

NIH STRIDES

Accelerating biomedical research

Workflow for Bulk RNA-seq and AI-enabled Visualizations

August 11, 2026 and August 13, 2026



NIH STRIDES

Accelerating biomedical research

Part 1: Workflow for Bulk RNA-seq

Aug 11, 2026



Part 1 Agenda

Part 1: Workflow for Bulk RNA-seq

- Introduction to STRIDES and Cloud Lab
- What is RNA-seq ?
- Workflow for RNA-seq analysis



NIH STRIDES

Accelerating biomedical research

STRIDES Initiative Overview

STRIDES is the NIH...
Science and Technology
Research Infrastructure
for
Discovery,
Experimentation, and
Sustainability Initiative



STRIDES leverages **partnerships with commercial cloud service providers** (CSPs) to enable broader access to data and reduce barriers to using the cloud working with large, complex data sets.

STRIDES Enables access to the most **advanced, cost-effective computational infrastructure, tools, and services** through CSP partnerships, empowering NIH-funded researchers to accelerate research outcomes.



STRIDES provides **discounts on cloud services, professional services, and training**, & potential collaborative engagements for NIH intramural and extramural programs/projects that want to use data in the cloud.

What is Cloud Lab?

A no-cost, 90-day (or less) program to try Amazon Web Services, Google Cloud, or Microsoft Azure in an NIH-approved environment. It is not a production research environment or a place to work with sensitive data. Services help trainees grow their cloud skills:



Receive Account and Credits

Use your account for up to 90 days or until you use your \$500 of free credits, whichever is first



Access Tutorials

Across 3 cloud providers that demonstrate how to run real AI, bioinformatics, and data science workflows



Support

One-on-one support from a team of technical and bioinformatics NIH experts



Learn More!

See [Overview of NIH Cloud Lab & GitHub Resources](#), scan the QR code or visit cloud.nih.gov/resources/cloudlab.

NIH STRIDES

Accelerating biomedical research

NIH Use Cases

Evaluate Utility & Cost

Provides an easy route to evaluate the cloud's utility/cost for a project without major time or financial commitments.

Develop New Tools

Allows experienced teams to prototype new architectures and evaluate software and hardware combinations.

Learn New Skills

Simplifies access to tools and cloud environments that participants can use for training purposes.

Cloud Lab Lowers Barriers to Cloud Adoption

Challenges to Cloud Adoption

Acquisition is complex and time-consuming.

Security protections are unclear.

Costs are unpredictable.

Training deficit is substantial.

Cloud Lab Lowers Barriers to Entry

Eligible users can be onboarded to Cloud Lab and receive free cloud credits (up to **\$500**).

Periodic scanning is conducted for PHI and PII data.

Cost guardrails are implemented.

100+ tutorials have been developed, with more coming.

How Cloud Lab Works:



Sign up through [NIH Cloud Lab](#)



Receive cloud credits to learn in a secure cloud environment



Access tutorials in GitHub for use on AWS, Azure, and GCP

How is a Cloud Lab Account Different from a STRIDES Account?

Cloud Lab Account

Account Feature	Description
Choice of Cloud Service Providers	AWS, Google Cloud, Microsoft Azure
Funding	Up to \$500 of free credits
Duration	Up to 90 days
Data Limitations	No sensitive data
Use Cases	• Evaluate whether cloud is a good fit for your project • Benchmark costs • Build proofs of concept • Prototype new architectures • Supplement Cloud Training

STRIDES Account

Account Feature	Description
Choice of Cloud Service Providers	AWS, Google Cloud, Microsoft Azure
Funding	Use your own funding
Duration	Determined by your needs and budget
Data Limitations	If your data contains PHI, a business associate agreement (BAA) is required
Use Cases	• The use cases for Cloud Lab • Projects/tools: ○ With sensitive data, ○ With ≥ 1 user, and/or ○ That will last longer than 90 days

Publicly Available Interactive Tutorials Across All Three CSPs!

Cloud Lab offers a [**suite of interactive tutorials**](#) designed to help participants perform viable research in the cloud. Open to all who wish to gain skills, not just Cloud Lab participants!

Scientific

- Accessing Open Datasets
- AI/ML
- BLAST
- FHIR
- Genomic Variant Calling
- GWAS
- Long Read Genomics
- Medical Image Segmentation
- Proteomics
- RNAseq
- scRNAseq
- Radiogenomics
- SRA Data Interaction
- VCF Query with SQL
- ...and more

GenAI

- Model Training
- Guidance & Best Practices
- Introduction to tools across all platforms
- Leverage Large Language Models (LLMs)
- Chatbot Implementation
- Retrieval- Augmented Generation (RAG)

General Cloud Ops

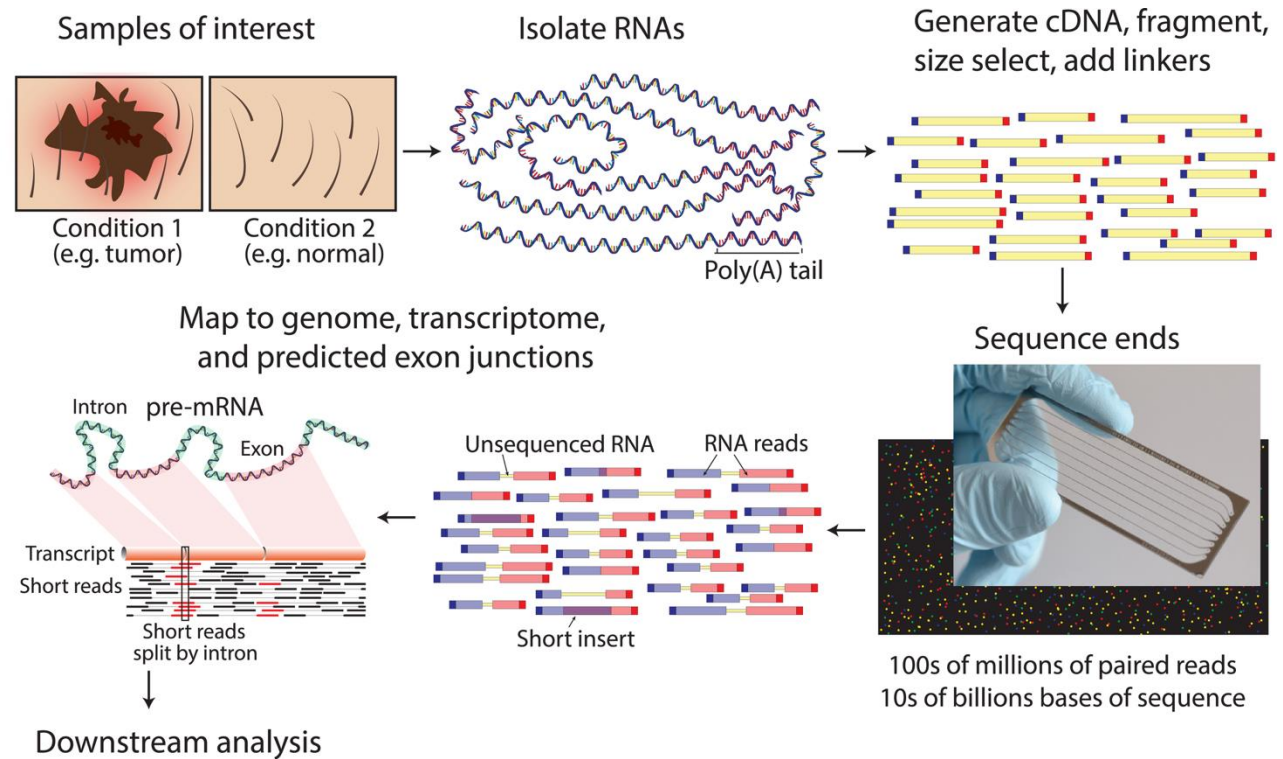
- Access Marketplace Offerings
- Access Public Datasets
- Command Line Tools / SDKs
- Disk Images
- Ingest and Store Data
- Introduction to Cloud
- Jupyter Notebooks
- Serverless Functionality
- Virtual Machines

Computing & Code

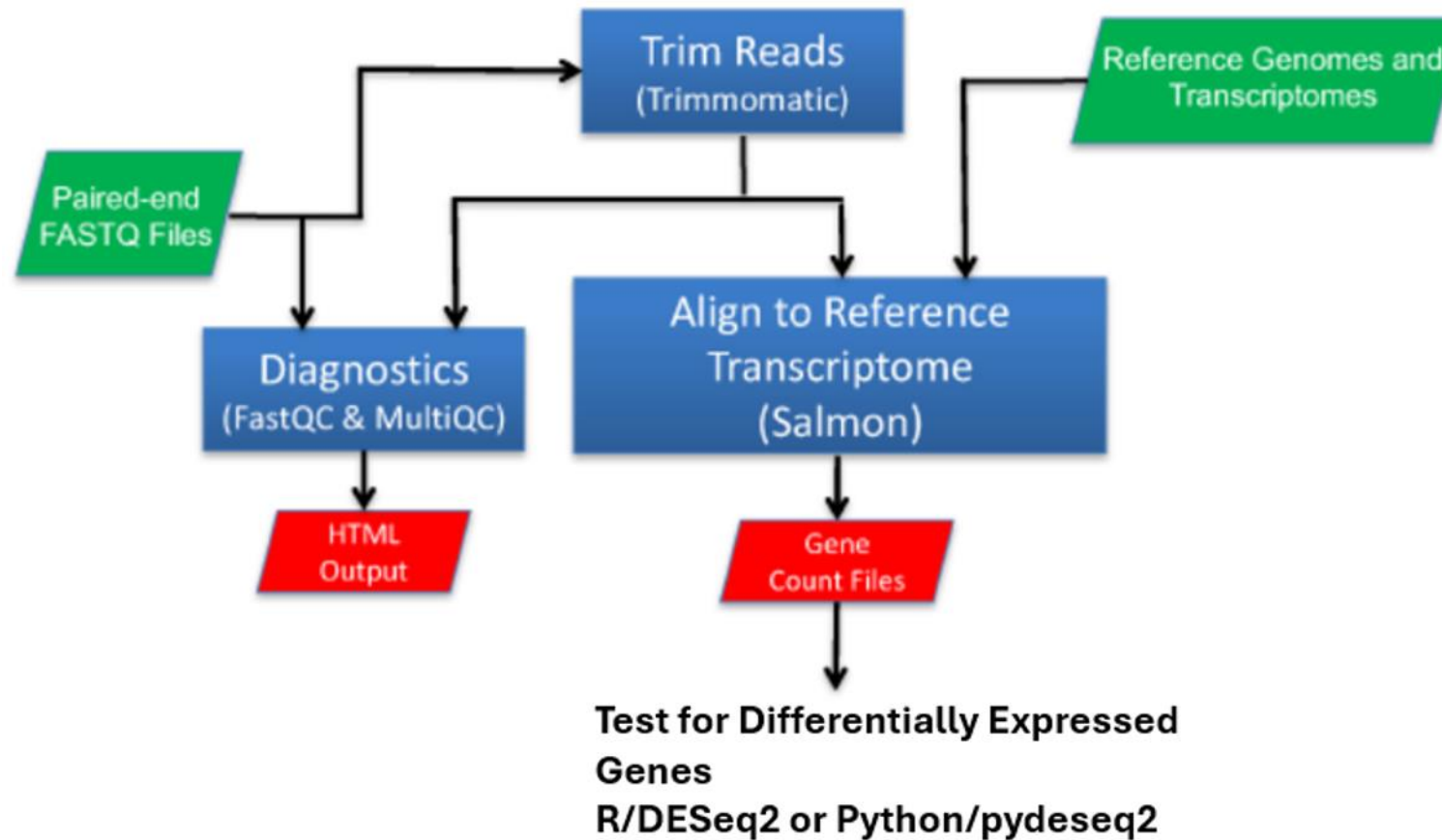
- Conda Environments
- Container Registries
- Git Repos
- HPC Clusters
- Kubeflow
- Serverless Functionality

What is RNA-seq?

- RNA-seq, short for RNA sequencing, is a method for sequencing an entire set of RNA molecules.
- This method involves isolating RNA molecules in the tissue or cell sample, copying those RNAs into DNA and then sequencing the resulting DNA molecules.
- RNA-seq helps researchers determine which segments of DNA have been transcribed into RNA and how much a gene is expressed.



RNA-seq Analysis workflow



Links to Aid in Analyzing RNA-seq

1. https://github.com/STRIDES/NIHCloudLabAzure/blob/main/notebooks/GenAI/initiate_jupyter_notebook/InitiateJupyterNotebook.md
2. [NIHCloudLabAzure/notebooks/rnaseq-myco-tutorial-main/Bulk_RNA-Seq_pipeline.ipynb at main · STRIDES/NIHCloudLabAzure](#)

NIH STRIDES

Accelerating biomedical research

For more information, please visit cloud.nih.gov
or send an email to STRIDES@nih.gov.
