


Using BLAST to Compare Protein Sequences: Influenza Vaccination

Basic Local Alignment Search Tool

 U.S. National Library of Medicine

NCBI National Center for Biotechnology Information

BLAST®

Home Recent Results


Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

QuickBLASTP
Try [QuickBLASTP](#) for a fast protein search of nr.
Tue, 23 May 2017 13:00:00 EST
[More BLAST news...](#)

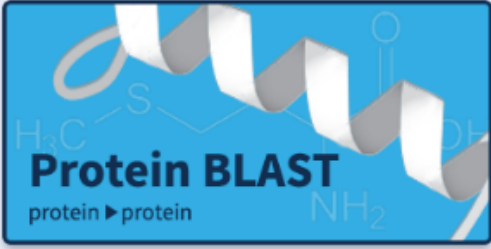
Web BLAST



Nucleotide BLAST
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide



Protein BLAST
protein ► protein

BLAST Genomes

Human Mouse Rat Microbes

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

A Tool for Comparing Primary Sequence Information

Compare Two or More
Sequences to One Another

Compare a Sequence “Query”
to an NCBI Database

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0

[Alignments](#) [Download](#) [Graphics](#)

☐ [Father BRCA1 DNA Sequence](#)

Alignments

		Description
Query 1	GTGTACAAGTTTGCAGAAAAACACCATCACTTTAACTAATCTAATTACTGAAGAGACT	60
49893 1	60
Query 61	ACTCATGTTGTTATGAAAACAGATGCTGAGTTTGTGTGTGAACGGACACTGAAATATTTT	120
49893 61	120
Query 121	CTAGGAATTCGGGAGGAAAATGGGTAGTTAGCTATTTCTGGGTGACCCAGTCTATTAAA	180
49893 121	180
Query 181	GAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGGAGATGTGTCATGGAAGA	240
49893 181	240
Query 241	AACCACCAAGGTCCTAAGCGAGCAAGAGAATCCAGGACAGAAAGATCTTCAGGGGGCTA	300
49893 241	300
Query 301	GAAATCTGTTGCTATGCGCCCTTCACCAACATGCCACAGATCAACTGGAATGGATGGTA	360
49893 301	360
Query 361	CAGCTGTGTGGTGCTTCTGTGGTGAAGGAGCTTTCATCAITCACCCITGGCAGGTGTC	420
49893 361	420
Query 421	CACCCAATTGTGTTGTGTCAGCCAGATGCCTGACAGAGGACAATGGCTTCCATGCAATT	480
49893 421	480
Query 481	GGGCAGATGTGTGAGGCACCTGTGTTGACCCGAGAGTGGGTGTTGGACAGTGTAGCACTC	540
49893 481	540
Query 541	TACCAAGTCCAGGAGCTGGACACCTACCTGATACCCAGATCCCCACAGCCACTACTGA	600
49893 541	600

BLAST Basic Local Alignment Search Tool

Query ID: id11559
Description: none
Molecule type: nucleic acid
Query Length: 1542

Database Name: nr
Description: All GenBank+EMBL+DDBJ+PDB sequences (but not EST, STS, GSS, environmental samples or phage C, I or L HTGS sequences)
Program: BLASTN 2.2.3+ [Blast](#)

Other reports: [Search Summary](#) [Taxonomy reports](#) [Sequence logos](#)

Graphic Summary

Distribution of 100 Blast Hits on the Query Sequence

Click over to see the details, click to show alignments

Descriptions

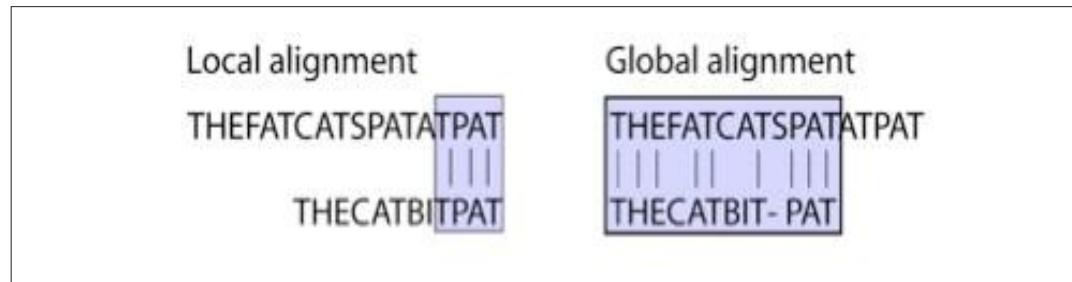
Legend for links to other resources: [UniProt](#) [GenBank](#) [EMBL](#) [DDBJ](#) [PDB](#) [RefSeq](#) [Ensembl](#) [NCBI](#)

Accession	Description	Hit #	Score	Ident	E-Value	Max	Unit
U00006.1	Gorilla gorilla mitochondrial DNA, complete genome	236	2762	100%	0.0	100%	
U00006.1	Gorilla gorilla mitochondrial DNA, complete genome	232	2772	100%	0.0	99%	
U00006.1	Gorilla gorilla mitochondrial genome	212	2715	100%	0.0	99%	
U00006.1	Pen troglodytes chromosome 7 clone CH01-275010, complete sequence	2107	2107	99%	0.0	91%	
U00006.1	Pen troglodytes BAC clone CH01-250214 from chromosome 7, complete sequence	2107	2107	99%	0.0	91%	
U00006.1	Pen troglodytes mitochondrial DNA, complete sequence	2107	2107	99%	0.0	90%	
U00006.1	Pen troglodytes valens mitochondrial DNA, complete genome	2107	2114	99%	0.0	90%	
U00006.1	Pen troglodytes troglodytes BAC clone CH01-250214, complete genome	2111	2111	99%	0.0	90%	

Terminology

- **Query Sequence:**
 - Same root as “question”
 - Sometimes called a “reference sequence”
 - The sequence to which other sequences are compared
 - Independent or control variable.
- **Subject Sequence:**
 - The sequence being compared
 - Dependent or experimental variable
- **BLAST Scores**
 - **Max Score, Total Score**
 - **Query Coverage**
 - **Percent Identity**

The “Local” in Local Search Tool



- Sequences are broken into “words”
- Defaults vary by program (11-28 bases or amino acids)
- BLAST looks for word matches, then tries to extend those words
- Proteins:
 - BLAST Scoring Table contains the probability of each amino acid being replaced with another amino acid (ex: ELVIS vs. ELVES)
- Nucleotides:
 - Plus points where Nucleotides match, minus points where they do not match [default is adjustable]

How to BLAST

Comparing Two or More Sequences to One Another

The screenshot shows the 'Align Sequences Nucleotide BLAST' interface. It has tabs for 'blastn', 'blastp', 'blastx', 'tblastn', and 'tblastx'. The 'blastn' tab is selected. The main section is titled 'Enter Query Sequence' and 'Enter Subject Sequence'. Both sections have a text area for 'Enter accession number(s), gi(s), or FASTA sequence(s)', a 'Clear' button, and a 'Query subrange' or 'Subject subrange' section with 'From' and 'To' input fields. Below these are 'Or, upload file' and 'Browse...' buttons, and a 'Job Title' field. A checkbox 'Align two or more sequences' is checked. At the bottom, there is a 'Program Selection' section with radio buttons for 'Highly similar sequences (megablast)', 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)'. A 'BLAST' button is at the bottom left.

Compare a Sequence “Query” to an NCBI Database

The screenshot shows the 'Basic Local Alignment Search Tool' interface. It has tabs for 'blastn', 'blastp', 'blastx', 'tblastn', and 'tblastx'. The 'blastn' tab is selected. The main section is titled 'Enter Query Sequence' and 'Enter Subject Sequence'. Both sections have a text area for 'Enter accession number, gi, or FASTA sequence', a 'Clear' button, and a 'Query subrange' or 'Subject subrange' section with 'From' and 'To' input fields. Below these are 'Or, upload file' and 'Browse...' buttons, and a 'Job Title' field. A checkbox 'Align two or more sequences' is unchecked. At the bottom, there is a 'Choose Search Set' section with a 'Database' dropdown menu. The dropdown is open, showing a list of databases including 'Nucleotide collection (nr/nt)', 'Genomic plus Transcript', 'Reference mRNA sequences (refseq_ma)', 'Reference genomic sequences (refseq_genomic)', 'NCBI Genomes (chromosome)', 'Expressed sequence tags (est)', 'Non-human, non-mouse ESTs (est_others)', 'Genomic survey sequences (gss)', 'High throughput genomic sequences (HTGS)', 'Patent sequences (pat)', 'Protein Data Bank (pdb)', 'Human ALU repeat elements (alu_repeats)', 'Sequence tagged sites (dbsts)', 'Whole-genome shotgun reads (wgs)', and 'Environmental samples (env_nt)'. The 'Database' dropdown is set to 'Nucleotide collection (nr/nt)'.

Available Genomes & Databases

BLAST Assembled Genomes

Find Genomic BLAST pages:

Enter organism name or id—completions will be suggested

GO

- ▣ [Human](#)
- ▣ [Mouse](#)
- ▣ [Rat](#)
- ▣ [Cow](#)
- ▣ [Pig](#)
- ▣ [Dog](#)

- ▣ [Rabbit](#)
- ▣ [Chimp](#)
- ▣ [Guinea pig](#)
- ▣ [Fruit fly](#)
- ▣ [Honey bee](#)
- ▣ [Chicken](#)

- ▣ [Zebrafish](#)
- ▣ [Clawed frog](#)
- ▣ [Arabidopsis](#)
- ▣ [Rice](#)
- ▣ [Yeast](#)
- ▣ [Microbes](#)

The screenshot shows the NCBI BLAST web interface. At the top, there's a navigation bar with 'Home', 'Recent Results', 'Saved Strategies', and 'Help'. Below this, the 'NCBI/BLAST/blastn suite' is selected. The 'Standard Nucleotide BLAST' dropdown menu is open, displaying a list of databases. The 'Nucleotide collection (nr/nt)' is highlighted in blue. Other databases listed include 'Genomic plus Transcript', 'Human genomic plus transcript (Human G+T)', 'Mouse genomic plus transcript (Mouse G+T)', 'Reference RNA sequences (refseq_ma)', 'Reference genomic sequences (refseq_genomic)', 'RefSeq Representative genomes (refseq_representative_genomes)', 'NCBI Genomes (chromosome)', 'Expressed sequence tags (est)', 'Genomic survey sequences (gss)', 'High throughput genomic sequences (HTGS)', 'Patent sequences(pat)', 'Protein Data Bank (pdb)', 'Human ALU repeat elements (alu_repeats)', 'Sequence tagged sites (dbsts)', 'Whole-genome shotgun contigs (wgs)', 'Transcriptome Shotgun Assembly (TSA)', '16S ribosomal RNA sequences (Bacteria and Archaea)', and 'Sequence Read Archive (SRA)'. The 'Nucleotide collection (nr/nt)' is also shown at the bottom of the dropdown menu.

BLAST Results Homepage

blast.ncbi.nlm.nih.gov/Blast.cgi

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

BLAST® » blastp suite-2sequences » results for RID-6APBG4YZ11N

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 A/Michigan/45/2015 (H1N1)-2017-Present Vaccine

RID 6APBG4YZ11N Search expires on 04-25 22:06 pm [Download All](#) ▾

Program Blast 2 sequences [Citation](#) ▾

Query ID Icl|Query_325414 (amino acid)

Query Descr A/Michigan/45/2015 (H1N1)-2017-Present Vaccine

Query Length 469

Subject ID Icl|Query_325416 and 5 more subject(s) (amino acid)

Subject Descr [See details](#) ▾

Filter Results

Percent Identity to

E value to

Query Coverage to

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

Multiple subjects information

Icl Query_325416	A/California/07/2009(H1N1)-2011-2016 Vaccine	469
Icl Query_325417	A/Brisbane/59/2007(H1N1)-2008-2010 Vaccine	462
Icl Query_325418	Patient A11	469
Icl Query_325419	Patient A12	469
Icl Query_325420	Patient A13	469
Icl Query_325421	Patient A17	469

[Download](#) ▾ [Select columns](#) ▾ Show 100 ▾ ?

[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Patient A17	956	956	100%	0.0	99.79%	469	Query_325421

Feedback

BLAST Results: Descriptions & Scores

Subject Descr [See details ▼](#)

Subject 2807

Length

Descriptions

Graphic Summary

Alignments

Sequences producing significant alignments

Download ▼

Select columns ▼

Show 100 ▼



☒ select all 6 sequences selected

[Graphics](#)

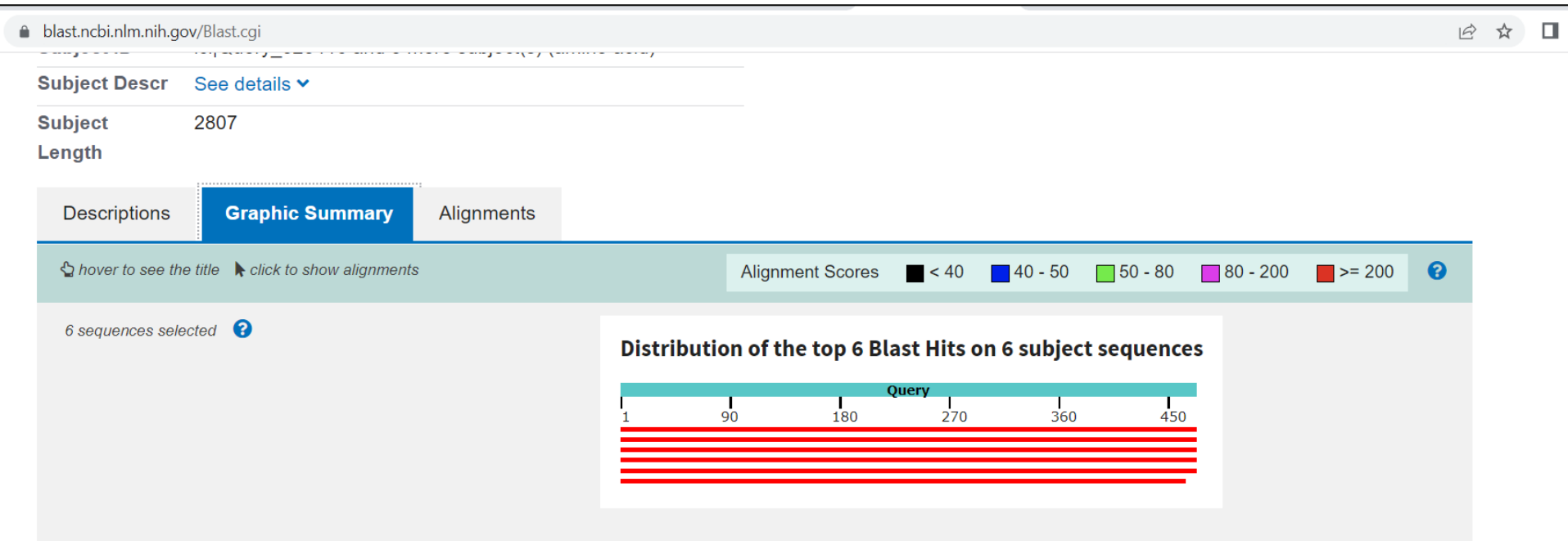
[Distance tree of results](#)

[Multiple alignment](#)

[MSA Viewer](#)

	Description ▼	Scientific Name ▼	Max Score ▼	Total Score ▼	Query Cover ▼	E value ▼	Per. Ident ▼	Acc. Len ▼	Accession
<input checked="" type="checkbox"/>	Patient A17		956	956	100%	0.0	99.79%	469	Query_325421
<input checked="" type="checkbox"/>	Patient A13		956	956	100%	0.0	99.79%	469	Query_325420
<input checked="" type="checkbox"/>	Patient A12		956	956	100%	0.0	99.79%	469	Query_325419
<input checked="" type="checkbox"/>	Patient A11		956	956	100%	0.0	99.79%	469	Query_325418
<input checked="" type="checkbox"/>	A/California/07/2009(H1N1)-2011-2016 Vaccine		941	941	100%	0.0	97.01%	469	Query_325416
<input checked="" type="checkbox"/>	A/Brisbane/59/2007(H1N1)-2008-2010 Vaccine		773	773	98%	0.0	80.74%	462	Query_325417

BLAST Results: Graphic Summary – Multiple Sequence Alignment of Sample Sequences



BLAST Results: Graphic Summary – Identifying a Sequence in the NCBI Database

blast.ncbi.nlm.nih.gov/Blast.cgi

An official website of the United States government [Here's how you know](#)

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

BLAST® » blastp suite » results for RID-6AZ66FHP013

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 A/Michigan/45/2015 (H1N1)-2017-Present

RID 6AZ66FHP013 [Search expires on 04-26 00:37 am](#) [Download All](#)

Program BLASTP [Citation](#)

Database nr [See details](#)

Query ID IclQuery_87755

Description A/Michigan/45/2015 (H1N1)-2017-Present Vaccine

Molecule type amino acid

Query Length 469

Other reports [Distance tree of results](#) [Multiple alignment](#) [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](#)

Compare these results against the new Clustered nr database [BLAST](#)

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

Sequences producing significant alignments [Download](#) [Select columns](#) [Show](#) 100 [?](#)

☒ select all 100 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/India/Pun151245/2015(H1N1))]	Influenza A virus (A/India/Pun151245/2015(H1N1))	957	957	100%	0.0	100.00%	469	AKE37534.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/North Dakota/01/2016(H1N1))]	Influenza A virus (A/North Dakota/01/2016(H1N1))	957	957	100%	0.0	99.79%	469	ANE27703.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Sistan/57/2015(H1N1))]	Influenza A virus (A/Sistan/57/2015(H1N1))	957	957	100%	0.0	99.79%	469	AMA65394.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Montana/03/2016(H1N1))]	Influenza A virus (A/Montana/03/2016(H1N1))	957	957	100%	0.0	99.79%	469	ANE27769.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Viroinia/15/2016(H1N1))]	Influenza A virus (A/Viroinia/15/2016(H1N1))	957	957	100%	0.0	99.79%	469	ANM90834.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/India/DRDE_GWL989/2015(H1N1))]	Influenza A virus (A/India/DRDE_GWL989/2015(H1N1))	957	957	100%	0.0	99.79%	469	ANC28534.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Indiana/01/2016(H1N1))]	Influenza A virus (A/Indiana/01/2016(H1N1))	957	957	100%	0.0	99.79%	469	ANE28407.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/California/123/2016(H1N1))]	Influenza A virus (A/California/123/2016(H1N1))	957	957	100%	0.0	99.79%	469	AOZ86020.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Idaho/43/2015(H1N1))]	Influenza A virus (A/Idaho/43/2015(H1N1))	957	957	100%	0.0	99.79%	469	ANE27318.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Rhode Island/03/2016(H1N1))]	Influenza A virus (A/Rhode Island/03/2016(H1N1))	957	957	100%	0.0	99.79%	469	ANE29500.1
<input checked="" type="checkbox"/>	neuraminidase [Influenza A virus (A/Czech Republic/7/2016(H1N1))]	Influenza A virus (A/Czech Republic/7/2016(H1N1))	957	957	100%	0.0	99.79%	469	AMP44463.1

Pairwise Comparisons & Default Alignment Format

Descriptions

Graphic Summary

Alignments

Alignment view

Pairwise

Restore defaults

6 sequences selected

Download

Graphics

Patient A17
Sequence ID: **Query_325421** Length: **469** Number of Matches: **1**

Range 1: 1 to 469

Next Match

Previous Match

Score	Expect	Method	Identities	Positives	Gaps
956 bits(2470)	0.0	Compositional matrix adjust.	468/469(99%)	469/469(100%)	0/469(0%)
Query 1	MNPNQIITIGSICHTIGMANLILQIGNIISIWVSHSIQIGNQSQIETCNQSVITYENNT	60			
Sbjct 1	MNPNQIITIGSICHTIGMANLILQIGNIISIWVSHSIQIGNQSQIETCNQSVITYENNT	60			
Query 61	WVNQTYVNISNTNFAAGQSVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFIREF	120			
Sbjct 61	WVNQTYVNISNTNFAAGQSVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFIREF	120			
Query 121	FISCSPLCERTFFLTQGALLNDKHSNGTIKDRSPYRTLMSCPIGEVPSPYNSRFESVAWS	180			
Sbjct 121	FISCSPLCERTFFLTQGALLNDKHSNGTIKDRSPYRTLMSCPIGEVPSPYNSRFESVAWS	180			
Query 181	ASACHDGINLWTIGISGPDGAVAVLYNGIITDTIKSWRNILRTQSEACVNGSCFT	240			
Sbjct 181	ASACHDGINLWTIGISGPDGAVAVLYNGIITDTIKSWRNILRTQSEACVNGSCFT	240			
Query 241	IMTDGPGDQASYKIFRIEKGKIKSVEMKAPNYHYEECSYPSSEITCVCRDNIHGSN	300			
Sbjct 241	IMTDGPGDQASYKIFRIEKGKIKSVEMKAPNYHYEECSYPSSEITCVCRDNIHGSN	300			
Query 301	RPWVSFNQLEYQMGYICSGVFGDNPRPNKDTGSCGPVSNANGVKGFsfkyGNGVNIIG	360			
Sbjct 301	RPWVSFNQLEYQMGYICSGVFGDNPRPNKDTGSCGPVSNANGVKGFsfkyGNGVNIIG	360			
Query 361	RTKSISRKGFEMIWDPNWGTGTDNKFSIKQDIVGINEWSGYSGSFVQHPELTGLDCIRP	420			
Sbjct 361	RTKSISRKGFEMIWDPNWGTGTDNKFSIKQDIVGINEWSGYSGSFVQHPELTGLDCIRP	420			
Query 421	CFWVELIRGRPEENTIWTSGSSISFCGVNSDVTGWSWPDGAELPFTIDK	469			
Sbjct 421	CFWVELIRGRPEENTIWTSGSSISFCGVNSDVTGWSWPDGAELPFTIDK	469			

Download

Graphics

Patient A13
Sequence ID: **Query_325420** Length: **469** Number of Matches: **1**

Range 1: 1 to 469

Next Match

Previous Match

Score	Expect	Method	Identities	Positives	Gaps
956 bits(2470)	0.0	Compositional matrix adjust.	468/469(99%)	469/469(100%)	0/469(0%)
Query 1	MNPNQIITIGSICHTIGMANLILQIGNIISIWVSHSIQIGNQSQIETCNQSVITYENNT	60			
Sbjct 1	MNPNQIITIGSICHTIGMANLILQIGNIISIWVSHSIQIGNQSQIETCNQSVITYENNT	60			
Query 61	WVNQTYVNISNTNFAAGQSVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFIREF	120			

Reformatting Results Creates Custom Views: Query-Anchored with Dots for Identities

The screenshot displays a web-based alignment viewer interface. At the top, there are three tabs: 'Descriptions', 'Graphic Summary', and 'Alignments', with 'Alignments' being the active tab. Below the tabs, there is a section for 'Alignment view' with a dropdown menu. The dropdown menu is open, showing several options: 'Query-anchored with dots for identities' (selected), 'Pairwise', 'Pairwise with dots for identities', 'Query-anchored with letters for identities', 'Flat query-anchored with dots for identities', and 'Flat query-anchored with letters for identities'. To the right of the dropdown, there is a 'Line length' input field set to '60', a 'Restore defaults' button, and a 'Download' button. Below the dropdown, the alignment results are displayed in three sections, each with a 'Download' button and navigation links ('Next', 'Previous', 'First Range').

Query range 1:

Query	1GNQSQIETCNQSVITYENNT	60
Query_325421	1	60
Query_325420	1	60
Query_325419	1	60
Query_325418	1	60
Query_325416	1V.....I.....L.....N.....	60
Query_325417	1SIA..IIS.M.....A.....T.S.NNTGI...RI.....S.	60

Query range 2: 61 to 120

Query	61	WVNQTYVNISNTNFAAGQSVSVSVKLAGNSSLCVPVSGWAIYSKDNSVRIGSKGDVFIREF	120
Query_325421	61	120
Query_325420	61	120
Query_325419	61	120
Query_325418	61	120
Query_325416	61	120
Query_325417	61	...H.....N...VV...EDKT...T.....SI.....T...I.....	120

Query range 3: 121 to 180

Query	121	FISCSPLECRTFFLTQGALLNDKHSNGTIKDRSPYRTLMSCPIGEVPSPYNSRFESVAWS	180
Query_325421	121	180
Query_325420	121	180
Query_325419	121	180
Query_325418	121	180
Query_325416	121	180
Query_325417	121H.....V.....A.....L..A.....K.....	180

Examples of Query Sequences

- **Genetic Testing:** Comparing a patient's *BRCA1* gene sequence to a sequence known *not* to contain cancer-causing mutations
- **DNA Barcoding:** Comparing your DNA sequence to all of the DNA sequences in the NCBI database to identify your unknown sample
- **Evolution:** Compare the human sickle cell allele [hemoglobin genotype AS] to the hemoglobin gene from other species to predict if other species get sickle cell disease

BLAST Scores: Query Coverage

Query Coverage: Where does the match start and where does it end?

Which alignment has the highest query coverage?

1. ATGGATACGT
TGAGATGATC

2. ATGCCGACAG
AGGGCAACAG

3. ATGGATAAGT
TGGGATGATC

4. ATGCCGATTC
AGGGCAACAG

BLAST Scores: Query Coverage

Query Coverage: Where does the match start and where does it end?

Which alignment has the highest query coverage?

1. ATGGATACGT

TGAGATGATC

Query Coverage = $3/10 = 30\%$

2. ATGCCGACAG

AGGGCAACAG

Query Coverage = $10/10 = 100\%$

3. ATGGATAAGT

TGGGATGATC

Query Coverage = $6/10 = 60\%$

4. ATGCCGATTC

AGGGCAACAG

Query Coverage = $7/10 = 70\%$

BLAST Scores: Percent Identity

Percent Identity: Where the query and subject sequences align, how much of the alignment is identical?

Which alignment has the highest percent identity?

(= Number of matches / Length of Query Coverage)

1. ATGGATACGT

TGAGATGATC

Query Coverage = 3/10 = 30%

2. ATGCGACAG

AGGCAACAG

Query Coverage = 10/10 = 100%

3. ATGGATAAGT

TGGGATGATC

Query Coverage = 6/10 = 60%

4. ATGCGATTC

AGGCAACAG

Query Coverage = 7/10 = 70%

BLAST Scores: Percent Identity

Percent Identity: Where the query and subject sequences align, how much of the alignment is identical?

Which alignment has the highest percent identity?

(= Number of matches / Length of Query Coverage)

1. ATGGATACGT

TGAGATGATC

Query Coverage = 3/10 = 30%

Percent Identity = 3/3 = 100%

2. ATGCCGACAG

AGGCAACAG

Query Coverage = 10/10 = 100%

Percent Identity = 7/10 = 70%

3. ATGGATAAGT

TGGGATGATC

Query Coverage = 6/10 = 60%

Percent Identity = 5/6 = 83%

4. ATGCCGATTC

AGGCAACAG

Query Coverage = 7/10 = 70%

Percent Identity = 4/7 = 57%

BLAST Scores: Use Together

Look at query coverage and percent identity together

Which alignment shows the best match?

1. ATGGATACGT

TGAGATGATC

Query Coverage = $3/10 = 30\%$

Percent Identity = $3/3 = 100\%$

3. ATGGATAGT

TGGGATATC

Query Coverage = $6/10 = 60\%$

Percent Identity = $5/6 = 83\%$

2. ATGCGACAG

AGGCAACAG

Query Coverage = $10/10 = 100\%$

Percent Identity = $7/10 = 70\%$

4. ATGCGATTC

AGGCAACAG

Query Coverage = $7/10 = 70\%$

Percent Identity = $4/7 = 57\%$

BLAST Scores: E-Values

E-Value: How likely is our result? Could it be due to chance?

An E-value is essentially the same as a p-value.

$p = 0.05$ means that 5% of the time (1 in 20) you will see this result even if it's not "true" (it's due to random chance)

Descriptions

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0






[Alignments](#) [Download](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	pLemon-YFP What is the likelihood that this result is due to chance?	486	486	100%	4e-180	98%	47194
<input type="checkbox"/>	mTomato-RFP	70.9	70.9	100%	2e-19	28%	47195
<input type="checkbox"/>	mGrape1-CFP	62.8	62.8	95%	1e-16	25%	47196
<input type="checkbox"/>	pLime-GFP	487	487	99%	1e-180	99%	47197
<input type="checkbox"/>	pBlueberry-BFP	489	489	100%	0.0	98%	47198
<input type="checkbox"/>	mTangerine1.5	70.1	70.1	100%	3e-19	29%	47199
<input type="checkbox"/>	mCherry-RFP	71.6	71.6	100%	8e-20	29%	47200
<input type="checkbox"/>	mOrange-OFP	74.7	74.7	100%	8e-21	29%	47201

BLAST Scores Comparing Patients & Vaccine Strains of Influenza

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

Alignments  Download  Graphics  Distance tree of results  Multiple alignment 													
Description							Max score	Total score	Query cover	E value	Ident	Accession	
<input type="checkbox"/> Patient D17							956	956	100%	0.0	99%	Query_25663	
<input type="checkbox"/> Patient D13							956	956	100%	0.0	99%	Query_25662	
<input type="checkbox"/> Patient D12							956	956	100%	0.0	99%	Query_25661	
<input type="checkbox"/> Patient D11							956	956	100%	0.0	99%	Query_25660	
<input type="checkbox"/> Patient C17							956	956	100%	0.0	99%	Query_25659	
<input type="checkbox"/> Patient C13							956	956	100%	0.0	99%	Query_25658	
<input type="checkbox"/> Patient C12							956	956	100%	0.0	99%	Query_25657	
<input type="checkbox"/> Patient C11							956	956	100%	0.0	99%	Query_25656	
<input type="checkbox"/> Patient B17							956	956	100%	0.0	99%	Query_25655	
<input type="checkbox"/> Patient B13							956	956	100%	0.0	99%	Query_25654	
<input type="checkbox"/> Patient B12							956	956	100%	0.0	99%	Query_25653	
<input type="checkbox"/> Patient B11							956	956	100%	0.0	99%	Query_25652	
<input type="checkbox"/> Patient A17							956	956	100%	0.0	99%	Query_25651	
<input type="checkbox"/> Patient A13							956	956	100%	0.0	99%	Query_25650	
<input type="checkbox"/> Patient A12							956	956	100%	0.0	99%	Query_25649	
<input type="checkbox"/> Patient A11							956	956	100%	0.0	99%	Query_25648	
<input type="checkbox"/> A/California/07/2009(H1N1)-2011-2016 Vaccine	2011-2016 Vaccine						941	941	100%	0.0	97%	Query_25646	
<input type="checkbox"/> A/Brisbane/59/2007(H1N1)-2008-2010 Vaccine	2008-2010 Vaccine						773	773	98%	0.0	81%	Query_25647	

BLAST Alignment

Current Vaccine Strain

Influenza Strain
Infecting
Classroom
Patients

2011-2016 Vaccine
2008-2010 Vaccine

Take-Home: If the patients had been vaccinated they probably wouldn't have gotten sick!

Query	241	IMTDGPSDGGQASYKIFRIEKGKIIKSVEMKAPNYHYEECSYPDSSEITCVCRDNIHGSN	300
Query_25663	241	300
Query_25662	241	300
Query_25661	241	300
Query_25660	241	300
Query_25659	241	300
Query_25658	241	300
Query_25657	241	300
Query_25656	241	300
Query_25655	241	300
Query_25654	241	300
Query_25653	241	300
Query_25652	241	300
Query_25651	241	300
Query_25650	241	300
Query_25649	241	300
Query_25648	241	300
Query_25646	241	V.....N.....V.....N.....F.....TGIVM.....	300
Query_25647	241NKA.....K.....VT..I..LN...F.....TGIVM.....	300
Query	301	RPWVSFNQNLEYQMGYICSGVFGDNPRPNDKTGSCGPVSSNGANGVKGFsfkyGNGVWIG	360
Query_25663	301V.....	360
Query_25662	301V.....	360
Query_25661	301V.....	360
Query_25660	301V.....	360
Query_25659	301V.....	360
Query_25658	301V.....	360
Query_25657	301V.....	360
Query_25656	301V.....	360
Query_25655	301V.....	360
Query_25654	301V.....	360
Query_25653	301V.....	360
Query_25652	301V.....	360
Query_25651	301V.....	360
Query_25650	301V.....	360
Query_25649	301V.....	360
Query_25648	301V.....	360
Query_25646	301I.....I.....E..GE...N..TVD.....Y..D.....	360
Query_25647	301D..I.....E..GE...N..TVD.....Y..D.....	360
Query	361	RTKSISSRKGFEMIWDPNGWGTGTDNKFfSIKQDIVGINESGYSGSFVQHPELTGLDCIRP	420
Query_25663	361	420
Query_25662	361	420
Query_25661	361	420
Query_25660	361	420
Query_25659	361	420
Query_25658	361	420
Query_25657	361	420
Query_25656	361	420
Query_25655	361	420
Query_25654	361	420
Query_25653	361	420
Query_25652	361	420
Query_25651	361	420
Query_25650	361	420
Query_25649	361	420
Query_25648	361	420
Query_25646	361N.....N.....	420
Query_25647	361NRL.....N..SD..V...V..A..TD.....	420

Distance Tree of Results

