VAN ANDEL INSTITUTE®

GENOMICS CORE

Sequencing Policies

Please thoroughly review these policies before submitting samples. The submitter is responsible for reviewing this document before submission to verify their compliance with the current policies. If you have any questions regarding this policy or any other aspect of next-generation sequencing, please contact us at genomics@vai.org.

Sample Submission:

- Before submitting an initial project, VAI researchers must meet with Genomics Core Staff to discuss sequencing goals and details of starting material.
- Projects are received on a first-come, first-served basis. They will not be entered into the project queue until samples are in the possession of the Genomics Core Facility and a project submission form is completed on our website (https://vari.my.site.com/submission/s/login/). Please contact the Genomics Core directly if any extenuating circumstances are associated with a project being submitted.
- Once samples are entered into the queue, the core will provide a tentative project timeline. We strive to honor this timeline; however, issues beyond our control may delay the run. In the event of a delay longer than a week, the Core will email the customer to inform them of an updated run date.
- Dropped-off samples must be prepared according to VAI specifications unless the Core Director has granted previous permission. Submission requirements are summarized in this policy's 'Minimum and Optimal Sample Requirements' section.
- After samples are processed, Core personnel will initiate a CrossLab request for the project. Please review and authorize the project if everything looks correct.

Sample Delivery and Labeling:

- Projects with <24 samples should be submitted in 1.5mL microcentrifuge tubes. Projects with ≥ 24 samples should be submitted in a full-skirted 96-well PCR plate (such as Eppendorf twin.tec). Do not submit samples in strip tubes; you will be asked to transfer your samples and resubmit.
- If submitting in a PCR plate, place samples in column orientation (A1, B1, C1, etc.).
- Tubes/plates must be labeled with the Project Number from our submission portal (PRxxxxxx) and Sample Name (matching what is submitted in the portal).
- Sample names will be no greater than 14 characters and must be composed only of alphanumeric characters. Any spaces, dashes, underscores, or special characters will be removed as they are incompatible with the downstream software.

Protected Health Information:

VAI Genomics Core does not accept samples or documentation that contain PHI. If PHI is transferred incidentally, the Core will inform the submitting lab within five business days of becoming aware of the disclosure and destroy any identifiable PHI (including samples and ePHI).

Minimum and Optimal Sample Requirements:

You will be informed via email if the following initial QC requirements are not met. You may elect to either resubmit samples of sufficient quality and quantity or proceed with the existing samples. Should you proceed with samples that do not meet our requirements, the samples will be run AS IS, and we make no assurances on the quality of downstream data. For library prep options not listed or further clarification, please visit the Genomics Core SharePoint site or contact Genomics@vai.org with specific questions.

gDNA for short read sequencing:

- 100 ng – 1 ug of high quality, high molecular weight gDNA in 10 mM Tris pH 8.0 or water.
- Quantify DNA by fluorometric methods if possible. If quantified by spectrophotometer, A_{260}/A_{280} should be ~ 1.8. Deviations from this number suggest inaccurate quantification.
- Material should be >10 kb
- Column-based cleanups or ethanol precipitation are *strongly* recommended prior to drop-off, especially for organic extractions.

HMW DNA for long-read sequencing:

- ≥ 2 ug of high quality, high molecular weight gDNA in 10 mM Tris pH 8.0, water, or low TE buffer (EDTA ≤ 0.1 mM).
- Perform an RNase digestion to remove RNA prior to submission.
- Quantify DNA by fluorometric methods if possible, as these are the most accurate.
- Spectrophotometer measurements should be as follows: A_{260}/A_{280} ~1.8 and A_{230}/A_{260} ~2.0-2.2. Deviations from this number suggest inaccurate quantification and the presence of contaminants.
- Material should be >25 kb. Shorter fragments will lower sequencer output and reduce overall read lengths.
- Column-based cleanups or ethanol precipitation are *strongly* recommended prior to drop-off, especially for organic extractions.
- Low molecular weight, incorrectly quantified, and/or contaminated DNA (e.g., salt, EDTA, protein, organic solvents) can significantly impact downstream processes, including sequencing runs.

ChIP/Cut&Run:

- 1 ng minimum, 30 ng optimal, post IP DNA sheared to 200-600 bp. DNA sheared to >1000 bp will not be accepted, as it is incompatible with the sequencer.
- Samples < 1 ng material cannot have quality assessed via electrophoresis, and we cannot guarantee successful library generation from this material.
- Quantify DNA by fluorometric methods if possible. Do not quantify IPs by spectrophotometer.
- qPCR on a known target is recommended to verify enrichment, as is submitting an input sample as a normalization control.

RNA:

- 100 ng - 1 ug of **total** RNA or 10 ng – 100 ng of mRNA in 10 mM Tris pH 8.0 or water
- Low input (1 – 10 ng) options are available for both total and mRNA library preps; we strongly recommend contacting the Genomics Core prior to RNA isolation to ensure compatibility.
- A_{260}/A_{280} and A_{260}/A_{230} ratios should be ≥1.8 in the range of 1.8 to 2.2. Ratios ≤1.8 may indicate protein and/or organic contamination
- Column-based cleanup is *strongly* recommended, especially for organic extractions and assaying ribosomal peak integrity on an Agilent Bioanalyzer PicoChip.
- Prokaryotic total RNA must be ribosomally reduced by submitting lab prior to sample submission.

Lab Prepared Libraries:

- At least 10 uL of library at ≥2 ng/ul, measured by fluorometry only, in 10 mM Tris pH 8.0 or water. EDTA should be avoided as it retards downstream sequencing reactions.
- For libraries requiring higher read depth, such as full flow cell lanes, more material may be needed.
- Provide all index *SEQUENCES*, as the same index number from different library preparation kits does not always refer to the same sequence. The vendor's sequences are usually either in an appendix or a separate supporting table.

- Index sequences should be provided in the following orientation regardless of the intended sequencer to be used (contact the Genomics Core if you are unsure of the proper orientation):
  - i7: 'i7 Bases for Sample Sheet'
  - i5: 'Reverse Orientation' – some vendors will list which instruments use a specific orientation; the Reverse Orientation will have NovaSeq 6000 v1.5 listed.
- Projects containing libraries with incorrect indices provided will be subject to a $240 charge for additional processing and will cause delays in data release, including possible resequencing.
- Libraries containing UMIs should note the length and location of these sequences on the project page. The VAI Genomics Core is not responsible for UMIs not sequenced due to the absence of notification or inaccurate notification of UMI placement within the library.
- If possible, please provide the expected size of your library. A gel image is not necessary but can be helpful.
- Charges for lab-prepared libraries will include the cost of pre-sequencing electrophoresis, fluorometric quantification, and qPCR QC.

Core Deliverables:

If your data does not meet these criteria, or if you have other questions/concerns regarding your data, please contact us at genomics@vai.org to address your concerns.

Short Read Data:

- A minimum of the requested number of reads for the entire project, 80% bases Q30 (99.9% accuracy) or greater, delivered in FASTQ format.
- Reads will not be trimmed for adapter read-through.
- Data are delivered in FASTQ format within five business days of run finish. Further analysis may be contracted separately with the Bioinformatics and Biostatistics Core (BBC).
- Libraries that do not meet these specifications will be re-run at the earliest possible date; however, there may be a wait for a run that will meet or exceed the requested parameters. Libraries may not be run on the same read length flowcell for the sake of expediency; however, reads will be trimmed to the requested read length for data release.

Long Read Data:

- Data will be delivered in both FASTQ and POD5 format, along with the run report.
- Q-score filtering thresholds will be left at instrument default (8). Passed reads are delivered; failed reads can be supplied upon request.
- All barcoded libraries are demultiplexed. The barcode is part of the read, and barcode trimming is turned off so that any resulting data will contain a barcode sequence.
- We strive to deliver data that meets or exceeds instrument specification metrics. However, there are no guarantees of output, read length, or read counts. Many factors can affect these metrics, such as species, DNA length, DNA quality, contaminants, library prep type, GC content, etc.

Deliverable Exceptions:

While we will make every effort to create successful sequencing runs with high-quality data output, we cannot be responsible for improperly prepared samples. Therefore, we make **no guarantees** on the quantity and quality of data generated from sequencing if:

- The user prepares libraries, as we have minimal control over quality.
- Samples that do not meet VAI minimum submission requirements are submitted.
- Custom barcoding methods are used. Barcodes placed at the 5'end of the sequence with no redundancy are especially prone to read errors.
- Libraries use a shorter than the recommended +8+8 barcoding strategy.
- Species are sequenced for which there is no reference genome available.

Data Storage:

- Deliverable sequence data will be stored in an HPC download directory accessible by your lab.
- Data for labs outside of VAI will be stored on a lab-specific directory shared on Globus.
- Due to space considerations, the Genomics Core will remove files after two weeks; please be prompt when downloading your data.
- Raw binary output from each sequencing run will be stored by the VAI Genomics Core for 60 days. If you wish to have a copy of this data, you **must** contact the core manager at genomics@vai.org within one month of the beginning of your sequencing project to facilitate the transfer.

Leftover Samples:

Any remaining samples will be stored for 45 days after project completion; labs are welcome to collect any leftover materials during this period. A notification of disposal will be sent via email both two weeks and one day before disposal, after which samples will be discarded if the lab has not contacted the Genomics Core. Please get in touch with genomics@vai.org for pick-up arrangements. Libraries generated by the VAI Genomics Core will be stored indefinitely.