

Find a Gene Project

Renee Zuhars (PID: A17329856)

Table of contents

Question 1:	1
Question 2	2
Question 3	3
Question 4	5

Question 1:

Beginning my search, I knew I wanted to limit my organism to some kind of fungus, because they are very understudied and I think their diversity and versatility are fascinating!

I decided to narrow my search to those proteins that help the fungus *Ophiocordyceps Unilateralis* “zombify” a host insect by taking over the hosts’ neurological systems, eventually killing the host.

protein name: serine/threonine-protein kinase MAK, partial

species: *Ophiocordyceps Unilateralis*

accession number: ADI72911.1

function: The role of MAK-like kinases in this species is to induce behavioral changes in the host by interfering with Mitogen- Activated Protein Kinase signaling pathways. (ChatGPT)

```
knitr::include_graphics("ophiocordyceps.png")
```



Question 2

To attempt to find a homologous protein, I inputted the accession number into NCBI's tblastn search, using the est database, and did not add any limits or restrictions.

My BLAST results were as followed:

```
knitr::include_graphics("original protein blast results.png")
```

Job Title

gb|ADI72911.1|

RID

17SCCHF016 Search expires on 05-03 05:03 am [Download All](#)

Program

TBLASTN [Citation](#)

Database

est [See details](#)

Query ID

ADI72911.1

Description

serine/threonine-protein kinase MAK, partial [Ophiocordyci ...

Molecule type

amino acid

Query Length

139

Other reports

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

Select columns

Show

☒ select all

24 sequences selected

[GenBank](#)

[Graphics](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA sequence	Zymoseptoria tritici	156	156	91%	3e-46	59.84%	589	AW180074.1
<input checked="" type="checkbox"/>	ApulSEQ16326 Aureobasidium pullulans pBluescript (EcoRI-XhoI) Aureobasidium pullulans cDNA clone ApulSEQ...	Aureobasidium p...	124	124	91%	2e-33	56.59%	726	DY862280.1
<input checked="" type="checkbox"/>	CATN5183 fwd CATN Nectria haematococca mpVI Sporulation (PDB) 77-13-4 Mycelia 24 hour culture [Nectria] ha...	Fusarium vanettenii	100	100	91%	1e-23	48.03%	721	GE229289.1
<input checked="" type="checkbox"/>	Asn_08688 Aspergillus niger cBluescript (EcoRI-XhoI) Aspergillus niger cDNA clone Asn_08688, mRNA sequence	Aspergillus niger	99.0	99.0	74%	2e-23	54.29%	670	DR707892.1
<input checked="" type="checkbox"/>	FQ083057 Botrytis tulipae B19901 Botrytis tulipae cDNA, mRNA sequence	Botrytis tulipae	99.0	99.0	91%	2e-23	48.91%	656	FQ083057.1
<input checked="" type="checkbox"/>	FQ105848 Botryotinia ficarinarum CBS17663 Botryotinia ficarinarum cDNA, mRNA sequence	Botryotinia ficaria...	95.5	95.5	91%	2e-22	48.18%	574	FO105848.1
<input checked="" type="checkbox"/>	FQ948614 Botryotinia fuckeliana 2 days old mycelia in rich medium cDNA library Botrytis cinerea cDNA, mRNA se...	Botrytis cinerea	95.1	95.1	91%	2e-22	48.18%	563	FQ948614.1

I decided to focus on the first result, “MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA sequence”.

Question 3

Here is some information about the homolog I am looking into:

FASTA format sequence, translated using EMBOSS Transeq: AW180074.1_1
MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA sequence
RQLSVNSQGNHYAEIHRQEAERALVGASALKSPTGSQRESFFSHLRKRAR-
RLSGRNSGVI TPSMDAMETSAGCVPWAANKQTTFDTHSIASAAADPSSDPNFAEL-
DRALQSVRYSLDAAA NATQQARKPTNRVVEQPSLKRHHSLPHGVRHKTNPTTVY-
HDEH*STPRAADTRPPTKKKN SRRSHELASASRTAFSX

AW180074.1_2 MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA sequence
DNYQSTHRAITTPKFTGRKLSVLWLAQALSSHRLAAKEKASSLICARGREDF-
PAATQVSS HLQWMLWKPALGAFLGLLTNKPSTPTRSRPQPPIRHQTPISLSWIVHCK-
VYDTAWMPPR TRLNKLGSLRTALSNHHSVTTRFLTALDTRPTQPPYTTTSTEARHEQPIQDPRRRRRI
LDEVMNSAHLAARRSR

AW180074.1_3 MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA sequence
TIISQLTGQSLRRNSPAGSACSGWRKRSQVTDWQPKRKLSSAQEGEKTFRPQLRCHH
TFNGCYGNQRWVRS LGCQTNHLRHPLDRVCRSRSVIRPQFRAGSCTAKCTIQPGCRRE

RDSTSEAYEPRSATIIIEASPLASSRRTQDQPNHRIPRRALKHATSSRYKTPDEEEEF
*STKS*TQRISPHGV LX*

AW180074.1_4 MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137
5', mRNA sequence RERRAARCAEFMTSSRILLRRGSCIGCSWRASVLVVYG-
GWVGLVSN AVRKR VVTLQ*W LLNYAVRRLPSLLSRVRGGIQAVSYTLQCTIQL-
SEIGVRIGCGRRDRVGVEGGLFVSS PRNAPSAGFHSIH*R***CDDT**VAAGKSSRPLAQMREEAF
VNFGVVIALVDLS

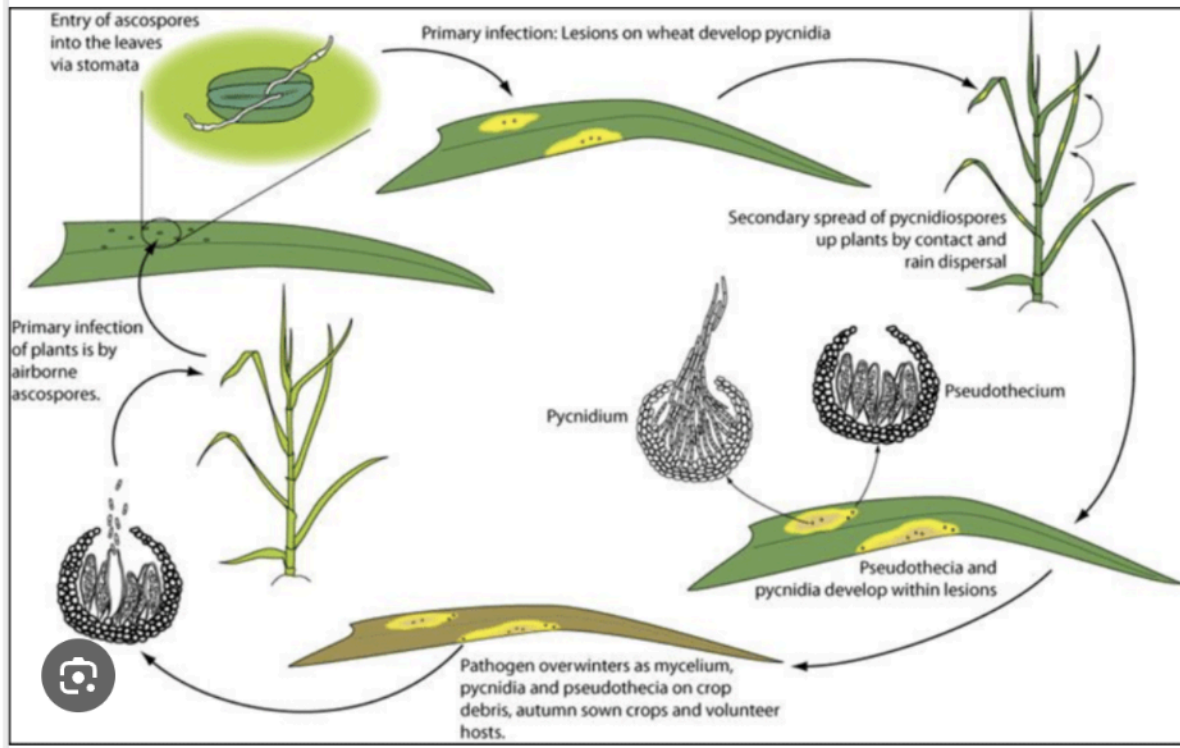
AW180074.1_5 MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA
sequence SRTPCGEMR *VHDFVENSSSSSGVLYRLLVACFSARRGIRWLGWSCVRREEASGDASMM*
VAQLRGSASLVESRRRHPGCIVHFAVHDP AQRNWGLMTDRLRQTRSSGCRRWFVCQ
*PKERTQRWFP HPLKV**HLSCGRKVFSPSCADERRSFLFGCQSVTERLRQPEHAQLPA*
GEFRRSDCPVSLIIVX

AW180074.1_6 MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA
sequence ENAVRRDALSSLRREFFFFVGGLVSAARGVLQCSSWYTVVGLVLCLTPGSEWR**FNDG**
CSTTRFVGFLACVAFAAASRLYRTLCSARSSSAKLGSDDGSAADAIEWVSKVVCLLAA
QGTHPALVSIASIEGVMTPELRPESLLALLRREKKLSLWLPVGDLRALAPTRARSASCR
*ISA**LPCELTDNCR*

Name: MgA0137f MgA Library Zymoseptoria tritici cDNA clone MgA0137 5', mRNA se-
quence

Species Derived from: Zymoseptoria tritici : this is a pathogenic fungus that attacks wheat
plants. It is resistant to multiple fungicides, and causes septoria leaf blotch.

```
knitr::include_graphics("zymoseptoria tritici.png")
```



Question 4

To determine if my protein is novel, I used the “blastp” tool and NCBI’s nr database and ran the above FASTA format sequence through it.

My results were as followed:

```
knitr::include_graphics("novel protein blast results.png")
```

Job Title

AW180074.1_1 MgA0137f MgA Library Zymoseptoria

RID

17TB565R013 Search expires on 05-03 05:19 am [Download All](#) ▾

Results for

1:|cl|Query_8618976 AW180074.1_1 MgA0137f MgA Library Zymosep ▾

Program

BLASTP [Citation](#) ▾

Database

nr [See details](#) ▾

Query ID

|cl|Query_8618976

Description

AW180074.1_1 MgA0137f MgA Library Zymoseptoria tritic ...

Molecule type

amino acid

Query Length

197

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"/>	serine/threonine protein kinase_CMGC_family [Zymoseptoria tritici IPO323]	Zymoseptoria tritici IPO323	306	306	79%	7e-96	96.77%	777	XP_003856578.1
<input checked="" type="checkbox"/>	unnamed protein product [Zymoseptoria tritici ST99CH_3D7]	Zymoseptoria tritici ST99CH_3D7	305	305	79%	4e-95	96.77%	839	SMQ46477.1
<input checked="" type="checkbox"/>	CMGC/RCK/MAK protein kinase [Zymoseptoria brevis]	Zymoseptoria brevis	297	297	78%	2e-92	94.77%	777	KJY01546.1
<input checked="" type="checkbox"/>	putative serine/threonine-protein kinase_active [Septoria linicola]	Septoria linicola	215	215	75%	2e-61	69.59%	788	KAI5369935.1

There is no match with 100% identity!