

Tutorial for SRA-BLAST:

Search for SRA datasets based on specific topics/conditions

<https://www.ncbi.nlm.nih.gov/sra>

For example, searching for meRIP gives 2295 entries

Limiting by Human, gives 1215 entries

Limiting by RNA-seq, gives 521 entries, etc

Task: Search for presence of TERT_R4 in specific SRA sets: the position of m6A site is highlighted in grey

>TERT_R4_42bp

GGACAACCCGGGGACCGCGCCTCACTCACCTGCACGTGACAG

Selected SRA sets: HeLa meRIP-seq

- Filters activated: RNASeq. [Clear all](#)

Select item 96383361.

[GSM4216728: HeLa hypoxia 24h meRIP-seq; Homo sapiens; RIP-Seq](#)

1 ILLUMINA (HiSeq X Ten) run: 12.3M spots, 3.7G bases, 1.3Gb downloads

Accession:

[SRX7354763](#)

Select item 96383342.

[GSM4216726: HeLa hypoxia 12h meRIP-seq; Homo sapiens; RIP-Seq](#)

1 ILLUMINA (HiSeq X Ten) run: 11.7M spots, 3.5G bases, 1.3Gb downloads

Accession:

[SRX7354761](#)

Select item 96383323.

[GSM4216724: HeLa hypoxia 6h meRIP-seq; Homo sapiens; RIP-Seq](#)

1 ILLUMINA (HiSeq X Ten) run: 11.8M spots, 3.5G bases, 1.3Gb downloads

Accession:

[SRX7354759](#)

Select item 96383304.

[GSM4216722: HeLa normoxia meRIP-seq; Homo sapiens; RIP-Seq](#)

1 ILLUMINA (HiSeq X Ten) run: 15.8M spots, 4.7G bases, 1.7Gb downloads

Accession:

[SRX7354757](#)

TERT: [NM_198253](#)

CLPTML1: [NM_030782](#)

Tool: SRA BLAST

https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&BLAST_SPEC=SRA&SHOW_DEFAULTS=on

1. Open SRA BLAST link, enter target fasta sequence in “Enter Query Sequence” box.

The screenshot shows the SRA BLAST interface. At the top, there's a navigation bar with 'BLAST® » blastn suite' and links for 'Home', 'Recent Results', 'Saved Strategies', and 'Help'. Below this is a 'blastn' tab and the title 'Sequence Read Archive Nucleotide BLAST'. The main section is 'Enter Query Sequence', which includes a text box for 'Enter accession number(s), gi(s), or FASTA sequence(s)' containing a sample FASTA sequence, a 'Query subrange' section with 'From' and 'To' fields, and an 'Or, upload file' section with a 'Choose File' button and 'No file chosen' text. There's also a 'Job Title' field. Below this is the 'Choose Search Set' section, which has a text box for 'Enter an SRA accession (experiment, study, or submission), title, the scientific name or tax id.' and an 'Add organism' button. The 'Program Selection' section has radio buttons for 'Optimize for' with options: 'Highly similar sequences (megablast)', 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)' (which is selected). There's also a 'Choose a BLAST algorithm' link. At the bottom, there's a 'BLAST' button and a checkbox for 'Show results in a new window'.

2. enter SRX number set on “Choose Search Set” section, can add multiple SRA by click “Add organism”

This screenshot shows the same SRA BLAST interface, but with the 'Choose Search Set' section expanded. It shows a list of SRA experiment sets (SRX) with their accession numbers and titles. The list includes: 'SRX7354763 (Homo sapiens taxid:9606; run:SRR10677735)', 'SRX7354761 (Homo sapiens taxid:9606; run:SRR10677733)', 'SRX7354759 (Homo sapiens taxid:9606; run:SRR10677731)', and 'SRX7354757 (Homo sapiens taxid:9606; run:SRR10677729)'. There's an 'Add organism' button next to each entry. The 'Program Selection' section remains the same, with 'Somewhat similar sequences (blastn)' selected. The 'BLAST' button and 'Show results in a new window' checkbox are still at the bottom.

3. Select “somewhat similar sequences (blastn)” and hit **BLAST**
4. The result will show on next page with multiple tables, select “See details” can show relation between SRX number and SRR number

Job Title

Nucleotide Sequence

RID

NDC8N3U1016

Search expires on 10-25 23:36 pm

Download All

Program

BLASTN

Citation

Database

SRA

See details

Filter Results

Percent Identity

to

E value

to

Query Coverage

to

Filter

Reset

SRA Blast search set information

SRX7354763	SRR10677735
SRX7354761	SRR10677733
SRX7354759	SRR10677731
SRX7354757	SRR10677729

Descriptions

Graphic Summary

Alignments

Sequences producing significant alignments

Download

Manage columns

Show 100

select all

26 sequences selected

Graphics

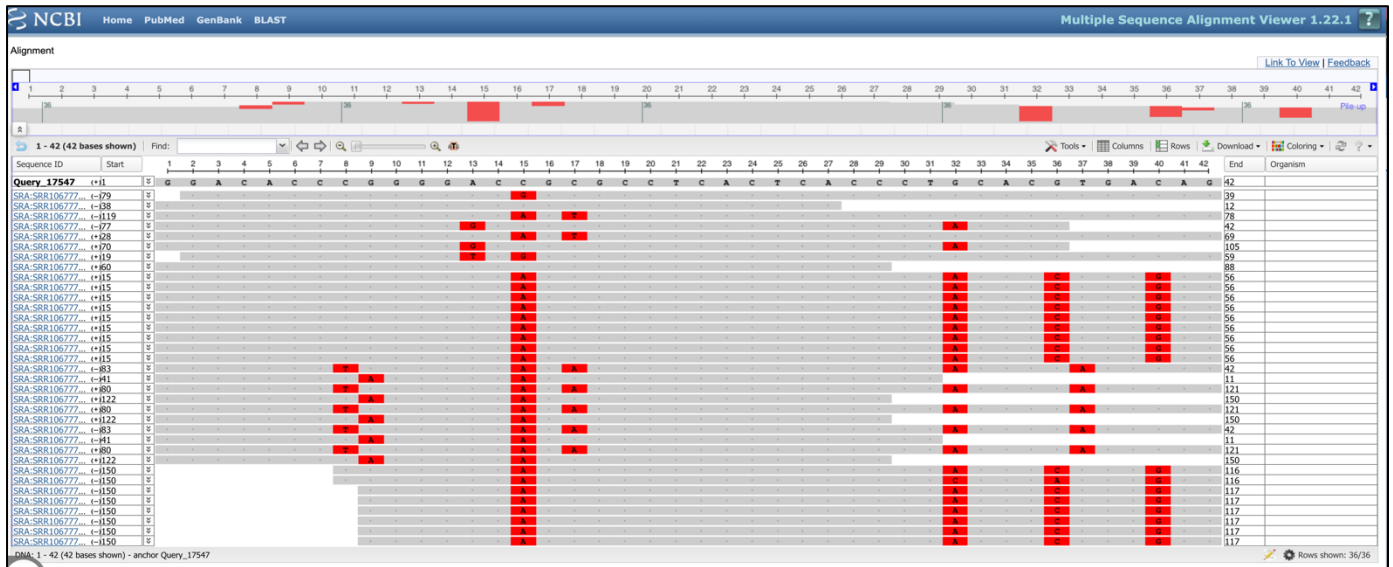
Distance tree of results

MSA Viewer

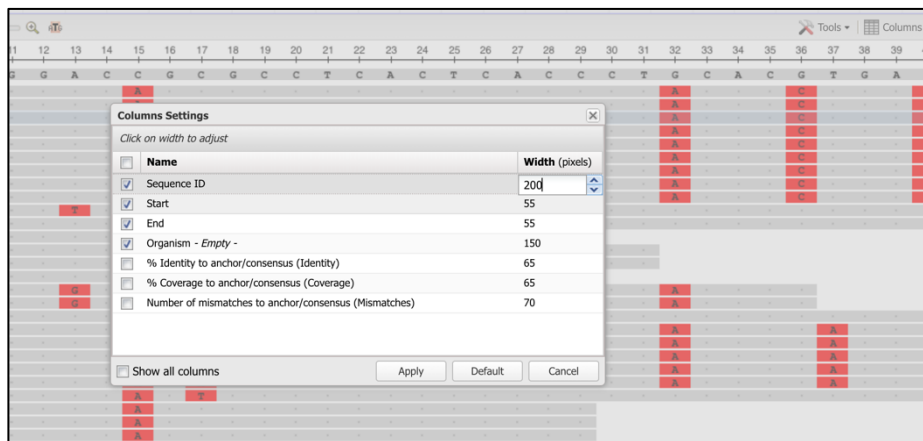
	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	10847216	70.7	120	100%	8e-11	97.56%	150	SRA:SRR10677735.10847216.2
<input checked="" type="checkbox"/>	2102563	68.0	125	100%	3e-10	95.24%	150	SRA:SRR10677733.2102563.2
<input checked="" type="checkbox"/>	2102563	68.0	125	100%	3e-10	95.24%	150	SRA:SRR10677733.2102563.1
<input checked="" type="checkbox"/>	10847216	66.2	119	100%	1e-09	95.12%	150	SRA:SRR10677735.10847216.1
<input checked="" type="checkbox"/>	12178527	59.0	59.0	100%	1e-07	90.48%	150	SRA:SRR10677735.12178527.1
<input checked="" type="checkbox"/>	11006052	59.0	59.0	100%	1e-07	90.48%	150	SRA:SRR10677735.11006052.1
<input checked="" type="checkbox"/>	7729281	59.0	59.0	100%	1e-07	90.48%	150	SRA:SRR10677735.7729281.1

Also check the Graphic Summary and Alignment Tabs

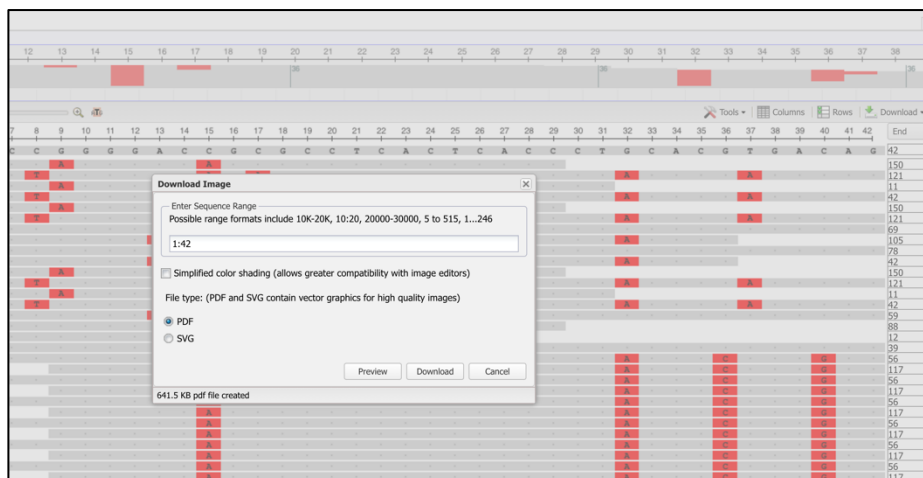
- To view all the alignment, click “MSA Viewer” on the right side of page. For example (https://www.ncbi.nlm.nih.gov/projects/msaviewer/?anchor=0&coloring=diff&key=NCID_1_20641650_130.14.18.128_9147_1666626263_588037759_0MetA0_S_NC_PhyloTree&columns=d:200,b:55,x:17,aln,e:55,o:150) This link is copied from select “Link to View” on the top right.



- It can change the size of the column text to see all the information by selecting “Columns” and type desired Width (Pixels) in the box and hit Apply.



- To output pdf format of alignment, select “Download” button and select range and type of output format (PDF or SVG). Click “preview” to save the file.



8. The output will look like:



9. Conclusions:

- TERT-R4 sequences are detectable in PolyA sequences (read the protocol for sample preps):

For meRIP-seq, the libraries were constructed by Truseq Stranded mRNA Sample Prep Kit (Illumina) according to the manufacturer's instructions and quantified by BioAnalyzer High Sensitivity DNA chip, and then deeply sequenced on the Illumina HiSeq X10 to generate 150-bp paired-end reads. For RNA-seq library, mRNA enrichment, cDNA synthesis, adaptor addition, circularization, PCR amplification and library examination were performed on the BGISEQ 500 at Beijing Genome Institute (BGI; Shenzhen, China)

- Read the reporting paper:
[Wang YJ et al.. "Reprogramming of m⁶A epitranscriptome is crucial for shaping of transcriptome and proteome in response to hypoxia.", *RNA Biol*, 2021 Jan;18\(1\):131-143
<https://pubmed.ncbi.nlm.nih.gov/32746693/>](https://pubmed.ncbi.nlm.nih.gov/32746693/)
- TERT_R4 expression is increased on hypoxia compared to normoxia conditions in HeLa cells