

Homework Assignment #4. BLAST

Implement nucleotide random sequence generator and test it with BLAST algorithm.

Database was chosen as “Nucleotide collection (nr/nt)”

Choose Search Set

Database

☐ Human genomic + transcript ☐ Mouse genomic + transcript ☒ Others (nr etc.):

Nucleotide collection (nr/nt)

Algorithm was chosen as “blastn”

Program Selection

Optimize for

☐ Highly similar sequences (megablast)

☐ More dissimilar sequences (discontiguous megablast)

☒ Somewhat similar sequences (blastn)

Choose a BLAST algorithm

Run #1 Sequence length = 100

NIH U.S. National Library of Medicine
National Center for Biotechnology Information

BLAST® » blastn suite » results for RID-UK5WW45T014

Home Recent Results Saved Strategies Help

[< Edit Search](#) [Save Search](#) [Search Summary](#)

[How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title Nucleotide Sequence

RID UK5WW45T014 [Search expires on 10-19 18:44 pm](#) [Download All](#)

Program BLASTN [Citation](#)

Database nt [See details](#)

Query ID lc|Query_9611

Description None

Molecule type dna

Query Length 100

Other reports [Distance tree of results](#) [MSA viewer](#)

Filter Results

Organism only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity to **E value** to **Query Coverage** to

[Filter](#) [Reset](#)

Descriptions [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

Sequences producing significant alignments [Download](#) [Manage Columns](#) [Show](#) 100

☒ select all 3 sequences selected

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	Trichobilharzia regenti genome assembly scaffold: TRE_scaffold000720	43.7	43.7	26%	2.2	96.15%	LL000720.1
<input checked="" type="checkbox"/>	Candidatus Methanoplasma termitum strain MpT1, complete genome	43.7	43.7	28%	2.2	93.33%	CP010070.1
<input checked="" type="checkbox"/>	Diplodia corticola fungal specific transcription factor domain-containing protein (BKCO1_27000129), partial mRNA	41.9	41.9	35%	7.7	85.71%	XM_020273537.1

Run #2 Sequence length = 1000

NIH

U.S. National Library of Medicine
National Center for Biotechnology Information

Log in

BLAST® » blastn suite » results for RID-UK648TJ6015

HomeRecent ResultsSaved StrategiesHelp

Edit Search

Save Search

Search Summary

How to read this report?

BLAST Help Videos

Back to Traditional Results Page

Job TitleNucleotide Sequence

RIDUK648TJ6015

Search expires on 10-19 18:48 pm

Download All

ProgramBLASTN

Citation

Databasent

See details

Query IDlcl|Query_54355

DescriptionNone

Molecule typedna

Query Length1000

Other reportsDistance tree of resultsMSA viewer

Filter Results

Organism

only top 20 will appear

Type common name, binomial, taxid or group name

Add organism

Percent Identity

E value

Query Coverage

to

to

to

Filter

Reset

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Manage Columns

Show100

select all8 sequences selected

GenBankGraphicsDistance tree of results

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	Diphyllobothrium latum genome assembly, scaffold: DILT_scaffold0036577	46.4	46.4	3%	2.7	93.33%	LL608355.1
<input checked="" type="checkbox"/>	Oryzias latipes strain HNI chromosome_18	45.5	45.5	3%	9.3	90.62%	CP020796.1
<input checked="" type="checkbox"/>	Zebrafish DNA sequence from clone DKEY-109I21 in linkage group 10, complete sequence	45.5	45.5	5%	9.3	82.35%	BX511008.19
<input checked="" type="checkbox"/>	Xenopus laevis paralamin 2 L homeolog (palm2.L), transcript variant 2, mRNA	44.6	44.6	3%	9.3	88.24%	NM_001371546.1
<input checked="" type="checkbox"/>	PREDICTED: Xenopus laevis A-kinase anchoring protein 2 L homeolog (akap2.L), transcript variant X2, mRNA	44.6	44.6	3%	9.3	88.24%	XM_018248508.1
<input checked="" type="checkbox"/>	PREDICTED: Xenopus laevis A-kinase anchoring protein 2 L homeolog (akap2.L), transcript variant X1, mRNA	44.6	44.6	3%	9.3	88.24%	XM_018248502.1
<input checked="" type="checkbox"/>	Pyrenophora tritici-repentis Pt-1C-BFP conserved hypothetical protein, mRNA	44.6	44.6	2%	9.3	93.10%	XM_001930448.1
<input checked="" type="checkbox"/>	Xenopus laevis paralamin 2 L homeolog (palm2.L), transcript variant 3, mRNA	44.6	44.6	3%	9.3	88.24%	NM_001094918.1

Run #3 Sequence length = 10000

NIH

U.S. National Library of Medicine
National Center for Biotechnology Information

Log in

BLAST® » blastn suite » results for RID-UK6725FY014

HomeRecent ResultsSaved StrategiesHelp

Edit Search

Save Search

Search Summary

How to read this report?

BLAST Help Videos

Back to Traditional Results Page

Job TitleNucleotide Sequence

RIDUK6725FY014

Search expires on 10-19 18:50 pm

Download All

ProgramBLASTN

Citation

Databasent

See details

Query IDlcl|Query_9231

DescriptionNone

Molecule typedna

Query Length10000

Other reportsDistance tree of resultsMSA viewer

Filter Results

Organism

only top 20 will appear

Type common name, binomial, taxid or group name

Add organism

Percent Identity

E value

Query Coverage

to

to

to

Filter

Reset

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Manage Columns

Show100

select all9 sequences selected

GenBankGraphicsDistance tree of results

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	Coregonus sp. 'balchen' genome assembly, chromosome: 38	54.5	54.5	0%	0.19	88.64%	LR664381.1
<input checked="" type="checkbox"/>	PREDICTED: Delphinapterus leucas WD repeat domain 73 (WDR73), mRNA	50.0	50.0	0%	2.3	93.75%	XM_022564029.2
<input checked="" type="checkbox"/>	PREDICTED: Monodon monoceros WD repeat domain 73 (WDR73), transcript variant X2, mRNA	50.0	50.0	0%	2.3	93.75%	XM_029239304.1
<input checked="" type="checkbox"/>	PREDICTED: Monodon monoceros WD repeat domain 73 (WDR73), transcript variant X1, mRNA	50.0	50.0	0%	2.3	93.75%	XM_029239303.1
<input checked="" type="checkbox"/>	Microbacterium sp. SGAir0570 chromosome, complete genome	48.2	48.2	0%	7.9	89.47%	CP027929.1
<input checked="" type="checkbox"/>	Lateolabrax maculatus chromosome Lm22	48.2	48.2	0%	7.9	89.47%	CP027283.1
<input checked="" type="checkbox"/>	Lactobacillus parabuchneri strain FAM21731, complete genome	48.2	48.2	0%	7.9	93.55%	CP018796.1
<input checked="" type="checkbox"/>	Microbacterium nautovilla strain CC3, complete genome	48.2	48.2	0%	7.9	89.47%	CP018134.1