

Data Sequence untuk BLAST Search

MK Bioinformatika ATA 2023/2024

KELAS C



Identitas Mahasiswa

Nama: Bimo Hadi Permana

NPM: 2306157103

Kelas: C Bioinformatika

Absensi: 20

Tugas Individu

- 1. Pilih data unknown sequence.
- 2. Kirim data sequence ke NCBI BLAST Search Tool.
- 3. Buat beberapa opsi dengan click: "Exclude" dan "Limit to"
- 4. Buatlah Hasil Identifikasi berdasarkan hasil BLAST Search sesuai dengan beberapa opsi yang dipilih (Butir no.3). Buatlah Kesimpulan apakah terdapat perbedaan hasil identifikasi?
- 5. Buatlah Laporan Hasil identifikasi *Unknown sequence* tersebut dalam bentuk PPT, tuliskan Nama mahasiswa, NPM, dan beri keterangan kelas Bioinformatika. Beri nama file dengan: **No. Absen_Kelas_Nama Lengkap_Tugas 1.**
- 6. Upload file di tempat yang telah disediakan di EMAS.

1. QUERY sequence name: _____ 2. Search Set: 3. BLAST Search Result: Query length: _____ bp • Gen: Top Hit: _____, dengan persentase identitas _____%. Max Score: ____; Total Score: ____: Query cover: ____%; E value: Organisme: ______.

SEQUENCE 1

Pencarian sequence dengan BLASTN

>Sequence-1

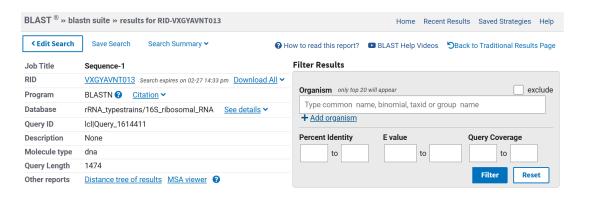
AGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGGAAAGGCCCCTTCGGGGGGTACTCGAGCGGCGAACGGGTGAGTAACA CGTGAGCAACCTGCCCCTGACTCTGGGATAAGCCTGGGAAACCGGGTCTAATACCGGATATGACCCTTCATCGCATGGTGTGAGGGTGGAAA GTTTTTCGGTTGGGGATGGGCTCGCGGCCTATCAGCTTGTTGGTGGGGTGATGGCCTACCAAGGCGACGACGGGTAACCGGCCTGAGAGGG CGACCGGTCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGGCGGAAGCCTGACGC AGCGACGCCGCGTGAGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGACGAAGCTAACGTGACGGTACCTGCAGAAGAAGCG CCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGTTTGTT GCGTCCGTCGTGAAAGCCCACGGCTTAACTGTGGGTCTGCGGTGGATACGGGCAGACTAGAGGCAGGTAGGGGAGCATGGAATTCCCGGT GTAGCGGTGAAATGCGCAGATATCGGGAGGAACACCGGTGGCGAAGGCGGTGCTCTGGGCCTGTACTGACGCTGAGGAGCGAAAGCGTGG GGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGTTGGGCGCTAGGTGTGGGGTCCTTCCACGGATTCCGCGCCCGTAGCTA ACGCATTAAGCGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGGGGCACATGTT GCTTAATTCGACGCAACGCGAAGAACCTTACCAAGGCTTGACATCGCCGGAAAACTCGCAGAGATGCGGGGGTCCTTTTGGGCCGGTGACAG GTGGTGCATGGCTGTCGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCGTTCCATGTTGCCAGCACGTAGTG GTGGGGACTCATGGGAGACCGCCGGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTATGTCTTGGGCTGCAAAC ATGCTACAATGGCCGGTACAGAGGGCTGCGATACCGTGAGGTGGAGCGAATCCCTTAAAGCCGGTCTCAGTTCGGATCGAAGTCTGCAACTC GACTTCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCCCCCGTCACGTCA CAAGGTAGCCGTACCGGAAG GTGC

- 1. QUERY sequence name: Sequence-1
- 2. Search Set: 16S ribosomal RNA seq; sequences from type material; and exclude: models (XM/XP), uncultured/environmental sample sequences
- 3. BLAST Search Result:
- Query length: <u>1474</u> bp
- Gen: 16S ribosomal RNA
- Top Hit: <u>Actinomadura vinacea</u> strain NBRC 14688 16S ribosomal RNA, partial sequence, dengan persentase identitas 98,85 %.
- Max Score: <u>2628</u>; Total Score: <u>2628</u>: Query cover: <u>100%</u>; E value: <u>0.0</u>
- Organisme: <u>Bacteria; Bacillati; Actinomycetota; Actinomycetes;</u> <u>Streptosporangiales; Thermomonosporaceae; Actinomadura; Actinomadura vinacea strain NBRC 14688</u>.

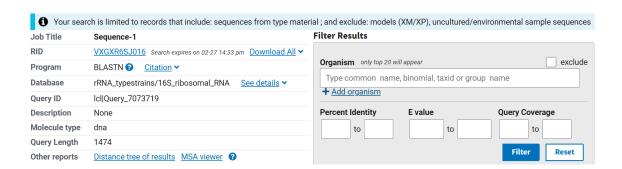
Actinomadura vinacea strain NBRC 14688 dengan persentase identitas 98,85 %.

- QUERY sequence name: <u>Sequence-1</u>
- 2. Search Set: 16S ribosomal RNA seq; sequences from type material; and exclude: -
- 3. BLAST Search Result:
- Query length: <u>1474</u> bp
- Gen: 16S ribosomal RNA
- Top Hit: Actinomadura vinacea strain NBRC 14688 16S ribosomal RNA, partial sequence, dengan persentase identitas 98,85%.
- Max Score: 2628; Total Score: 2628: Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Actinomycetota; Actinomycetes; Streptosporangiales; Thermomonosporaceae; *Actinomadura*; *Actinomadura*; *vinacea* strain NBRC 14688.

- 1. QUERY sequence name: Sequence-1
- 2. Search Set: 16S ribosomal RNA seq; tanpa type material, exclude: -
- 3. BLAST Search Result:
- Query length: 1474 bp
- Gen: 16S ribosomal RNA
- Top Hit: *Actinomadura* sp. HAT23 16S ribosomal RNA gene, partial sequence, dengan persentase identitas 100%.
- Max Score: 2723; Total Score: 2723; Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Actinomycetota; Actinomycetes; Streptosporangiales; Thermomonosporaceae; *Actinomadura*; *Actinomadura* sp. HAT23.



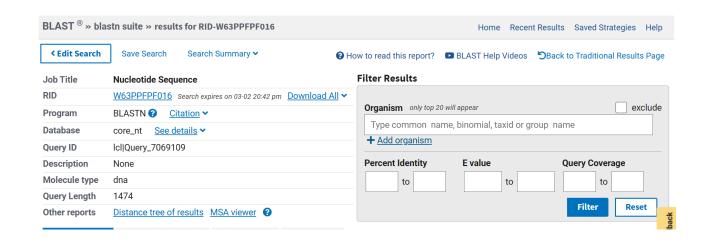




		Description	Scientific Name	Max Score	Total Score	Query Cover	E value ▼	Per. Ident	Acc. Len	Accession
5	4	Actinomadura vinacea strain NBRC 14688 16S ribosomal RNA, partial sequence	Actinomadur	2628	2628	100%	0.0	98.85%	1471	NR_112737.1
	4	Actinomadura montaniterrae strain CYP1-1B 16S ribosomal RNA, partial sequence	Actinomadur	2590	2590	100%	0.0	98.37%	1488	NR_151939.1
5	/	Actinomadura barringtoniae strain GKU 128 16S ribosomal RNA, partial sequence	Actinomadur	2586	2586	100%	0.0	98.31%	1497	NR_178624.1
5	/	Actinomadura jiaoheensis strain NEAU-Jh1-3 16S ribosomal RNA, partial sequence	Actinomadur	2584	2584	100%	0.0	98.31%	1511	NR_146004.1
6	/	Actinomadura sporangiiformans strain NEAU-Jh2-5 16S ribosomal RNA, partial sequence	Actinomadur	2584	2584	100%	0.0	98.31%	1511	NR_146003.1
	/	Actinomadura geliboluensis strain A8036 16S ribosomal RNA, partial sequence	Actinomadur	2575	2575	100%	0.0	98.17%	1490	NR_109059.1

Kiri: Tanpa filter

Kanan: Filter exclude (seluruh) dan sequence from type material



	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
✓	Actinomadura sp. HAT23 16S ribosomal RNA gene, partial sequence	Actinomadura sp. HAT23	2723	2723	100%	0.0	100.00%	1474	KT210129.1
	Actinomadura sp. 21ATH 16S ribosomal RNA gene, partial sequence	Actinomadura sp. 21ATH	2717	2717	100%	0.0	99.93%	1485	KT876994.1
~	Actinomadura rugatobispora JCM 3366 DNA, complete genome	Actinomadura rugatobispora	2660	13290	100%	0.0	99.25%	12447268	<u>AP029191.1</u>
	Actinomadura sp. strain 7C222 16S ribosomal RNA gene, partial sequence	Actinomadura sp.	2628	2628	100%	0.0	98.85%	1480	KX928701.1
✓	Actinomadura vinacea strain NBRC 14688 16S ribosomal RNA, partial sequence	Actinomadura vinacea	2628	2628	100%	0.0	98.85%	1471	NR_112737.1

Hasil tanpa filter (no type material dan exclude)

KESIMPULAN QUERY SEQUENCE-1

Data exclude model menggunakan data cloning hasil lab dan untuk data uncultured/environmental biasanya menggunakan data hasil lingkungan seperti barcoding dan metagenomik. Type material memiliki peran intrinsic dalam penamaan spesies.

Pada exclude (pengecualian), data model menggunakan data yang panjang sedangkan untuk uncultured/environmental 300 bp saja sehingga banyak nukleotida yang dibuang.

Hasil yang ditunjukkan pada sequence-1 dengan filter (sequences from type material; and exclude: models (XM/XP), uncultured/environmental sample sequences) menunjukkan hasil yang sama dengan tanpa filter tapi dengan type material, yaitu dengan persentase identitas 98,85%. Arti dari hal tersebut yaitu adanya indikasi sequence cocok pada Actinomadura vinacea strain NBRC 14688, namun terdapat 1% perbedaan pada hasil query dengan subject database.

Hasil lain pada tanpa type material dan exclude, yaitu 100% pada *Actinomadura* sp. HAT23. Hasil tersebut mengindikasikan bahwa *sequence* berasal dari organisme yang belum dijelaskan penamaan spesiesnya, sehingga dapat muncul apabila tidak menunjukkan *type material*.

SEQUENCE 2

Pencarian sequence dengan BLASTX

>Sequence-2

ATGTCCCGCCTTATATATCGAACAGAATTGGGTATATTTGCTGCCTGTCTGGCCCTTGGTACCCTGCTCTTCTTTACGGTGTATCAGACCCGACCT GTCCAGGCAGGTAAAGAAGAGGACAAGGTGCAGGCCTGGCAGACGACGGCTGATCTGAGCACCCATCTGGCTCCACAGGCCAATAAAACG TTTACGACTGGCAGTGGGTCAGACGCGACAAAGATCACGGTTGATGAGAATCAGCAGTATCAACAGATGGACGGCTTTGGTGCCTCGTTTACTG ATTCTTCGGCCTGGCTGGTCTCCAGTAAAATGGATACTAAGCAACGCGATACGTTAATGAAAAATCTTTTTGATGCTCATCAAGGTATTGGCCTGAA CTTTTTACGCCAGCCAATGGGTGCCTCAGATCTGACACGGCCAGCACCCGAGGTCGGTGAGTATTCGTATGACGATATGCCACAGGGTCAGA CTGATCCTAATCTGGATCATTTCTCCATTGACCACGATACGAGTTATATCATTCCGGTACTGCAGCAGGCGCTGAAAGAGAATAAAGATATCAAGAT TATGGCCTCCCCATGGAGCGCACCTGGTTGGATGAAGTCCAGCGGCTCGATGGAAGGGGGAACCCTCAATAGTAGCGCCTATACGGCCTAT GCTAACTACTTTGTGAAATTCATTCAAGCCTATCAGGCTCAGGGTCTCCCCATCTATGCGGTCACACCGCAGAATGAACCACTCTATGTGCCTTC CGGCTATCCAGGCATGTCTTTTCCTGCTAAAGATGAAACTACCTTTATCAGGGACTACCTGGGACCAGCAATGGTGAGCAATGGCCTGAGCACT AAAATTCTGGGTTATGATCATAACTGGGATCAGCCAGGGTATCCAACCACCATTCTGAGTGATCCTGGCACCACACCATATGCCACAGGTACCG CCTGGCATTGCTATGGCGGGTCCGTCAGTGCTCAGACGCCGATCCATAATGCGTATCCAGACAAAGACACCTATGAGACCGATTGTTCTGGTG GTCAGTGGGAAGGCAGCAATGGCTTTGCCAATACGATGGATCTCCTGATCGGCACCACGCGCAACTGGAGCAAGTCAGTAGTTCGCTGGGGA CACTGTCACCTACAATGGCGATTACTATGGTCTGGGACAGGCCAGCAAGTTTGTCCAACCCGGTGCCTATCGTGTCGCATCCAGTTCTGGCCT CAATGGCATTAAAGATGTCGCCTTTAAGAATGCGGATGGTTCGAAGGTGCTGGTGACGTATAATTCGTCTTCCGACCCACAAACGTTTGATGTGC AGTGGGGCGATAAATGGTTCGCTGACACCCTGCCCGCAGGCGCAGCCACAACCTTCAAATGGACTGGCAAACAATCTACCAATAACCACAGT GCGCTGAATCGCACTGGTTGGACCACTTCCACCAGCGCTACACCCGCCAGCACCAATAGCGCGGCCAATGTGCTGGATGATGATTATAC TGCAGGGAGTGACGGCGACTATCCACACGGTTATCAGGTATCTGTCTCGGATGATGGGGCCAACTGGAGCAATCCTGTTGCCAGTGGCAATG GTCCATTTATGAAATGAATGTGTATCCGTAA

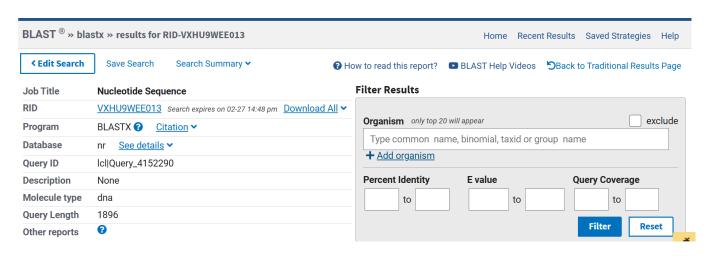
- 1. QUERY sequence name: Sequence-2
- 2. Search Set: Standard database (nr etc.), exclude: non-redundant RefSeq proteins (WP)
- 3. BLAST Search Result:
- Query length: 1896 bp (631 asam amino)
- Protein: Glucosylceramidase
- Top Hit: glucosylceramidase [*Dictyobacter vulcani*], dengan persentase identitas 100%.
- Max Score: 1269; Total Score: 1269: Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Chloroflexota; Ktedonobacterales; Dictyobacter; Dictyobacter vulcani.

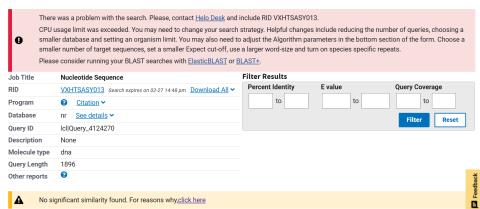
- 1. QUERY sequence name: Sequence-2
- 2. Search Set: Standard database (nr etc.), exclude: models (XM/XP) dan non-redundant RefSeq proteins (WP)
- 3. BLAST Search Result:
- Query length: 1896 bp (631 asam amino)
- Protein: Glucosylceramidase
- Top Hit: glucosylceramidase [*Dictyobacter vulcani*], dengan persentase identitas 100%.
- Max Score: 1269; Total Score: 1269: Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Chloroflexota; Ktedonobacterales; Dictyobacter; Dictyobacter vulcani.

- 1. QUERY sequence name: Sequence-2
- 2. Search Set: Standard database (nr etc.), exclude: -
- 3. BLAST Search Result:
- Query length: 1896 bp (631 asam amino)
- Protein: discoidin domain-containing protein
- Top Hit: discoidin domain-containing protein [*Dictyobacter vulcani*], dengan persentase identitas 100%.
- Max Score: 1269; Total Score: 1269: Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Chloroflexota; Ktedonobacterales;
 Dictyobacter; Dictyobacter vulcani

- 1. QUERY sequence name: Sequence-2
- 2. Search Set: Standard database (nr etc.), exclude: models (XM/XP), non-redundant RefSeq proteins (WP) dan Uncultured/environmental
- 3. BLAST Search Result:

tidak terbaca





	Description	Scientific Name	Max Score		Query Cover	E value	Per. Ident	Acc. Len	Accession
~	discoidin domain-containing protein [Dictyobacter vulcani]	Dictyobacter vu	1269	1269	100%	0.0	100.00%	631	WP_233097673.1
~	discoidin domain-containing protein [Ktedonobacteraceae bacterium]	Ktedonobactera	796	892	100%	0.0	61.43%	777	HEX6555940.1
/	discoidin domain-containing protein [Ktedonobacteraceae bacterium]	Ktedonobactera	784	883	100%	0.0	61.02%	777	MBA2395303.1
/	discoidin domain-containing protein [Ktedonobacteraceae bacterium]	Ktedonobactera	776	881	100%	0.0	59.22%	776	MBO0779299.1
~	discoidin domain-containing protein [Chloroflexota bacterium]	Chloroflexota b	768	877	100%	0.0	59.32%	779	MDQ2903005.1
/	discoidin domain-containing protein [Ktedonobacteraceae bacterium]	Ktedonobactera	764	873	100%	0.0	59.63%	777	HEX6477783.1

Kiri: Tanpa filter exclude;

KESIMPULAN QUERY SEQUENCE-2

Untuk pembuktian data mana yang membuat hasil tidak sesuai kapasitas (parameter seharusnya exclude seluruh opsi dan non-exclude). Saya menemukan masalah pada saat menggunakan filter uncultured/environmental yang menyiratkan bahwa antara sequence terlalu panjang (631 asam amino), dalam jumlah besar memiliki query sequence pada satu pencarian dan query sequence memiliki kompleksitas rendah atau berulang tanpa filtering dan masking.

Sequence-2 dapat ditemukan dengan menggunakan keywords Refseq untuk prokariotik bakteri non-redundant protein record (WP_). Pada data hasil yang sudah difilter dapat ditemukan pada exclude XM/XP dan WP, sedangkan pada hasil dengan uncultured tidak dapat ditemukan karena CPU yang tidak memadai dan permasalahan lain seperti kompleksitas rendah apabila uncultured/environmental dikecualikan.

Solusi dari CPU exceeded yaitu spesifikasi parameter, cut-off sequence, menggunakan spesies repeat sequence dan menggunakan ElasticBLAST atau BLAST+.

Terdapat hasil perbedaan saat mencari dengan exclude (all filter) tidak dapat dicari (CPU exceeded) dan non-exclude ditermukan discoidin domain-containing protein [*Dictyobacter vulcani*] dengan persentase identitas 100%.

KESIMPULAN QUERY SEQUENCE-2

Pada hasil menggunakan exclude RefSeq mendapatkan hasil protein glucosylceramidase dengan organisme sama dengan discoidin domain-containing protein (ddc), yaitu *Dictyobacter vulcani*. Hasil lebih merujuk ke protein ddc dibandingkan dengan glucosylceramidase karena merupakan sequence referensi. Lalu hasil seperti ini dapat terjadi pada struktur protein memiliki kesamaan urutan asam aminonya (*identical protein*

group).



Terdapat hasil perbedaan saat mencari dengan exclude (all filter) tidak dapat dicari (CPU exceeded) dan non-exclude ditermukan discoidin domain-containing protein [*Dictyobacter vulcani*] dengan persentase identitas 100%.

SEQUENCE 3

Pencarian sequence dengan BLASTP

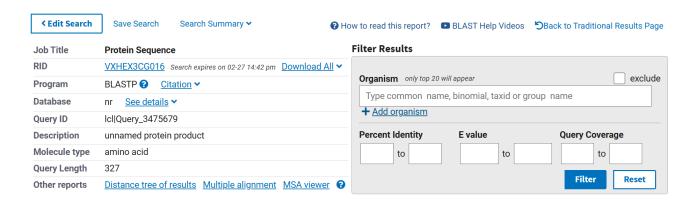
>Sequence-3

MLLAGDIGGTKTNLAVFSSKDELRKPLYAKKFPSAQYTTLAALVTDFLSEVDLPIDRAVFGVAGPVLEGKAKITNLPWVMEENQLQSALKIPTIRLINDL AATAQSIPALEPADLHTLNAGEPMKNGTMSVVAPGTGLGEAFLVWDGSKYAIYPSEGGHADFAPTNAFEVGLLVYMLERLPHVSYEHVCSGIGLPNI YAYIKESGMFVEPEWLSEKLANAADRTPVIADGAMAAEPAPICMAALKSFAAILGAEAGNMAIKVLSTGGVYLGGGIPPRILPFLESDDFMRAFRNKG RFSNMLGNVPVHVILRPDAGLIGAAAYGFNML

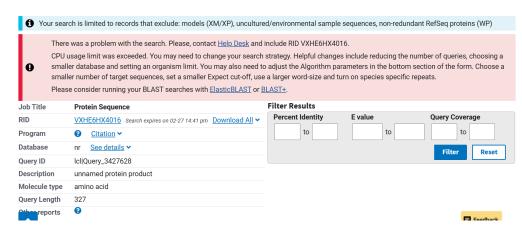
- 1. QUERY sequence name: Sequence-3
- 2. Search Set: Database non-redundant protein sequences (nr); exclude: -
- 3. BLAST Search Result:
- Query length: 327 asam amino
- Protein: glucokinase
- Top Hit: glucokinase [*Tengunoibacter tsumagoiensis*], dengan persentase identitas 100%.
- Max Score: 663; Total Score: 663: Query cover: 100%; E value: 0.0
- Organisme: Bacteria; Bacillati; Chloroflexota; Ktedonobacteria;
 Ktedonobacterales; Dictyobacteraceae; Tengunoibacter; Tengunoibacter tsumagoiensis

- 1. QUERY sequence name: Sequence-3
- 2. Search Set: Database non-redundant protein sequences (nr), exclude models (XM/XP), uncultured/environmental sample sequences, non-redundant RefSeq proteins (WP)
- 3. BLAST Search Result:

tidak terbaca



	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
~	glucokinase [Tengunoibacter tsumagoiensis]	Tengunoibacter tsumagoiensis	663	663	100%	0.0	100.00%	327	WP_126583343.1
\checkmark	glucokinase [Dictyobacter formicarum]	Dictyobacter formicarum	472	472	100%	6e-164	67.94%	354	GHO83574.1
\checkmark	glucokinase [Dictyobacter formicarum]	<u>Dictyobacter formicarum</u>	471	471	100%	1e-163	67.94%	340	WP_201361243.1
\checkmark	glucokinase [Dictyobacter kobayashii]	Dictyobacter kobayashii	470	470	99%	1e-163	67.55%	340	WP_126556085.1
\checkmark	glucokinase [Dictyobacter alpinus]	<u>Dictyobacter alpinus</u>	467	467	100%	3e-162	65.09%	338	WP_126630717.1
\checkmark	glucokinase [Dictyobacter aurantiacus]	Dictyobacter aurantiacus	461	461	99%	5e-160	65.68%	340	WP_126601200.1



Kiri: tanpa filter exclude;

KESIMPULAN QUERY SEQUENCE-3

Pada data hasil sama dengan sequence-3, dengan hasil yang tidak dapat ditemukan pada data dengan exclude seluruh opsi (models (XM/XP), uncultured/environmental sample sequences, non-redundant RefSeq proteins (WP)). Lalu apabila filter exclude tidak dipilih, maka hasil mengarah pada *Tengunoibacter tsumagoiensis* dan seratus persentase identitas (100%) dengan protein glucokinase.

Sudah dicoba oleh pelaku (penulis) pada beberapa percobaan, hasil yang sulit untuk dicari apabila terdapat exclude uncultured/environmental. Hal tersebut dikarenakan oleh hasil data metagenomik yang banyak dan dapat berkorelasi dengan sequence-3. Masalah lain seperti sequence spesifik yang hanya ada di *Tengunoibacter tsumagoiensis dan seratus persentase identitas* (100%), sedangkan organisme lain pada produk protein tidak ada yang highly identic.

- 1. OR169734
- 2. AB008413
- Screenshot hasil pencarian identitas
- Cantumkan FASTA filenya

1. OR169734

Screenshot hasil pencarian identitas

```
Euconnus simplex isolate MSC_11760 cytochrome c oxidase subunit I (COX1)
gene, partial cds; mitochondrial
GenBank: OR169734.1
FASTA Graphics
Go to: ✓
LOCUS
                                  421 bp
                                           DNA
                                                   linear INV 27-JUN-2023
DEFINITION Euconnus simplex isolate MSC 11760 cytochrome c oxidase subunit I
           (COX1) gene, partial cds; mitochondrial.
ACCESSION OR169734
VERSION
           OR169734.1
KEYWORDS
SOURCE
           mitochondrion Euconnus simplex
 ORGANISM Euconnus simplex
           Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta;
           Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
           Staphyliniformia; Staphylinidae; Scydmaeninae; Euconnus.
REFERENCE 1 (bases 1 to 421)
 AUTHORS Caterino, M.S. and Recuero, E.
          Caterino, M.S. and Recuero E. (in review) Shedding light on dark
           taxa in high Appalachian leaf litter - assessing patterns of
           endemicity using large scale, voucher-based barcoding. INSECT
           CONSERVATION & DIVERSITY
           Unpublished
REFERENCE 2 (bases 1 to 421)
          Caterino, M.S. and Recuero, E.
           Direct Submission
 JOURNAL Submitted (12-JUN-2023) Plant & Environmental Sciences, Clemson
           University, 277 Poole Ag. Center, Clemson, SC 29634, USA
COMMENT
           ##Assembly-Data-START##
```

Fasta File

2. AB008413

Screenshot hasil pencarian identitas

Aspergillus zonatus gene for 18S rRNA, partial sequence

```
GenBank: AB008413.1
FASTA Graphics
Go to: ♥
LOCUS
           AB008413
                                   1733 bp
                                                      linear PLN 13-FEB-1999
DEFINITION
           Aspergillus zonatus gene for 18S rRNA, partial sequence.
           AB008413
ACCESSION
VERSION
           AB008413.1
           18S ribosomal RNA.
KEYWORDS
           Penicilliopsis zonata
 ORGANISM Penicilliopsis zonata
           Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina;
           Eurotiomycetes; Eurotiomycetidae; Eurotiales; Aspergillaceae;
           Penicilliopsis.
REFERENCE
           Nikkuni, S., Nakajima, H., Hoshina, S.I., Ohno, M., Suzuki, C.,
           Kashiwagi, Y. and Mori, K.
 TITLE
           Evolutionary relationships among Aspergillus oryzae and related
           species based on the sequences of 18S rRNA genes and internal
           transcribed spacers
 JOURNAL
          J. Gen. Appl. Microbiol. 44 (3), 225-230 (1998)
           12501432
           2 (bases 1 to 1733)
REFERENCE
 AUTHORS
           Nikkuni,S.
           Direct Submission
          Submitted (22-OCT-1997) Savuki Nikkuni, Japan International
           Research Center for Agricultural Sci.; Ohwashi 1-2, Tsukuba,
           Ibarakki 305, Japan (E-mail:nikkuni@jircas.affrc.go.jp,
           Tel:0298-38-6362, Fax:0298-38-6652)
                    Location/Qualifiers
FEATURES
                    1 1733
```

```
FEATURES
source
1..1733
/organism="Penicilliopsis zonata"
/mol_type="genomic DNA"
/isolate="IFO8817"
/db_xref="taxon:41063"
<1..1733
/product="18S ribosomal RNA"
```

2. AB008413

Fasta file

>AB008413.1 Aspergillus zonatus gene for 18S rRNA, partial sequence AAAGATTAAGCCATGCATGTCTAAGTATAAGCAATTTATACGGTGAAACTGCGAATGGCTCATTAAATCA GTTATCGTTTATTTGATAGTACCTTGCTACATGGATACCTGTGGTAATTCTAGAGCTAATACATGCTGAA AACCTCGACTTCGGAAGGGGTGTATTTATTAGATAAAAAACCAATGCCCCTCGGGGCTCCTTGGTGATTC **ATAATAACTAAACGAATCGCATGGCCTTGCGCCGGCGATGGTTCATCAAATTTCTGCCCTATCAACTTT** CGATGGTAGGATAGTGGCCTACCATGGTGGCAACGGGTAACGGGGAATTAGGGTTCGATTCCGGAGAGGG AGCCTGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGCGCGCAAATTACCCAATCCCGATTCGGGGA GGTAGTGACAATAAATACTGATACGGGGCTCTTTTGGGTCTCGTAATTGGAATGAGTACAATTTAAATCC CTTAACGAGGACAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTAT ATTAAAGTTGTTGCAGTTAAAAAGCTCGTAGTTGAACCTTGGGTCTGGCCGGTCCGCCTCACCGCG AGTACTGGTCCGGCTGGACCTTTCCTTCTGGGGAACTCCATGGCCTTCACTGGCCGTGGGGGGAACCAGG ACTTTTACTGTGAAAAAATTAGAGTGTTCAAAGCAGGCCTTTGCTCGAATACATTAGCATGGAATAATAG AATAGGACGTGCGGTTCTATTTTGTTGGTTTCTAGGACCGCCGTAATGATTAATAGGGATAGTCGGGGGC AAGGATGTTTCATTAATCAGGGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCTTA ACCATAAACTATGCCGACTAGGGATCGGACGGTGTTCTTTTTATGACCCGTTCGGCACCTTACGAGAAAT CAAAGTTTTTGGGTTCTGGGGGGGGTATGGTCGCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACC ACAAGGCGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTCACCAGGTCCAGACAAAATAAGG ATTGACAGATTGAGAGCTCTTTCTTGATCTTTTGGATGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTG **ATTTGTCTGCTTAATTGCGATAACGAACGAGACCTCGGCCCTTAAATAGCCCGGTCCGCATTTGCGGGCC** GCTGGCTTCTTAGGGGGACTATCGGCTCAAGCCGATGGAAGTGCGCGGCAATAACAGGTCTGTGATGCCC TTAGATGTTCTGGGCCGCACGCGCGCTACACTGACAGGGCCAGCGAGTACATCACCTTGGCCGAAAGGTC TGGGTAATCTTGTTAAACCCTGTCGTGCTGGGGATAGAGCATTGCAATTATTGCTCTTCAACGAGGAATG CCTAGTAGGCACGAGTCATCAGCTCGTGCCGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTAC TACCGATTGAATGGCTCAGTGAGGCCTTCGGACTGGCTCAGGAGGGTTGGCAACGACCCCCCAGAGCCGG AAAGTTGGTCAAACTCGGTCATTTAGAGGAAGTAAAAGTCGTAACAAGGTTTC

References

- 1. NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information, Nucleic Acids Research, Volume 43, Issue D1, 28 January 2015, Pages D6–D17, https://doi.org/10.1093/nar/gku1130
- 2. Stephen F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schäffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.
- 3. Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.
- 4. Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb Miller (2000), "A greedy algorithm for aligning DNA sequences", J Comput Biol 2000; 7(1-2):203-14.