NEW YORK GENOME CENTER® | Collaboration Research Discovery

Bulk RNA Sequencing

NYGC offers comprehensive bulk RNA sequencing services to support a broad range of transcriptomic research applications. These services include sample quality control, library preparation, sequencing, bioinformatics analysis, and secure cloud-based data delivery. Our highly automated, GCLP-compliant laboratory workflows and validated analysis pipelines generate high-quality, reproducible results that are compatible with major genomic research initiatives. Our scientific team works with you during project planning to select the library preparation workflow best suited to your sample type, RNA quality, and research objectives. Standard bioinformatics analysis pipelines are available for germline, somatic, and patient-derived xenograft (PDX) RNA sequencing projects, with customized analysis options available for specialized research needs.

We offer two stranded RNA sequencing library preparation workflows:

- **Poly(A)-Selected mRNA Sequencing** – Designed for high-quality RNA samples, this workflow enriches for polyadenylated messenger RNA (mRNA) to provide robust gene expression profiling of protein-coding genes and long intergenic non-coding RNAs (lincRNAs).
- **Total RNA Sequencing** – This workflow uses ribosomal RNA (rRNA) depletion (optimized for human, mouse, and rat samples) to capture a broader range of RNA species. It is the preferred option for lower-quality or partially degraded RNA (typically RIN < 8) and is well suited for challenging sample types.

For RNA isolated from **whole blood**, NYGC recommends globin depletion to reduce abundant globin transcripts and maximize sequencing efficiency. Globin depletion is available as part of our Total RNA workflow, and our team will review your samples to recommend the most appropriate library preparation strategy.

For **formalin-fixed, paraffin-embedded (FFPE)** samples, NYGC offers a specialized RNA sequencing workflow optimized for degraded RNA. This approach incorporates targeted enrichment of coding regions to improve data quality and support robust gene expression analysis from archived tissue specimens commonly used in translational and oncology research.

From project planning through data delivery, NYGC combines optimized laboratory workflows with validated bioinformatics pipelines to deliver high-quality, research-ready transcriptomic data.

RNA Submission Requirements - Bulk RNAseq

The following are the standard minimum RNA input requirements intended to maximize sequencing performance and data quality. If your samples fall outside these specifications, please contact us before submission. In many cases, we can recommend alternative workflows or evaluate samples on a project-specific basis.

Poly(A)-Selected Stranded mRNA Sequencing

- **Input amount:** Preferred 2 µg total RNA (minimum 500 ng)
- **RNA quality:** DNase-treated RNA with a **RIN ≥ 8**, assessed by an Agilent Bioanalyzer or equivalent
- **Sample volume:** 50–100 µL in nuclease-free water
- **Quantification:** RiboGreen or equivalent fluorescence-based assay

Total Stranded RNA Sequencing

- **Input amount:** Preferred 500 ng total RNA (minimum 100 ng)
- **RNA quality:** DNase-treated RNA with a **RIN ≥ 7**, assessed by an Agilent Bioanalyzer or equivalent
- **Sample volume:** 15–30 µL in nuclease-free water
- **Quantification:** RiboGreen or equivalent fluorescence-based assay

FFPE RNA Sequencing

- **Input amount:** Minimum 400 ng total DNase-treated RNA
- **Sample volume:** 25 µL in nuclease-free water
- **Quantification:** RiboGreen or equivalent fluorescence-based assay

Standard Germline/Non-Somatic Analysis & Deliverables - Bulk RNA Sequencing

NYGC provides comprehensive bioinformatics analysis as part of our RNA sequencing services, delivering high-quality, research-ready data through extensively tested analysis pipelines. Our standard RNA sequencing pipeline includes high-quality read alignment to the human reference genome, comprehensive sequencing and alignment quality control, gene- and transcript-level expression quantification, normalization, and differential gene expression analysis (when appropriate). The pipeline incorporates widely adopted, community-standard bioinformatics tools that have been extensively validated and are routinely used throughout the genomics research community. For studies involving complex experimental designs, our bioinformatics team has extensive experience analyzing large cohorts, multiple experimental conditions, longitudinal studies, and datasets requiring correction for technical batch effects. In addition to our standard workflows, customized bioinformatics analyses are available to address specific scientific questions or project requirements.

Analysis includes:

- Read alignment to the human or mouse reference genome
- Comprehensive sequencing and alignment quality control
- Duplicate marking and alignment quality assessment
- Gene-level expression quantification and annotation
- Normalized gene expression matrices
- Differential gene expression analysis (when appropriate for the study design)

Deliverables:

- Raw sequencing data
- Aligned sequence files (BAM/BAI)
- Sequencing and alignment quality control reports
- Raw and normalized gene expression count matrices
- Regularized expression matrices for downstream analyses

Standard Somatic Analysis & Deliverables - Bulk RNA Sequencing

Our standard somatic RNA sequencing analysis pipeline builds upon our validated RNA sequencing workflows to provide comprehensive characterization of tumor transcriptomes. The pipeline includes high-quality read alignment, sequencing and alignment quality control, gene- and transcript-level expression quantification, normalization, differential gene expression analysis (when appropriate), and detection of biologically relevant gene fusions.

For cancer research projects that include both tumor RNA sequencing and whole-genome sequencing (WGS), NYGC's Computational Biology team can integrate these complementary datasets to provide a more comprehensive view of tumor biology. Combined DNA and RNA analyses enable researchers to investigate genomic alterations alongside their functional consequences, supporting studies of gene expression, structural variants, fusion events, and other molecular features that contribute to cancer development and progression.

Analysis includes:

- Read alignment to the reference genome
- Comprehensive sequencing and alignment quality control
- Gene- and transcript-level expression quantification
- Gene annotation and normalized expression analysis
- Detection of gene fusion events

Deliverables:

- Raw sequencing data (FASTQ, upon request)
- Aligned sequence files (BAM/BAI)
- Sequencing and alignment quality control reports
- Raw and normalized gene expression count matrices
- Fusion detection reports

Standard Patient-Derived Xenograph (PDX) Analysis & Deliverables - Bulk RNA Sequencing

NYGC's specialized PDX RNA sequencing pipeline is designed to accurately distinguish human-derived tumor transcripts from mouse background, enabling reliable downstream analysis of tumor gene expression.

Analysis includes:

- Separation of human and mouse sequencing reads
- Alignment of human-derived reads to the reference genome
- Comprehensive sequencing and alignment quality control
- Gene expression quantification and annotation
- Detection of gene fusion events

Deliverables:

- Raw sequencing data (FASTQ)
- Filtered alignment files (human-derived reads)
- Sequencing and alignment quality control reports, including human/mouse read proportions
- Gene expression count matrices
- Fusion detection reports

Frequently Asked Questions – Bulk RNA Sequencing

What information is needed to start a project?

To begin planning a RNAseq project, we typically request:

- Number and type of samples
- Sample source and preservation method
- Desired sequencing coverage
- Required timeline
- Data delivery requirements
- Any planned downstream analyses

How are projects initiated?

Once project scope and service requirements have been finalized with your NYGC Project Manager, NYGC will provide a formal service quote outlining the agreed-upon deliverables, pricing, and project details. To begin project execution and submit samples, NYGC requires a signed copy of the quote and an approved purchase order (PO). For large-scale projects, additional discussions regarding customized terms and conditions may be accommodated.

What is the expected turnaround time?

All turnaround times can be found on our 'main research services page'.

What is a standard submission size?

Our standard bulk RNAseq projects typically begin with a minimum of 50 to 100 samples depending on the sequencing depth. For larger studies, including projects with 500+ samples, we work with PIs to develop a sequencing and delivery plan using rolling batches to optimize turnaround time and data availability. Volume discounts can apply.

For smaller projects, please contact us at ProjectManagement@nygenome.org to discuss project-specific options.

How should RNA be submitted?

For U.S.-based submissions, once a project has been quoted and a purchase order has been received, we provide NYGC-barcoded tubes in a matrix rack, an electronic sample manifest, and an overnight return shipping label for sample submission. Note that the tube orientation is critical, and tubes should not be rearranged.

Your Project Manager will coordinate international submissions, including shipping requirements and documentation.

What RNA input requirements do you recommend?

We recommend submitting high-quality RNA that meets our standard quantity and quality requirements. Specific requirements may vary depending on sample type and requested sequencing depth. Our team will review your sample information during project planning to confirm suitability before submission.

Will you check RNA quality and quantity before library preparation?

Yes. All projects include an initial quality control (QC) assessment to confirm RNA quantity and evaluate RNA quality. QC results, including RNA integrity metrics, will be shared with you for review before proceeding with library preparation.

Do you have experience working with RNA from whole blood?

Yes. Because whole blood RNA often contains high levels of globin transcripts, which can reduce the efficiency of sequencing and limit the detection of other biologically relevant transcripts, we generally recommend incorporating a globin depletion step when appropriate. Globin depletion can improve overall data quality and increase sensitivity for downstream gene expression analyses. Currently, globin depletion is available as part of our bulk Total RNA sequencing workflow. Our team can help determine whether this option is appropriate based on your sample type and research objectives.

How do you handle low-quality RNA or FFPE samples?

For degraded or lower-quality RNA samples, we recommend Total RNA sequencing rather than poly(A)-selected mRNA sequencing. Total RNA workflows are generally more compatible with fragmented RNA, while poly(A) enrichment relies on intact poly(A) tails, which may be compromised in lower-quality samples. To maximize data quality from challenging FFPE samples, NYGC employs a specialized RNA sequencing workflow that combines optimized library preparation with a reverse transcriptase engineered for RNA sequencing. Our Project Management team can help determine the most appropriate RNA sequencing workflow based on your sample quality and research objectives.

What sequencing depth options are available?

We offer a range of sequencing depths depending on study goals. Our team can help determine the appropriate coverage based on your research objectives, sample type, and downstream analysis needs.

What quality metrics will be provided?

We provide sequencing and QC metrics to help evaluate data quality, including metrics related to sequencing performance, alignment, coverage, and sample quality.

What species are supported?

Our standard analysis pipelines support human, mouse, and PDX samples. Library prep kits typically are tested only on mammalian samples. Projects on other species may be accepted and would require custom analysis.

How is data returned?

We support data delivery to a variety of client-managed endpoints and storage environments. Before delivery, we can perform a test transfer to confirm connectivity and ensure efficient transfer of large sequencing datasets.

What data formats are provided?

Data delivery formats can be customized based on project requirements. Common deliverables include standard alignment files (BAM) and associated quality metrics. Additional analysis outputs can be discussed during project planning.

Do you provide secondary analysis?

Yes. Our standard sequencing services include secondary analysis options, as described in the Analysis and Deliverables sections above. These services provide processed data and analytical outputs designed to support downstream research needs. For projects requiring additional customization, our bioinformatics team offers consulting services to address specific research questions, develop tailored analytical approaches, and generate publication-ready visualizations and reports.

For more collaborative engagements involving deeper scientific partnership, custom analyses, or method development, collaboration agreements may be established. Please contact our

Project Management team at ProjectManagement@nygenome.org to discuss your project needs.

Can samples be submitted anonymously/de-identified?

Yes. **Samples submitted for research sequencing services must be de-identified prior to submission.** Sample identifiers should not contain personal health information (PHI) or other identifying information. As part of our sample intake process, submitted sample information is reviewed for compliance with these requirements. If identifying information is detected, the sample record will be flagged and the submitting party will be notified so that the information can be corrected. Any records containing unintended identifying information will be removed and handled according to established data management procedures.

For clinical sequencing services, please contact us at https://www.nygenome.org/clinical-genetics/.

How are samples tracked throughout the workflow?

NYGC uses a comprehensive laboratory information management system (LIMS) and integrated informatics infrastructure to provide secure, end-to-end tracking of every sample throughout the sequencing process. From initial sample receipt through library preparation, sequencing, quality control, and data delivery, each sample is tracked with complete process traceability. Our production laboratory has used the LIMS platform since 2015, with integrations across laboratory workflows, instrumentation, robotics, sequencing platforms, and reporting systems. This enables automated tracking of sample status, workflow progression, and quality metrics while supporting thousands of samples across hundreds of active projects. Sequencing operations are supported by additional monitoring and reporting tools that provide real-time visibility into sequencing performance, data generation, quality control metrics, and workflow status. Automated pipelines track key processing steps, including demultiplexing, alignment, quality assessment, and issue resolution, ensuring that potential challenges are identified and addressed promptly. Before data delivery, NYGC's Project Management and Computational Biology teams review sequencing performance, coverage metrics, and project-specific quality requirements to confirm that deliverables meet established standards. This integrated approach provides clients with confidence that their samples are handled securely, tracked throughout the workflow, and delivered with reliable, high-quality results.