



Understanding SRA Search Results

Sequence Read Archive Submissions Staff

SRA Response Page Links

The Entrez SRA search response page provides a list of links to SRX/ERX/DRX accessions. What are these accessions and where do these links go?

These accessions are for experiments (SRX) from various studies associated with the search terms you used. The SRX/ERX/DRX links go to an experiment report page that provides:

- Links to each run in the experiment
- Experiment design information
- Submission accession and submitter name
- Study summary and abstract (if available)
- Experiment sample (SRS) information
- Links to a list of experiments (SRX)
- Library information (if available)
- Platform information
- Processing information (base calls and quality scores)
- Spot descriptor information
- Links to related information

Note: If you choose to download files, install the “[Aspera Connect](#)” plug-in to transfer files at a significantly faster rate than ftp.

Each SRX entry in the Entrez SRA search response page contains a list of SRR/ERR/DRR accessions that are also links. What are these accessions and where do these links go?

These accessions are for runs for that experiment (SRR). The links go to the Run Browser page for each run.

The Run Browser page provides you with specific information about the run, including read and intensity data, and **will allow you to download data from a particular run experiment or study.**

You can also **use the Run Browser to search the spots of a particular read for a specific sequence:**

1. Enter sequence data into the “Filter” text box.
2. Click the “Find” button.
3. If your sequence is found, a list of the spots that contain the sequence entered will appear just below the grey box containing the “Find Spots” query, as will links to runs that contain the sequence data you are looking for.m
4. You can also use the “Filter” to search for the name of a spot or the name of a sample pool member.

If you want to download the data you find using the run browser's filter, click the "Filtered Download" button. The response page allows you to download the data in a selected format (filtered, clipped, FASTA and FASTQ).

Confused About which Accession Number to Use

When I use Entrez SRA to search for data in SRA, which link on the search response page do I use to see the data ?

When you get the response page from an [SRA Entrez search](#), you will see a list of experiment accessions (SRX) from studies associated with the search terms you used. The SRX/ERX/DRX links go to an experiment report page that provides:

- Links to each run (SRR) in the experiment
- Experiment design information
- Submission accession and submitter name
- Study summary and abstract (if available)
- Experiment sample (SRS) information
- Links to a list of experiments (SRX)
- Library information (if available)
- Platform information
- Processing information (base calls and quality scores)
- Spot descriptor information
- Links to related information

Note: If you choose to download files, install the "[Aspera Connect](#)" plug-in to transfer files at a significantly faster rate than ftp.

When I use SRA Search browser to search for an SRA submission, which link on the search response page do I use to see the data?

When you get the response page from the SRA "[Search for Related objects](#)", the most helpful link is the [SRA Studies link](#). The link will provide you with a list of links to all the studies associated with the search terms you used. Once you are on the response page, click on the SRP accession number for the study of interest to go to the study summary page. The study summary page includes:

- Submission accession and submitter name
- Study abstract (if available)
- A text description of the study (if available).
- Links to associated NCBI resources (BioProjects, GEO, WGS, dbGaP, PubMed, etc.).
- Links to associated external resources.
- Links to all experiments in study
- Link to Entrez summary documentation for all experiments.
- Links to download study reads

Note: If you choose to download files, install the "[Aspera Connect](#)" plug-in to transfer files at a significantly faster rate than ftp.

Why does SRA have so many different accession types? A single GenBank ID number accesses all the information you need for a particular molecule.

An SRA submission accession number does not represent a single [object](#) the way a GenBank accession number represents a single object (a sequence) submitted to GenBank. The SRA submission accession is a

modular data construct that groups together a number of objects, and therefore has no specific response page.

For example, the GenBank accession [NM_001048036.2](#) represents a single object: a GenBank sequence submission (in this case, a RefSeq mRNA from the domestic dog). An SRA (submission) accession number (e.g. [SRA010122](#)), doesn't represent a single object – it is a reference number that represents 5 distinct data types that are independently reusable.

Examples of repeated and independent use of SRA Accessions:

1. A submitter can reference (link to) a sample (SAMN accession) multiple times from any experiment they submit. A good case in point for this example is HapMap samples, some of which have been referenced (linked to) in experiments from 5 or 6 different submitting centers (see [SAMN00004417](#)).
2. The set of variables captured in an Experiment (SRX) accession is similar to what a biologist might think of as a “library”, and includes the nucleic acid type extracted from the sample, as well as the selection type applied to enrich the intended sequencing target. This information can be referenced (linked to) by any run (SRR) produced using the set of variables described by that SRX.
3. The SRP accession, which provides all the metadata describing a particular study, references (links to) all the experiments (SRX) contained within the study, while each of the SRX accessions for that study all reference the same SRP accession. Since the SRP accession ultimately references (links to) all 5 data types in a study, it can be used as a starting point to access any of the data in that study.

Benefits of SRA modular data construction to the submitter:

- It is a **time-saver** since the submitter can directly reference (link to) existing data or metadata for a study, analysis, sample, experiment, or run from any point within a submission, instead of having to physically enter the data.
- If a particular **instance** of any one of the 5 data types in a particular study is not yet available, once the submitter creates a particular instance of a data type, that data can be referred to again and again in the current submission or in a future submission.

Benefits of SRA modular data construction to the user:

- It is a **time-saver** since the user can look a single instance of any one of the 5 data types from a particular study without having to download the data for the entire study.
- Once the metadata for a particular experiment (SRX) or sample (SRS) is generated, there is **uniformity of description** within the Sequence Read Archive for that particular experiment or sample, since the metadata description can be used over and over again in the same submission or in other submissions.
- The user can **make logical connections between studies** when different studies use the same samples. Because the samples are linked to the experiments, which in turn are interlinked to studies, runs and analyses, it is easy for the user to compare data between studies.

What do the different SRA accessions represent?

There are 6 different SRA accession types:

Accession Prefix	Accession Name	Definition	Example
SRA	SRA submission accession	The submission accession represents a virtual container that holds the objects represented by the other five accessions and is used to track the submission in the archive.	Since the SRA accession number is an artificial packaging construct, there is no example available since the SRA accession number has no specific response page
SRP	SRA study accession	A Study is an object that contains the project metadata describing a sequencing study or project. Imported from BioProject.	HTML
SRX	SRA experiment accession	An Experiment is an object that contains the metadata describing the library, platform selection, and processing parameters involved in a particular sequencing experiment.	HTML
SRR	SRA run accession	A Run is an object that contains actual sequencing data for a particular sequencing experiment. Experiments may contain many Runs depending on the number of sequencing instrument runs that were needed.	HTML
SRS	SRA sample accession	A Sample is an object that contains the metadata describing the physical sample upon which a sequencing experiment was performed. Imported from BioSample.	HTML
SRZ	SRA analysis accession	An analysis is an object that contains a sequence data analysis BAM file and the metadata describing the sequence analysis.	

The first letter of the accession prefix shows which **INSDC** archive the data originated from (S = **NCBI-SRA**, E = **EMBL-SRA**, D = **DDBJ-SRA**):

Data originating from EMBL-EBI:

Accession Prefix	Accession Name
ERA	ERA submission accession
ERP	ERA study accession
ERX	ERA experiment accession
ERR	ERA run accession
ERS	ERA sample accession
ERZ	ERA analysis accession

Data originating from DDBJ:

Accession Prefix	Definition
DRA	DRA submission accession
DRP	DRA study accession
DRX	DRA experiment accession

Table continued from previous page.

Accession Prefix	Definition
DRR	DRA run accession
DRS	DRA sample accession
DRZ	DRA analysis accession

Matching Search Results to SRA Accessions Provided in Publication

How do I know which accessions mentioned in a paper correspond to specific data mentioned in the paper (e.g. which accession represents replicate 1? replicate 2? Illumina? SOLiD?).

The only way you can determine what the SRA accessions provided in a publication actually represent in terms of the data discussed in the publication is to contact the author of the publication and ask.

Currently, there is no requirement for authors to use a specific SRA accession in their publications. If there was, we would recommend that authors use the SRP (study) accession, which would provide the reader/user with a complete overview of the study and a set of links to all the data from that study.

If you are a submitter/author whose paper describes many experiments with identical information, we suggest that you describe how your experiments differ from one another or how the samples used in the experiments differ from each other. If you describe your experiments/samples in this way, a reader/user will be able to more easily differentiate the samples and libraries used in the experiments.

Search Results All Look the Same

Why do all the experiments for a single study all look the same? For example, SRP001156 has 9 different SRX accessions, but when I click on each, the description is the same. Is there a difference between the SRX accessions?

Many of the experiments do have similarities, but you can see some differences if you look closely at the samples used in the experiments as well as the number of runs, spots and base calls. If you still have trouble differentiating between experiments, contact the submitting author. Since SRA is a raw data archive (not curated), submitting authors have the opportunity to provide descriptions of the experiments when they submit, and it is up to them to provide a description that will differentiate one experiment from another in the same study.