strategically selected genomes. Philos. Trans. R. Soc. London, Ser. B **376**(1825). doi:ARTN 2020016110.1098/rstb.2020.0161.

Statistics Canada. 2024. Table 32-10-0107-01 Aquaculture, production and value [dataset]. DOI: https://doi.org/10.25318/3210010701-eng. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210010701

Sturch, W.H., and D'Aloia, C.C. 2023. Hierarchical genetic structure in a direct-developing whelk (*Buccinum undatum*) throughout the western North Atlantic. Can. J. Fish. Aquat. Sci. **80**(2): 261-272. doi:10.1139/cjfas-2022-0138.

Takeuchi, T., Kawashima, T., Koyanagi, R., Gyoja, F., Tanaka, M., Ikuta, T., Shoguchi, E., Fujiwara, M., Shinzato, C., Hisata, K., Fujie, M., Usami, T., Nagai, K., Maeyama, K., Okamoto, K., Aoki, H., Ishikawa, T., Masaoka, T., Fujiwara, A., Endo, K., Endo, H., Nagasawa, H., Kinoshita, S., Asakawa, S., Watabe, S., and Satoh, N. 2012. Draft Genome of the Pearl Oyster *Pinctada fucata*: A Platform for Understanding Bivalve Biology. DNA Res. **19**(2): 117-130. doi:10.1093/dnares/dss005.

Van Wyngaarden, M. 2025. “mallory-vw/ModiolusRADseq: Initial Publication. *Zenodo*. Available at: doi: 10.5281/zenodo.14962722. (Accessed April 5, 2025).

Vasimuddin, M., Misra, S., Li, H., and Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. *In* 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). 20-24 May 2019 2019. pp. 314-324.

Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., and Schatz, M.C. 2017. GenomeScope: fast reference-free genome profiling from short reads. Bioinformatics **33**(14): 2202-2204. doi:10.1093/bioinformatics/btx153.

Wang, X., Xu, W., Wei, L., Zhu, C., He, C., Song, H., Cai, Z., Yu, W., Jiang, Q., Li, L., Kun, W., and Feng, C. 2019. Nanopore Sequencing and De Novo Assembly of a Black-Shelled Pacific Oyster (*Crassostrea gigas*) Genome. Frontiers in Genetics **10**: 1211. doi:10.3389/fgene.2019.01211.

Weisenfeld, N.I., Yin, S., Sharpe, T., Lau, B., Hegarty, R., Holmes, L., Sogoloff, B., Tabbaa, D., Williams, L., Russ, C., Nusbaum, C., Lander, E.S., MacCallum, I., and Jaffe, D.B. 2014. Comprehensive variation discovery in single human genomes. Nat. Genet. **46**(12): 1350-1355. doi:10.1038/ng.3121.

Wildish, D.J., Fader, G.B.J., Lawton, P., and MacDonald, A.J. 1998. The acoustic detection and characteristics of sublittoral bivalve reefs in the Bay of Fundy. Cont. Shelf Res. **18**(1): 105-113. doi:Doi 10.1016/S0278-4343(98)80002-2.

Wilson, B.R., Brown, C.J., Sameoto, J.A., Lacharite, M., Redden, A.M., and Gazzola, V. 2021. Mapping seafloor habitats in the Bay of Fundy to assess megafaunal assemblages associated with *Modiolus modiolus* beds. Estuarine Coastal Shelf Sci. **252**: 107294. doi:ARTN 10729410.1016/j.ecss.2021.107294.

Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., and Yorke, J.A. 2013. The MaSuRCA genome assembler. Bioinformatics **29**(21): 2669-2677. doi:10.1093/bioinformatics/btt476.

TABLES

Table 1 – Summary of horse mussel samples collected, extracted, and screened for DNA quality and concentration.

Site Name	Population ID	No. Tissue Samples	No. Samples Extracted	No. Samples with [DNA] ≥ 8 ng/uL	No. Samples with suitable ('good') DNA quality	No. Samples Selected for RAD-seq[1]
29 West	29W	100	97	97	97	81
Briar Island	BI	95	95	94	86	68
Bioherm	Bio	100	97	87	97	86
Georges Bank	GB	6	6	5	6	0[2]
Grand Manan	GM	100	97	94	96	84
Spencer's Island	SI	100	97	96	97	74
Eastern Shore Islands	ESI	78	78	67	78	64
Island Cove	IC	35	35	20	35	18
Total		614	602	560	592	480

[1] Criteria for selection of samples used in RAD-seq libraries was (DIN ≥ 6.9, Conc. ≥ 8 ng/μL).
[2] These were excluded from the final set of libraries because of the small sample size.

Table 2 – Distribution of samples between libraries.

	Collection Site							
Library	IC	ESI	SI	Bio	GM	BI	29W	Total Per Library
1	0	3	11	5	11	7	11	48
2	1	6	5	9	12	5	10	48
3	3	6	5	10	7	10	7	48
4	2	6	11	9	7	7	6	48
5	0	1	11	12	12	3	9	48
6	1	4	9	7	5	11	11	48
7	1	6	10	6	11	6	8	48
8	2	1	9	10	6	11	9	48
9	3	18	1	11	6	3	6	48
10	5	13	2	11	7	5	5	48
Total Per Site	18	64	74	90	84	68	82	480

Table 3 – RAD library quality metrics and labelling indexes. We used the standard dual index primer set from New England Biolabs, with Universal i5 primer and unique i7 primer.

RAD-Seq Library	No. DNA Samples per library	NEBNext i7 Index	i7 Index Barcode Sequence	Average Size (bp)	Concentration (nM)
1	48	1	ATCACG	510	12.74
2	48	2	CGATGT	526	12.91
3	48	3	TTAGGC	507	25.68
4	48	4	TGACCA	532	14.10
5	48	5	ACAGTG	548	18.79
6	48	6	GCCAAT	524	18.91
7	48	7	CAGATC	573	10.29
8	48	8	ACTTGA	521	16.91
9	48	9	GATCAG	533	10.21
10	48	10	TAGCTT	514	7.38

Table 4 – Overview of the microsatellite PCR primers used, their performance (resultant product concentration) and the duplex panel in which they were run. Refer to Methods and Appendix 3 for PCR conditions. Note: size ranges may differ if additional populations are genotyped.

Locus	F Primer	Size range (bp)	Repeat Motif	Annealing Temp (°C)	Panel	Conc. (µM)	Dye	Result	Source
MM13	F-CACAGCCTCCTGGTCACAATA R-TGGCGTGTTATTCTAGCAAATG	150-200	GAA	57	1	2	6FAM	Success	Gormley et al., 2015
MM20	F-AATTGCTCACTTGGCGTAAAAC R-TGGAAATGGAGAGACAGATCCT	180-248	TCA	57	1	1	SUN	Success	Gormley et al., 2015
MM2	F-CTCCGCTATGTTTGACCATGTA R-TCCACACCGAGTAACAAATCAG	116-317	CA	57	2	4	6FAM	Failed processing	Gormley et al., 2015
MM30	F-CACACAAGACAGGCCAGATAGA F-GAAGAATCCCCACAAACACATT	147-183	CA	57	2	2	SUN	Success	Gormley et al., 2015
MM_pp19	F-GGTCGTTCCCTTTGACATGAACCC R-AAACATCTTTCGCACCCGTTTGCC	348-389	AT	60	3	2	6FAM	Success	Mackenzie et al., 2022
MM_pp15	F-TGAGGTTAGTGAAAACTAATTGAGCAACCC R-CGTTTCAGATTCTCCTTACAATTTGCC	357-369	ATT	60	3	-	SUN	Failed testing	Mackenzie et al., 2022
MM_pp27	F-TTTACTGAGTTCACACTGTTTTGCC R-GCATCATATGTTACCCGTTCCC	310-326	AT	60	4	2	6FAM	Removed - null alleles	Mackenzie et al., 2022
MM_pp07	F-TCCAGGTATTTTAGTTCCAGAGATAGGG R-GATTATTCATCTTGGAGCCATTGCC	304-308	CGG	60	4	1	SUN	Success	Mackenzie et al., 2022
MM_pp17	F-TCTTACAGATTCGGGATTGTGAACCG R-TCAACTTCAATCTTTTGGCCTTATCGG	235-260	AC	60	5	2	6FAM	Removed - null alleles	Mackenzie et al., 2022
MM_pp37	F-CCGTTGTGGATTTGTGAGAATACGC R-GCGACTTAGTTCCACGCTTTTATTACGG	227-272	AT	60	5	1	SUN	Success	Mackenzie et al., 2022
MM_pp24	F-TTTTCCTTTCTCTCTCCGCATTCGC R-TGCTACCAAGGTTGTAACGAGATTCCC	292-309	AT	60	6	2	6FAM	Success	Mackenzie et al., 2022
MM_pp05	F-ACACCAAATTTAGCCCCTTTAGGC R-AAAACCAAATGTTCCACCTAACCC	266-292	TCG	60	6	-	SUN	Failed testing	Mackenzie et al., 2022

Table 5 – Locus-specific number of individuals (N), number of alleles (N_a), allelic richness (Ar), observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), and number of populations will null alleles for nine microsatellite loci amplified in 95 individual horse mussels.

Locus	N	N_a	Ar	H_o	H_e	F_{IS}	Populations with Null Alleles
MM13	95	5	2.66	0.52	0.50	-0.04	1
MM20	89	12	3.70	0.61	0.67	0.15	3
MM30	95	12	2.99	0.52	0.53	0.05	3
MM_pp07	92	7	2.92	0.45	0.53	0.20	3
MM_pp17	85	13	4.96	0.39	0.84	0.55	8
MM_pp19	95	6	2.81	0.53	0.54	-0.02	2
MM_pp24	95	13	3.58	0.60	0.67	0.09	2
MM_pp27	93	9	2.75	0.25	0.46	0.45	7
MM_pp37	95	5	2.85	0.56	0.62	0.08	4

FIGURES

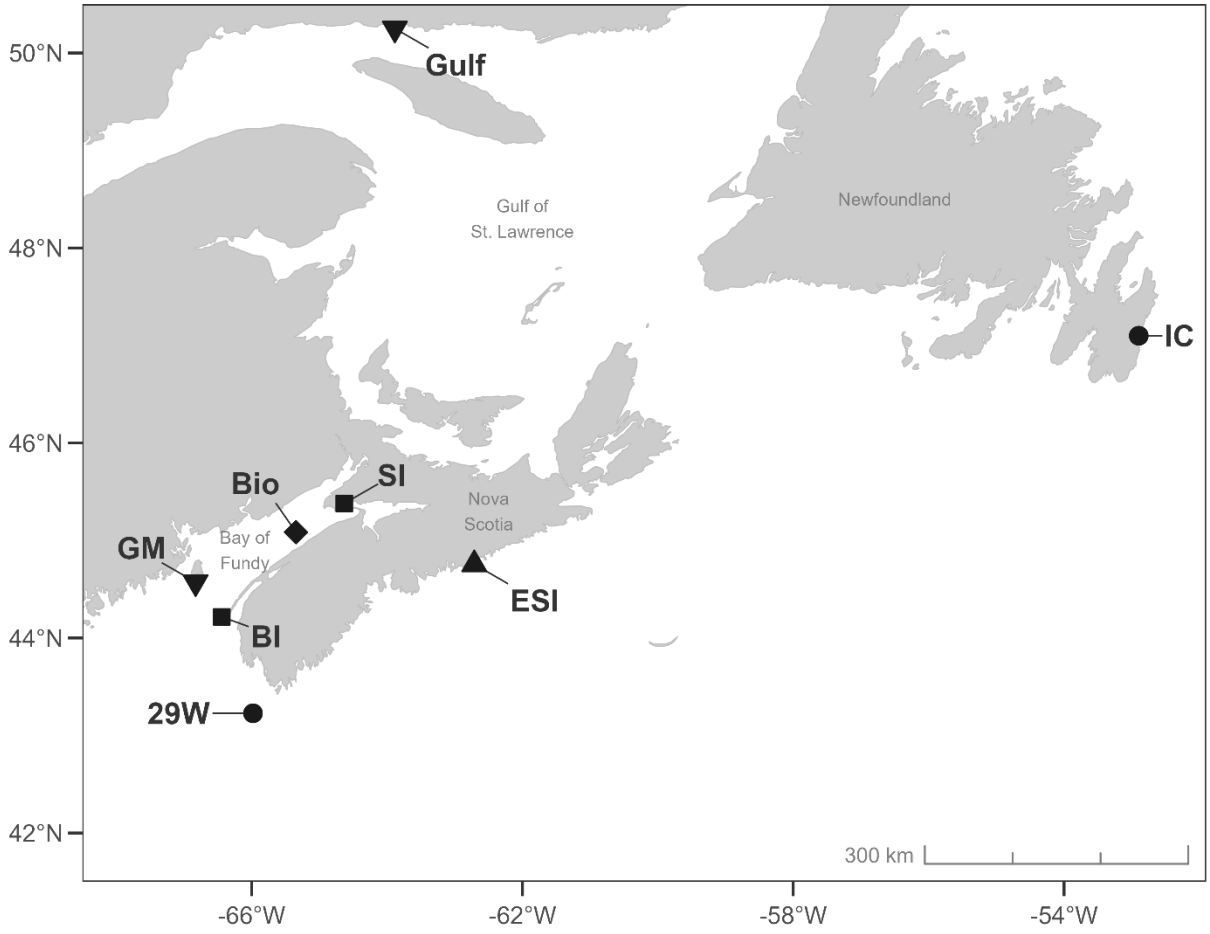


Figure 1 – Map of eight horse mussel collection locations in the Northwest Atlantic Ocean

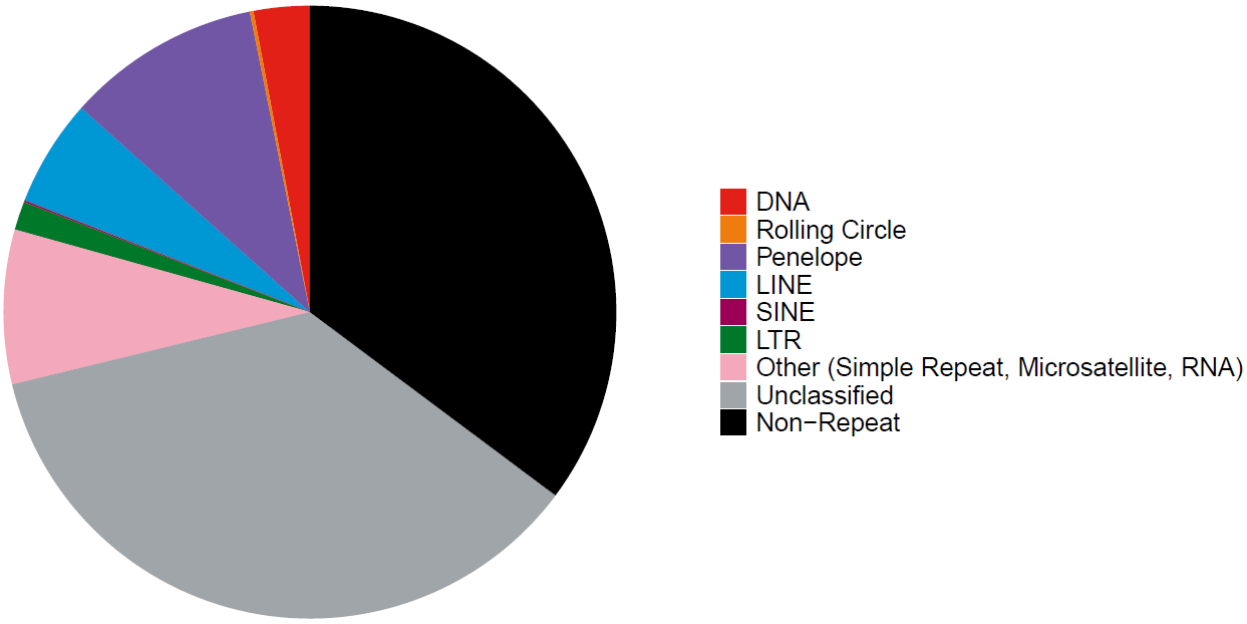


Figure 2 – Proportion of different transposable element types found in the genome of the horse mussel.

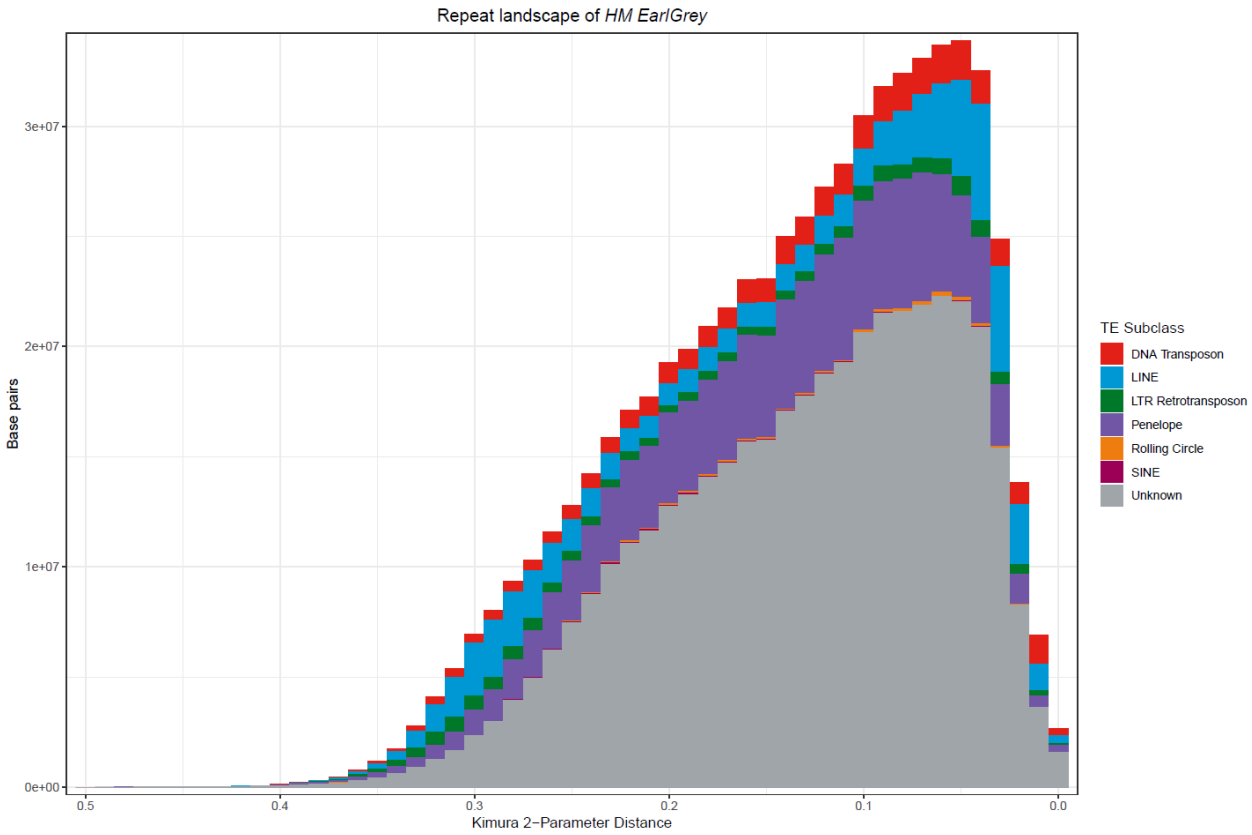


Figure 3 - Repeat landscape plot for the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs.

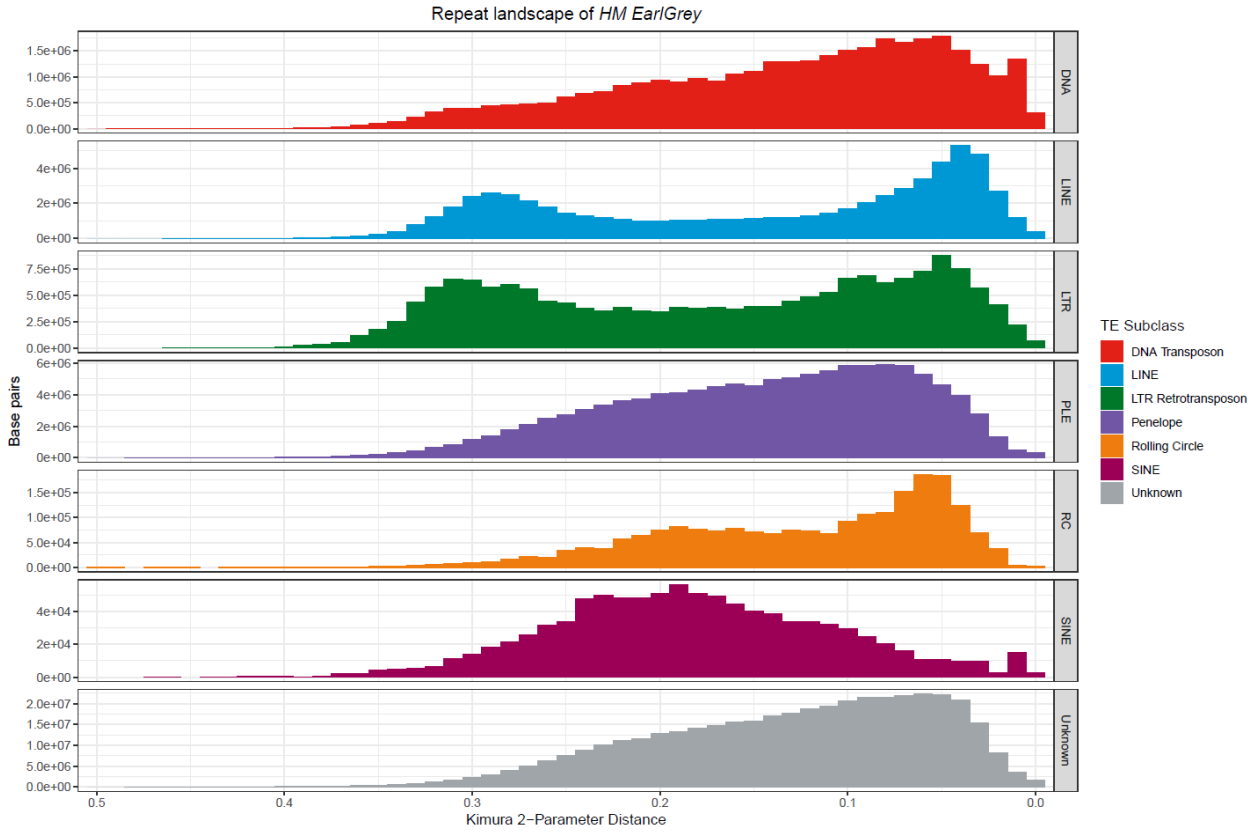


Figure 4 – Repeat landscape plot for different types of transposable elements identified in the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs.

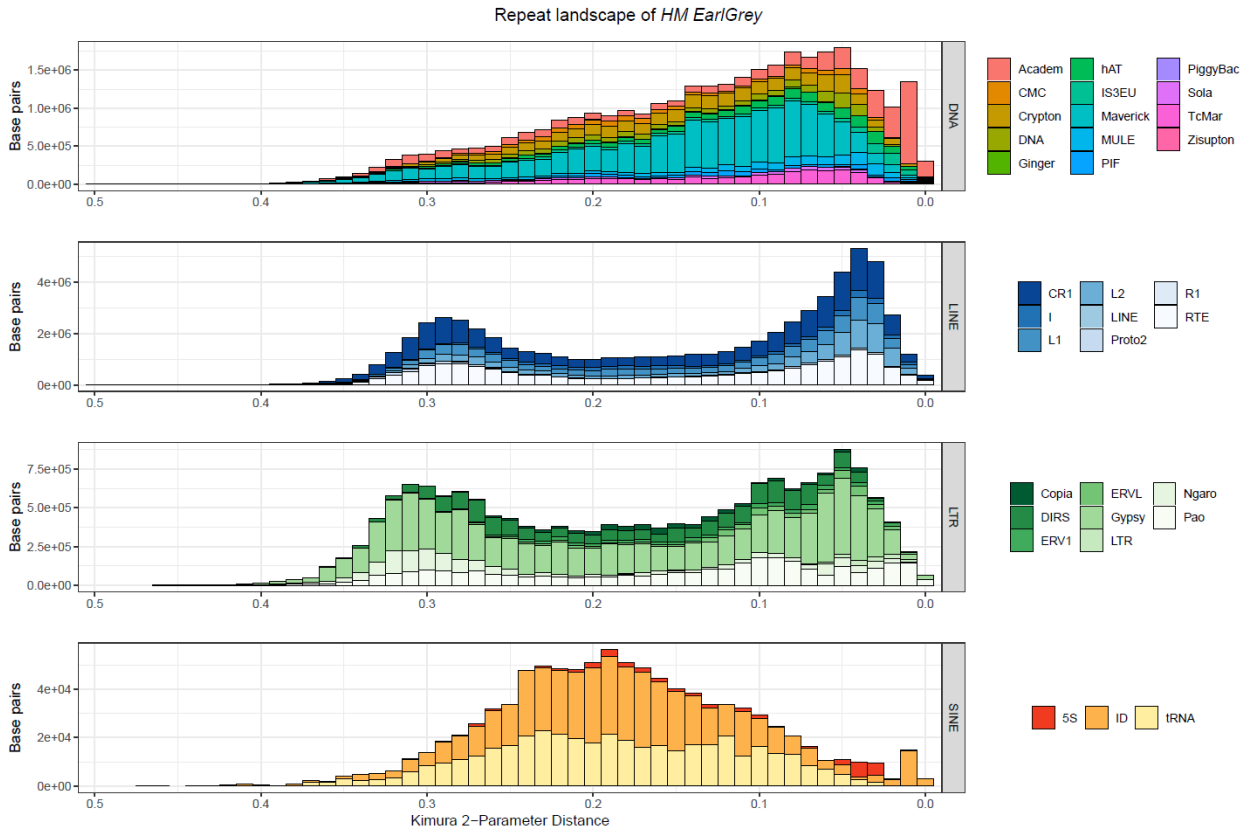


Figure 5 – Detailed repeat landscape plot for different types and sub-types of transposable elements identified in the genome of the horse mussel. The x-axis is the Kimura 2-Parameter Distance from the consensus sequence for the transposable element type and the y-axis is the amount of genome annotated in base pairs.

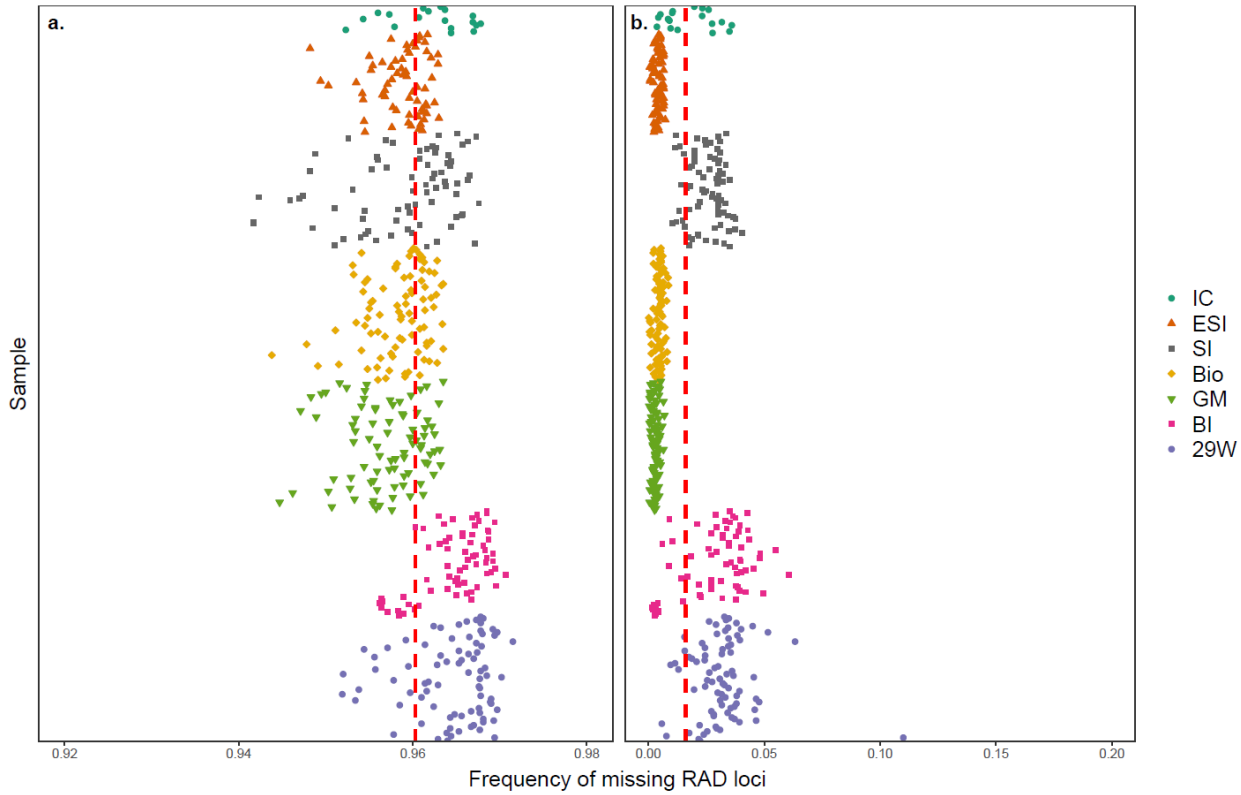


Figure 6 - Frequency of missing RAD loci per sample (a) before and (b) after filtering for RAD loci present in seven populations and 75% of individuals per population

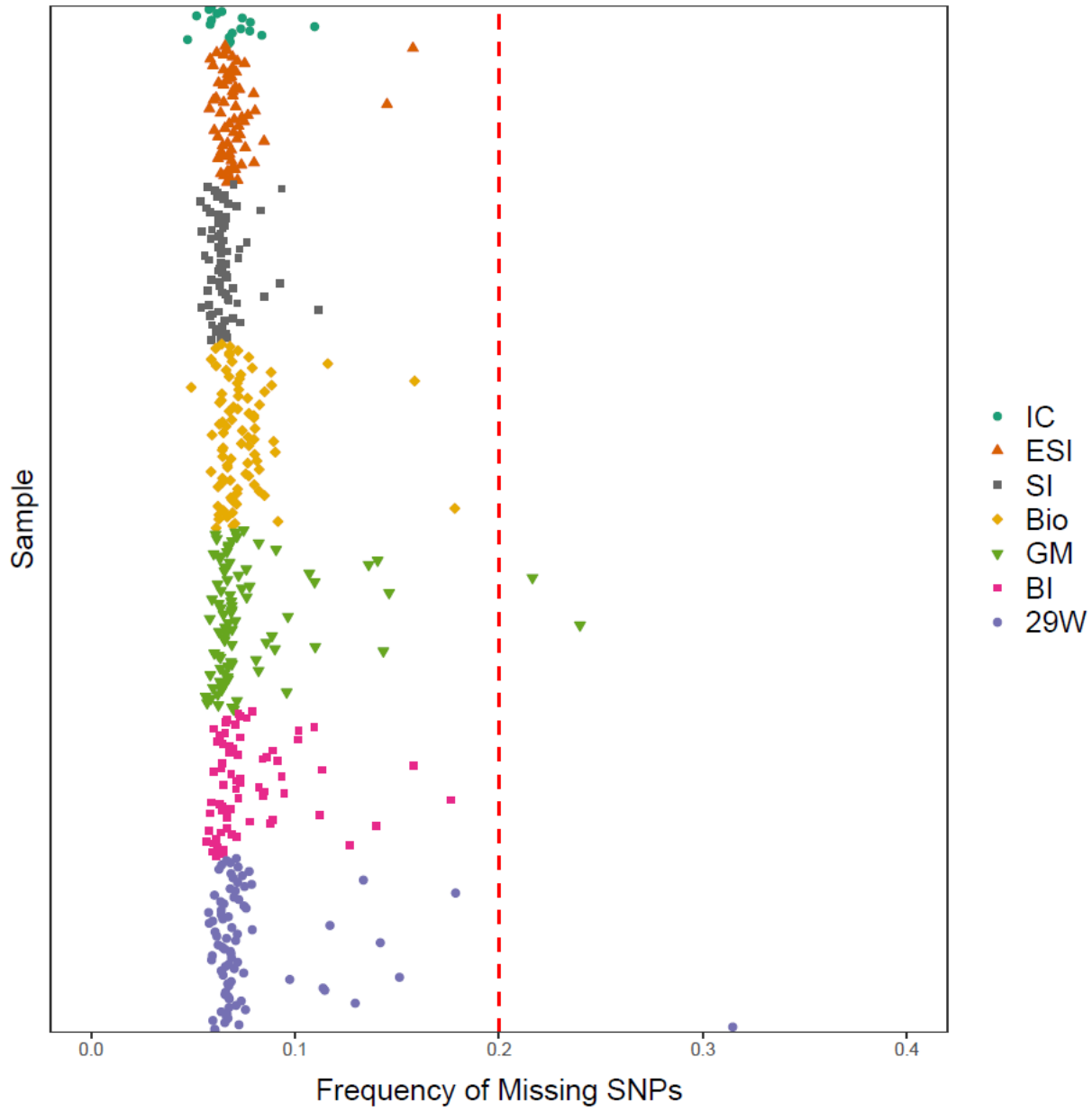


Figure 7 - Frequency of missing SNPs per sample after retaining one random SNP per RAD locus and filtering for a MAF of 0.01

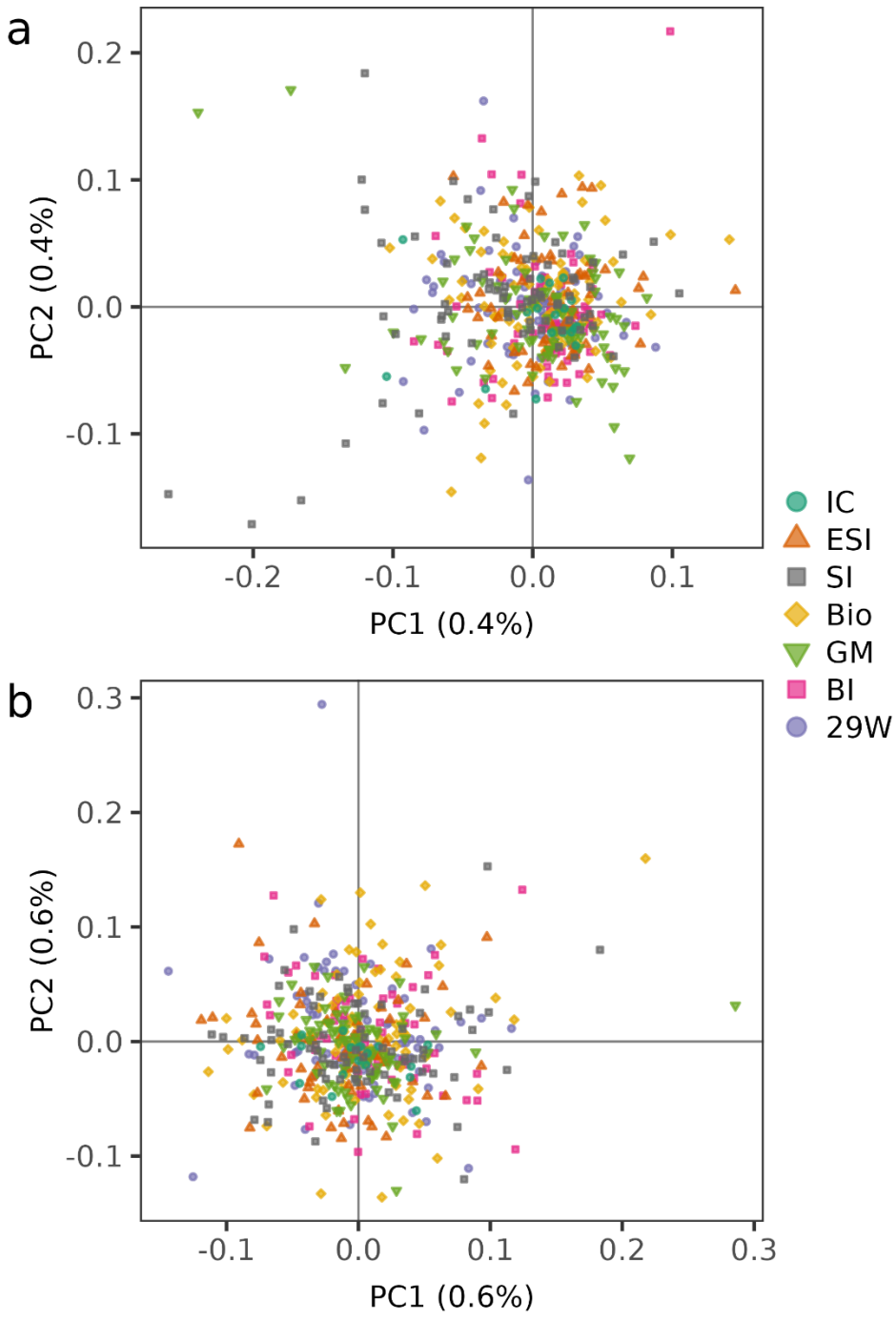


Figure 8 Principal components analysis of genotypes for 472 horse mussels using (a) 8414 genome-wide SNPs from sequencing reads aligned to the horse mussel genome and (b) 3500 SNPs detected de novo from sequenced RAD tags, with one random SNP retained per RAD tag

APPENDICES

APPENDIX 1 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR DNA EXTRACTION AND RAD-SEQ (RADCAP), WITH MODIFICATIONS FOR HORSE MUSSEL

The objective of this document is to record the relevant steps used in extracting horse mussel DNA and preparing RAD-seq libraries for sequencing. It is also likely suitable for use in other mollusc species.

Safety Considerations

General

Personnel should have proper training from a lab manager or other experienced individuals.

All relevant training, instruction, safe work procedures and chemical SDS should be read prior to work.

DNA Extraction

Read SDS for QIAGEN 96 DNeasy Plate Blood & Tissue kit components and absolute ethanol.

Wear lab coat and gloves when handling QIAamp DNA Micro kit components and DNA samples.

DNA Quantification and Normalization:

Read SDS for Quant-iT PicoGreen Reagent in DMSO.

Wear gloves when handling Quant-iT PicoGreen dsDNA Assay reagent in DMSO (Thermo Fisher Scientific).

PCR

The PCR hood unit has a UV light source. As a safety feature, the workstation is designed not to operate the UV light when the door is open. Do not over-ride the safety latch.

Loading Plate Preparation

Read SDS for HiDi Formamide. Wear appropriate gloves when handling.

Equipment

General

- Pipettes (RAININ, Mettler Toledo – various catalog numbers)
- Repeater Xstream positive displacement pipette (Eppendorf catalogue # 4987000118)
- BMP51 or BMP53 labeller (Brady – various label tapes)

DNA Extraction

- QIAcube HT (QIAGEN)
- Bead sterilizer (FST250 or Steri250) (Fine Science Tools, catalogue # 18000-45))
- Incubator (Fisher Scientific Model 624D)
- TissueLyserII (QIAGEN catalogue # 85300)
- TissueLyser 3 mm 96 bead dispenser (QIAGEN catalogue # 69975)

- ThermoMixer C with 0.5 mL, 1.5 mL, and PCR Smart blocks, with ThermoTop (Fisher Scientific catalogue # 05-412-504, 05-412-505, 05-412-512)

DNA Quantification and Normalization

- Eppendorf epMotion 5075 liquid handling robot
- Synergy H1 Microplate Reader (Agilent)
- Qubit 3 (Invitrogen)

PCR

- PCR machines (e.g., Eppendorf Mastercycler Pro with standard 96-well block)
- Eppendorf MixMate® (Eppendorf Catalog No. 2231000804) or Scientific Instruments Vortex Genie 2 (VWR catalogue # 14216-188)

DNA Qualification and Assay Screening

- Agilent 4200 TapeStation System (Agilent Part Number G2991BA)

DNA shearing

- BioRuptor (Diagenode catalogue # UCD-200-TS)
- 650 W Blender (e.g. Black & Decker 12 speed model “Cyclone”)

RAD sequencing library preparation

- ThermoMixer C with 0.5 mL, 1.5 mL, and PCR Smart blocks, with ThermoTop (Fisher Scientific catalogue # 05-412-504, 05-412-505, 05-412-512
- DynaMag-2 (ThermoFisher catalogue #12321D)

Supplies

General

- Brady deep well plate labels (Newark/Element 14 catalogue # M-81-488)
- Brady Top and Side Vial Labels (Brady Canada catalogue # M-120-492)
- Filtered pipette tips: 20 µL, 200 µL, 1000 µL (RAININ – various catalogue #s)
- Microcentrifuge tubes: 0.5 mL, 1.0 mL, 2.0mL (various brands and catalogue #s)

DNA Extraction

- DNeasy 96 Blood & Tissue Kit (QIAGEN catalogue # 69581)
- Forceps (Student Standard Pattern Forceps 11.5 cm – Fine Science Tools catalogue # 91100-12)
- 100 mL plastic beakers
- Paper towel
- Tungsten carbide beads, 3 mm (QIAGEN catalogue # 69997)

DNA Quantification

- Corning 96 well half area black microplate (VWR catalogue # 29444-314)

DNA Normalization

- Eppendorf semi-skirted LoBind plate (Fisher Scientific catalogue # E0030129512)
- epTIPS Motion reloads 50 µL (Eppendorf catalogue # 0030014430)

PCR

- Eppendorf skirted LoBind Plates (Fisher Scientific catalogue # E951020401)

- Eppendorf Biopur 0.5 mL Combitip Plus (Fisher Scientific catalogue # 13683720)
- Clear Adhesive PCR film (Applied Biosystems catalogue # 4306311)

DNA and Assay Screening (Agilent TapeStation)

- 96-well plates and foil seals (Agilent catalogue #s 5042-8502, 5067-5154)
- Strip tubes and strip caps (Agilent catalogue #s 401428, 401425)
- Pipette tips for the TapeStation (Agilent catalogue # 5067-5599)
- Eppendorf Biopur 0.2 mL Combitip Plus (Fisher Scientific catalogue # 13-683-719)

RAD sequencing library preparation

- 0.5 mL Bioruptor tubes (for DNA shearing)

Reagents and Solutions

DNA Quantification

- Quant-IT Picogreen dsDNA assay kit (ThermoFisher catalogue # P11496)

DNA Normalization

- 10mM Tris (pH 8.0):
    - Preparation (25 mL):
        - 250 µL 1M Tris (pH 8.0)
        - 24.75 mL Ultrapure H_2O
    - Note: can also use prepared Buffer EB instead of 10mM Tris (pH 8.0) (QIAGEN, catalogue # 19086)

DNA and Assay Screening on the Agilent TapeStation

- Agilent Genomic DNA Assay Kit (includes Genomic DNA ScreenTape, Genomic DNA Buffer, and Genomic DNA Ladder) (Agilent, Screen Tape order # 5067-5365, Genomic DNA Reagents order # 5067-5366)
- Agilent HS D1000 ScreenTape Assay Kit (includes HS D1000 ScreenTape, HS D1000 Buffer, and D1000 Ladder) (Agilent catalogue #s 5067-5584, 5067-5585)

RAD Library preparation

<u>Adapter preparation</u>

- See Appendix 1 – Table 1 for sequence of example adapter primers. (100 nmol synthesis from IDT with desalt purification)

<u>Restriction digest</u>

- *Sbf*I HF 2500 U (New England Biolabs catalogue # R3624L)
- NEBuffer 4 (New England Biolabs catalogue # B7004S (5 mL))

<u>Adapter Ligation</u>

- rATP (100 mM) (also known as ATP) (Fisher Scientific catalogue # PR-E6011)
- T4 DNA Ligase HC (New England Biolabs catalogue # M0202M)

<u>Pool and Purify samples</u>

- Agencourt AMPure XP beads (VWR Scientific catalogue # A63881, 60 mL)
- Dynabeads M-280 Streptavidin (Invitrogen catalogue # 11206D, 10 mL)

Library prep for Illumina

- NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB catalogue # E7370S)
- NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) (NEB catalogue # E7335S)

PCR enrichment of Adaptor-ligated DNA

Adapter preparation

- 10X Buffer AB (Adapter Binding)
    - Composition:
        - 500 mM NaCl
        - 100 mM Tris (pH 8.0)
    - Preparation (5mL):
        - 4 mL Ultrapure H_2O
        - 0.5 mL 5M NaCl
        - 0.5 mL 1M Tris-HCl (pH 8.0)

Pooling and Purifying samples

- Low TE
    - Composition:
        - 10 mM Tris (pH 8.0)
        - 0.1 mM EDTA (pH 8.0)
    - Preparation (5mL):
        - 50 µL 1M Tris (pH 8.0)
        - 1 µL 0.5M EDTA (pH 8.0)
        - 4949 µL Ultrapure H_2O

Isolation of RAD-tagged DNA

- 2X Binding and Wash Buffer:
    - Composition:
        - 10 mM Tris (pH 7.5)
        - 1 mM EDTA (pH 8.0)
        - 2 M NaCl
    - Preparation (15mL):
        - 150 µL 1 M Tris-HCl (pH 7.5)
        - 30 µL 0.5 M EDTA (pH 8.0)
        - 6 mL 5 M NaCl
        - 8.82 mL Ultrapure H_2O
- 1X Binding and Wash Buffer (Make two of these, one for room temp and one for 56 °C):
    - Composition:
        - 5 mM Tris (pH 7.5)
        - 0.5 mM EDTA (pH 8.0)
        - 1 M NaCl
    - Preparation (5mL):
        - 2.5 mL 2X Binding and Wash Buffer
        - 2.5 mL Ultrapure H_2O

Buffer for diluting Adaptor in NEBNext Ultra II DNA Library Prep Kit

- 10 mM Tris/10 mM NaCl (also referred to as Tris/NaCl)
    - Composition
        - 10 mM Tris (pH 7.5)

        - 10 mM NaCl
    - Preparation (5 mL)
        - 50 µL 1 M Tris (pH 7.5)
        - 10 µL 5 M NaCl
        - 4940 µL Ultrapure H_2O

Clean-up of PCR Reaction

- 0.1X TE
    - Composition
        - 1 mM Tris
        - 0.1 mM EDTA
    - Preparation (5mL)
        - 5 µL 100X TE (Sigma)
        - 4995 µL Ultrapure H_2O

Procedures

Tissue digest

Note: Digestion is to be completed the day (ideally the afternoon) before DNA extraction

Note: Read DNeasy 96 extraction manual and QIACube HT manual before undertaking tissue digestion

1. Turn on incubator and set to 56 °C.
    a. If precipitates present in the provided Buffer ATL, place bottle in incubator until use.
2. In Excel, plan the layout of the plate (*i.e.,* which samples are going into which well).
    a. Each extraction should be given a unique identifier to prevent mistakes.
    b. Each 96-well plate can hold 95 samples plus one extraction negative.
3. Unwrap a set of collection tubes and rack from DNeasy 96 Blood & Tissue kit, then label the edge of each strip of tubes with the extraction plate name (or some identifier)
4. Prepare supplies to dissect the tissue – forceps, scissors and sterilization equipment.
    a. Turn on glass bead sterilizer to heat up
    b. Add Ultrapure water to 2 x 100 mL beakers (fill to ~80 mL mark), label one “Dirty” and one “Rinse”. Add Reagent ethanol (80%) to a third 100 mL beaker labelled “80% ethanol”.
    c. Sterilize 24 forceps and 24 scissors to start (otherwise you will have to sterilize between samples):
        i. Rinse in the “Dirty” beaker of Ultrapure water (change the water when beaker becomes dirty from use)
        ii. Heat in the bead sterilizer for ~ 20-30 s.
        iii. Cool off in the “Rinse” beaker of clean Ultrapure water.
        iv. Dip them in the beaker of 80% alcohol.
        v. Transfer to clean empty beaker to air dry.
5. Prepare your work area for tissue sample:
    a. Lay out some paper towels to work on and have more to replace as needed.
    b. Arrange samples and collection tube strips per plate layout.
    c. Transfer one strip of tubes to an empty rack at time to minimize cross contamination between tubes.
6. Using a sterile 5mL Combitip (or other multichannel pipette), add 200 µL Buffer ATL to each collection tube and recap tubes.

7. Using a sterile set of forceps and scissors, dissect off a ~20 mg piece of tissue from the sample vial and transfer to collection tube, ensuring tissue is submerged in the ATL. Place the used scissors and forceps in the “Dirty” beaker with Ultrapure water.
8. Repeat tissue excision and transfer to collection tube until strip of 8 has been completed.
    a. Use a new sterile set of forceps and scissors for each sample.
    b. Cap the strip of samples (you may choose to transfer completed strip to another rack for convenience)
9. Sterilize scissors and forceps per step 4 as needed.
10. Repeat for all the strips of sample collection tubes until complete for the plate(s) you are going to extract.
    a. After the last tissue samples complete, incubate at room temperature for 30 minutes to rehydrate tissue.
11. Just before 30 min incubation is complete, prepare the ATL + Proteinase K mixture in a 25 mL tube, mixing gently by inversion:

Reagent	**1x**	**100x**
ATL buffer	180 µL	18 mL
Proteinase K	20 µL	2 mL
Total	**200µL**	**20 mL**

12. After 30 minutes, use a P1000 multichannel pipette to remove ATL buffer from all samples. Discard tips between samples.
13. Using the TissueLyser Plate Bead Dispenser for 3 mm beads, add one 3 mm tungsten carbide bead to each tube.
14. Using a 5 mL Combitip, add 200 µL of ATL + Proteinase K to each sample collection tube.
    a. Cap each strip with new caps.
15. Clamp the plates in the TissueLyserII and homogenize the samples two times for 10 s at 20 cycles per second.
    a. Reverse the orientation of the plate of tubes in the clamping plates between the 10 s mixes.
16. Centrifuge the collection microtubes briefly ~5 s using the “short” button to collect the tissue and liquid at the bottom of the tubes.
17. Place the plate of tubes in the 56 °C incubator. Put a weight on the top of the tubes to ensure the caps stay sealed.
    a. (optional) After 1 hour remove the tubes from the incubator, being VERY careful as the caps may pop off after they are warm, then mix the samples gently by inversion. Centrifuge the collection microtubes briefly ~5 s using the “short” button to collect the tissue and liquid at the bottom of the tubes.
    b. Place the plate of tubes back in the 56 °C incubator. Put a weight on the top of the plate to help ensure the caps stay on.
18. Let the samples incubate overnight.
19. Return to original sample tubes to storage.

DNA Extraction

Note: This should be completed in the morning following overnight tissue digest

Note: Read DNeasy 96 extraction manual and QIACube HT manual before undertaking DNA extraction.

1. As necessary, prepare any DNeasy 96 Blood & Tissue kit reagents .
    a. Example: Add ethanol to the AL buffer (for extractions from tissue samples)
    b. Loosen the cap of the AE buffer and place in the 70 °C incubator to warm up.
2. Remove plate of sample collection tubes from the incubator. Ensure that the incubated sample collection tubes are properly sealed to avoid leakage during shaking.
    a. Flick each strip of tubes to check if the sample is digested.
    b. If the samples are digested, centrifuge them briefly (~5 s) using the “short” button and proceed with the extraction.
    c. If the samples aren’t digested, add 20 µL proteinase K, then put them back in the incubator at 56 °C for a while, checking every hour until digested.
3. Using a multichannel pipette, transfer digested samples out of the tubes (and off of the beads) and into a new S-block in the corresponding well locations.
    a. Avoid any undigested bits in the samples to avoid samples clogging the extraction filters.
4. Turn-on the QIACube HT computer, then turn on the QIACube HT.
5. Launch QIACube HT software and create or load a run. Note: see software manual for details on creating run templates.
6. Calibrate any plasticware that is not calibrated (marked with a red exclamation mark) in the corner of the plasticware position on the deck layout.
    a. This step will not be required each time. It will be required if plasticware is in a new position, or if the QIACube HT has gone through preventative maintenance.
7. Click on the wizard icon or choose Wizards → Vacuum Extraction Wizard
8. Check that the standard options are selected:
    a. Under “How are the samples to be introduced into the extraction procedure?” make sure “My samples are preloaded into an empty lysis block” is selected.
    b. Under “Options” the following should be listed:
        i. How many reagents are you lysing with (pre capture)? = 1 (AL)
        ii. How many reagents are you washing with (post capture)? = 2 (AW1 & AW2)
        iii. How many reagents are you eluting with? = 1 (AE)
9. Click “Next” on the Wizard window to “Table setup” screen, and check that the standard deck setup is listed.
10. Click “Next” on the Wizard window to be taken back to the “Configuration (1)” page.
11. Under “Configure run options” leave “Turn HEPA filter on automatically” checked.
12. Under “Extraction will be performed on all columns marked red” indicate the rows on the plate that will be extracted (the default is a full plate).
13. Click “Next” on the Wizard window (“Configuration (2)” screen).
14. Make sure the standard configuration of reagents is shown.
15. Click “Next” on the Wizard window.
16. On the “Load Pre Capture Reagent 1” window make sure the following standard options are listed:
    a. Under “Select the reagent (lysis or equivalent) and volume:
        - Reagent = AL/EtOH
        - Reagent volume = 410µL
    b. Volume of sample to be used?
        - Sample volume = 200µL
    c. How often should the wells be mixed after injection of this reagent?
        - Number of times to mix = 0
            - NOTE: the samples will be mixed with the AL/E in the next window
    d. Wait period after combining lysis/sample before continuing
        - Incubation: Timed pause = 00:00:00

- The remaining options for this category will be greyed out

17. Click “Next” on the Wizard window.
18. On the “Load Capture Plate” window make sure the following standard options are listed:
    a. Under “Total amount of lysed sample to be loaded into capture plate”
        i. Load volume = 600 µL
    b. Under “Vacuum”
        i. Use vacuum at the end of each step? = checked
        ii. Vacuum run-time = 00:10:00 (10 minutes)
        iii. Vacuum pressure = 40 kPa
        iv. Wait for user confirmation before turning vacuum off? = checked
            1. NOTE: This step allows you to visually check that the samples have loaded onto the column before proceeding with the extraction.
    c. Under “Mix Options”
        i. Premix lysed samples (in lysis plate) before loading into capture plate? = checked
        ii. Premix iterations = 10 times
            1. NOTE: This is the mix of the samples with the AL/E before loading on the columns
19. Click “Next” on the Wizard window.
20. On the “Wash Step 1” window of the wizard make sure the standard parameters are set:
    a. Under “Select the reagent and volume”
        i. Reagent = AW1
        ii. Reagent volume = 500 µL
    b. Under “Wait period after loading reagent before proceeding?”
        i. Incubation time = 00:00:00
    c. Under “Vacuum?”
        i. Use vacuum at the end of this step? = checked
        ii. Vacuum run-time = 00:05:00 (5 minutes)
        iii. Vacuum pressure = 40 kPa
        iv. Wait for user confirmation before turning vacuum off? Checked
            1. NOTE: This allows you to check that the AW1 buffer has gone all the way through. You will have to come back during the run to continue the program.
            2. NOTE: If there are wells where the liquid hasn’t gone through, you can use a pipette tip (standard pipette tip, not the QIACubeHT tips) to remove any obstruction
    d. Under “Iterations?”
        i. Perform this process = 1 times
21. Click “Next” on the Wizard window.
22. On the “Wash Step 2” window of the wizard make sure the standard parameters are set:
    a. Under “Select the reagent and volume”
        i. Reagent = AW2
        ii. Reagent volume = 500 µL
    b. Under “Wait period after loading reagent before proceeding?”
        i. Incubation time = 00:00:00
    c. Under “Vacuum?”
        i. Use vacuum at the end of this step? = checked
        ii. Vacuum run-time = 00:15:00 (15 minutes)
        iii. Vacuum pressure = 40 kPa

iv. Wait for user confirmation before turning vacuum off? Unchecked
    1. NOTE: If you want to double check that the AW2 buffer has gone all the way through, you can check this box. If selected (checked) you will have to come back during the run to continue the program.

d. Under “Iterations?”
    i. Perform this process = 1 times

23. Click “Next” on the Wizard window.
24. On the “Dry Sample” window of the wizard, make sure the standard parameters are set:
    a. Under “Vacuum”
        i. Use vacuum/pause? = checked
        ii. Vacuum run-time/ Pause period = 00:10:00 (10 minutes)
        iii. Vacuum pressure= 40 kPa
        iv. Wait for user confirmation before proceeding? = unchecked
        v. Ignore Feedback? = checked
25. Click “Next” on the Wizard window.
26. On the “Elution Step 1” window of the wizard, make sure the standard parameters are set:
    a. Under “Select the reagent and volume”
        i. Reagent = AE
        ii. Regent volume = 110 µL
            1. NOTE: There is some volume lost to the column, the additional AE will help compensate.
        iii. User Pause before loading reagent? = checked
            1. NOTE: This allows us to load warm AE onto the deck before the elution
    b. Under “Incubation”
        i. Wait period after loading buffer before proceeding? = 00:05:00 (5 minutes)
    c. Under “Vacuum”
        i. Use vacuum at the end of this step? = checked
        ii. 2 Step Vacuum = checked
        iii. Vacuum run time: First block = 00:04:00 (4 minutes); Second block = 00:00:30 (30 seconds)
        iv. Vacuum Pressure: First block = 40 kPa; Second block = 0kPa
            1. NOTE: The first vacuum step is 4 minutes at 40 kPa and the second vacuum step is 30 seconds at 0kPa (no vacuum).
        v. Wait for user confirmation before continuing? = unchecked
        vi. Confirmation? = unchecked
        vii. Ignore Feedback? = checked
    d. Under “Post-run”
        i. Perform this step = 2 times
            1. NOTE: This indicates to do two 100 µL elutions (actually 110µL elutions,
27. Click “Next” on the Wizard window. This gives you the pre-processing report.
28. Click “Finish” on the Wizard window. This takes you back to the deck layout.
29. Load the QIAcube deck.
    a. Setup the vacuum chamber
        i. Check the O-ring on the channeling adapter. If necessary, add some grease to the O-ring
        ii. Load the channeling adapter into the left position of the vacuum chamber Push it down firmly to ensure the spigot is properly seated in the drain.

iii. Place the riser block (EMTR) into the right-hand position of the vacuum chamber (this is the elution chamber).
    1. NOTE: The raised corner goes in the upper right position. This adapter has a very close fit in the elution chamber. Make sure that it goes in straight and is all the way to the bottom
iv. Ensure the silicon gasket is in the transfer carriage. Put the transfer carriage on the lip of the vacuum chamber over the channeling adapter (left side of the vacuum chamber).
    1. NOTE: Make sure this is all the way to the left.
v. Place the DNeasy plate on the transfer carriage
vi. Place the elution tubes RS (blue plate that comes with the DNeasy 96 Blood & Tissue kit) on the EMTR riser block (right-hand side of the elution chamber).
    1. NOTE: The raised corner of the riser block will match with the notched corner of the blue plate holding the elution tubes.
    2. NOTE: The eluted samples will be transferred to an Eppendorf LoBind deep well plate, so a temporary label is sufficient on these elution tubes.

b. Add the solutions. The deck layout indicates the position and volume of solutions to add. Do not add the AE at this time – it will be added to the deck just before the elutions are done. For a full plate the solution volumes are:
    i. AL/EtOH = 42.36 mL (aim for 43 mL) in a 70 mL reservoir.
    ii. AW1 = 55 mL in a 170 mL reservoir.
    iii. AW2 = 55 mL in a 170 mL reservoir.
    iv. AE = 22.2 mL (aim for 23 mL) in a 70 mL reservoir.
        1. NOTE: Make sure that the reservoirs are seated flat and that the lids are on (to minimize evaporation).

c. Add the tips.
    i. Place a box of tips in each of B2 and C2.
    ii. If the box of tips is not full, indicate in the software which rows of tips are missing – click the row(s) (they will be highlighted in red), right click “Set selected tips to Unavailable” – the circles will then be white. Conversely, if there are tips in rows marked as not being present, but actually are in the box, mark those as being Available.
        1. NOTE: The “Tip Info” window on the right will indicate if you have enough tips to execute the program. The table indicates the number of tips available; the number required and if there are “Enough Tips?” (Yes/No).

d. Add the samples.
    i. Place the S-block with the samples in the B1 position.
        1. NOTE: Make sure the A1 position is in the top left. Make sure to remove the mat from the plate.

30. Check that the waste bottle (located with the pump below the QIACube HT) has sufficient space for the waste. The waste bottle has a 4 L capacity and will register as being “full” at ~2 L.
31. Remove the red cover from the waste chute.
32. Remove the lid from the waste tip container.
33. Close the lid of the QIACube HT, then Click the green arrow () to start the run.
34. Make sure to return to continue the program:

a. After the samples are loaded onto the DNeasy plate (~ 26 minutes + 10 min vacuum)
b. If you have selected to visually check after the AW1 (~11 min) or AW2 (~22 min) washes.
c. Before the elution, to add the pre-heated AE buffer. (~45 min after the samples finish loading)

35. When the run is complete you will also get a post run report. You should save a copy of this report along with the extraction plate information.
36. Remove your eluted samples from the QIACube HT and cover the tubes.
37. Transfer your eluted DNA to a labelled Eppendorf LoBind deep well plate using the epMotion 96.
38. Clean up the QIAcube after use:
    a. Unused buffers can be poured back into the stock bottles – except AE (should be disposed)
    b. Reservoirs should be cleaned with Ultrapure water between uses. Allow the reservoirs to dry between uses. If you want to do back-to-back runs on the QIACube HT, have two sets of labelled reservoirs.
    c. Put the lid on any partially filled tip racks. Recycle any empty tip racks.
    d. Discard the S-block, any used plates, and any used tips
    e. Rinse the QIACube HT waste tip collection container with reverse osmosis (RO) water to remove the salts.
    f. Remove the tip chute and was with diluted LiquiNox (diluted 1:100). Rinse it with RO water and dry with paper towel. Replace the tip chute back on the QIACube HT.
    g. Remove the channeling adapter and wash with diluted LiquiNox (diluted 1:100) and rinse with RO water. Dry the channelling adapter with a kimwipe.
    h. Remove the carriage and its gasket. Wash with diluted LiquiNox (diluted 1:100) and rinse with RO water. Dry with a kimwipe and put the gasket back on the carriage.
    i. Clean any reagent spilled on the instrument deck with a damp paper towel.
39. If you are doing another extraction you can proceed from step 6.a, otherwise proceed with shutting down the instrument:
    a. Close the software. This will launch the instrument cleaning window.
    b. Select “Run UV cleaning operation before closing?”
    c. Using the checklist in the wizard, make sure the deck is ready for shutdown.
    d. Pour 20 mL of RO water into the left (waste) side of the vacuum chamber.
        1. NOTE: The Instrument Cleaning window must be open before pouring the water into the vacuum chamber. Starting the wizard closes the vacuum line to the elution chamber. This prevents the wash water from flowing into the elution chamber and draining out the air bleed hole onto the bench.
40. Close the lid of the QIACube HT and press “Next” on the Instrument Cleaning window.
41. Wait for the UV cycle to complete (15 min).
42. Turn off the QIACube HT, then shut down computer.

DNA Quality Screening on the Agilent TapeStation

1. Use the Genomic DNA ScreenTape Kit. Refer to the Agilent Genomic DNA ScreenTape Assay Quick Guide for 4200 TapeStation System.
    a. Combine 1 µL of gDNA with 10 µL of the Genomic DNA Sample Buffer in a 96-well plate (Agilent) provided for the TapeStation.

b. Apply foil seal to the plate, mix for 1 min on the IKA vortex at 2000 rpm. Centrifuge for 1 min at 1500 rpm.
    i. Check for bubbles. If further vortexing and centrifugation do not remove the bubbles, carefully remove the foil seal and break the surface tension with a pipette tip.
c. Load the samples onto the instrument. Follow instructions in the Quick Guide.

2. In general, criteria for good quality DNA are DNA Integrity Numbers (DINs) of around 6.0 or higher *and* nice clearly defined bands with little to no smearing in the lane. See Appendix 1 – Figure 1 for example images.

DNA Quantification

NOTE: This uses the Quant-iT PicoGreen dsDNA Kit

1. Prepare your sample plate layout. 6 wells are required for the standard curve and blank. Standards can be run on the same plate as samples (if space allows) or on a separate plate (if space does not allow).
2. Thaw the Quant-iT PicoGreen in the dark at room temperature.
3. Prepare a 5-point standard curve of λDNA:
    a. First prepare a 2 ng/µL DNA standard (Standard 1) by diluting 6 µL of the 100 µg/mL stock of λDNA in 294 µL of TE. Mix by vortex.
    Note: Pre-wet the tips before dispensing to improve accuracy.
    b. Then prepare the remaining 4 standards as serial dilutions of the 2 ng/µL standard as follows:
        i. Standard 2 (1.5 ng/µL): 150 µL of the 2 ng/µL standard in 50 µL TE
        ii. Standard 3 (1.0 ng/µL): 100 µL of the 1.5 ng/µL standard in 50 µL TE
        iii. Standard 4 (0.5 ngµL): 50 µL of the 1.0 ng/µL standard in 50 µL TE
        iv. Standard 5 (0.05 ngµL):10 µL of the 0.5 ng/µL standard in 90 µL TE
4. To a black 96-well plate, add 49 µL TE to each well required for the samples based on your plate layout (not standards or blank).
5. Add 50 µL of TE into the well for the blank.
6. Make a 1:100 dilution working stock of PicoGreen in TE. Each well requires 50 µL of working stock, so calculate for the number of wells needed + 5%. Mix well using the vortex.

Reagent	**1 well**	**105 wells**
PicoGreen	1 µL	105 µL
TE	49 µL	5145 µL
Total	**50 µL**	**5250 µL**

7. Add 50 µL of the PicoGreen working stock to all wells (samples, standard curve, blank) using a repeater pipette.
8. Vortex gDNA for 30 s at 1000 rpm using MixMate, then centrifuge for 1 min at 1500 rpm.
9. Add 1 µL DNA from each sample to the appropriate well of the black plate according to your layout.
10. Pipette 50 µL of each standard into the appropriate wells of the plate according to your layout. Each well (standards, samples, blank) should now contain 100 µL.
11. Without covering the plate, mix gently on the plate shaker (the speed will vary depending on the proper setting for the type of plate shaker being used, make sure nothing spills out of the wells) for 30 s to 1 min.
12. Incubate at room temperature in the dark for 2 to 5 minutes (longer is OK).
13. Turn on the Synergy H1 MicroPlate Reader and associated computer.

14. Load the plate into the Synergy H1 MicroPlate Reader, refer to manual for standard operation.

RAD Library preparation

- Two libraries can be processed in one batch
- In this example, libraries consist of 48 individuals

DNA dilution

Note: It is recommended you read the epMotion robot manual before proceeding.

1. Using epBlue software, load/design your protocol for making dilution plate
    a. Aim to dilute gDNA to 10 ng/µL in 45 µL total
    b. Volumes of EB buffer and gDNA will depend on DNA concentrations determined in DNA Quantification step.
2. Vortex gDNA in plate with MixMate at 1000 rpm for 30s
3. Centrifuge plate at 1500 rpm for 1 min to bring sample off plate seal (silicone mat)
4. Arrange gDNA plate, EB buffer reservoir w/ EB, and destination Lobind PCR plate on epMotion stage
5. Start epMotion program
6. After epMotion program is complete
    a. Seal and label dilution plate, then store in 4 °C until next steps (or -20 °C for longer storage)
    b. Seal gDNA plate with silicone mat, then move to -20 °C for storage

Adapter preparation

1. Make 50 µM dilutions from the 100 µM stock primers using EB Buffer.
    a.

Reagent	**1X**
100 µM stock primer	25 µL
Buffer EB	25 µL
Total	**50 µL**

2. Combine complementary adapters (*i.e.,* the matching RAD-Cap Primer Barcode forward and reverse) as follows in 1.5 mL LoBind tubes:
    a. Note: The adapters need to be a final concentration of 50 nM (0.05 µM)

Reagent	**1X**	**Concentration**
10X Buffer AB	100 µL	1X
Primer 1F (50 µM)	1 µL	0.05 µM
Primer 1R (50 µM)	1 µL	0.05 µM
Ultrapure H_2O	898 µL	
Total	**1000 µL**	

3. Add 1 L of RO water to a 2 L beaker
4. Bring the RO water to a rolling boil on the hot plate
5. Turn off the hot plate

6. Float the tubes in the RO water until it cools to room temperature (approximately 6 hours). Use a thermometer to monitor the temperature.
    a. The tops of the tubes may pop open, so monitor the tubes especially in the beginning.
    b. After 30 minutes of cooling, you can remove the beaker from the hot plate and allow it to cool on the bench (this reduces the chance of scorching the bench)
    c. The adapters do not need to be removed from the beaker immediately when the RO water cools to room temperature (around 22-23 °C), but should be stored at -20 °C as soon as possible.
7. The resulting adapters are 0.1 µM in 1X AB.
8. Make working plate of 48 prepared adapters by adding 40 µL of each prepared adapter to a skirted PCR plate and seal with tape.
9. Store the prepared adapters and working plate at -20 °C.

Restriction Digest

1. Turn on the Thermomixer C with a PCR 96 heat block and let it warm to 37 °C. Load the following protocol:
    a. 37 °C 1 hr
    b. 80 °C 20 min (inactivate the enzyme)
    c. 21 °C Hold
2. Prepare the restriction digest master mix with *Sbf*I-HF (NEB) in 1.5 mL tube on chill beans:

Reagent	**1x (per sample)**	**50x (master mix)**	**Final Concentration in reaction**
10X NEBuffer 4	2.5 µL	125 µL	1X
Ultrapure H_2O	1.5 µL	75 µL	
*Sbf*I-HF (20U/µL)	1 µL	50 µL	10X
gDNA (10ng/µL*)	20 µL	------	200 ng (total)
Total	**25 µL**	**250 µL**	

+Note: Ali et al. (2016) digested 200 ng in a 12 µL reaction. This amount is scaled up for 25 µL digest to allow for evaporation and 2 attempts at library preparation from one digest.

*Note: used DNA samples normalized to 10 ng/µL for most libraries

3. Add 5 µL of Master Mix to the required wells of a 96-well plate (per your plate layout).
    a. Use a DNA LoBind plate.
    b. Keep the plate on an Eppendorf freezer block while preparing the restriction digest.
4. Add 20µL DNA to the appropriate wells (per plate layout).
5. Seal the plate with PCR film, then mix the plate on the MixMate at 1650 rpm for 30 seconds.
6. Centrifuge the plate in the freezer block at 1500 rpm for 1 minute.

7. Place the plate in the ThermoMixer C and run the loaded program. It takes 1 hr and 20 min.
    c. Another option since the ramp speed is slow – Use two ThermoMixer C with one set to 37 °C and another to 80 °C and manually switch between the two once the time is up.
8. Continue on to ligation immediately after this is complete.

Adaptor Ligation

1. Load RAD-Cap Ligation protocol on the Thermomixer C. The program consists of:
    a. 20 °C 1 hr
    b. 65 °C 20 min (inactivate the enzyme)
    c. 21 °C Hold
2. Prepare ligation master mix on chill beans:

Reagent	**1X (per sample)**	**50X (master mix)**	**Final Concentration in reaction**
Ultrapure H_2O	1.28 µL	64 µL	
10X NEBuffer 4	0.4 µL	20 µL	1X
rATP (100 mM)	0.16 µL	8.0 µL	1 mM
T4 DNA Ligase (2000 kU/µL)	0.16 µL	8.0 µL	20 kU
Total	**2 µL**	**100 µL**	

3. Add 2 µL of ligation master mix to the appropriate wells of a new Eppendorf LoBind PCR plate on an Eppendorf freezer block.
4. Add 12 µL of the *Sbf*I digested DNA samples to the appropriate wells.
5. Add 2 µL of annealed *Sbf*I biotinylated RAD barcode adapters (0.1 µM) to each sample.
    a. For samples pooled into the same library, each sample will need to have a different barcode (see Appendix 1 – Table 1).
6. Mix the plate on the MixMate at 1650 rpm for 30 s.
7. Centrifuge the plate at 1500 rpm for 1 min on the ThermoBlock.
8. Place the plate in the ThermoMixer C and run RAD-Cap Ligation program.
    a. It will take 1 hr and 20 min to run
9. Mix the plate on the MixMate at 1650 rpm for 30 s (*i.e.*, 96 PCR setting).
10. Centrifuge the plate at 1500 rpm for 1 min.
    a. You can stop at this step if necessary, store samples in the refrigerator (4˚C).

Pool and Purify samples

1. Prepare reagents prior to start:
    a. Allow Agencourt AMPure XP beads come to room temperature prior to use.
    b. Prepare **fresh** 80% ethanol using Ultrapure H_2O and absolute ethanol
    c. Gently vortex the AMPure XP bottle to resuspend the beads. It should appear homogenous and consistent in colour.

2. Pool 5µL of each *Sbf*I digested, adapter-ligated sample per library into a LoBind 1.5 mL labelled tube.
    a. Can use epMotion robot for automated pooling
    b. Store the remaining sample (*Sbf*I digested, biotinylated RAD barcode adapter ligated DNA in 96-well plate) at -20 °C. The library preparation step can be started again from this point, if required.
3. Gently vortex the Agencourt AMPure XP bottle to resuspend the beads. It should appear homogenous and consistent in colour.
4. Transfer **equal amount** of Agencourt AMPure XP beads into the pooled samples.
    a. For our example, there are 48 individuals/samples within one library, so the volume should be ~240 µL, but measure by pipette to get exact volume of beads to add.
5. Mix via pipette 10 times. The colour of the solution should appear homogenous after mixing.
6. Let the solution incubate at room temperature for 5 min.
7. Place the tube on the magnetic rack for 2-5 mins to separate the beads from the solution
    a. Note: Wait for the solution to be clear before proceeding to the next step. In this first wash the solution may not clear completely, this is normal.
8. Leaving the tube on the magnet, remove the supernatant using a pipette and discard.
    a. Note: Do not disturb the “pellet” of separated magnetic beads.
9. Leaving the tube on the magnet, dispense 250 µL of <u>fresh</u> 80% absolute ethanol to the tube and incubate for 30s at room temperature.
10. Remove the ethanol using a pipette and discard.
    a. Note: The beads are not drawn out easily when in alcohol, so it is not necessary to leave any supernatant behind.
11. Repeat steps 9-10 for another ethanol wash.
12. Remove tube from the magnet, and then add 205 µL of Low TE (refer to Reagents and Solutions; Pooling and Purifying Samples for prep.) and mix via pipetting 10 times.
13. Incubate for 2 min at room temperature.
14. Place tube on magnet for 1 min to separate the beads from the solution.
15. Once the solution is clear, transfer the supernatant to a new 0.5 mL Eppendorf tube.
    a. Keep samples at 4 °C. Do not freeze to avoid additional freeze/thaw cycles.

<u>DNA shearing</u>

Note: Only two libraries at a time should be run through the Sonicator. Adding tubes from more than two libraries may have a negative impact on the sonication. The first few times you undertake this protocol, it may also be advisable to run one library at a time until you have confirmed you are getting the correct fragment sizes.

1. Pre-chill a 2 L beaker and the Sonicator (BioRuptor) tube holder in a refrigerator.
2. Procure ice.
3. Add 4 x 250 mL Nalgene Beakers of the cubed ice to the BioRuptor water bath.
    a. Cover the ice, to the water bath fill line, with RO water and allow the water bath to cool for 30 minutes. During this incubation prepare libraries for sonication:
        i. Divide the pooled, purified sample libraries into two 0.5 mL BioRuptor tubes, each containing 100 µL (designated tubes A and B for each library).
        ii. Transfer the BioRuptor tubes with the libraries to the BioRuptor tube holder in the fridge.
        iii. Balance with BioRuptor tubes containing equal volumes of MilliQ water.
    b. When there is 5 minutes left in the cooling of the water bath, prepare crushed ice

i. Add 4 x 250 mL Nalgene beakers of cubed ice to a blender, with enough RO water to cover the ice.
ii. Transfer the pre-chilled 2 L beaker to the ice bucket and put it on the remaining ice.
iii. Use the “Pulse” button on the blender and blend the ice/water mix until you have a slurry of crushed ice
iv. Transfer the slurry of crushed ice to the 2 L beaker and keep on ice.

4. When the time is up on the pre-chilling, remove the cubed ice from the BioRuptor water bath using a 250 mL Nalgene beaker. Leave the chilled water.
5. Add 1 x 250 mL Nalgene beaker of crushed ice to the BioRuptor water bath.
    a. Fill the beaker with ice and drain the water. Pack the crushed ice into the beaker.
6. Use a small spatula or scoopula to break-up the crushed ice and evenly distribute it.
7. Adjust the RO water level to the fill line. If you need to add RO water, use the pre-chilled ice RO water from the crushed ice bucket
8. Transfer the BioRuptor tube holder to the BioRuptor mechanical lid (fits on top of the water bath lid – be sure the rungs of both match up).
    a. If you need to place the tube holder + tubes on the benchtop before adding them to the water bath, put them on the chill beans to keep them cold.
9. Sonicate half the DNA (one tube) for 10 minutes at 30 sec ON/ 15sec OFF (Tube A), and the other half for 10 minutes at 30 sec ON/ 30 sec OFF (Tube B). This is to reduce variability in fragment size ranges between libraries.
    a. Make sure the control panel is turned on and the timer is set correctly. Turn dial to 10 min.
    b. If running more than one library at a time, load tube 1A (from library 1 tube A) and tube 2A (from library 2, tube A) in the tube holder. Balance with two Bioruptor tubes (100 µL of MilliQ water each).

    Note: These settings were determined empirically using a subset of good quality DNA samples with high concentrations and throughout the process of producing the RAD libraries. The Bioruptor can produce variable results.

10. Remove the tube holder return it to the refrigerator and leave the samples there while you clean up the BioRuptor water bath and associated pieces
11. Disassemble the tube holder.
    a. Return the blank tubes to the rack
    b. Put the tube holder in the drying rack
    c. Vortex and centrifuge the sonication product
12. If you have more sonication runs that day, put the BioRuptor sample holder and the 2 L beaker back in the refrigerator to chill.
13. Combine the sonication products (Tube A and B) for each library into a separate labelled 1.5mL Eppendorf LoBind tube.
14. Store the combined sonication products for each library in the fridge (4 ˚C) until the sonication step is completed for all libraries.

<u>DNA shearing verification</u>

1. For each library, run 2 µL on a HS D1000 ScreenTape Assay on the 4200 TapeStation
    a. Sample Preparation
        i. Allow HS D1000 Sample Buffer and HS D1000 Ladder to warm to room temperature for 30 minutes.

ii. Use Agilent strip tubes and add 2 µL of the HS D1000 Sample buffer to position A1, then add 2 µL of HS D1000 Ladder to the same tube.
iii. Add 2 µL of the HS D1000 Sample buffer to position B1 and then add 2 µL of sonicated library pool to the same tube.
iv. If running more than one library, continue through position C1, D1, etc.
v. Cap the strip tubes.
vi. Mix the strip tubes at 2000 rpm for 1 minute on the IKA MS3 vortex.
vii. Centrifuge the strip tubes for 30 s at 1500 rpm.

b. Instrument preparation:
   i. Install the tip rack and remove lid.
   ii. Setup the software
      1. On the image of the plate, select the wells to be run
         a. The Ladder well in the strip tube will be automatically labelled.
         b. On the right side there will be a list of the wells to be run
            i. Select the "…" to the right of "Description"
            ii. Select the option to import your sample IDs
      2. Check under the "Required For Run" section. It will list the number of ScreenTapes, tips and the amount of ladder required. Make sure you have loaded sufficient supplies.
      3. Add in the lot information for the ScreenTape and reagents
      4. In Notes: put the Sonicated Verification Library #s being run.
      5. In Prefix: Put your project identifier.
      6. Start the run.

2. The TapeStation results should show smearing in the lane, and you won't see distinct bands.
   a. You are looking for the DNA ***smear*** concentrated around 300-500 bp (see **Appendix 1 – Figure 2.**
3. If the sonication product is not around 300-500 bp, repeat the sonication step starting from the multiplexed library (the purified/pooled ligation samples). Adjust the sonication settings as required:
   a. If the sonication product was too large (>300-500 bp), increase the number of cycles. This can be done by decreasing the sonication "off" time (<u>total time must not exceed 10 minutes</u>). Do not reduce the "off" time below 15 sec because a lot of heat is produced by sonication and sufficient time must be given for it to cool off between the "on" times.
   b. If the sonication product was too small (<300 bp), reduce the number of cycles. This can be done by increasing the sonication "off" time (<u>total time must not exceed 10 minutes</u>).
   c. Use 5 s increments to adjust the "off" time.
   d. Aim for a mean fragment size between 300-500 bp.
4. Export the .csv file with the sample IDs and image and the report, include all options in the pdf (It is likely you won't need this, but if there are issues identified later, it is good to have for reference.).
5. If you have to stop, put the libraries in the freezer (-20 °C).

<u>Isolation of RAD-tag DNA fragments</u>

1. Prepare reagents prior to start

a. Make two 10 mL tubes of 1X binding wash buffer. Heat one tube of the 1X binding wash buffer to 56 °C (refer to Reagents and Solutions: Isolation of RAD-tagged DNA fragments for recipe).
b. Make 250 µL 1X NEBuffer 4 (from 10X stock) per library using Ultrapure water in a 1.5mL tube.

2. Turn on the Thermomixer C and let it warm to 37 °C. Load the following protocol:
    a. 37 °C 1 hr
3. Resuspend Dynabeads M-280 Streptavidin for at least 30 s - 1 min on the vortex.
4. Transfer 30 µL of Dynabeads into a new 1.5m L LoBind tube.
5. Place tube on magnet and remove supernatant using a pipette.
6. Wash the beads with 2X binding and wash buffer (refer to Reagents and Solutions: Isolation of RAD-tagged DNA fragments for recipe):
    a. Remove from magnet and add 100 µL of 2X binding and wash buffer.
    b. Pipette up and down 10x to mix.
    c. Place the tube on a magnet for at least 1 min and discard the supernatant.
    Note: The solution should be clear before discarding the supernatant.
7. Repeat step 6 for a total of two washes with 2X binding and wash buffer.
8. Resuspend the Dynabeads in 200 µL of 2X binding and wash buffer.
9. Add the sheared DNA library (combined Tubes A and B from DNA Shearing, 200 µL) to the washed Dynabeads (create the bead/DNA mixture).
10. Incubate the bead/DNA mixture for 20 min at room temperature and mix with gentle occasional mixing (*e.g.,* gently flicking).
11. Place the tube on the magnet for 2-3 min and discard the supernatant.
    Note: The solution should be clear before discarding the supernatant.
12. Wash the bead/DNA mixture with room temperature 1X binding and wash buffer:
    a. Remove tube from magnet and resuspend the bead/DNA mixture with 150 µL of room temperature 1X binding and wash buffer.
    b. Place bead/DNA mixture on magnet for 1 min or until the solution is clear and discard the supernatant.
13. Repeat step 12, for a total of two washes with room temperature 1X binding and wash buffer.
14. Wash the bead/DNA mixture with 56 °C 1X binding and wash buffer:
    a. Remove the tube from magnet and resuspend bead/DNA mixture with 150 µL of the 56 °C 1X binding and wash buffer.
    b. Place bead/DNA mixture on magnet for 1 min or until the solution is clear and discard the supernatant.
15. Repeat step 14, for a total of two washes with 56 °C 1X binding and wash buffer.
16. Wash the bead/DNA mixture with 1X NEBuffer 4:
    a. Remove the tube from magnet and resuspend bead/DNA mixture with 100 µL 1X NEBuffer 4.
    b. Place bead/DNA mixture on a magnet for 1 min or until the solution is clear and discard the supernatant.
17. Repeat step 16, for a total of two washes with 1X NEBuffer 4.
18. Prepare *Sbf*I mix as follows:

Reagent	**1X**	**2X + extra**	**Final concentration in Reaction**
Ultrapure H_2O	34 µL	69.7 µL	

10X NEBuffer 4	4 µL	8.2 µL	1X
*Sbf*I-HF (20 U/µL)	2 µL	4.1 µL	40U

19. Remove tube from magnet and add 40 µL of the *Sbf*I mix.
20. In the Thermomixer C, incubate the bead/DNA mixture at 37 °C for 1 hr.
21. Place the tube on the magnet for 1 min or until the solution is clear and then transfer the supernatant to a new LoBind 1.5 mL tube.
22. **If you have to stop here, put the samples in the freezer (-20 °C).**

Purify RAD-tag pooled DNA samples

1. Prepare reagents prior to start:
    a. Allow Agencourt AMPure XP beads come to room temperature prior to use.
    b. Prepare **fresh** 80% ethanol using Ultrapure H_2O and absolute ethanol (350 µL per library).
    c. Gently vortex the AMPure XP bottle to resuspend the beads. It should appear homogenous and consistent in colour.
2. Add equal amount (40 µL) of AMPure XP beads into the RAD-tagged pooled samples.
3. Mix via pipette 10 times. The colour of the solution should appear homogenous after mixing.
4. Let the solution incubate at room temperature for 5 min.
5. Place the tube on the magnet for ~2 min to separate the beads from the solution
    Note: Wait for the solution to be clear before proceeding to the next step
6. Leaving the tube on the magnet, remove the supernatant (clear solution) and discard.
    Note: Do not disturb the magnetic beads
7. Wash the bead/DNA mixture with **fresh** 80% ethanol:
    a. Leaving the tube on the magnet, dispense 150 µL of **fresh** 80% ethanol to the tube and incubate for 30 s at room temperature.
    b. Remove the ethanol via pipette and discard.

    Note: Do not disturb the magnetic beads, more easily done this time
8. Repeat step 7 for total of two washes with **fresh** 80% ethanol.
9. Remove tube from the magnet, and then add 60 µL of low TE and mix via pipetting 10 times (refer to Reagents and Solutions; Purify RAD-tag pooled DNA samples for prep.).
10. Incubate for 2 min at room temperature.
11. Place tube on magnet for 1 min to separate the beads from the solution.
12. Once the solution is clear, transfer the supernatant to a new LoBind 0.5 mL tube.
    a. Can store in 4 °C for short term (few days) or -20 °C longer term (days to weeks)

Quantification of RAD-tag pooled DNA samples

Note: This uses the Qubit dsDNA HS Kit

1. Quantify **5 µL** of the library using the Qubit dsDNA HS Assay following kit instructions.

NEBNext Ultra II DNA Library Prep for Illumina

Note: Starting material should be 500 pg – 1 µg DNA. If the DNA volume post shearing is less than 50 µL, add Low TE to a final volume of 50 µL.

NEBNext End Prep (End Repair)

1. Turn on the Thermomixer C. Load the following protocol:
    a. 20 °C for 30 min, shake at 300 rpm
    b. 65 °C for 30 min, shake at 300 rpm
    c. Hold at 4 °C
2. For each library combine the following in a 0.5 mL tube:

Component	**Cap colour**	**Volume**
RAD-tag pooled DNA		50 µL
NEBNext Ultra II End Prep Reaction Buffer	Green	7 µL
NEBNext Ultra II End Prep Enzyme Mix	Green	3 µL
Total Volume		**60 µL**

3. Set your 200 µL pipette to 50 µL and pipette the entire volume up and down at least 10 times.
4. Briefly spin the tubes to bring down all the liquid.
5. Put the tubes in the ThermoMixer C and run the loaded protocol from Step 1.
6. Proceed immediately to the Adaptor Ligation step.
   Note: It is helpful to have a second ThermoMixer C loaded with the next program to avoid waiting for the temperature to ramp up/down.

Adaptor Ligation

1. If the total amount of DNA in the RAD-tag pooled DNA is less than 5 ng than the NEBNext adaptor for Illumina (provided at 15 µM) needs to be diluted 25-Fold (1:25) using Tris/NaCl (pH 7.5-8) to give a working concentration of 0.6 µM (refer to Reagents and Solutions; Buffer for diluting Adaptor in NEBNext Ultra II for prep.).
    a. The NEBNext adaptors are found in the NEBNext oligos kit.
2. Load the Adaptor Ligation program on the ThermoMixer C
    a. 20 °C for 15 minutes with no heated lid @ 300 rpm
3. For each library combine the following, directly into the End Prep reaction mixture:
    a. Note: Mix the Ultra II Ligation Master Mix by pipetting up and down several times prior to adding to the reaction.

Component	Cap Colour	Volume
End Prep reaction mixture		60 µL
NEBNext Adaptor for Illumina (diluted 1:25)	Red	2.5 µL
NEBNext Ultra II Ligation Master Mix	Red	30 µL
NEBNext Ligation Enhancer	Red	1 µL
Total		93.5 µL

*Be sure to wait until the ThermoMixer C is at the proper temperature before combining all components.

4. Set a 200 µL pipette to 80 µL and then pipette the entire volume up and down at least 10 times to mix thoroughly. Perform a quick spin to collect all liquid from the sides of the tube.
    a. Note: the NEBNext Ultra II Ligation KMaster Mix is very viscous. Care should be taken to ensure adequate mixing of the ligation reaction, as incomplete mixing will result in reduced ligation efficiency.
5. Put the tubes in the ThermoMixer C and run program Adaptor Ligation.
6. Remove the tubes from the ThermoMixer C.
7. **This is a safe stopping point, if you are not proceeding directly to the next step, store the reactions at -20 °C.**

8. Load the Adaptor 2 program on the ThermoMixer C and add 3 µL of USER Enzyme (red cap) to the ligation mixture.
    a. 37 °C for 15 minutes with heated lid @ 300 rpm
9. Vortex and centrifuge to bring the liquid to the bottom of the tube.
10. Put the tubes in the ThermoMixer and run program Adaptor 2
11. **This is a safe stopping point, if you are not proceeding directly to the next step, store the reactions at -20 °C.**

<u>Clean-up of Adaptor-ligated DNA (For input ≤ 50 ng)</u>

1. Prepare reagents prior to start:
    a. Allow Agencourt AMPure XP beads come to room temperature prior to use.
    b. Prepare **fresh** 80% ethanol using Ultrapure H_2O and absolute ethanol
    c. Gently vortex the AMPure XP bottle to resuspend the beads. It should appear homogenous and consistent in colour.
2. Move 96.5 µL of product from the Adapter Ligation step to 1.5 mL tubes.
    a. Measure exactly how much volume is transferred (may be < 96.5 µL)
3. Add 0.9X volumes of AMPure XP beads to sample volume

e.g., if sample volume is 96.5 µL, add 87 µL beads

4. Mix by pipetting up and down 10 times. Be careful to expend all of the liquid out of the tip during the last mix.
5. Incubate the samples on the bench for 5 minutes at room temperature.
6. Place the tube on the magnetic stand for 5 minutes. Remove and discard the supernatant.
7. Add 200 µL of **fresh** 80% absolute ethanol while the tube is on the magnet. Incubate at room temperature for 30 seconds. Then carefully remove and discard the supernatant. Be careful not to disturb the beads (they contain the DNA targets).
8. Repeat step 7 once, for a total of two washes. Make sure to remove the last traces of ethanol. If necessary, spin down tube briefly and place back on magnet and remove any traces of ethanol using a p10 tip.
9. Air dry the beads for up to 5 min, usually 2-3 min (with the tube on the magnet and the lid open).
    a. Do not over-dry the beads. This may result in lower recovery of DNA target. Elute the samples when the beads are still dark brown and glossy looking, but when all visible liquid has evaporated. When the beads turn lighter brown and start to crack, they are too dry.
10. Remove the tube from the magnet. Elute the DNA from the beads by adding 17 µL of Qiagen Buffer EB (10 mM Tris-HCl).
11. Mix well by pipetting up and down 10 times.
12. Incubate for at least 2 minutes at room temperature.
13. Place the tube on the magnet. After 5 minutes (or when solution is clear), transfer 15 µL to a breakaway PCR plate – this is the adaptor ligated DNA fragments used as the target for the PCR. Or if not proceeding to the PCR enrichment step than put in 0.5 mL LoBind tube instead.
14. **This is a safe stop point. The libraries can be stored at -20 °C.**

PCR enrichment of Adaptor-ligated DNA

Note: Verify that the concentration of the index oligos are 10 µM

1. Sign-up for a PCR machine.
2. Ensure the correct (NEBNext Ultra II PCR) program has been setup on the machine:
    a. 98 °C for 30 seconds
    b. # cycles of:
        i. 98 °C for 10 seconds
        ii. 65 °C for 1 min 15 seconds (75 seconds)
    c. 65 °C for 5 minutes
    d. 10 °C hold

 # cycles based on amount of DNA input:

5 ng = 7-8 cycles
1 ng = 9-10 cycles
0.5 ng = 10-11 cycles

3. To each well add:

Component	Tube Colour	Volume
Adaptor Ligated DNA fragments		15 µL
NEBNext Ultra II Q5 Master Mix	Blue	25 µL
Index Primer/i7 Primer	Blue	5 µL
Universal PCR primer/i5 Primer	Blue	5 µL
Total		50 µL

4. Set a 200 µL pipette to 40 µL and pipette up and down 10 times to mix.
5. Briefly spin the plate to bring down the liquid
6. Put the plate in the PCR machine and run the program.

Cleanup of PCR Reaction

1. Prepare reagents prior to start:
    a. Allow Agencourt AMPure XP beads come to room temperature prior to use.
    b. Prepare **fresh** 80% ethanol using Ultrapure H_2O and absolute ethanol
    c. Gently vortex the AMPure XP bottle to resuspend the beads. It should appear homogenous and consistent in colour.
2. Spin down the PCR products, then transfer the products (~50 µL*) to 1.5 mL LoBind labelled tubes.
    a. *Measure the exact volume of PCR product transferred with pipette.
3. Add 0.9X volumes of AMPure XP beads to sample volume

e.g., if sample volume is 50 µL, add 45 µL beads

4. Mix by pipetting up and down at least 10 times.
5. Incubate the samples for 5 minutes at room temperature.
6. Place the tubes on the magnet for ~5 minutes (or when the solution is clear), carefully remove and discard the supernatant.
    a. Note: Be careful not to disturb the beads that contain the target DNA.
7. Add 200 µL of freshly prepared 80% ethanol to each tube while on the magnet.
8. Incubate for 30 seconds.
9. Remove and discard the supernatant. Be careful not to disturb the beads.
10. Repeat wash steps 7-9. Make sure all the all ethanol has been removed.
11. Air dry the bead for up to 5 min, but 2-3 min is usually good.
    a. Note: Do not over dry the beads. This will result in lower DNA recovery.
    b. Beads should still be dark brown and glossy but visible liquid has evaporated. When the beads turn lighter brown and start to crack, they are too dry.
12. Remove the tubes from the magnet and add 33 µL of 0.1X TE (Refer to Reagents and Solutions; Clean-up of PCR reaction for prep.).
13. Mix by pipetting up and down 10 times.
14. Incubate for 2 minutes at room temperature.

15. Put the tube on the magnet for 2-3 min (or when the solution is clear), carefully transfer 30 µL to a new 0.5 mL tube.
16. **If not proceeding directly to library preparation verification, store samples at -20 °C.**

Library Preparation Verification:

1. TapeStation check
    - Run 2 µL test amplified library on the 4200 TapeStation using the High Sensitivity D1000 ScreenTape (see DNA Shearing Verification above)
    - Verify that the smear of the DNA is around 300 to 500bp but should have a longer average fragment length than post-shearing DNA as adaptor and index sequences have been added (see Appendix 1 – Figure 3). There will be some variation in smear size library to library.
    - Highlight the smear region using the Tapestation analysis software "Edit Regions" function to identify the average fragment size of the library
        - This is the size in bp that you will use to calculate final library concentration (in nM) and may be needed to provide on sequencing centre.

Quantify Libraries using qPCR

1. Quantify the library using NEBNext Library Quant kit for Illumina with 6 standards. Followed instruction manual (see below):
    - Thaw and Mix Reagents
        - Thaw the NEBNext Library Quant Master Mix and NEBNext Library Quant Primer Mix. Ensure mixing of NEBNext Library Quant Primer mix by vortexing for 5 seconds. Centrifuge briefly and place on ice.
        - Thaw NEBNext Library Quant DNA Standards 1-6. Mix well by inverting 3-5 times. Centrifuge briefly and place on ice.
        - Thaw the NEBNext Library Quant Dilution Buffer (10X). Mix well by vortexing for 10 seconds. Centrifuge briefly and place on ice.
    - Preparation of the NEBNext Library Quant Master Mix (+ primers)
        - NOTE: This is done once per kit.
        - Prepare the NEBNext Library Quant Master Mix (+ primers) by adding 100 µL NEBNext Library Quant Primer Mix to a tube of NEBNext Library Quant Master Mix (1.5 mL) and mix by vortexing for 10 seconds.
        - Record the date the primers were added.
        - Mix can be stored at -20 °C for seven months and 30 freeze/thaw cycles.
        - Add 20 µL of Low ROX to the Library Quant Master Mix tube.
    - Prepare NEBNext Library Quant Dilution Buffer (1X)
        - Prepare the NEBNext Library Quant Dilution Buffer (1X) by making a 1:10 dilution of the 10X buffer in Ultrapure water. Prepare sufficient buffer for the desired number of libraries to be quantitated, allowing 1.4 mL for each library.
    - Prepare Library dilutions
        - Prepare an initial 1:1,000 dilution of each library sample by adding 1 µL library sample to 999 µL NEBNext Library Quant Dilution Buffer (1X)
        - Add 10 µL of the 1:1,000 dilution to 90 µL NEBNext Library Quant Dilution Buffer (1X) to create 1:10,000 dilution

- Add 10 µL of the 1:10,000 dilution to 90 µL NEBNext Library Quant Dilution Buffer (1X) to create 1:100,000 dilution.

Ingredient	1:1,000 (µL)	1:10,000 (µL)	1:100,000 (µL)
Library	1	10	10
NEBNext Library Quant Dilution Buffer (1X)	999	90	90
Total:	1000	100	100

- Prepare qPCR Assay

  Note: Each DNA standard, library sample and no-template control (NTC) should be done in triplicate.

  - Prepare DNA standards and diluted library samples:
    - Add 16 µL per well NEBNext Library Quant Master Mix (+ primers)
    - Add 4 µL DNA standard or library dilution
  - Prepare NTC
    - Add 16 µL per well NEBNext Library Quant Master Mix (with primers)
    - Add 4 µL NEBNext Library Quant Dilution Buffer (1X)
  - Mix reactions via pipetting at least 5x. Try to minimize bubbles in plate wells, but 1-2 bubbles per well will be removed by heating and not affect results.
  - Centrifuge for 5 min at 2500-3000 rpm.
- Setting up qPCR run in the QuantStudio Real-Time PCR software file
  - Instrument: QuantStudio 7
  - Block used: Fast 96-well
  - Type of Experiment: Standard Curve
  - Reagents: SYBR Green Reagents
  - Properties: Fast
  - Assign standards, libraries, NTCs in triplicate to the plate well
    - Task U = libraries
    - Task S = standards, define quantity
    - Task N = NTC
  - Run Method
    - Initial denaturation = 95 °C for 1 min
    - For 35 cycles:
      - Denaturation = 95 °C for 15 s
      - Extension = 63 °C for 45 s

Appendix 1 References

Omar A Ali and others, RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping, Genetics, Volume 202, Issue 2, 1 February 2016, Pages 389–400

Agilent Genomic DNA ScreenTape Assay Quick Guide for 4200 TapeStation System.

Agilent HS D1000 ScreenTape Assay Quick Guide for 4200 TapeStation System.

Agilent HS D1000 ScreenTape Assay Quick Guide for 4200 TapeStation System.

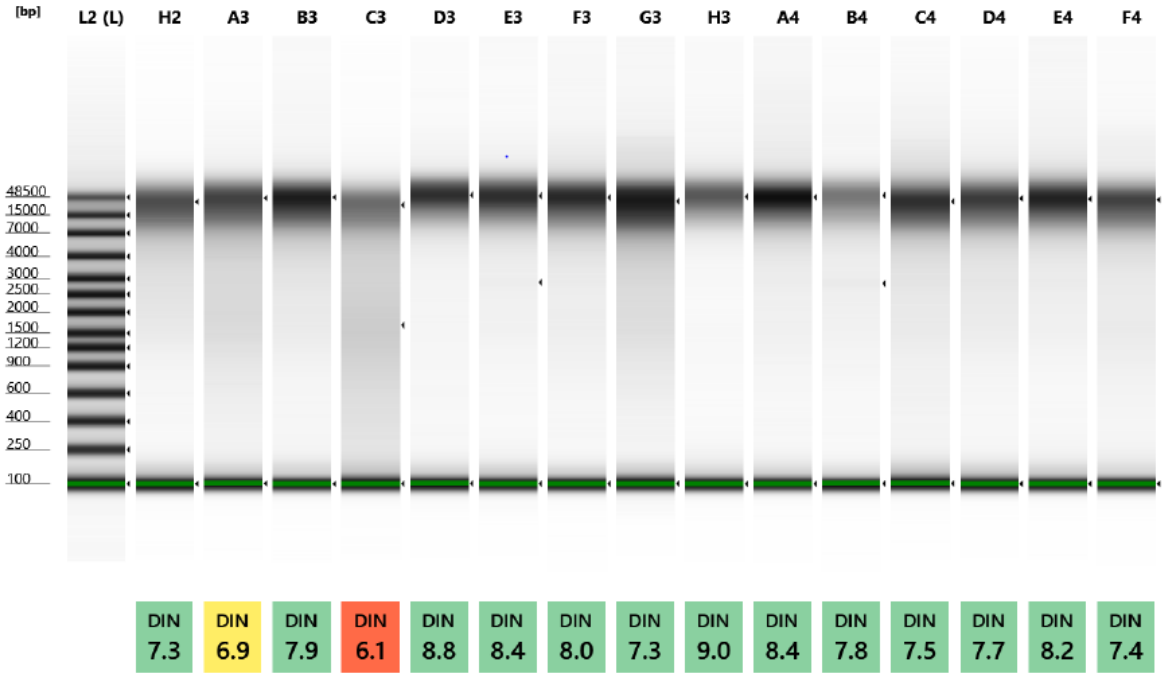
DNeasy 96 Blood & Tissue kit Handbook, QIAGEN

NEBNext Ultra II DNA Library Prep kit for Illumina Instruction Manual, New England BioLabs, version 6.1, May 2020

Appendix 1 - Figures

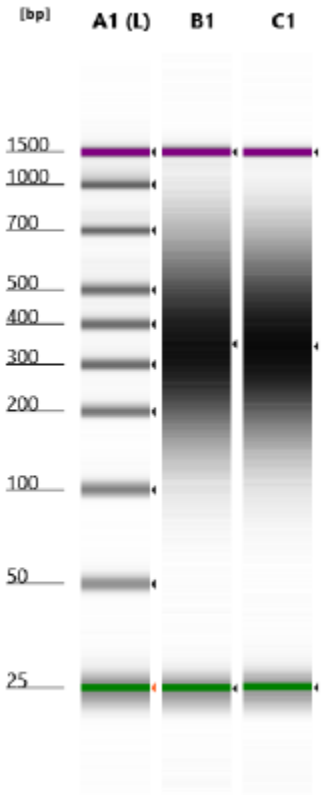
Appendix 1 – Figure 1.

Examples of gDNA quality criteria for RAD-seq, using the Agilent TapeStation. Samples with ideal DNA quality are those with a clean, well-defined band at high molecular weight (top of lane) with little or no DNA smearing (e.g., lanes B3, D3, E3, etc.). Acceptable quality gDNA have a DIN of around 6.5 or higher. DNA quality of lower grades (e.g., DIN ~/< 6.0) are not acceptable for RAD-seq (e.g., lane C3).



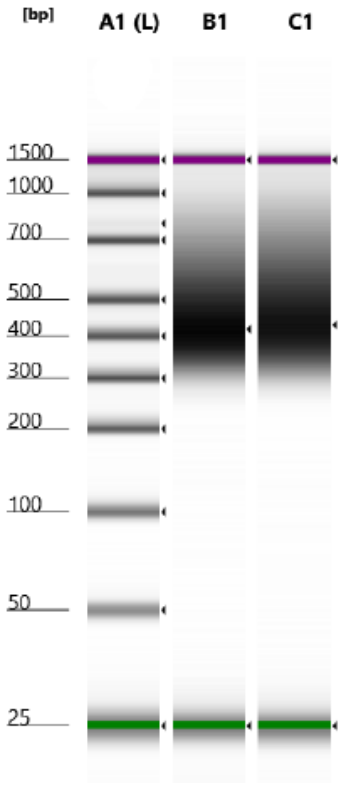
Appendix 1 – Figure 2.

Examples of post-shearing adaptor-ligated DNA, using the Agilent TapeStation. The fragmented DNA should appear as a smear of varying fragment lengths, but most concentrated (intense) between 300-500 bp in length. Aim for most abundant fragment length (arrows on right side of lanes) to be between 300-400 bp. For the DNA samples below, peak fragment lengths were found to be 352 bp (lane B1) and 345 bp (lane C1), which is a good result.



Appendix 1 – Figure 3.

Examples of final prepared libraries, using the Agilent TapeStation. The completed library DNA should appear as a smear of varying fragment lengths, but most concentrated (intense) between 350-550 bp in length. The most abundant fragment length (arrows on right side of lanes) should be between 350-500 bp as the library preparation adds length to the post-shearing DNA. The final library purification step should remove DNA fragments < 300 bp, so low molecular weight fragments should not be present.



Appendix 1 – Table 1.

RAD-Cap Barcode Primer (Barcode) sequences (use standard desalting). Primers were selected from Ali et al. 2016 to be used in this project (B1-B50). Note "5Biosg" on the forward sequence is biotin followed by 6-bp filler nucleotides (black) followed by the SbfI restriction site (green; used to release the fragments from the streptavidin beads), the 8-bp barcode (blue), and lastly the other half of the SbfI site (red; the overhang).

Name	Forward	Name	Reverse
RAD-Cap B1F	/5Biosg/GTACGTCCTGCAGGACAAGCTATGCA	RAD-Cap B1R	/5Phos/TAGCTTGTCCTGCAGGACGTAC
RAD-Cap B2F	/5Biosg/GTACGTCCTGCAGGCGCTGATCTGCA	RAD-Cap B2R	/5Phos/GATCAGCGCCTGCAGGACGTAC
RAD-Cap B3F	/5Biosg/GTACGTCCTGCAGGAGTGGTCATGCA	RAD-Cap B3R	/5Phos/TGACCACTCCTGCAGGACGTAC
RAD-Cap B4F	/5Biosg/GTACGTCCTGCAGGACCTCCAATGCA	RAD-Cap B4R	/5Phos/TTGGAGGTCCTGCAGGACGTAC
RAD-Cap B5F	/5Biosg/GTACGTCCTGCAGGATCCTGTATGCA	RAD-Cap B5R	/5Phos/TACAGGATCCTGCAGGACGTAC
RAD-Cap B6F	/5Biosg/GTACGTCCTGCAGGCAAGACTATGCA	RAD-Cap B6R	/5Phos/TAGTCTTGCCTGCAGGACGTAC
RAD-Cap B7F	/5Biosg/GTACGTCCTGCAGGCACTTCGATGCA	RAD-Cap B7R	/5Phos/TCGAAGTGCCTGCAGGACGTAC
RAD-Cap B8F	/5Biosg/GTACGTCCTGCAGGCCTCCTGATGCA	RAD-Cap B8R	/5Phos/TCAGGAGGCCTGCAGGACGTAC
RAD-Cap B9F	/5Biosg/GTACGTCCTGCAGGCGACTGGATGCA	RAD-Cap B9R	/5Phos/TCCAGTCGCCTGCAGGACGTAC
RAD-Cap B10F	/5Biosg/GTACGTCCTGCAGGCTGAGCCATGCA	RAD-Cap B10R	/5Phos/TGGCTCAGCCTGCAGGACGTAC
RAD-Cap B11F	/5Biosg/GTACGTCCTGCAGGGACTAGTATGCA	RAD-Cap B11R	/5Phos/TACTAGTCCCTGCAGGACGTAC
RAD-Cap B12F	/5Biosg/GTACGTCCTGCAGGGAGCTGAATGCA	RAD-Cap B12R	/5Phos/TTCAGCTCCCTGCAGGACGTAC
RAD-Cap B13F	/5Biosg/GTACGTCCTGCAGGGCGAGTAATGCA	RAD-Cap B13R	/5Phos/TTACTCGCCCTGCAGGACGTAC
RAD-Cap B14F	/5Biosg/GTACGTCCTGCAGGGGTGCGAATGCA	RAD-Cap B14R	/5Phos/TTCGCACCCCTGCAGGACGTAC

RAD-Cap B15F	/5Biosg/GTACGTCCTGCAGGGTACGCAATGCA	RAD-Cap B15R	/5Phos/TTGCGTACCCTGCAGGACGTAC
RAD-Cap B16F	/5Biosg/GTACGTCCTGCAGGTAGGATGATGCA	RAD-Cap B16R	/5Phos/TCATCCTACCTGCAGGACGTAC
RAD-Cap B17F	/5Biosg/GTACGTCCTGCAGGTCTTCACATGCA	RAD-Cap B17R	/5Phos/TGTGAAGACCTGCAGGACGTAC
RAD-Cap B18F	/5Biosg/GTACGTCCTGCAGGTGAAGAGATGCA	RAD-Cap B18R	/5Phos/TCTCTTCACCTGCAGGACGTAC
RAD-Cap B19F	/5Biosg/GTACGTCCTGCAGGTGGCTTCATGCA	RAD-Cap B19R	/5Phos/TGAAGCCACCTGCAGGACGTAC
RAD-Cap B20F	/5Biosg/GTACGTCCTGCAGGTTCACGCATGCA	RAD-Cap B20R	/5Phos/TGCGTGAACCTGCAGGACGTAC
RAD-Cap B21F	/5Biosg/GTACGTCCTGCAGGAAGGACACTGCA	RAD-Cap B21R	/5Phos/GTGTCCTTCCTGCAGGACGTAC
RAD-Cap B22F	/5Biosg/GTACGTCCTGCAGGAGGCTAACTGCA	RAD-Cap B22R	/5Phos/GTTAGCCTCCTGCAGGACGTAC
RAD-CAP B23F	/5Biosg/GTACGTCCTGCAGGAAACATCGTGCA	RAD-CAP B23R	/5Phos/CGATGTTTCCTGCAGGACGTAC
RAD-CAP B24F	/5Biosg/GTACGTCCTGCAGGCAGATCTGTGCA	RAD-CAP B24R	/5Phos/CAGATCTGCCTGCAGGACGTAC
RAD-CAP B25F	/5Biosg/GTACGTCCTGCAGGATGCCTAATGCA	RAD-CAP B25R	/5Phos/TTAGGCATCCTGCAGGACGTAC
RAD-CAP B26F	/5Biosg/GTACGTCCTGCAGGAACAACCATGCA	RAD-CAP B26R	/5Phos/TGGTTGTTCCTGCAGGACGTAC
RAD-CAP B27F	/5Biosg/GTACGTCCTGCAGGACGCTCGATGCA	RAD-CAP B27R	/5Phos/TCGAGCGTCCTGCAGGACGTAC
RAD-CAP B28F	/5Biosg/GTACGTCCTGCAGGACTATGCATGCA	RAD-CAP B28R	/5Phos/TGCATAGTCCTGCAGGACGTAC
RAD-CAP B29F	/5Biosg/GTACGTCCTGCAGGAGATCGCATGCA	RAD-CAP B29R	/5Phos/TGCGATCTCCTGCAGGACGTAC
RAD-CAP B30F	/5Biosg/GTACGTCCTGCAGGAGCAGGAATGCA	RAD-CAP B30R	/5Phos/TTCCTGCTCCTGCAGGACGTAC
RAD-CAP B31F	/5Biosg/GTACGTCCTGCAGGATTGAGGATGCA	RAD-CAP B31R	/5Phos/TCCTCAATCCTGCAGGACGTAC
RAD-CAP B32F	/5Biosg/GTACGTCCTGCAGGCAACCACATGCA	RAD-CAP B32R	/5Phos/TGTGGTTGCCTGCAGGACGTAC

RAD-CAP B33F	/5Biosg/GTACGTCCTGCAGGCATACCAATGCA	RAD-CAP B33R	/5Phos/TTGGTATGCCTGCAGGACGTAC
RAD-CAP B34F	/5Biosg/GTACGTCCTGCAGGCCAGTTCATGCA	RAD-CAP B34R	/5Phos/TGAACTGGCCTGCAGGACGTAC
RAD-CAP B35F	/5Biosg/GTACGTCCTGCAGGCTCAATGATGCA	RAD-CAP B35R	/5Phos/TCATTGAGCCTGCAGGACGTAC
RAD-CAP B36F	/5Biosg/GTACGTCCTGCAGGGCCACATATGCA	RAD-CAP B36R	/5Phos/TATGTGGCCCTGCAGGACGTAC
RAD-CAP B37F	/5Biosg/GTACGTCCTGCAGGGCTCGGTATGCA	RAD-CAP B37R	/5Phos/TACCGAGCCCTGCAGGACGTAC
RAD-CAP B38F	/5Biosg/GTACGTCCTGCAGGGGAGAACATGCA	RAD-CAP B38R	/5Phos/TGTTCTCCCCTGCAGGACGTAC
RAD-CAP B39F	/5Biosg/GTACGTCCTGCAGGGTCGTAGATGCA	RAD-CAP B39R	/5Phos/TCTACGACCCTGCAGGACGTAC
RAD-CAP B40F	/5Biosg/GTACGTCCTGCAGGGTGTTCTATGCA	RAD-CAP B40R	/5Phos/TAGAACACCCTGCAGGACGTAC
RAD-CAP B41F	/5Biosg/GTACGTCCTGCAGGTATCAGCATGCA	RAD-CAP B41R	/5Phos/TGCTGATACCTGCAGGACGTAC
RAD-CAP B42F	/5Biosg/GTACGTCCTGCAGGTCCGTCTATGCA	RAD-CAP B42R	/5Phos/TAGACGGACCTGCAGGACGTAC
RAD-CAP B43F	/5Biosg/GTACGTCCTGCAGGAATCCGTCTGCA	RAD-CAP B43R	/5Phos/GACGGATTCCTGCAGGACGTAC
RAD-CAP B44F	/5Biosg/GTACGTCCTGCAGGATAGCGACTGCA	RAD-CAP B44R	/5Phos/GTCGCTATCCTGCAGGACGTAC
RAD-CAP B45F	/5Biosg/GTACGTCCTGCAGGCCGACAACTGCA	RAD-CAP B45R	/5Phos/GTTGTCGGCCTGCAGGACGTAC
RAD-CAP B46F	/5Biosg/GTACGTCCTGCAGGCCTCTATCTGCA	RAD-CAP B46R	/5Phos/GATAGAGGCCTGCAGGACGTAC
RAD-CAP B47F	/5Biosg/GTACGTCCTGCAGGCGGATTGCTGCA	RAD-CAP B47R	/5Phos/GCAATCCGCCTGCAGGACGTAC
RAD-CAP B48F	/5Biosg/GTACGTCCTGCAGGCTAAGGTCTGCA	RAD-CAP B48R	/5Phos/GACCTTAGCCTGCAGGACGTAC
RAD-CAP B49F	/5Biosg/GTACGTCCTGCAGGGAACAGGCTGCA	RAD-CAP B49R	/5Phos/GCCTGTTCCCTGCAGGACGTAC
RAD-CAP B50F	/5Biosg/GTACGTCCTGCAGGGATGAATCTGCA	RAD-CAP B50R	/5Phos/GATTCATCCCTGCAGGACGTAC

APPENDIX 2 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR MINION NANOPORE LONG-READ SEQUENCING, WITH MODIFICATIONS FOR HORSE MUSSEL

Note: This protocol was developed for MinION Flow Cell R10.4.1 and the Ligation Sequencing Kit V14. It may require modification based on changes to manufacturer's protocols since it was developed in April 2023.

Safety Considerations

General

Personnel should have proper training from a lab manager or other experienced individuals.

All relevant training, instruction, safe work procedures and chemical SDS should be read prior to work.

Equipment

General

- Aluminum Beads (refrigerated; and their associated freezer pack)
- Pipettes
    - Rainin 2 µL (Mettler Toledo catalogue# 17014393)
    - Rainin 20 µL (Mettler Toledo catalogue# 17014392)
    - Rainin 200 µL (Mettler Toledo catalogue# 17014391)
    - Rainin 1000 µL (Mettler Toledo catalogue# 17014382)
- Personal microfuge
- Vortex
- Rotating mixer (hula mixer)
- Magnet for AMPure XP beads
- ThermoMixer or thermocycler

Supplies

General

- Eppendorf DNA LoBind 1.5 mL tubes (Fisher Scientific catalogue#
- 0.5 mL thin wall tubes
- Tips
    - Rainin 20 µL filter tips (Mettler Toledo catalogue# 30389274)
    - Rainin 200 µL filter tips (Mettler Toledo catalogue# 30389276)
    - Rainin 1000 µL filter tips (Mettler Toledo catalogue# 30389272)

Kits

- Control Expansion Kit (Oxford Nanopore catalogue# EXP-CTL001)
- Flow Cell Wash Kit (Oxford Nanopore catalogue# EXP-WSH004)
- Flow Cell (Oxford Nanopore catalogue# FLO-MIN114)
- Ligation Sequencing Kit V14 (Oxford Nanopore catalogue# SQK-LSK114)
- NEBNext Companion Module Ligation Sequencing (NEB catalogue# E7180S)

Reagents

- Ultrapure Water (Invitrogen) (Thermo Fisher Scientific catalogue# 10977023)

- Absolute ethanol

Procedures

NOTE: This procedure has been created using OxfordNanopore documents from the Nanopore Community site, https://community.nanoporetech.com/. In addition to instruction pages, the site has many useful protocol videos for different MinION applications.

If this is your first time running the MinION, the Lambda Control Experiment should be run to practice library prep, flow cell loading, run monitoring, and analysis

Check the Flow Cell - Approx. 20 minutes

This should be done within 3 months of purchasing the flow cell.

NOTE: Before beginning this procedure, make sure the MinION has been plugged in to the computer, MinKNOW has been installed, and the Configuration Test Cell check has been run (video instructions are available on the Nanopore Community site). You will need a valid Nanopore Community account to do this, which you can register for at https://community.nanoporetech.com/.

1. Open the MinION Mk1B/GridION lid and insert the MinION Flow Cell.
    a. Insert the flow cell in the MinION Mk1B/GridION by sliding the flow cell under the clip. Firmly press down on the flow cell to ensure good thermal and electrical contact.
2. Connect the assembled MinION Mk1B and flow cell to the host computer. Once successfully plugged in, you will see a light and hear the fan.
3. Open MinKNOW software on the computer and log in using your community credentials.
4. Select the sequencing device connected to the computer. The Sequencing Overview page should show the flow cell has not had any checks carried out.
5. Navigate to the start homepage and select 'Flow Cell Check'.
6. When you see the flow cell type and flow cell IDs recognised, click 'Start' to begin.
7. You will be automatically navigated to the Sequencing Overview page. A loading bar will be displayed under the flow cell during the checks.
    a. The flow cell check should take a few minutes.
8. The quality of the flow cell will be shown as one of the three outcomes:
    a. Yellow exclamation mark: The number of sequencing pores is below warranty. Take the flow cell out of the device, re-insert it and run a flow cell check again. If the flow cell is still below warranty, contact support@nanoporetech.com
    b. Green tick: The number of sequencing pores is above warranty.
    c. Question mark: A flow cell check has not been run on the flow cell during this MinKNOW session.
9. **NOTE**: The indicator of quality will only remain visible during a MinKNOW session. Once the MinKNOW session has ended, the status of the flow cell will be erased.
10. Once the check is complete, the flow cell can be returned to its package and stored in the fridge until use.

Prepare gDNA library

- Oxford Nanopore recommends the following quality criteria for gDNA sequencing:
    - Purity as measured using Nanodrop
        - OD 260/280 of 1.8
        - OD 260/230 of 2.0–2.2
    - Average fragment size >30 kb

- Input mass 1 µg (100–200 fmol)
    - Use Qubit/Picogreen assay, not nanodrop
- If your DNA does not meet these requirements, further purification steps may be needed. Shorter fragment libraries can be run on the MinION, but read lengths will be reduced.

DNA Quality Screening on the Agilent TapeStation

1. Use the Genomic DNA ScreenTape Kit. Refer to the Agilent Genomic DNA ScreenTape Assay Quick Guide for 4200 TapeStation System.
    a. Combine 1 µL of gDNA with 10 µL of the Genomic DNA Sample Buffer in a 96-well plate (Agilent) provided for the TapeStation.
    b. Apply foil seal to the plate, mix for 1 min on the IKA vortex at 2000 rpm. Centrifuge for 1 min at 1500 rpm.
        i. Check for bubbles. If further vortexing and centrifugation do not remove the bubbles, carefully remove the foil seal and break the surface tension with a pipette tip.
    c. Load the samples onto the instrument. Follow instructions in the Quick Guide.
2. In general, criteria for good quality DNA are DNA Integrity Numbers (DINs) of around 6.0 or higher *and* nice clearly defined bands with little to no smearing in the lane. See Appendix 2 – Figure 1 for example images.

DNA Quantification

1. Quantify your sample using the Qubit dsDNA HS Assay following kit instructions.

DNA Repair and End-Prep – Approx. 1 hour

1. Thaw DNA Control Sample (DCS) and at room temperature, spin down, mix gently by pipetting, and place on ice.
    a. OxfordNanopore recommends using the DNA Control Sample (DCS) in your library prep for troubleshooting purposes. However, you can omit this step and make up the extra 1 µl with your sample DNA.
2. Prepare the NEBNext FFPE DNA Repair Mix and NEBNext Ultra II End Repair / dA-tailing Module reagents in accordance with manufacturer's instructions, and place on ice. ******DO NOT** vortex the FFPE DNA Repair Mix or Ultra II End Prep Enzyme Mix.****
    a. For optimal performance, NEB recommend the following:
        i. Thaw all reagents on ice.
        ii. Flick and/or invert the reagent tubes to ensure they are well mixed.
        iii. The Ultra II End Prep Buffer and FFPE DNA Repair Buffer may have a little precipitate. Allow the mixture to come to room temperature and pipette the buffer up and down several times to break up the precipitate, followed by vortexing the tube for 30 seconds to solubilise any precipitate.
        iv. The FFPE DNA Repair Buffer may have a yellow tinge and is fine to use if yellow.
3. Prepare the DNA in Ultrapure water:
    a. Transfer 1-3 µg input DNA into a 1.5mL Eppendorf DNA LoBind tube.
    b. Adjust the volume to 47 µL with Ultrapure water.
    c. Mix thoroughly by pipetting up and down, or by flicking the tube.
    d. Spin down briefly in a microfuge.
4. In a 0.2 mL thin-walled PCR tube, mix the following:
    a. Between each addition, pipette mix 10 - 20 times.

Reagent	Volume
DNA from Step 3	47 µL
DNA CS (optional)	1 µL
NEBNext FFPE DNA Repair Buffer	3.5 µL
NEBNext FFPE DNA Repair Mix	2 µL
Ultra II End-prep reaction Buffer	3.5 µL
Ultra II End-prep enzyme Mix	3 µL
Total	60 µL

5. Ensure the reaction is thoroughly mixed by gently pipetting and spin down briefly.
6. Using a thermocycler, incubate at:
    a. 20 °C 5 min
    b. 65 °C 5 min
7. Resuspend the AMPure XP Beads (AXP) by vortexing.
8. Transfer the DNA sample to a clean 1.5 mL Eppendorf DNA LoBind tube.
9. Add 60 µL of resuspended the AMPure XP Beads (AXP) to the end-prep reaction and mix by flicking the tube.
10. Incubate on rotating mixer:
    a. 10 rpm, 5 minutes, room temperature
11. Prepare 500 µL of fresh 80% ethanol in Ultrapure water.

Reagent	Volume
Ultrapure water	100 µL
Absolute ethanol	400 µL
Total	500 µL

12. Spin down the sample and place the tube on the magnet for 2-5 mins to separate the beads from the solution
    a. **NOTE**: Wait for the solution to be clear before proceeding to the next step.
13. Leaving the tube on the magnet, remove the supernatant (clear solution) and discard.
    a. **NOTE**: Do not disturb the “pellet” of separated magnetic beads.
14. Wash 2X
    a. Leaving the tube on the magnet, dispense 200 µL of <u>fresh</u> 80% absolute ethanol to the tube.
    b. Aspirate out the ethanol and discard.
15. Remove tube from magnet, spin down, and place back on the magnet. Pipette off any residual ethanol. With the tube open, allow to air-dry for ~30 s, but do not let the pellet dry to the point of cracking.
16. Remove the tube from the magnetic rack and resuspend the pellet in 61 µL of Ultrapure water. Incubate:
    a. 2 minutes at room temperature.
17. Place the tube on the magnet for 2-5 min, or until the solution is clear and colourless.
18. Transfer 61 µL of eluate into a clean 1.5 mL Eppendorf DNA LoBind tube.

<u>Adapter Ligation and Cleanup – Approx. 45 minutes</u>

1. Spin down the Ligation Adapter (LA) and Quick T4 Ligase, and place on ice.
2. Thaw Ligation Buffer (LNB) at room temperature, spin down and mix by pipetting. Due to viscosity, vortexing this buffer is ineffective. Place on ice immediately after thawing and mixing.
3. Thaw the Elution Buffer (EB) at room temperature, mix by vortexing, spin down and place on ice.

4. Thaw either Long Fragment Buffer (LFB) or Short Fragment Buffer (SFB) at room temperature, mix by vortexing, spin down and place on ice.
    a. Depending on the wash buffer (LFB or SFB) used, the clean-up step after adapter ligation is designed to either enrich for DNA fragments of >3 kb, or purify all fragments equally.
        i. To enrich for DNA fragments of 3 kb or longer, use Long Fragment Buffer (LFB)
        ii. To retain DNA fragments of all sizes, use Short Fragment Buffer (SFB)
5. In a 1.5 mL Eppendorf DNA LoBind tube, mix the following in order:
    a. Between each addition, pipette mix 10 - 20 times.

Reagent	**Volume**
Repaired and end-prepped DNA	60 µL
Ligation Buffer (LNB)	25 µL
NEBNext Quick T4 DNA Ligase	10 µL
Ligation Adapter (LA)	5 µL
Total	100 µL

6. Ensure the reaction is thoroughly mixed by gently pipetting and spin down briefly.
7. Incubate:
    a. 10 minutes at room temperature.
8. Resuspend the AMPure XP Beads (AXP) by vortexing.
9. Add 40 µL of resuspended the AMPure XP Beads (AXP) to the end-prep reaction and mix by flicking the tube.
10. Incubate on rotating mixer:
    a. 10 rpm, 5 minutes, room temperature
11. Spin down the sample and place the tube on the magnet for 2-5 mins to separate the beads from the solution
    a. **NOTE**: Wait for the solution to be clear before proceeding to the next step.
12. Leaving the tube on the magnet, remove the supernatant (clear solution) and discard.
    a. **NOTE**: Do not disturb the “pellet” of separated magnetic beads.
13. Wash 2X
    a. Add 250 µL of either Long Fragment Buffer (LFB) or Short Fragment Buffer (SFB) to the tube. Flick the beads to resuspend, spin down, then return the tube to the magnetic rack and allow the beads to pellet.
    b. Remove the supernatant and discard.
14. Remove tube from magnet, spin down, and place back on the magnet. Pipette off any residual buffer. With the tube open, allow to air-dry for ~30 s, but do not let the pellet dry to the point of cracking.
15. Remove the tube from the magnetic rack and resuspend the pellet in 25 µL of Elution Buffer (EB). Spin down and incubate:
    a. 10 minutes at room temperature (or 37 °C for high-molecular weight fragments)
16. Place the tube on the magnet for 10 min, or until the solution is clear and colourless.
17. Transfer 25 µL of eluate into a clean 1.5 mL Eppendorf DNA LoBind tube.

Quantify Cleaned Library – Approx. 10 minutes

1. Quantify **1 µL** of the cleaned DNA library using the Qubit dsDNA HS following kit instructions.

Check DNA library size using the TapeStation Genomic Kit – Approx. 10 minutes

1. Use the Genomic DNA ScreenTape Kit. Refer to the Agilent Genomic DNA ScreenTape Assay Quick Guide for 4200 TapeStation System.
    a. Combine 1 µL of gDNA with 10 µL of the Genomic DNA Sample Buffer in a 96-well plate (Agilent) provided for the TapeStation.
    b. Apply foil seal to the plate, mix for 1 min on the IKA vortex at 2000 rpm. Centrifuge for 1 min at 1500 rpm.
        i. Check for bubbles. If further vortexing and centrifugation do not remove the bubbles, carefully remove the foil seal and break the surface tension with a pipette tip.
    c. Load the samples onto the instrument. Follow instructions in the Quick Guide.

Dilute Final Library – Approx. 10 minutes

1. Use the NEB calculator (https://nebiocalculator.neb.com/#!/dsdnaamt) to determine the amount of DNA in the library in fmol using the average size and total DNA available.
2. Split your library into three libraries of 10-20 fmol in 12 µL using Elution Buffer (EB). Store on ice until ready to load the flow cell.
    a. If the DNA yield is not sufficient for three separate libraries of 150 ng, we recommend making up the first library to 150 ng and splitting the remaining library across the last two libraries.
    b. We recommend washing and reloading your flow cell twice and increasing run time to 100 hours to accommodate the flow cell reloading.
3. OxfordNanopore recommends loading 10-20 fmol of this final prepared library onto the R10.4.1 flow cell.
    a. Loading more than 20 fmol of DNA can reduce the rate of duplex read capture. Dilute the library in Elution Buffer if required.
4. Library storage recommendations
    a. OxfordNanopore recommend storing libraries in Eppendorf DNA LoBind tubes at 4 °C for short term storage or repeated use, for example, reloading flow cells between washes.
    b. For single use and long term storage of more than 3 months, OxfordNanopore recommends storing libraries at -80 °C in Eppendorf DNA LoBind tubes.

Run gDNA Library on MinION

Prime and Load the Flow Cell – Approx 30 minutes

1. Start by watching the video on Flow Cell Priming and Loading, available at https://community.nanoporetech.com/nanopore_learning/lessons/sars-cov-2-priming-and-loading-your-flow-cell
    a. This explains loading and working with the flow cell better than any written instructions.
2. Thaw the Sequencing Buffer (SB), Library Beads (LIB), Flow Cell Tether (FCT) and one tube of Flow Cell Flush (FCF) at room temperature. Mix by vortexing and spin down.
3. OPTIONAL: For optimal sequencing performance and improved output on MinION R10.4.1 flow cells (FLO-MIN114), OxfordNanopore recommends adding Bovine Serum Albumin (BSA) to the flow cell priming mix at a final concentration of 0.2 mg/mL.
    a. **NOTE**: They do not recommend using recombinant BSA.
4. To prepare the flow cell Priming Mix, add the following reagents directly to the tube of Flow Cell Flush (FCF), and mix by inverting the tube and pipette mix at room

temperature:

Reagent	With BSA	Without BSA
Flow Cell Flush (FCF)	1170 µL	1170 µL
Bovine Serum Albumin (BSA) at 50 mg/mL	5 µL	---
Flow Cell Tether (FCT)	30 µL	30 µL
Total	1205 µL	1200 µL

5. Open the MinION device lid and slide the flow cell under the clip. Press down firmly on the flow cell to ensure correct thermal and electrical contact.
6. Slide the flow cell priming port cover clockwise to open the priming port.
7. After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles:
    a. Set a P1000 pipette to 200 µL
    b. Insert the tip into the priming port
    c. Turn the wheel until the dial shows 220-230 µL, to draw back 20-30 µL, or until you can see a small volume of buffer entering the pipette tip.
    d. **NOTE**: Visually check that there is continuous buffer from the priming port across the sensor array. Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µL, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores
8. Slowly load 800 µL of the priming mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Wait for 5 minutes. During this time, prepare the library for loading by following the steps below.
9. Thoroughly mix the contents of the Library Beads (LIB) by pipetting.
    a. The Library Beads (LIB) tube contains a suspension of beads. These beads settle very quickly. It is vital that they are mixed immediately before use.
10. In a new tube, prepare the library for loading as follows:

Reagent	Volume
Sequencing Buffer (SB)	37.5 µL
Library Beads (LIB) mixed immediately before use	25.5 µL
DNA library (10-20 fmol)	12 µL
Total	75 µL

11. Complete the flow cell priming:
    a. Gently lift the SpotON sample port cover to make the SpotON sample port accessible.
    b. Slowly load 200 µL of the priming mix into the flow cell priming port (not the SpotON sample port), avoiding the introduction of air bubbles.
12. Mix the prepared library gently by pipetting up and down just prior to loading.
13. Add 75 µL of the prepared library to the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop flows into the port before adding the next.
14. Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port, and replace the MinION device lid.

Sequence

1. Open MinKNOW software on the computer and log in using your community credentials.

    a. **NOTE**: The computer must have 1TB of space free on an SSD for the complete run.
2. Plug in the MinION.
3. On the MinKNOW Connection Manager page, select the sequencing device and click “Start Sequencing”.
4. Type in the experiment name, sample ID, and choose flow cell type from the drop down menu. If you’ve already saved your settings as a template, select “Load settings from template”. If not, then “Continue to Kit Selection”.
5. Select the kit used from the kit selection menu.
    a. For kits with V14 chemistry, select the speed for the sequencing run. Currently, there are two speed options available for Fast, HAC and SUP basecallers:
        i. 260 bps for highest accuracy
        ii. 400 bps for highest output (default speed)
6. To keep default settings (recommended unless you have a specific reason to change them), click “Skip to final review”. If you would like to change any defaults, select “Continue to run options”.
    a. Select run options for run time and minimum read length.
        i. To enable Short Fragment Mode (SFM), select the preferred minimum read length, from as short as 20 bp. This directs the software to write sequencing files from the minimum size read length selected.
    b. Select “Continue to analysis”.
    c. Choose basecalling, barcoding, and alignment options.
        i. To use alignment during sequencing, select the toggle and upload an alignment reference file as a .fasta or minimap index file.
    d. Select “Continue to Output”.
    e. Select the output data location, format and filtering options.
        i. Use the checkboxes to save output data as FAST5, POD5, FASTQ and/or BAM files. FAST5 and FASTQ are on by default.
        ii. Filtering options can be used to determine which reads are classed into pass or fail files. These options may also be used to determine which predefined reads, read lengths and Qscore during basecalling can be split out in some live graphs.
    f. Select “Continue to final review”, which is where you will be if you use the default settings (recommended by OxfordNanopore).
7. Once settings have been reviewed and adjusted if needed, click “Start”.
8. You will be automatically navigated to the Sequencing Overview when sequencing starts.
    a. From here, you can see a progress bar below the flow cell to show the progression of the sequencing script.
    b. In some cases, the device will take a few minutes before starting a sequencing run if an alignment reference file is used. A progress bar will show the progress before sequencing begins.
    c. Flow cell health will be displayed after the first pore scan.
9. Select the flow cell to open the quick view to check the number of active pores reported in the pore scan are similar (within 10-15%) to those reported at the end of the Flow Cell Check.
    a. If there is a significant reduction in the numbers, restart MinKNOW.
    b. If the numbers are still significantly different, close down the host computer and reboot.

c. When the numbers are similar to those reported at the end of the Flow Cell Check, restart the experiment. There is no need to load any additional library after restart.

10. Once sequencing is complete, generate a run report.
    a. A run report containing information about the sequencing run and performance graphs can be generated manually by clicking Export run report and selecting which experiment to export to html.
    b. Run reports are automatically generated for MinION Mk1B running on Linux.
    c. A pore activity .csv file is also generated for every run.
    d. The report and .csv files are saved to the same folder as the .fast5 and .fastq read files.

Flush the Flow Cell – Approx. 1 hour

NOTE: To flush the flow cell and re-load a library (either the same library or a new library), complete "Flush the Flow Cell (~60 mins)" followed by "Prime and Load the Flow Cell – Approx. 30 minutes". To flush the flow cell and store is for use later, complete "Flush the Flow Cell – Approx. 1 hour" followed by "Store the Flow Cell – Approx 10 minutes".

- Over the course of a run, you may see an increase in pores in the "recovering" state. Pores can be reverted to the "active pore" state by washing the flow cell and re-loading the library. This will only work if pores are in the "recovering" state, not if they're in "saturated", etc.

1. Place the tube of Wash Mix (WMX) on ice. ***DO NOT** vortex the Wash Mix (WMX)***
2. Thaw one tube of Wash Diluent (DIL) at room temperature.
3. Mix the contents of Wash Diluent (DIL) thoroughly by vortexing, spin down briefly and place on ice.
4. In a clean 1.5 mL Eppendorf DNA LoBind tube, prepare the following Flow Cell Wash Mix:

Reagent	Volume
Wash Mix (WMX)	2 µL
Wash Diluent (DIL)	398 µL
Total	400 µL

5. Mix well by pipetting, and place on ice. ***DO NOT** vortex the tube***
6. Stop or pause the sequencing experiment in MinKNOW, and leave the flow cell in the device.
7. Before removing the waste fluid, ensure that the flow cell priming port cover and SpotON sample port cover are closed. Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the flow cell priming port and SpotON sample port are closed, no fluid should leave the sensor array area.
    a. **NOTE**: it is vital that the flow cell priming port and SpotON sample port are closed to prevent air from being drawn across the sensor array area, which would lead to a significant loss of sequencing channels.
8. Rotate the flow cell priming port cover clockwise so that the priming port is visible.
9. After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles:
    a. Set a P1000 pipette to 200 µL
    b. Insert the tip into the priming port

    c. Turn the wheel until the dial shows 220-230 µL, to draw back 20-30 µL, or until you can see a small volume of buffer entering the pipette tip
    d. **NOTE**: Visually check that there is continuous buffer from the priming port across the sensor array. Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µL, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.
10. Load 400 µL of the prepared Flow Cell Wash Mix into the flow cell priming port, avoiding the introduction of air.
11. Close the flow cell priming port and wait for 60 minutes.
12. Before removing the waste fluid a second time, ensure that the flow cell priming port cover and SpotON sample port cover are closed.
13. Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the flow cell priming port and SpotON sample port are closed, no fluid should leave the sensor array area.
14. To run a second library straight away, follow the instructions in “Prime and Load the Flow Cell – Approx. 30 minutes”. To store the flow cell for later use, follow the instructions in “Store the Flow Cell – Approx. 10 minutes”.

Store the Flow Cell – Approx 10 minutes

1. To store the flow cell for later use, thaw one tube of Storage Buffer (S) at room temperature.
2. Mix contents thoroughly by pipetting and spin down briefly.
3. Rotate the flow cell priming port cover clockwise so that the priming port is visible.
4. After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles:
    a. Set a P1000 pipette to 200 µL
    b. Insert the tip into the priming port
    c. Turn the wheel until the dial shows 220-230 µL, to draw back 20-30 µL, or until you can see a small volume of buffer entering the pipette tip.
    d. **NOTE**: Visually check that there is continuous buffer from the priming port across the sensor array. Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µL, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.
5. Slowly add 500 µL of Storage Buffer (S) through the flow cell priming port.
6. Close the priming port.
7. Using a P1000, remove all fluid from the waste channel through Waste port 1.
8. The flow cell can now be stored at 4-8 °C for a few weeks, but pore availability may decrease below useable levels.
9. When you wish to reuse the flow cell, remove the flow cell from storage, and allow it to warm to room temperature for ~5 minutes. A Flow Cell Check can be performed before loading the next library.

Flush the Flow Cell to return to OxfordNanopore - Approx 10 minutes

1. Follow this protocol if the flow cell is no longer useable (below the required number of pores, 800 for the MinION flow cell). You will need a flow cell returns box from Oxford Nanopore which you can request at https://community.nanoporetech.com/support/returns.

2. Remove the flow cell from the MinION and take it to Lab 411. Place sufficient absorbent material on the bench to take up approximately 4000 µL of flush waste, applied in 1000 µL aliquots.
3. Make sure the SpotON port cover is closed fully.
4. Wash 2X:
    a. Open the priming port cover, and using a pipette, slowly load 2x1 mL Ultrapure water via the priming port so that there is liquid filling the waste channel all the way to waste port 2.
    b. Close the priming port and empty the waste channel by withdrawing 2x1 mL waste from Waste port 2.
        i. **NOTE**: Make sure the priming port is closed while emptying the waste channel to ensure the array remains submerged.
5. Leave the priming port closed after the second wash and wipe away excess water with a kim-wipe.
6. Transfer the flow cell to the blister back included in the returns kit and place it in the returns box.
7. Close and seal the returns box and arrange pickup at https://community.nanoporetech.com/support/returns.

Monitor and Analyse the gDNA run

Flow Cell Health

- During a sequencing experiment, the MinKNOW Sequencing Overview page shows a flow cell icon with coloured bars. The bars represent the combined health of all pores in a flow cell, and indicate how well the flow cell is performing. The colours are:
    - Light green: sequencing
    - Dark green: open pore
    - Dark blue: pore recovering
    - Light blue: pore inactive

Experiment Summary Information

- The Experiments page displays summary information for all sequencing flow cells and device checks carried out on the device. From this page, the user is able to control specific runs and identify real-time information including flow cell health and reads, giving users real-time feedback for sequencing flow cells.
    - Run statistics: The total number of reads and bases produced across the experiment
    - Basecall statistics: There are two values for basecalled reads:
    - Basecalled reads as a percentage of the total reads produced across the experiment. This gives an indication as to the size of the queue for reads to be basecalled
    - Total number of reads basecalled across the experiment
    - Run time: The duration of the experiment
    - Temperature: The heatsink temperature of the selected position, which for sequencing should reach 34 °C with Kit 9-11, 30 °C with Kit 12 or 35 °C on Flongle. For Kit 14, temperature varies depending on basecalling speed, with higher temperatures for faster speeds and lower temperatures for slower basecalling speeds.
    - Voltage: The applied potential of the position at that point in time

Pore Occupancy

- A good library will be indicated by a higher proportion of light green channels in Sequencing than are in Pore. The combination of Sequencing and Pore indicates the number of active pores at any point in time. A low proportion of Sequencing channels will reduce the throughput of the run.
- Recovering indicates channels that may become available for sequencing again. A high proportion of this may indicate additional clean up steps are required during your library preparation.
- Inactive indicates channels that are no longer available for sequencing. A high proportion of these as soon as the run begins may indicate an osmotic imbalance.
- Unclassified are channels that have not yet been assigned one of the above classifications.
- Clicking on the Show Detailed button reveals a more detailed array of channel states:
    - Strand: the channel has strand
    - Adapter: the pore is sequencing the unligated sequencing adapter only. Reads will initially be classified as adapter until the DNA/RNA strand starts translocating through the pore and MinKNOW™ is able to reclassify the read
    - Single pore: the channel appears to show a single pore available for sequencing
    - Unavailable: the channel appears to show a pore that is currently unavailable for sequencing
    - Active feedback: the channel is reversing the current flow to eject the analyte
    - No pore from scan: the pore scan has not detected a pore in the well
    - Out of range 2: current level is between 10 and 9999 pA. Currently unavailable for sequencing
    - Possible multiple: the channel appears to show more than one pore. Unavailable for sequencing
    - Saturated: the channel has switched off due to current levels exceeding hardware limitations
    - Out of range 1: current level is between -5 and -9999 pA. Currently unavailable for sequencing
    - Zero: Current level is between -5 and 10 pA. Currently unavailable for sequencing

Pore Activity Plots

- The pore activity plot summarises the channel states over time.
- Each bar shows the sum of all channel activity in a particular amount of time. This time bucket defaults to 1 minute, and scales to 5 minutes automatically after 48 minutes. However, bucket size can be adjusted in the "Bucket size" box in Display Settings.
- The graph populates over time, and can be used as a way to assess the quality of your sequencing experiment, and make an early decision whether to continue with the experiment or to stop the run.

Read Length Histogram

- The cumulative histogram shows reads compared to bases. Use the options below to choose the axis legends:
    - Y-axis: Estimated bases or basecalled bases
    - X-axis: Read length or read counts
- Read count - this shows the number of reads vs read length. This enables you to understand how the read lengths vary in number and size.
- Read length - this shows the total number of bases vs the read length.

- The N50 value is presented (only for the whole set of passed reads) in the top left corner of the histogram.
- Reads that are outliers in terms of length can be removed from the graph by ticking the "Hide outliers" box below the histogram.
- Select "Split by read end" to view split reads and hover over for further information. This is useful for adaptive sampling which is further explained in the Adaptive Sampling info sheet.
    - Device changed MUX: The strand ended because there was a scheduled MUX change that interrupted the strand
    - Unblock voltage reversal: The strand ended by a downtick in the signal
    - Read completed: The strand ended naturally (there was an uptick in the signal, which is an indication of a pore)
    - Read became blocked: The strand ended because it was deemed of low quality and purposefully rejected
    - Adaptive sampling voltage reversal: The strand was rejected by adaptive sampling (typically happens after ~500 bp)

Cumulative output

- The cumulative output graph shows:
    - The number of bases that have been sequenced and basecalled
    - The number of reads that have been sequences and basecalled; and whether the reads have passed or failed the quality filters

Temperature and Bias Voltage graph

- Temperature vs time graph
    - The temperature graph gives a real-time representation of the heatsink temperature of the flow cell. If the temperature reading drifts out of the target zone, please consult Technical Services, otherwise the quality of your data may be compromised.
    - Please note, for our kits using Kit 14 chemistry, there will be some temperature fluctuation at the beginning of the run due to speed normalisation
- Bias voltage vs time graph
    - The bias voltage graph provides the running voltage in real-time. MinKNOW will automatically adjust the applied voltage and will naturally drift to lower voltages as the electrochemistry in the flow cell is depleted. This graph is useful for running a flow cell multiple times.
    - If you set the voltage for a subsequent run as the final running voltage of the previous run, then MinKNOW will find it easier to identify the appropriate running voltage.
    - You will notice drops in the voltage at regular intervals and these will correspond to the channel scans that are defaulted to occur every one and a half hours. Here, each channel will be scanned to look for its availability for sequencing. The common voltage is reversed before and after each channel assessment for clearer results.

Translocation speed and quality score graphs

- Translocation speed vs time
    - The translocation graph gives a real-time representation of the speed at which DNA/RNA strands pass through the pore. If the translocation speed drops below

this window, data quality and output may be compromised as strands take longer to move through the pore.
    - Please note, translocation speed will be dependent on the kit chemistry used and the speed selected for our kits using V14 chemistry.
- Quality score vs time
    - The Quality score graph gives a live representation of the median strand Q-score over time.
    - Modal quality score graph is available across the run. This graph presents both the spread of quality scores and the modal quality score across the run, and the minimum quality score set.
- Barcode read counts
    - The Barcode Read Counts graph shows the breakdown of barcoded reads, if barcoding was used for the experiment. The default view only shows reads that have passed the quality score filters. However, selecting the Display failed box will show all reads.
    - The X-axis of the histogram also has a zoom function using the scaled bar underneath. Use Reset to refocus the zoom bar and graph. Note that there is a small amount of cross-talk between barcodes. Some barcoded reads may appear on the graph even if the barcode in question was not used for the experiment.
- Alignment hits
    - The alignment hits graph will populate when alignment and basecalling is set up to run during sequencing. This bar graph shows the number of reads and bases that align to each of the entries in the user reference .fasta file or minimap index file. An entry in the reference file will only appear on the graph once a single read has aligned to it.
    - The X-axis of the histogram also has a zoom function using the scaled bar underneath. Use Reset to refocus the zoom bar and graph.
    - NOTE: If a .bed file is used, the alignment hits graph will still display hits from the reference .fasta file or precompiled minimap index file. Currently, the .bed file is not used to populate the alignment graphs. However, the .bed file alignment hits will be highlighted in the sequencing summary .txt file generated in the data folder.
- Alignment and barcode heatmap
    - The alignment and barcode heatmap is only available when alignment and demultiplexing is performed with basecalling during sequencing. The heatmap graph shows the alignment hits split per barcode. The colour gradient shows bases and reads which are the more popular barcode and alignment hit combinations.
- Traceviewer
    - The Traceviewer displays the current levels from individual channels. By default, it is set to show 10 channels. This number can be changed through the selection boxes beneath the viewer. Additional parameters that can be altered:
        - Time: The length of time plotted on one screen
        - Maximum: The highest current level to be shown on the y axis
- Pore scan
    - As the sequencing protocol starts, a pore scan begins before the sequencing. There are four groups of active pores, and the pore scan allows MinKNOW to pick the best-performing pores in each group, maximising the data output in the initial stages of the run. The software also instantly switches to a new channel in the group if a channel is in the “Saturated” state, or after ~5 minutes if a channel is “Recovering”.

Troubleshooting a run

Refueling the flow cell

1. If the DNA translocation speed drops below 300 bases per second, you may start to see a reduction in quality of data reflected in the Qscore. In this case, top up the flow cell with fuel, using the Flush Buffer (FB) from the Flow Cell Priming Kit.
    a. Remove one tube of Flush Buffer (FB) from the freezer and thaw by bringing to room temperature.
    b. Pause the experiment on the MinION Mk1B Flow Cell that is being refuelled.
        i. Navigate to Experiments, select the experiment running, and click “Pause”.
        ii. A dialogue box will open. Choose the flow cell to pause and click “Pause” to confirm.
    c. After opening the priming port, check for a small bubble under the cover. Draw back a small volume to remove the bubble (a few µL)
        iii. Set a P1000 pipette to 200 µL
        iv. Insert the tip into the priming port
        v. Turn the wheel until the dial shows 220-230 µL, or until you can see a small volume of buffer entering the pipette tip
    d. Visually check that there is continuous buffer from the priming port across the sensor array.
    e. Complete the flow cell refuelling:
        vi. Load 250 µL of the FB into the flow cell via the priming port (i.e. not the SpotON port), avoiding the introduction of air bubbles.
        vii. Close the priming port and replace the MinION Mk1B lid.
        viii. Unpause the experiment on the MinION Mk1B by clicking Resume.
        ix. (optional) Click Start pore scan and choose your flow cell, to pick a new set of the best channels for the remainder of the sequencing experiment

Pore activity plots

- The pore activity plot feature in the MinKNOW software can be used to judge the quality of your experiment. The pore activity plot shows the distribution of channel states over time, grouped by time chunks, or 'buckets'. The basic view shows the five main channel states:
    - Sequencing
    - Pore
    - Recovering
    - Inactive
    - Unclassified
- It is recommended to observe the pore activity plot populating over the first 30 min-1 hr of the sequencing run. By this time, the channel state distribution will give an indication whether the DNA/RNA library is of a good quality, and whether the flow cell is performing well.
- If Active Channel Selection is enabled during the run, the software instantly switches to a new channel in the group if a channel is in the “Saturated” or “Multiple” state, or after ~5 minutes if a channel is “Recovering”. This feature maximises the number of channels sequencing at the start of the experiment, however this may also result in an artificially high number of "Sequencing" or "Pore" channels in the pore activity plot. For this reason, OxfordNanopore recommends referring to the pore scan plot, which shows the true distribution of channel states at the point of the most recent pore scan.
- Good library

    - A good quality library will result in most of the pores being in the "Sequencing" state, and very few in "Pore", "Recovering" or "Inactive". A library that looks like this is likely to give a good sequencing output.
- Channel blocking
    - Under certain conditions (usually the presence of contaminants in the library), pores may become blocked and therefore unable to sequence. This manifests itself as a build-up of "Recovering" pores over time.
    - If, despite the channel blocking, the library is still producing a sufficient number reads to answer your biological question, you can carry on with the sequencing experiment.
    - Otherwise, stop the sequencing run in MinKNOW. Then wash out the library from the flow cell following “Flush the Flow Cell – Approx. 1 hour”. Then prepare another library and load it on the flow cell.
- Osmotic imbalance
    - If the plot shows a high number of 'Inactive' channels building up over time, this could indicate that the channels or membranes have been damaged, for example by air bubbles, osmotic imbalance, or the presence of detergents or surfactants in the library.
    - Recommendation: check the channel panel
        - If the Inactive channels are all grouped in one part of the flow cell, this could indicate an air bubble that has been introduced during flow cell flushing or library loading. If the remaining channels are still sequencing, it is possible to carry on with the run. Do not try to move the air bubble, as this can damage even more channels.
        - If the Inactive channels are distributed throughout the flow cell:
            - Check that the heat tape on the underside of the flow cell is intact.
            - Make sure that the input DNA is in either TE buffer or Ultrapure water, and that the buffer contains no detergents or surfactants.
            - Make a new batch of flow cell priming buffer (a mixture of Flush Buffer/Flow Cell Flush and Flush Tether/Flow Cell Tether). Flush the flow cell with the mixture, and load the library again.
- Low pore occupancy
    - If there was insufficient starting material, or some sample has been lost during library prep, or the sequencing adapters did not ligate well to the strand ends, the plot will show a high ratio of "Pore" to "Sequencing" states, meaning that only a limited number of pores are sequencing at any one time.
    - Recommendation:
        - Check the amount of DNA/RNA in your prepared library, for example by using the Qubit fluorometer. We recommend preparing a fresh library and reloading your flow cell with the recommended loading input from the relevant protocol
        - If your library is at a low concentration, prepare the library again using a higher amount of starting material.

Lambda Control Experiment

- The Lambda Control Experiment should be run at the beginning of an experiment and if troubleshooting is needed.

DNA Repair and End-Prep – Approx. 1 hour

1. Thaw DNA Control Sample (DCS) and Lambda DNA (LMD) at room temperature, spin down, mix gently by pipetting, and place on ice.
2. Prepare the NEBNext FFPE DNA Repair Mix and NEBNext Ultra II End Repair / dA-tailing Module reagents in accordance with manufacturer's instructions, and place on ice. ***__DO NOT__ vortex the __FFPE DNA Repair Mix__ or __Ultra II End Prep Enzyme Mix__.***
    a. For optimal performance, NEB recommend the following:
        i. Thaw all reagents on ice.
        ii. Flick and/or invert the reagent tubes to ensure they are well mixed.
        iii. The Ultra II End Prep Buffer and FFPE DNA Repair Buffer may have a little precipitate. Allow the mixture to come to room temperature and pipette the buffer up and down several times to break up the precipitate, followed by vortexing the tube for 30 seconds to solubilise any precipitate.
        iv. The FFPE DNA Repair Buffer may have a yellow tinge and is fine to use if yellow.
3. In a 0.2 mL thin-walled PCR tube, mix the following:
    a. Between each addition, pipette mix 10 - 20 times.

Reagent	**Volume**
Ultrapure water	27 µL
Lambda DNA	20 µL
DNA CS	1 µL
NEBNext FFPE DNA Repair Buffer	3.5 µL
NEBNext FFPE DNA Repair Mix	2 µL
Ultra II End-prep reaction Buffer	3.5 µL
Ultra II End-prep enzyme Mix	3 µL
Total	60 µL

4. Ensure the reaction is thoroughly mixed by gently pipetting and spin down briefly.
5. Using a thermocycler, incubate at:
    a. 20 °C 5 min
    b. 65 °C 5 min
6. Resuspend the AMPure XP Beads (AXP) by vortexing.
7. Transfer the DNA sample to a clean 1.5 mL Eppendorf DNA LoBind tube.
8. Add 60 µL of resuspended the AMPure XP Beads (AXP) to the end-prep reaction and mix by flicking the tube.
9. Incubate on rotating mixer:
    a. 10 rpm, 5 minutes, room temperature
10. Prepare 500 µL of fresh 80% ethanol in Ultrapure water.

Reagent	**Volume**
Ultrapure water	100 µL
Absolute ethanol	400 µL
Total	500 µL

11. Spin down the sample and place the tube on the magnet for 2-5 mins to separate the beads from the solution
    a. **NOTE**: Wait for the solution to be clear before proceeding to the next step.
12. Leaving the tube on the magnet, remove the supernatant (clear solution) and discard.
    a. **NOTE**: Do not disturb the "pellet" of separated magnetic beads.

13. Wash 2X
    a. Leaving the tube on the magnet, dispense 200 µL of fresh 80% absolute ethanol to the tube.
    b. Aspirate out the ethanol and discard.
14. Remove tube from magnet, spin down, and place back on the magnet. Pipette off any residual ethanol. With the tube open, allow to air-dry for ~30 s, but do not let the pellet dry to the point of cracking.
15. Remove the tube from the magnetic rack and resuspend the pellet in 61 µL of Ultrapure water. Incubate:
    a. 2 minutes at room temperature.
16. Place the tube on the magnet for 2-5 min, or until the solution is clear and colourless.
17. Transfer 61 µL of eluate into a clean 1.5 mL Eppendorf DNA LoBind tube.

Adapter Ligation and Clean-up – Approx. 45 minutes

1. Spin down the Ligation Adapter (LA) and Quick T4 Ligase, and place on ice.
2. Thaw Ligation Buffer (LNB) at room temperature, spin down and mix by pipetting. Due to viscosity, vortexing this buffer is ineffective. Place on ice immediately after thawing and mixing.
3. Thaw the Elution Buffer (EB) at room temperature, mix by vortexing, spin down and place on ice.
4. Thaw either Long Fragment Buffer (LFB) or Short Fragment Buffer (SFB) at room temperature, mix by vortexing, spin down and place on ice.
    a. Depending on the wash buffer (LFB or SFB) used, the clean-up step after adapter ligation is designed to either enrich for DNA fragments of >3 kb, or purify all fragments equally.
        i. To enrich for DNA fragments of 3 kb or longer, use Long Fragment Buffer (LFB)
        ii. To retain DNA fragments of all sizes, use Short Fragment Buffer (SFB)
5. In a 1.5 mL Eppendorf DNA LoBind tube, mix the following in order:
    a. Between each addition, pipette mix 10 - 20 times.

Reagent	**Volume**
Repaired and end-prepped DNA	60 µL
Ligation Buffer (LNB)	25 µL
NEBNext Quick T4 DNA Ligase	10 µL
Ligation Adapter (LA)	5 µL
Total	100 µL

6. Ensure the reaction is thoroughly mixed by gently pipetting and spin down briefly.
7. Incubate:
    a. 10 minutes at room temperature.
8. Resuspend the AMPure XP Beads (AXP) by vortexing.
9. Add 40 µL of resuspended the AMPure XP Beads (AXP) to the end-prep reaction and mix by flicking the tube.
10. Incubate on rotating mixer:
    a. 10 rpm, 5 minutes, room temperature
11. Spin down the sample and place the tube on the magnet for 2-5 mins to separate the beads from the solution
    a. **NOTE**: Wait for the solution to be clear before proceeding to the next step.
12. Leaving the tube on the magnet, remove the supernatant (clear solution) and discard.

    a. **NOTE**: Do not disturb the “pellet” of separated magnetic beads.
13. Wash 2X
    a. Add 250 µL of either Long Fragment Buffer (LFB) or Short Fragment Buffer (SFB) to the tube. Flick the beads to resuspend, spin down, then return the tube to the magnetic rack and allow the beads to pellet.
    b. Remove the supernatant and discard.
14. Remove tube from magnet, spin down, and place back on the magnet. Pipette off any residual ethanol. With the tube open, allow to air-dry for ~30 s, but do not let the pellet dry to the point of cracking.
15. Remove the tube from the magnetic rack and resuspend the pellet in 15 µL of Elution Buffer (EB). Spin down and incubate:
    a. 10 minutes at room temperature (or 37 °C for high-molecular weight fragments)
16. Place the tube on the magnet for 2-5 min, or until the solution is clear and colourless.
17. Transfer 15 µL of eluate into a clean 1.5 mL Eppendorf DNA LoBind tube.

Quantify Cleaned Library – Approx. 10 minutes

1. Quantify your library using the Qubit dsDNA HS Assay following kit instructions.

Check DNA library size using the TapeStation Genomic Kit – Approx. 10 minutes

2. Use the Genomic DNA ScreenTape Kit. Refer to the Agilent Genomic DNA ScreenTape Assay Quick Guide for 4200 TapeStation System.
    a. Combine 1 µL of gDNA with 10 µL of the Genomic DNA Sample Buffer in a 96-well plate (Agilent) provided for the TapeStation.
    b. Apply foil seal to the plate, mix for 1 min on the IKA vortex at 2000 rpm. Centrifuge for 1 min at 1500 rpm.
        i. Check for bubbles. If further vortexing and centrifugation do not remove the bubbles, carefully remove the foil seal and break the surface tension with a pipette tip.
    c. Load the samples onto the instrument. Follow instructions in the Quick Guide.

Dilute Final Library if Needed

1. Use the NEB calculator (https://nebiocalculator.neb.com/#!/dsdnaamt) to determine the amount of DNA in the library in fmol using the average size and total DNA available.
2. Using Elution Buffer (EB) if needed, dilute your library to contain 10-20 fmol in 12 µL. Store on ice until ready to load the flow cell.
3. OxfordNanopore recommends loading 10-20 fmol of this final prepared library onto the R10.4.1 flow cell.
    a. Loading more than 20 fmol of DNA can reduce the rate of duplex read capture. Dilute the library in Elution Buffer if required.
4. Library storage recommendations
    a. OxfordNanopore recommend storing libraries in Eppendorf DNA LoBind tubes at 4 °C for short term storage or repeated use, for example, reloading flow cells between washes.
    b. For single use and long term storage of more than 3 months, OxfordNanopore recommends storing libraries at -80 °C in Eppendorf DNA LoBind tubes.

Run Lambda Library on MinION – Approx. 7 hours

Priming and Loading the Flow Cell – Approx. 30 minutes

NOTE: Prepare and load the flow cell in the lab, then CAREFULLY bring the loaded MinION to the computer and plug it in.

1. Start by watching the video on Flow Cell Priming and Loading, available at https://community.nanoporetech.com/nanopore_learning/lessons/sars-cov-2-priming-and-loading-your-flow-cell
    a. This explains loading and working with the flow cell better than any written instructions.
2. Thaw the Sequencing Buffer (SB), Library Beads (LIB), Flow Cell Tether (FCT) and one tube of Flow Cell Flush (FCF) at room temperature. Mix by vortexing and spin down.
3. OPTIONAL: For optimal sequencing performance and improved output on MinION R10.4.1 flow cells (FLO-MIN114), OxfordNanopore recommends adding Bovine Serum Albumin (BSA) to the flow cell priming mix at a final concentration of 0.2 mg/mL.
    a. **NOTE**: They do not recommend using recombinant BSA.
4. To prepare the flow cell Priming Mix, add the following reagents directly to the tube of Flow Cell Flush (FCF), and mix by inverting the tube and pipette mix at room temperature:

Reagent	**With BSA**	**Without BSA**
Flow Cell Flush (FCF)	1170 µL	1170 µL
Bovine Serum Albumin (BSA) at 50 mg/mL	5 µL	---
Flow Cell Tether (FCT)	30 µL	30 µL
Total	1205 µL	1200 µL

5. Open the MinION device lid and slide the flow cell under the clip. Press down firmly on the flow cell to ensure correct thermal and electrical contact.
6. Slide the flow cell priming port cover clockwise to open the priming port.
7. After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles:
    a. Set a P1000 pipette to 200 µL
    b. Insert the tip into the priming port
    c. Turn the wheel until the dial shows 220-230 µL, to draw back 20-30 µL, or until you can see a small volume of buffer entering the pipette tip
    d. **NOTE**: Visually check that there is continuous buffer from the priming port across the sensor array. Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µL, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.
8. Slowly load 800 µL of the priming mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Wait for 5 minutes. During this time, prepare the library for loading by following the steps below.
9. Thoroughly mix the contents of the Library Beads (LIB) by pipetting.
    a. The Library Beads (LIB) tube contains a suspension of beads. These beads settle very quickly. It is vital that they are mixed immediately before use.
10. In a new tube, prepare the library for loading as follows:

Reagent	**Volume**

Sequencing Buffer (SB)	37.5 µL
Library Beads (LIB) mixed immediately before use	25.5 µL
DNA library (10-20 fmol)	12 µL
Total	75 µL

11. Complete the flow cell priming:
    a. Gently lift the SpotON sample port cover to make the SpotON sample port accessible.
    b. Slowly load 200 µL of the priming mix into the flow cell priming port (not the SpotON sample port), avoiding the introduction of air bubbles.
12. Mix the prepared library gently by pipetting up and down just prior to loading.
13. Add 75 µL of the prepared library to the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop flows into the port before adding the next.
14. Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port, and replace the MinION device lid.
15. Carefully bring the MinION with loaded flow cell to the computer being used to run the library.

Sequence – Approx. 6 hours

1. Open MinKNOW software on the computer and log in using your community credentials.
2. Plug in the MinION.
3. On the MinKNOW Connection Manager page, select the MinION and click “Start Sequencing”.
4. Type in the experiment name, sample ID, and choose flow cell type from the drop down menu. If you’ve already saved your settings as a template, select “Load settings from template”. If not, then “Continue to Kit Selection”.
5. Select the kit used from the kit selection menu.
6. Keep default settings (recommended by OxfordNanopore) and click “Skip to final review”.
7. Edit Experiment length – set to 6hrs for the Lambda Control Library, rather than 72hrs for a sample library.
8. Once settings have been reviewed and adjusted if needed, click “Start”.
    a. Videos on the Nanopore Community site explain the settings that can be changed, if needed.
9. You will be automatically navigated to the Sequencing Overview when sequencing starts.
    a. From here, you can see a progress bar below the flow cell to show the progression of the sequencing script.
    b. In some cases, the device will take a few minutes before starting a sequencing run if an alignment reference file is used. A progress bar will show the progress before sequencing begins.
    c. Flow cell health will be displayed after the first pore scan.
10. Select the flow cell to open the quick view to check the number of active pores reported in the pore scan are similar (within 10-15%) to those reported at the end of the Flow Cell Check.
    a. If there is a significant reduction in the numbers, restart MinKNOW.
    b. If the numbers are still significantly different, close down the host computer and reboot.

c. When the numbers are similar to those reported at the end of the Flow Cell Check, restart the experiment. There is no need to load any additional library after restart.

11. Once sequencing is complete, a run report is automatically created (on Linux).
    a. A pore activity .csv file is also generated for every run.
    b. The report and .csv files are saved to the same folder as the .fast5 and .fastq read files.

Monitor and Analyse the Lambda run

- The Lambda Control Experiment will give you a good example of what a successful sample library run looks like
- If you parameters fall within the expected results, you can move forward with library prep for your samples
- See “Monitoring and Analysing the gDNA run” for example plots and more information on what to monitor.

1. Once the sequencing run has started, select “Experiments” on the left pane in MinKNOW to see data about the run.
2. Several different plots can help monitor the run:
    a. Channels Panel
        i. Live status of each channel state
        ii. Expect high ratio of Sequencing to Available, and low amounts of Unavailable and Inactive pores
    b. Duty Time
        i. Channel states over time group by time chunks/buckets
        ii. Pore occupancy >50% indicates a good sequencing run
    c. Mux scan/Pore Scan
        i. Summary of all wells for all Mux scans throughout the run
        ii. Want to see pores stable over time
    d. Read length histogram
        i. Expect uniform distribution around 48.5 kB for Lambda DNA using the ligation sequencing kit
    e. Cumulative output
    f. Translocation speed
        i. Expect >300 bases per second, within the median target
    g. Qscore
        i. Expect within the median target
3. Once the run is complete, look at the FASTQ Control Experiment in EPI2ME
    a. Run the EPI2ME workflows and open the portal
    b. Examine the Reads Analysed and Qscore metrics
        i. Qscore around 11 is within what’s expected for a good run
    c. Examine the alignment to the Lambda reference genome
        i. Expect reads aligning across the full length of the genome.

Appendix 2 References

NEB calculator https://nebiocalculator.neb.com/#!/dsdnaamt

Oxford Nanopore Site https://community.nanoporetech.com/

APPENDIX 3 – AQUATIC BIOTECHNOLOGY LAB STANDARD OPERATING PROCEDURE FOR MICROSATELLITES, WITH MODIFICATIONS FOR HORSE MUSSEL

The objective of this document is to record the relevant steps used in generating microsatellite data from extracted horse mussel DNA and to allow for reproduction of these results if required. It is also likely suitable for use in other mollusc species.

Safety Considerations

General

Personnel should have proper training from a lab manager or other experienced individuals.

All relevant training, instruction, safe work procedures and chemical SDS should be read prior to work.

DNA Extraction

Read SDS for QIAamp DNA Micro kit components and absolute ethanol.

Wear lab coat and gloves when handling QIAamp DNA Micro kit components and DNA samples.

DNA Quantification and Normalization:

Read SDS for Quant-iT PicoGreen Reagent in DMSO.

Wear gloves when handling Quant-iT PicoGreen dsDNA Assay reagent in DMSO (Thermo Fisher Scientific).

PCR:

The PCR hood unit has a UV light source. As a safety feature, the workstation is designed not to operate the UV light when the door is open. Do not over-ride the safety latch.

Loading Plate Preparation:

Read SDS for HiDi Formamide. Wear appropriate gloves when handling.

Equipment

General

- Eppendorf 5810R Centrifuge (Eppendorf)
- Eppendorf 5430 Centrifuge (Eppendorf)
- Pipettes
- Rainin (Mettler-Toledo) and Eppendorf single channel pipettes (various catalogue numbers)
- Rainin (Mettler-Toledo) multi-channel pipettes (various catalogue numbers)
- Eppendorf Repeater Xstream or E3X (VWR catalogue # CA47800-214,or 75796-402)
- Mettler-Toledo Pipet-Lite Multi Pipette L8-10XLS+ (Mettler-Toledo catalogue # 17013802)
- Scientific Instruments Vortex Genie 2 (VWR catalogue # 1426-188).
- Eppendorf MixMate (VWR catalogue # CA14900-548)
- Fisher Scientific plate shaker (no cat # - discontinued model)
- BMP51 or BMP53 labeller (Brady, Newark/Element14 catalogue # 89T4134 or 78T0490 resp.)
- BMP21 labeller (Brady, Newark/Element14 catalogue # 53R2145)

- PCR Roller (VWR Scientific catalogue # 95042-898)

DNA Normalization

- epMotion 5075 Robot (Eppendorf model 5075LH) or epMotion 5070 robot (Mo) (Eppendorf model 5070)

PCR

- PCR machines (Eppendorf Mastercycler Pro with Silver block)

Genotyping

- AB 3500xl genetic Analyzer (with 24 capillary/50 cm array) (ThermoFisher Scientific Catalogue #4406016) – with accompanying computer

Supplies

General

- Brady labels (VWR various catalogue numbers)
- Reservoirs (VWR catalogue # 82026-356)
- Pipette tips (Ovation (VistaLab) and Eppendorf – various catalogue numbers)
- Flat silicon mats (Axygen's Axymat - VWR catalogue # CA10011-130)
- Tube for dilution(s), depending on volume select the appropriate sized tube from:
    - 1.5 mL tubes (VWR catalogue # 29443-259)
    - 5 mL tube (VWR catalogue # CA60819-728)
    - 15 mL tube (VWR catalogue # 21008-671)
    - 50 mL tube (VWR catalogue # 21008-671)
- Tips
    - 500 µL Combi-tip plus (VWR catalogue # CA21516-010)
    - Tips for single channel/multichannel pipettors (various catalogue numbers)

DNA Normalization

- epMotion reservoirs (30ml) (Eppendorf catalogue # 960051009) and holder
- epT.I.P.S Motion 50 µL (Eppendorf catalogue #s 0030014421, 0030014430(filter))
- epT.I.P.S Motion 10 µL (Eppendorf catalogue # 0030014545, 0030014510 (filter))
- Eppendorf LoBind twin.tec plates (Fisher Scientific catalogue # E0030129512 or VWR catalogue # CA10049-106) or equivalent (if supply issues occur)

PCR

- TwinTec Plate skirted (VWR catalogue # CA35920-170)
- Eppendorf PCR Cooler (VWR catalogue # CA62111-758)
- Eppendorf Biopur 0.5 mL Combitip plus (VWR catalogue # CA21516-146)
- Rainin (Terra-Rack) LTS 20 µL filter tips (Mettler-Toledo catalogue # 17014961)
- Rainin (Terra-Rack) LTS 200 µL filter tips (Mettler-Toledo catalogue # 17014963)
- Rainin (Terra-Rack) LTS 1000 µL filter tips (Mettler-Toledo catalogue # 17014967)
- MicroAmp Clear Adhesive film (Thermo Fisher Scientific catalogue # 4306311)

Loading Plate Preparation:

- MicroAmp Fast 96-well reaction plate, 0.1 mL (Life Technologies catalogue # 4346907)
- Eppendorf 0.5 mL combitip plus (VWR catalogu

- # CA21516-010)

Reagents

DNA Normalization

- 1M Tris pH 8.0 (Sigma Aldrich catalogue # T-3038), or Buffer EB (Qiagen, from kits)

PCR

- Type-it Microsatellite PCR kit (Qiagen catalogue # 206246 (2000 rxns))
- Specific horse mussel microsatellite primers (primers ordered from IDT)

Loading Plate Preparation

- HiDi Formamide (Thermo Fisher Scientific catalogue # 4311320)
- GeneScan 500 LIZ size standard (Thermo Fisher Scientific catalogue # 4322682)

Solutions

DNA Normalization

- 10 mM Tris (pH 8.0):
    - Preparation (25 mL):
        - 250 µL 1M Tris (pH 8.0)
        - 24.75 mL Ultrapure H_2O
    - Note: can also use prepared Buffer EB instead of 10 mM Tris (pH 8.0) (QIAGEN, catalog # 19086)

PCR

- Primer stock (100 µM preparation)
- 10X primer mixes (for multiplexed reactions)

Software

- Microsoft Excel, WORD, ACCESS, OneNote (Microsoft 365)
- GeneMapper 6.0 (Thermo Fisher Scientific)
- Gen 5 3.05 (BioTek)
- Eppendorf epBlue Client v. 40.09
- 3500 Series Data Collection Software 3 (Research Use Only version) (Thermo Fisher Scientific; 2018)

Procedures

DNA dilution – Approx. 1 hour

1. Select previously extracted samples to be genotyped.
2. If your samples are in different plates or tubes, note their location(s) and concentration(s).
3. Retrieve samples from storage (freezer), noting which samples have been found, what elution(s) are available and what volume (if any) remains in the elution plate wells.
4. Using epBlue software, load/design your protocol for making dilution plate
    a. Aim to dilute gDNA to 10 ng/µL in 30 µL total
    b. Volumes of EB buffer/Tris and gDNA will depend on your sample DNA concentrations.

5. Plate-based extractions:
    a. Vortex gDNA in plate with MixMate at 1000 rpm for 30s
    b. Centrifuge plate at 1500 rpm for 1 min to bring sample off plate seal (silicone mat)
    c. Arrange gDNA plate, EB buffer reservoir w/ EB, and destination Lobind PCR plate on epMotion stage.
6. Tube-based extractions:
    a. Thaw tubes, then vortex briefly and spin at 1500 rpm for 1 min.
    b. Arrange tubes in epMotion tube blocks, making sure the lids fit into the slot to hold them open.
    c. Arrange blocks, EB buffer reservoir, and LoBind 96-well plate on epMotion stage.
7. Start epMotion program.
8. Once finished, return your sample extractions (plate or tubes) to -20 °C for storage**.**

STORAGE: Seal and label dilution plate, then store in 4 °C until next steps (or -20 °C for longer storage)

Primer Rehydration – Approx. 30 minutes

1. Refer to Main Text Table 4 for example of primer details.
2. If primers are shipped dry and must be hydrated to 100 µM using Buffer AE. If they are shipped in solution, dilution may be required.
3. Centrifuge the tubes for 1 minute on max to ensure all the lyophilized primer is at the bottom of the tube.
4. For each primer determine the volume of liquid to add as follows:
    - From the primer QA document or from the label (sometimes it is hard to see) find the “nmol” amount. Multiply this by 10 and this is the amount of liquid to add to get 100 µM stocks
5. Using a new filter tip for each tube add the appropriate amount of Buffer AE based on your calculations.
6. Allow the primers to sit at least 30 min at room temp. They can also be kept in the fridge overnight.
7. Label enough 0.5 mL tubes for each primer (for aliquots) – each aliquot is 100 µL and you know the final volume of primer based on your calculation in #2. The labels should contain the following information:
    a. Side Label
        i. Supplier (i.e. Applied Biosystems, Sigma Genosys or IDT)
        ii. Full name of the primer (include the label at the beginning if applicable)
        iii. Concentration
        iv. Date
    b. Top Label
        i. Label (if applicable – if there isn’t one, leave this blank)
        ii. Primer name
        iii. Aliquot
8. Vortex the resuspended primer and centrifuge to make sure the liquid is mixed and brought to the bottom.
9. Add 100 µL of primer to each of the aliquot tubes.
10. Keep one of the 0.5 mL tubes of concentrated primer in the lab for your use. All other tubes should go to the -80 °C freezer for storage.

Primer Mix Preparation – Approx. 30 minutes

1. Create 10X primer mixes following Type-IT multiplex PCR kit instructions.

2. This PCR assumes 2 µM of each primer is required in in the final mix. If your PCR requires a different amount of primer, you will have to adjust accordingly.
3. Thaw the required 100 µM primer aliquots, vortex, and centrifuge to make sure the liquid is at the bottom of the tube.
4. Example for 6 duplex primer mixes, to a labelled 500 µL tube, add:
    a.

Component	**Volume (µL)**
Marker1_F (100 µM)	4
Marker1_R (100 µM)	4
Marker2_F (100 µM)	4
Marker2_R (100 µM)	4
Buffer AE	184
Total	200

5. Vortex the tube to ensure the primers are well mixed, centrifuge, and store at -20 °C until needed.

PCR – Approx. 2.5 hours

1. If necessary in your lab, sign up for a PCR machine.
2. Note the loci that will be used in each microsatellite panel and their associated sequences, labelling information, and original references.
3. Ensure the correct program has been setup on the machine. For this we will assume the following conditions are required:
    a. 95 °C for 5 minutes
    b. 28-35 cycles of**:**
        i. 95 °C for 30 seconds
        ii. 60 °C for 90 seconds
        iii. 72 °C for 30 seconds
    c. 60 °C for 30 minutes
    d. 4 °C hold
    - You will likely have to adjust and program your thermocycler to match the conditions required for successful amplification of your specific markers.
4. Place the tips you will require at your PCR workstation. Ensure your workstation is clean and contamination free.
    a. Bring in as many tips as necessary for plating your DNA.
5. If you have a UV sterilizer, run the UV cycle on the PCR workstation now. Take care to read the manual before operation and take all necessary precautions.
6. Take the required normalized DNA plate (10 ng/µL) out of the fridge or freezer.
7. Assemble the required reagents at your station. These will generally be in the freezer and will require time to thaw. These include:
    a. 2X Type-it multiplex PCR Master Mix
    b. Ultrapure H_2O
    c. 10X primer mixes
8. Turn on the required PCR cyclers. Select the appropriate programs for each PCR being run.
9. Label a PCR plate for PCR being run that day. Suggested information is your name, the date, an interpretable reference to the samples or plates being run and the primer(s)
10. Once the PCR workstation UV light cycle is complete, transfer all reagents to the PCR workstation. Bring your previously labelled PCR plates.
11. Label 1.5 mL tubes for your master mixes, one per microsatellite panel.

12. Vortex and spin down the reagents (in the PCR workstation) and the normalized DNA (outside the PCR workstation).
13. Standard PCR set-up:
    a. Prepare the PCR master mix(es):

Reagent	**1X**	**105X**
2X Type-it Multiplex PCR Master Mix	5.0 µL	525 µL
10X Primer mix (2uM each primer)	1.0 µL	105 µL
Ultrapure H_2O	2.0 µL	210 µL
DNA (10ng/µL)	2.0 µL	-
Total	**10.0 µL**	**840 µL**

    b. Vortex cocktail well before plating.
    c. Add the PCR MM (8 µL) to the plate using the Repeater Xstream and 0.5 mL Combitip (Sterile quality). Or another multi-channel pipette.
    d. In the PCR workstation, add 2 µL of the normalized DNA or Ultrapure H_2O (negative) to each well using a multichannel pipette.
        i. Be sure to discard tips after each row of samples is transferred. Try not to contaminate your samples.
        ii. Never put amplified DNA in the PCR workstation, but pre-PCR DNA is fine.
14. Seal the plate with PCR film using a roller.
15. Mix the plate on the MixMate at 1650 rpm for 30 s and spin the plate in the Eppendorf 5430 benchtop centrifuge with plate rotor plate rotor (use “short’ spin to 2000 rpm is sufficient). This ensures all the liquid is at the bottom of the plate wells.
16. Place the plates in the thermocyclers and run the program.
17. Once plates are in the cyclers, return normalized DNA to the fridge (in zip-bag). Return the PCR reagents to the freezer.
18. Clean your station.
19. Once the PCR program(s) are completed, spin plate(s) at 1500 rpm for 1 min.

STORAGE: Store PCR plate at 4 °C until next steps (or -20 °C for longer storage)

ABI Loading Plate Preparation

1. Remove the 96-well plate(s) of PCR product/dilutions of PCR products that you have previously prepared from the fridge or freezer. Thaw (if necessary), vortex, and centrifuge at 1500 rpm for 1min.
2. Thaw the appropriate amount of formamide for the plate(s) you are going to prepare.
    a. Lay out the previously labelled MicroAmp Fast 96-well reaction plate(s).

NOTE: Plates MUST be from Applied Biosystems or it will void the warranty on the array.

3. Prepare HiDi Formamide/GeneScan500 LIZ MM as follows:

Reagent	**1X**	**105X**	**210X**

HiDi Formamide	9.3 µL	976.5 µL	1953 µL
GeneScan500 LIZ	0.2 µL	21 µL	42 µL
PCR Product	0.5 µL	-	-

4. Vortex the Formamide/Ladder mixture and spin down.
5. Using a repeater pipette (or other multi-channel pipette), add 9.5µL of the Formamide/Ladder mixture, to each well of the ABI MicroAmp plate to which PCR product will be added.

NOTE: For any wells that do not contain PCR product but are required to complete an injection of 16 samples; 10µ L formamide/ladder mixture or just formamide must still be added. The ABI 3130xl picks up 16 samples at a time, if it attempts to ingest samples from empty wells, it can damage the machine.

6. Add 0.5 µL of PCR product to each well using a multichannel pipette and the 10 µL LTS tips.
    a. As you are pipetting, visually check the tips to ensure that they have drawn up the 0.5 µL of PCR product and after dispensing to ensure they have delivered the PCR product into the loading plate.
    b. Be sure to discard tips after each row of samples is transferred. Try not to contaminate your samples.
7. Place a silicone mat on the ABI plate and use roller to secure. Visually examine each well to ensure it is sealed properly and that well volumes are correct (10µL).
8. Seal the PCR plate with a silicone mat and use roller to secure. Place the PCR plate in a zip-lock bag and store at -20 ˚C until ABI run is completed.
9. Vortex the ABI plate and place in the Eppendorf 5810R centrifuge at 1500 rpm for 1 min.
10. Place the ABI plate in the refrigerator and incubate for 2+ hours.
11. Prepare your ABI sample sheet and save as .tsv.
12. Import the file into the Data Collection software on the ABI computer:
    a. The Data Collection Software should be open and the Dashboard on screen.
    b. On the left-hand side of the window click on “Library”. Click “Import” from the top task bar options.
    c. Confirm the successful importation by noting the Plate ID in the Library listing (current window).
    d. If using the ’MicroAmp 96 Fast Reaction’ plate, then the plate setting will need to be changed from the default from the ‘MicroAmp 96 Optical reaction plate’.
        i. Click on the plate name in the file list of the ‘Library’
        ii. Select ‘Define Plate Properties’
        iii. Ensure that ’96 **Fast’** is selected as the ‘96’ option is the default for this program.
        iv. ‘Save’ and ‘Close’ plate.
        v. Return to import procedure.

Load and Run ABI

1. Turn on ABI and log in. After the software has loaded, check that the buffers and POP-7 have not expired (else take time to replace or replenish as required).
2. Click the ‘Pre-heat’ button at the right of the ‘Dashboard’ window, to turn on the oven and laser.

3. Preheat Eppendorf 'Thermocycler C' unit(s) to 95 °C using the pre-set 10 minute denaturing program.
4. Remove the plate(s) you wish to run from the refrigerator.
    a. Pre-heat plates in Thermomixer or thermocycle. Each plate must be denatured at 95 °C for 10 min before being run on the ABI.
5. When denature is finished, remove pre-frozen (-20 °C) plate holders freezer.
6. Place the plate(s) immediately into the plate holder(s) and return to the freezer.
    a. Remove the plate(s) from the plate holder(s) and place carefully into the tray of pre-chilled 95-100% ethanol.
    b. Set a timer for 10 minutes and allow to 'freeze'.
7. After time is up, remove plate(s) from the ethanol tray and allow to sit for 10-20 seconds on a paper towel to absorb residual ethanol from the bottom of the plate(s).
8. Select the appropriate plate holder (i.e. 'Fast' or 'Optical') for the plate(s).
    a. The plate holder is in two parts: a blue bottom piece and a snap down white top piece.
9. Align the plate with the plate holder and carefully insert the plate firmly into the bottom (blue) portion of the holder.
10. Securely hold the unit and carefully remove the silicone flat mat; set aside on the paper towel.
    a. It is useful to pre-label the paper towel 'A' and 'B; positions to keep the mats for each plate separate.
11. Replace the flat mat with a new or clean plate septum (no visible gaps should be present in the "+" of each well when viewed from underneath).
12. Place the top cover over the bottom and snap down securely on all sides.
13. Press the 'Tray' button on the front of the 3500xL to move the plate tray into the loading position. Open the door once the tray has stopped moving.
14. Go to the 'Dashboard' and select "Quick Start' from the tool bar.
15. Close the door of the 3500xL and wait for the tray to re-position and the green light to come back on (it will be yellow when the tray is in motion).
16. Link the plate(s) one at a time as directed.
    a. If an error message appears, the software is detecting that the plate loaded (usually the '96 Fast Reaction Plate') does not match the plate definition for the import file (default setting is for the '96 Optical Plate').
        i. Cancel the plate load process and go back to 'Library' (select).
        ii. Click on the plate name in the file list.
        iii. Select 'Edit' > 'Plate Properties'>'96 Fast' (the default should be just '96' (= Optical plate)
        iv. Once the change has been made go to the top tool bar and select 'Save' and then "Close' plate.
        v. Go back to Step 15 and now it should load the plate without issue.
17. Once the plate(s) is/are linked, click on 'Run' button at the bottom of the page. The program will check the array, buffers and POP-7 for expiry and usage.
18. Run time per full plate is 2 hours and 40 minutes. Each injection takes 40 minutes.
19. Once all runs are completed, each plate run file (*.fsa) can be transferred to a computer with an appropriate genotyping software. In this example we will be using GeneMapper software (version = 6.0).
20. Rinse septum with Ultrapure water and place in folded KimWipe to dry on bench.

ABI Processing – Time required varies

1. If you have left the ABI to run overnight, the results (*.fsa files) can be collected now (See 19 above).
2. Move completed plates to the labeled a refrigerator for storage.
3. Score the microsatellite panels in GeneMapper. (Note – if you are using new markers, species, populations etc. you may need to prepare a new set of scoring bins. See software manual for more information)
4. NOTE: Ladder is GS500(-35,-250,-340)LIZ
    a. Ladder Peaks are: 50, 75, 100, 139, 150, 160, 200, 300, 350, 400, 450, 490, 500 bp
    b. Extra peaks may be visible in the ladder, these are often panel-specific junk DNA.
5. Analyze and score each run with your genotyping software. After scoring has completed, it is recommended that you visually inspect each allele call, and adjust as necessary. The amount of adjustment will generally decline as you run more samples/populations using each microsatellite-specific set of scoring bins as you will refine the bins each time.
6. Once the scoring has been completed, and you are satisfied with the quality of the ABI run (i.e. once scoring has been completed), the completed plates can be processed for disposal.

Appendix 3 References

QIAmp 96 DNA QIACube HT kit handbook (Qiagen)

QIAamp DNA Mini QIAcube Kit Handbook (Qiagen)

DNeasy Blood & Tissue Kit handbook (Qiagen)

Qiagen HT User manual (July 2013)

Operators manual, 2006 AirClean Systems 57880 Rev.4 (located in FL-141)

Mastercycler ep User's Manual