



## Searching for Data Generated by a Specific Next-Generation Sequencing Application

Created: August 22, 2011; Updated: March 18, 2014.

### Can I search SRA for data generated by a specific next-generation sequencing application like RNA-seq, ChIP-seq or CLIP-seq?

Yes you can by going to the [Entrez SRA](#) page, and entering “RNA-seq”, “CHIP-seq” or “CLIP-seq” in the search box at the top of the page and clicking the “Search” button.

If a submitter of RNA-seq, CHIP-seq, or CLIP-seq data didn't happen to use the term RNA-Seq, CHIP-seq or CLIP-seq anywhere in their submission, the above search will not find their records. If this is the case, you could try the following:

- Formulate your search using terms that would describe the data generated from that specific technology.

For example, If the application name ChIP-seq wasn't indexed in [Entrez](#) (it is indexed; the name is only being used for the purpose of this example), you could use the following search terms to find data generated by the ChIP-seq method:

- “protein”
- “chromatin”
- “protein” AND “chromatin”

#### Notes:

- The above search terms may not turn up data from the same ChIP-seq studies since SRA is a raw data archive (not curated) and can index only those terms that a submitter provides in the metadata of their submission.
- Depending on which terms the submitter used in their submission, the data generated by these searches may or may not contain the application term (e.g. ChIP-seq), so you may have to click on the “Study Summary” link within the experiment entry of interest to see if the submitter mentions the application name.
- The results for a search of this type will include more than just the data generated by the next-generation sequencing application, so you will have to review the search results yourself and look for data generated by the application of interest.
- If the above search strategy yields nothing, search [PubMed](#) using the application name and limit the search to the title (“ChIP-seq”[Title]) to find studies whose title contains the text: ChIP-seq. Open the PubMed record of interest in the resulting page. If SRA data was submitted for that publication, there will be a link to it in the PubMed record.

**I searched SRA for data generated by a new next-generation sequencing application. I used the name of that application as my search term, but found no records. What did I do wrong?**

If the sequencing technology is relatively new, it may not be indexed in SRA yet, so you won't get any results for your search. If this is the case, create your search using terms that would describe the data generated from that specific technology.

For example, If the application name ChIP-seq wasn't indexed in [Entrez](#) (it is indexed; it is only being used for the purpose of this example), you could use the following search terms to find data generated by the ChIP-seq method:

- “protein”
- “chromatin”
- “protein” AND “chromatin”

**Notes:**

1. The search terms you use may not give you all the SRA data associated with that search term since SRA is a raw data archive (not curated) can only index terms that are provided in the metadata of the submission.
2. Depending on which terms the submitter used in the text of their submission, the data generated by these searches may or may not contain the application term (See note number 1 above), so you may have to click on the “Summary” link within an experiment entry of interest to see if the submitter mentions the application name.
3. The results for a search of this type will include more than just the data generated by the next-generation sequencing application, so you will have to review the search results yourself for data generated by the application of interest (e.g. ChIP-seq).

If the above search strategy yields nothing, search [PubMed](#) using the application name and limit the search to the title (“ChIP-seq”[Title]) to find studies whose title contains the text: “ChIP-seq”. Open the PubMed records of interest in the resulting page. If SRA data was submitted for that publication, there will be a link to it in the PubMed record.

If a next-generation sequencing application is not indexed by name in SRA, please contact [sra@ncbi.nlm.nih.gov](mailto:sra@ncbi.nlm.nih.gov).