VAI Single-Cell Sequencing Policies

Please thoroughly review these policies before submitting samples. 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:

1. Before submitting an initial project, VAI researchers must meet with Genomics Core staff to discuss sequencing goals and details of starting material.
2. Projects are received on a first-come, first-served basis and will not be entered into the project queue until samples are in possession of the Genomics Core and a project submission form is completed on our website (https://vari.my.site.com/submission/s/login/). Please get in touch with the Genomics Core directly if any extenuating circumstances are associated with a project being submitted.
3. 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 project. In the event of a delay greater than a week, the Core will email the customer to inform them of an updated run date.
4. After samples are processed, Core personnel will initiate a CrossLab request for the project. Please review and authorize the project if everything looks correct.

PHI: VAI Genomics Core does not accept samples or documentation that contain PHI. In the unlikely event that PHI is transferred incidentally, the Core will inform the submitting lab within 5 business days of becoming aware of the disclosure and will subsequently destroy any identifiable PHI (including samples and ePHI).

Customer Responsibilities:

- Dropped-off samples must be prepared to VAI specifications unless the Core Director has granted previous permission. A summary of submission requirements can be found on pg. 3 of these policies.

VAI Deliverables:

- A minimum of the requested number of bases 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 5 business days of run finish. Further analysis may be contracted separately with the Bioinformatics and Biostatistics Core.
- Libraries that do not meet these specifications will be re-run as soon as possible; 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 length flowcell for expediency.

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:

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

If your data do 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

Minimum and Optimal Sample Requirements:

You will be emailed if the following initial QC requirements are unmet. You may 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 will 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 (internal users) or contact genomics@vai.org with specific questions. **All single-cell projects require an initial meeting with a Genomics Core associate before project initiation.**

Most single-cell projects will be sequenced by themselves; please contact the genomics core if you need help deciding which sequencer/flowcell will work best for your project.

10X Genomics FLEX samples

- FLEX probes are currently only available for the whole genome of mice or humans if working with an alternative species or specific mutation detection is required; please discuss this in the initial project meeting.
- We recommend using 10X Genomics' provided fixation kit and instructions for the best results. Please perform pre- and post-fixation and storage testing to gauge the cell/nuclei loss.
- We recommend using highly viable cell/nuclei suspensions (>80%); we request that the QC of cell/nuclei suspensions include images from the luna (flow core).
- Submission of at least 300,000 cells or 500,000 nuclei fixed and properly stored in 10X's storage buffer. Follow instructions for long-term storage at -80 C and submit frozen samples to the Core.
- 10X recommends 10 K sequencing reads/cell. Select a custom run length of 28x10x10x90.

10X Genomics single cell/nuclei samples (scRNA 3'or 5')

- Starting sample concentrations should be between 500-12000 live cells/nuclei per uL as counted by hemocytometer or automated cell counter. Cell viability should be >85%. We recommend doing preliminary QC imaging on the Luna.
- Plan to bring at least 100K cell total, in a max volume of 110 uL. Protocols exist for lower cell numbers if needed.
- Libraries can be made from 500–10,000 cells. Superloading is possible. Please discuss with Core staff before proceeding.
- Hashtagging and CITE/REAP seq can be accommodated; the core MUST know which antibody set is being used.
- 10X recommends 25K sequencing reads/cell + 5K reads/cell for CITE, VDJ, or Hash library; however, in our experience, 30–40K reads/cell produces a cleaner dataset. Select a custom run length of 26x10x10x90

10X Genomics single nuclei samples (ATAC or multiome)

- Starting sample concentration should be between 3,000–7,700 nuclei per uL in 10X's nuclei buffer provided by 10X genomics.
- Nuclei count and quality for ATAC can be counted by hemocytometer or automated cell counter.
- Nuclei count and quality for Multiome must be counted by automated cell counter such as the Luna located in the Flow Cytometry Core as well as the Cytoflex to look for small debris.
- For ATAC-seq, 10X recommends a sequencing depth of 25,000 read pairs per nucleus. Please select a custom read length of 50x8x16x50.
- For multiome projects, see above recommendations for the ATAC-seq libraries and calculate an additional 20,000 read pairs per nucleus for the gene expression libraries. Custom read length: 28x10x10x90

STORM-seq

- Please reach out to the Genomics Core to set up a meeting with the Flow Cytometry Core to go over plating requirements and sequencing for STORM-seq projects. Cells must be sorted into STORM lysis buffer provided by the Genomics Core.

Alternative single cell projects (parse, scale biosciences etc.)

- See meet with the Genomics Core to go over any specifics for any alternative single cell/nuclei project.

10X Genomics cDNA submission:

- If possible, please drop ship reagents from 10X directly to the Genomics Core. The Genomics Core takes no responsibility for reactions that fail if the Core uses reagents dropped off by the lab.
- Please complete the preps through the clean-up of the cDNA, and only stop at a safe stop that allows for storing the clean cDNA at -20 C for up to 4 weeks.
- It is the submitter's responsibility to ensure that the Core receives the cDNA with enough time to process the samples within the recommended time limits provided by 10X.
- The Genomics Core will QC the cDNA and inform the lab of the results before moving onto library prep.
- Primers for any additional VDJ, hashing, etc. must be provided and communicated clearly to the Genomics Core.

Lab Prepared Libraries:

- At least 10uL of library at ≥2ng/uL, measured by fluorometry only, in 10mM Tris pH 8.0 or water. EDTA should be avoided as it retards downstream sequencing reactions.
  - More material may be needed for libraries requiring higher amounts of reads, such as full flow cell lanes.
- Provide all index *SEQUENCES*, as the same index number from different library preparation kits do not always refer to the same sequence. Sequences are provided by the vendor, usually either in an appendix or separate supporting table.
- Index sequences should be provided in the following orientation regardless of intended sequencer to be used (contact the Genomics Core if you are unsure of 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 an expected size for your library. A gel image is not necessary but can be helpful.
- Charges for lab prepared libraries will include the cost of pre-sequencing Bioanalyzer, fluorometric quantification, and qPCR QC.

Sample Delivery and Labeling:

Samples should be submitted in 1.5mL microcentrifuge tubes, unless there are ≥24 in which case they should be submitted in a full skirted 96-well PCR plate (such as Eppendorf twin.tec). Do not submit samples in strip tubes or 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 labelled with the project number from our submission portal (PRxxxxxx) and sample name (matching what is submitted in the portal). Sample names will be 6-8 characters in length and must be composed only of alphanumeric characters. Any spaces, dashes, underscores, or special characters will be removed, as they are not compatible with the sequencing software.

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 share on Globus. Files will be removed by the Genomics Core after 2 weeks due to space considerations; please be prompt in 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 1 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, customers are welcome to collect any leftover materials during this period. A notification of disposal will be sent via email both 2 weeks and 1 day prior to disposal, after which samples will be discarded if the lab has not contacted the Genomics Core. Please contact genomics@vai.org for pick-up arrangement. Libraries generated by the VAI Genomics Core will be stored indefinitely.