ILVO

How to download the metadata file containing the BioSample accession and SRR accession numbers?

The metadata file containing the BioSample accession and SRR accession numbers, which can be downloaded from NCBI (National Center for Biotechnology Information). There are two methods for downloading the metadata.

1. From the search result page

   If you want to download the metadata from all the submissions in the BioProject, select the SRA database, enter the BioProject accession number in the Entrez query and click search. Then, click **Send to** on the top of the page, check the **File** radiobutton, and select **RunInfo** in the pull-down menu. This will generate a tabular **SraRunInfo.csv** file with metadata available for each Run within the BioProject of interest.
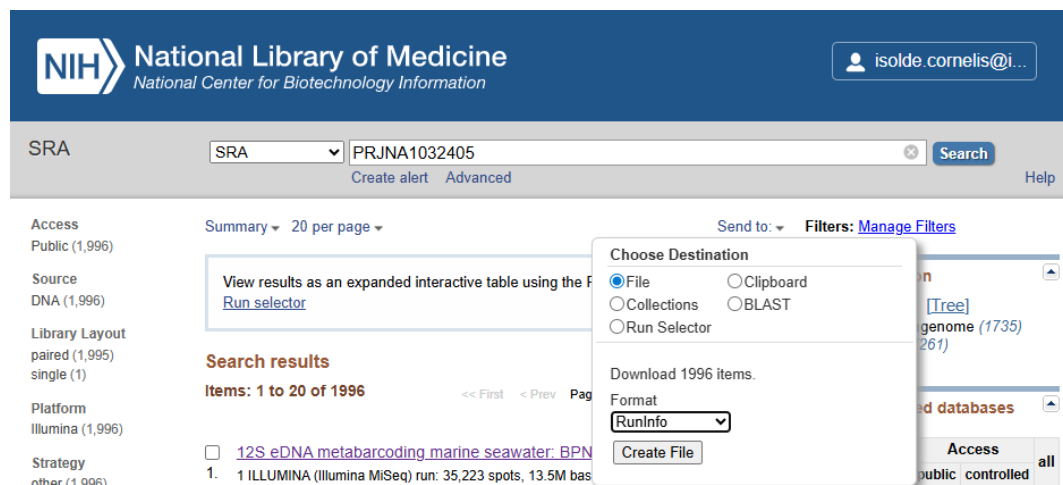
   If you only want the metadata from a single submission within the BioProject you can either filter in the SraRunInfo.csv file based on the Submission accession number, or select the SRA database, enter the Submission accession number in the Entrez query and click search. Then, click **Send to** on the top of the page, check the **File** radiobutton, and select **RunInfo** in the pull-down menu. This will generate a tabular **SraRunInfo.csv** file with metadata available for each Run within the submission of interest.

2. From Run Selector

   A slightly different set of metadata can be downloaded in a tab-delimited file from the Run Selector.

   To download metadata for each Run in your Entrez query (BioProject or Submission accession number (found in the SraRunInfo.csv file)):

- Click **Send to** on the top of the page, check the **Run Selector** radiobutton, and click the button **Go**.

- If necessary, refine your results by using various filters provided by the **Run Selector**'s interface.
- Click the **Metadata** button. This will generate a tabular **SraRunTable.csv** file with metadata available for each Run.

Attention: When directly using the Run Selector interface (Run Selector :: NCBI) it is not possible to enter the Submission accession number, only the BioProject accession number is accepted. Thus the metadata from all submissions within the same BioProject will be downloaded.

The **SraRunTable.csv** file downloaded through the Run Selector interface also contains the information about each run that was added to the Attributes table, this information is missing when downloading the **SraRunInfo.csv** file. The Script as used here uses the **SraRunTable.csv** file.