

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
GGACACCCGGGGACCGCGCCTCACTCACCTGCACGTGACAG

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](#)

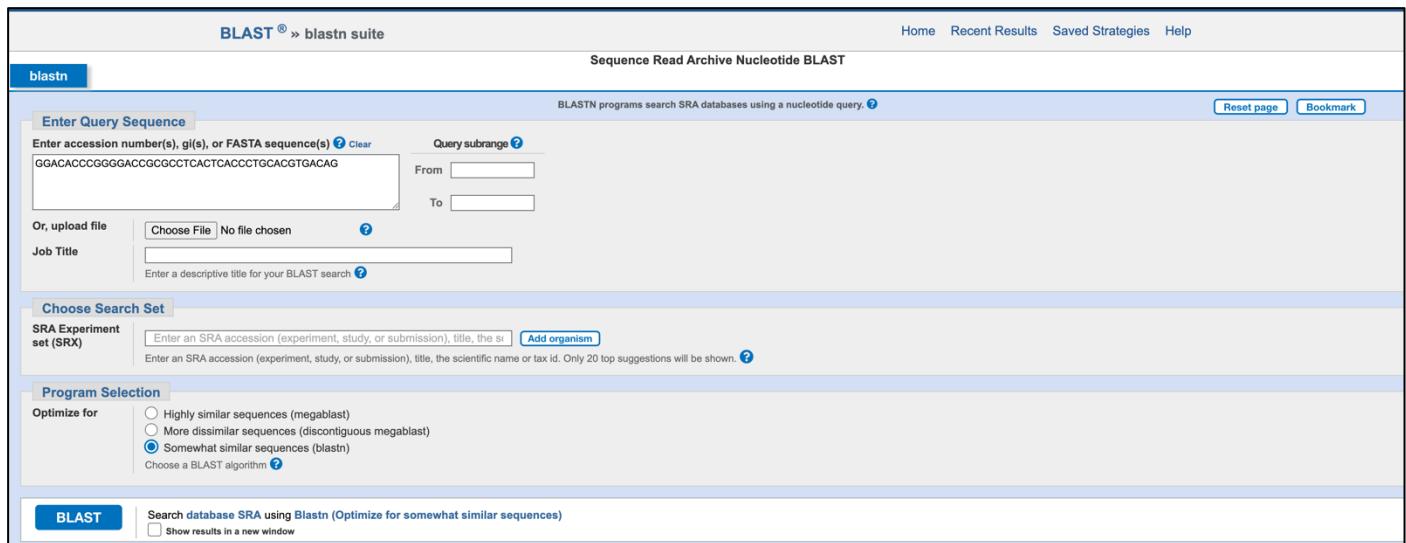
TERT: [NM_198253](#)

CLPTML1: [NM_030782](#)

Tool: SRA BLAST

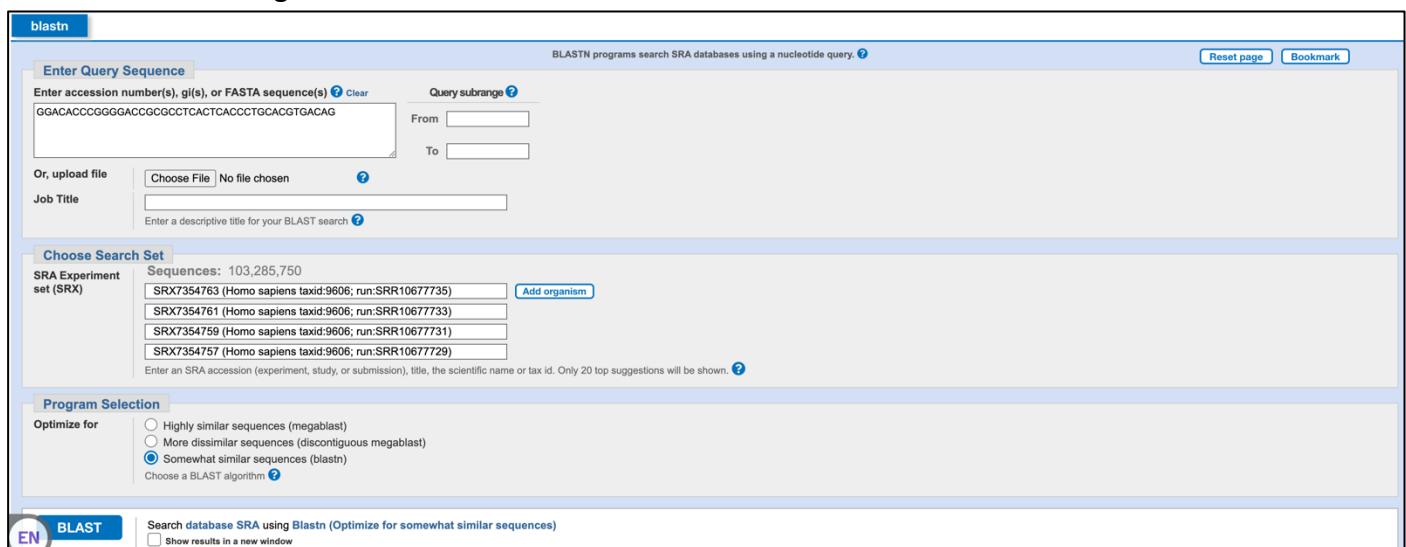
https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&BLAST_SPEC=SRA&SHOW_DEFAULTS=on

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



The screenshot shows the 'blastn' search interface. In the 'Enter Query Sequence' section, a target sequence 'GGACACCCGGGGACCGCGCCTCACTCACCTGACGTGACAG' is entered into the text input field. Below the input field are 'From' and 'To' date selection boxes. Underneath these are 'Or, upload file' and 'Job Title' fields. The 'Choose Search Set' section is visible, showing an 'SRA Experiment set (SRX)' input field with 'SRX' entered, and 'Add organism' and 'Search database SRA using Blastn' buttons. The 'Program Selection' section shows 'Optimize for' radio buttons for 'Highly similar sequences (megablast)', 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)', with 'Somewhat similar sequences (blastn)' selected. The 'BLAST' button is at the bottom.

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



The screenshot shows the 'blastn' search interface. In the 'Enter Query Sequence' section, a target sequence 'GGACACCCGGGGACCGCGCCTCACTCACCTGACGTGACAG' is entered. The 'Choose Search Set' section shows a list of SRX numbers: '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)'. The 'Add organism' button is highlighted. The 'Program Selection' section is identical to the first screenshot, with 'Somewhat similar sequences (blastn)' selected. The 'BLAST' button is 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](#) [?](#)

<input checked="" type="checkbox"/> 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

5. 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=NCI_D_1_20641650_130.14.18.128_9147_1666626263_588037759_0MetA0_S_NC_PhylT_ree&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.

Multiple Sequence Alignment Viewer 1.22.1

Link to View | Feedback

Alignment

1 - 42 (42 bases shown) | Find: | Tools | Columns | Rows | Download | Coloring | ?

Sequence ID | Start | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 | End | Organism

Query 17547

DNA: 1 - 42 (42 bases shown) - anchor Query_17547 | Rows shown: 36/36

6. 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.

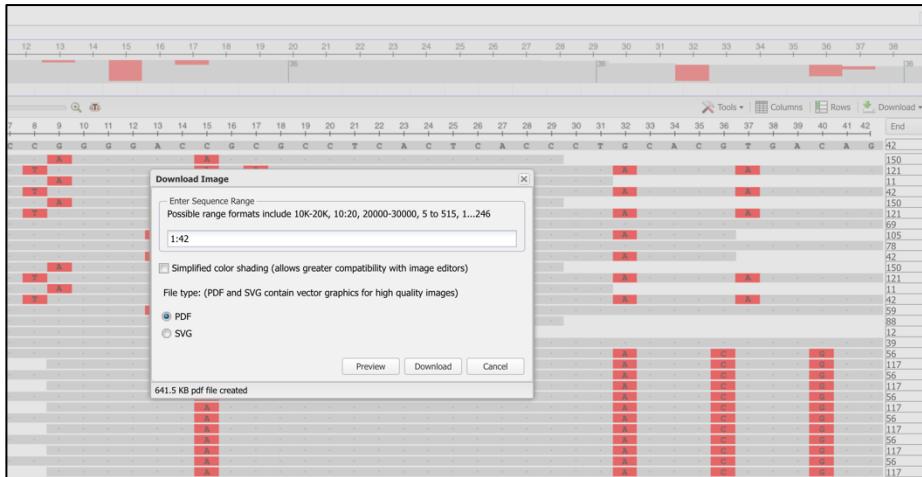
Columns Settings

Click on width to adjust

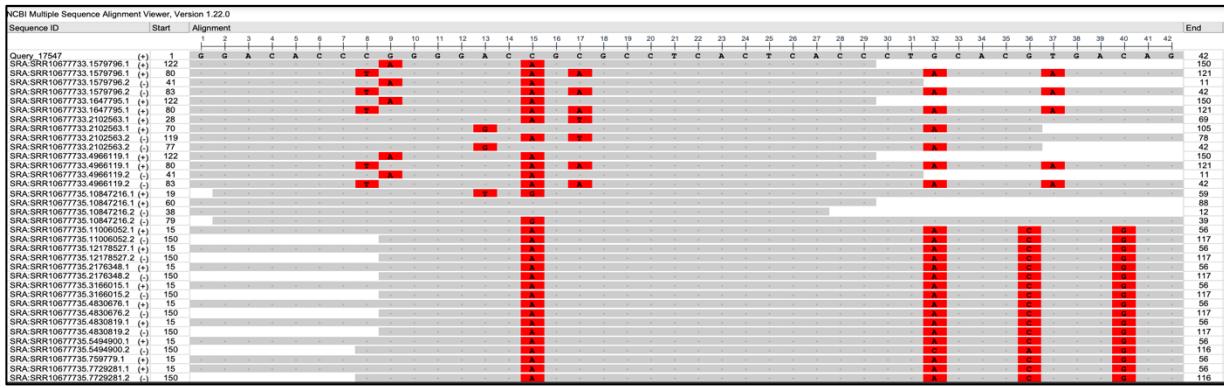
Name	Width (pixels)
Sequence ID	55
Start	55
End	150
Organism - Empty	65
% Identity to anchor/consensus (Identity)	65
% Coverage to anchor/consensus (Coverage)	65
Number of mismatches to anchor/consensus (Mismatches)	70

Show all columns | Apply | Default | Cancel

7. 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