AnVIL: SRA Data

January 30, 2025

Table of Contents

# About this Module

This module is part of a series of books for the Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL) of the National Human Genome Research Institute (NHGRI). Learn more about AnVIL by visiting <https://anvilproject.org> or reading the [article in Cell Genomics](https://www.sciencedirect.com/science/article/pii/S2666979X21001063).

## About SRA

This module will go over how to bring Sequence Read Archive (SRA) data into AnVIL.

You can check out the [SRA website](https://www.ncbi.nlm.nih.gov/sra) to learn more:

Sequence Read Archive (SRA) data, available through multiple cloud providers and NCBI servers, is the largest publicly available repository of high throughput sequencing data. The archive accepts data from all branches of life as well as metagenomic and environmental surveys. SRA stores raw sequencing data and alignment information to enhance reproducibility and facilitate new discoveries through data analysis.

It can be very exciting to learn how much data is at your fingertips! Once you have settled on some data to use, you’ll want to bring it into AnVIL if it’s not already there.

Navigate to the menu on the left to get started!

## Skills Level

*Genetics*

**Novice**: no genetics knowledge needed

*Programming skills*

**Novice**: no programming experience needed

## Learning Objectives

## AnVIL Collection

Please check out our full collection of AnVIL and related resources: <https://hutchdatascience.org/AnVIL_Collection/>

# 1 Student Guide

## 1.1 Activity One

You might want to create a student guide that contains a different subset of Rmd files from your book, or renders to a different output format (e.g. word document). You can specify the output and Rmd files that will be used for the student guide using the \_output.yml and \_bookdown.yml files in the student-guide directory.

## 1.2 Activity Two

Steps of the guide could go here.

# References