Tblastx geeft op alle 4 de sequenties geen hits.

**Informational Message: Failed to collect db stats for nr (entrez query: all [filter] NOT(txid2759 [ORGN]))**

**Informational Message: Query 'lcl|Query\_180794 HWI-M02942:21:000000000-ACNW4:1:1101:10943:2630/1' (# 1): Warning: GI or TI list filtering resulted in an empty database.**

**Informational Message: Query 'lcl|Query\_180795 HWI-M02942:21:000000000-ACNW4:1:1101:10943:2630/2' (# 2): Warning: GI or TI list filtering resulted in an empty database.**

**Informational Message: Query 'lcl|Query\_180796 HWI-M02942:21:000000000-ACNW4:1:1101:19821:2914/1' (# 3): Warning: GI or TI list filtering resulted in an empty database.**

**Informational Message: Query 'lcl|Query\_180797 HWI-M02942:21:000000000-ACNW4:1:1101:19821:2914/2' (# 4): Warning: GI or TI list filtering resulted in an empty database.**

Blastn geeft voor alle 4 de sequenties geen hits.

**Informational Message: Failed to collect db stats for nr (entrez query: all [filter] NOT(txid2759 [ORGN]))**

**Informational Message: Query 'lcl|Query\_195536 HWI-M02942:21:000000000-ACNW4:1:1101:10943:2630/1' (# 1): Warning: GI or TI list filtering resulted in an empty database.**

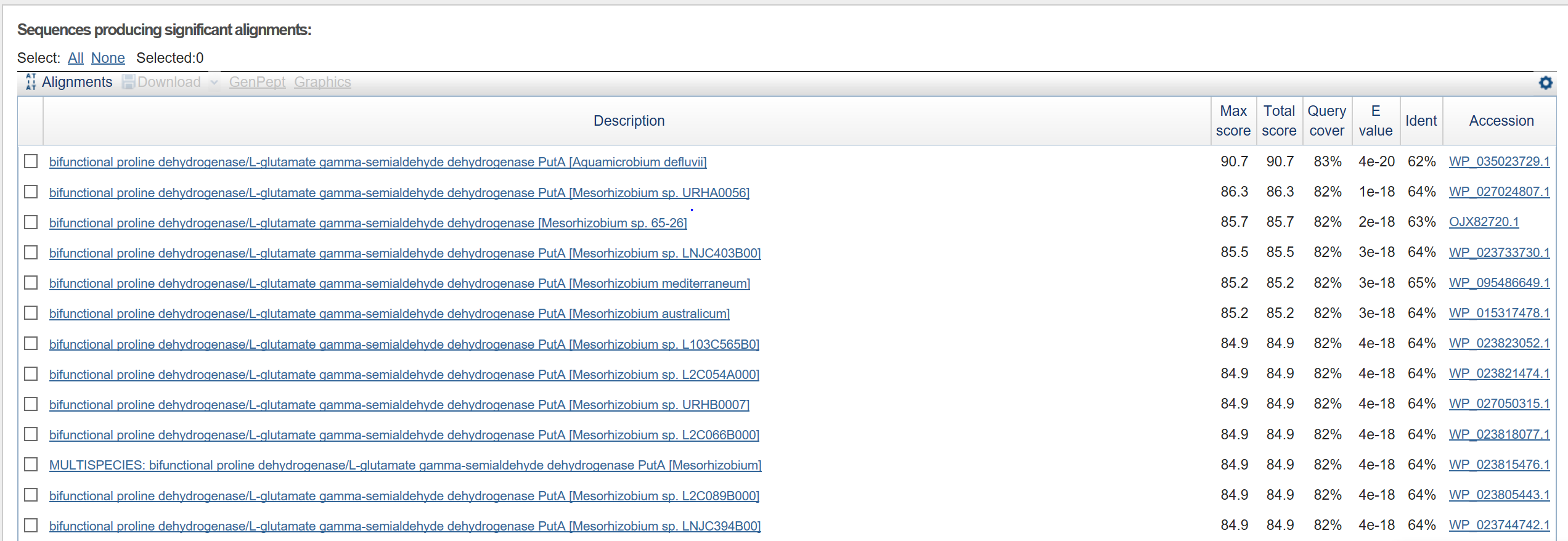
**Informational Message: Query 'lcl|Query\_195537 HWI-M02942:21:000000000-ACNW4:1:1101:10943:2630/2' (# 2): Warning: GI or TI list filtering resulted in an empty database.**

**Informational Message: Query 'lcl|Query\_195538 HWI-M02942:21:000000000-ACNW4:1:1101:19821:2914/1' (# 3): Warning: GI or TI list filtering resulted in an empty database.**

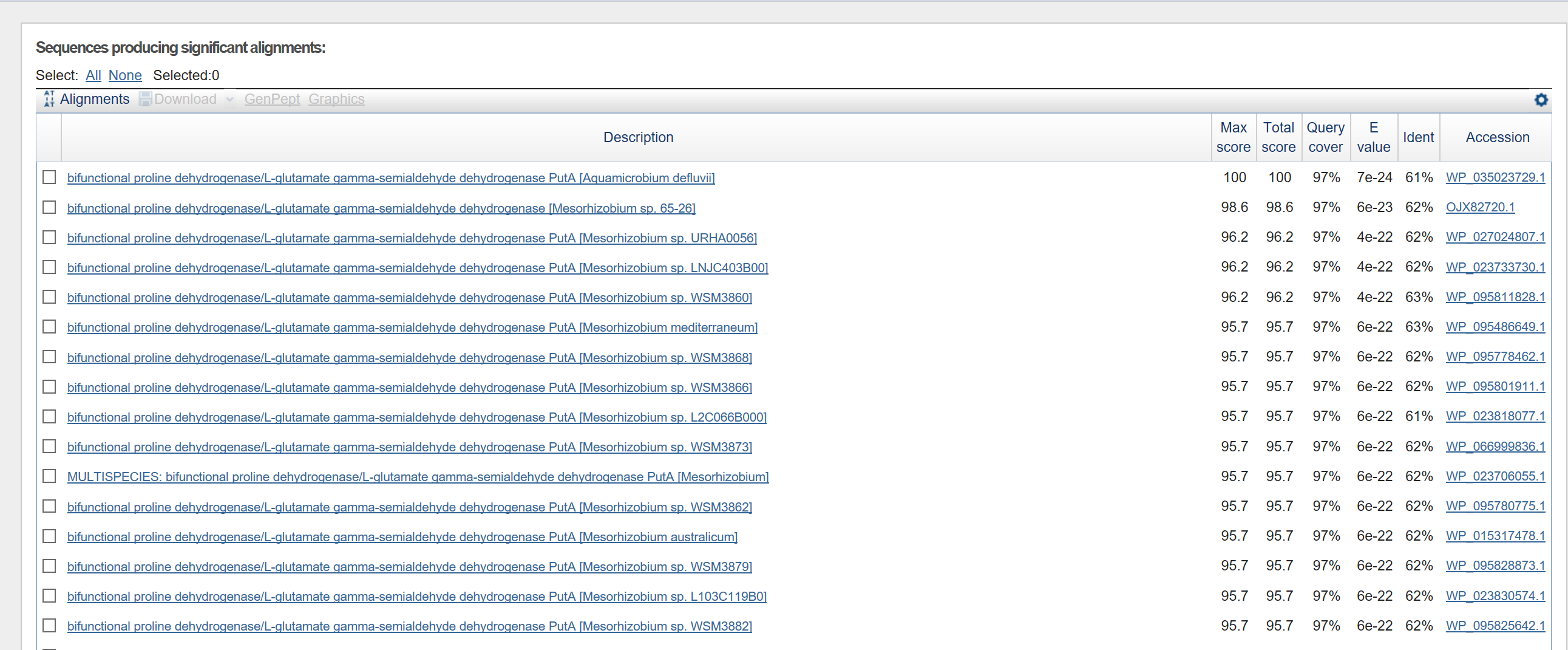
**Informational Message: Query 'lcl|Query\_195539 HWI-M02942:21:000000000-ACNW4:1:1101:19821:2914/2' (# 4): Warning: GI or TI list filtering resulted in an empty database.**

Blastx geeft voor alle 4 de sequenties een hit.

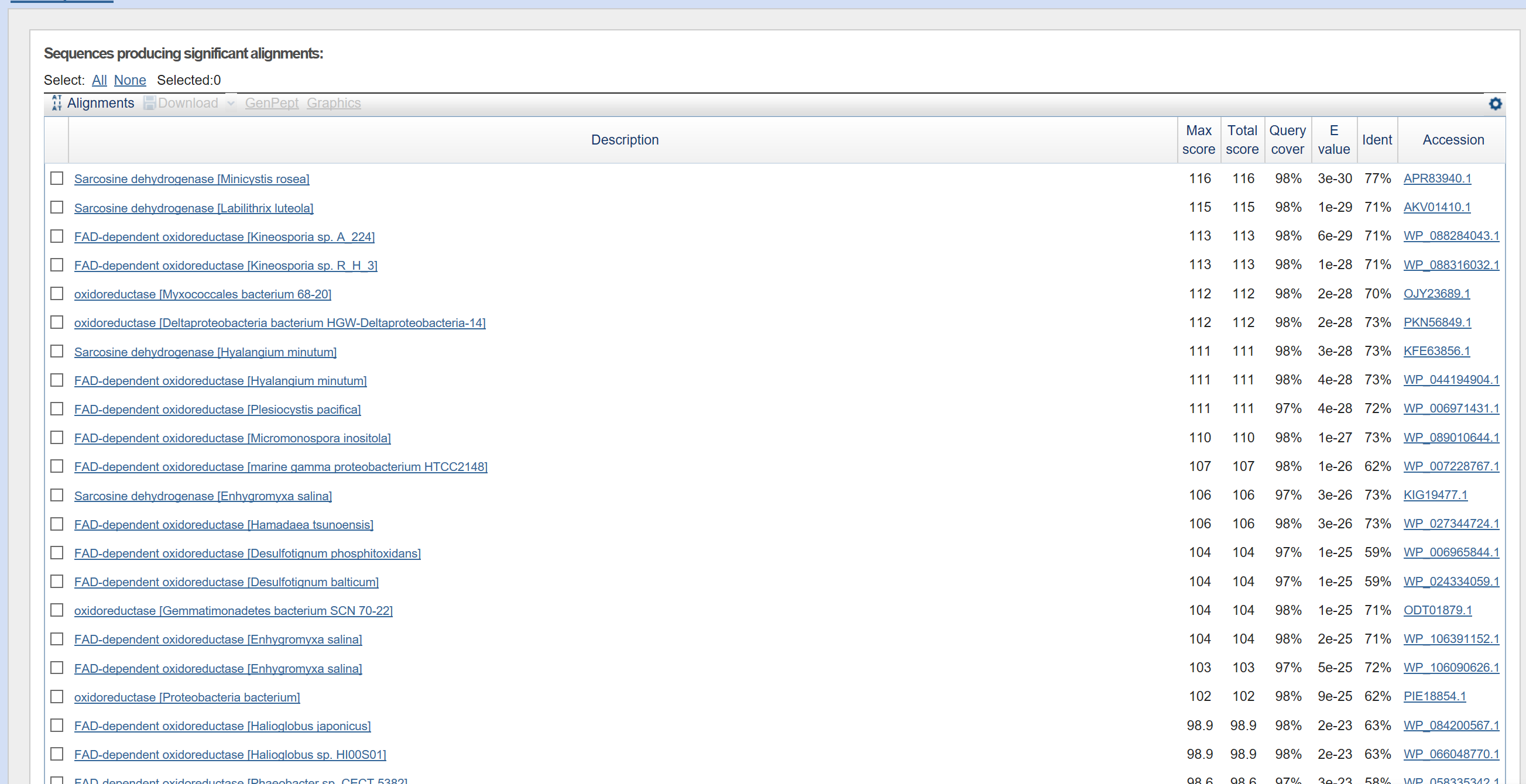
2630/1

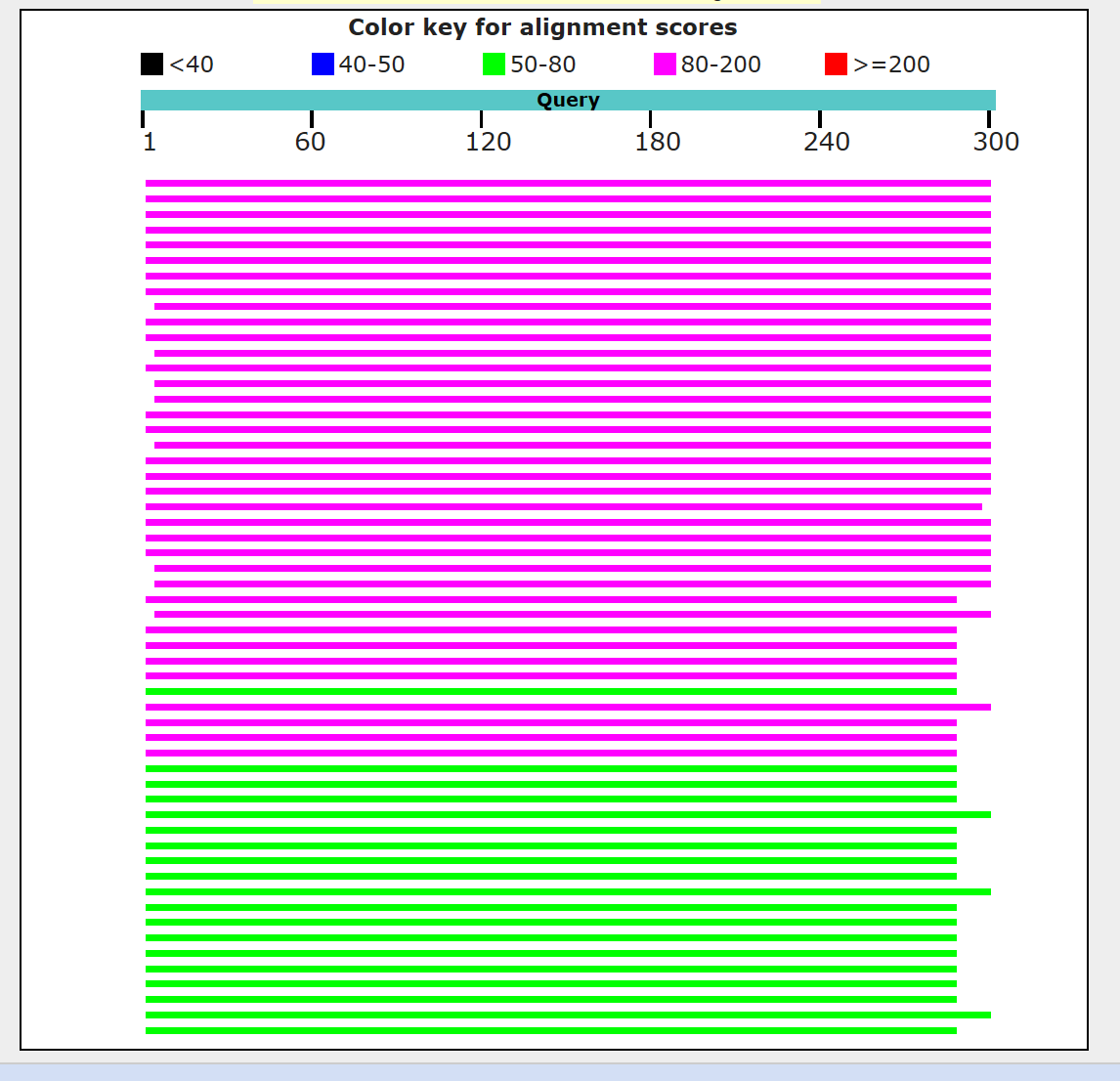
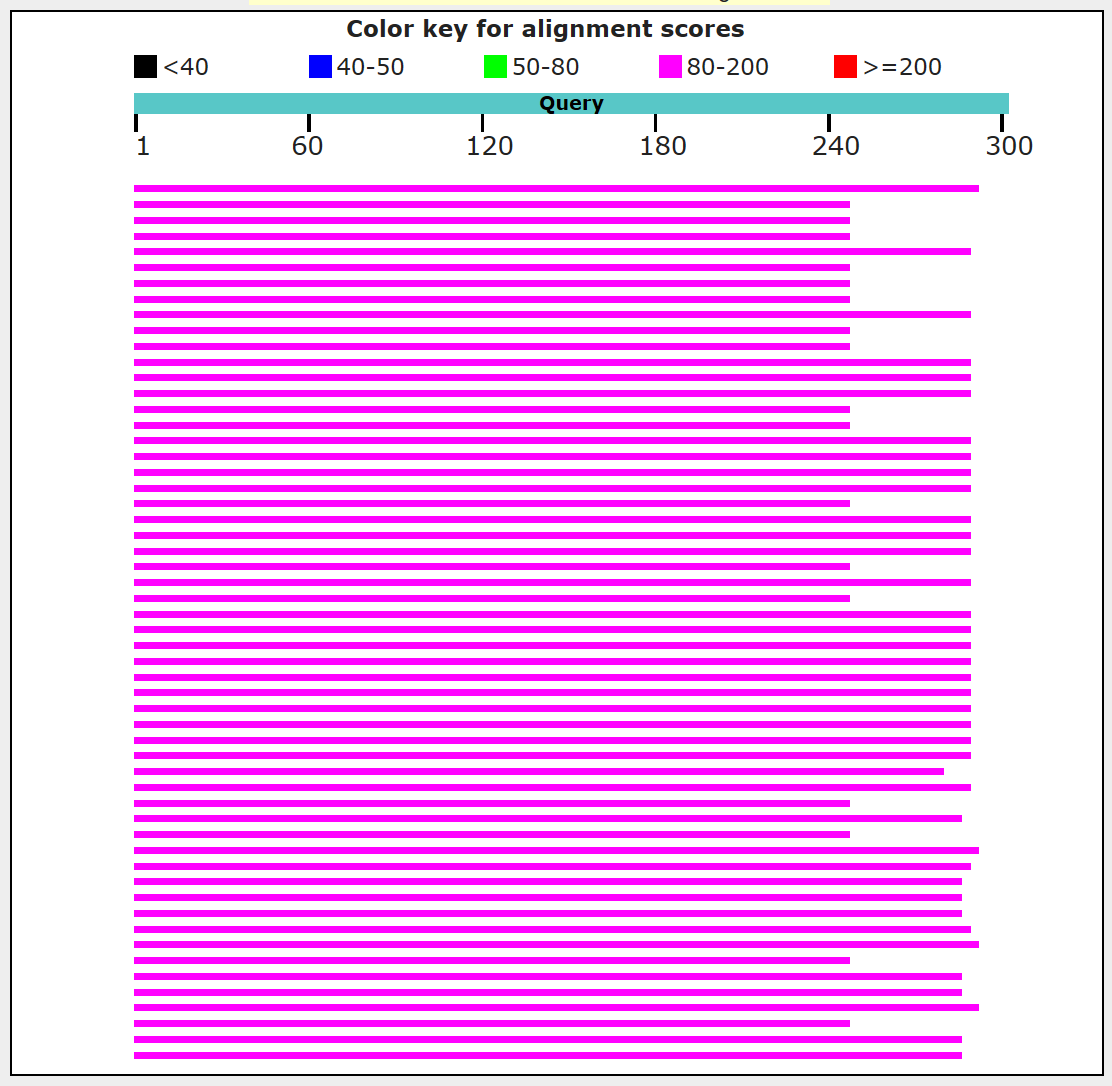


2630/1 2630/2

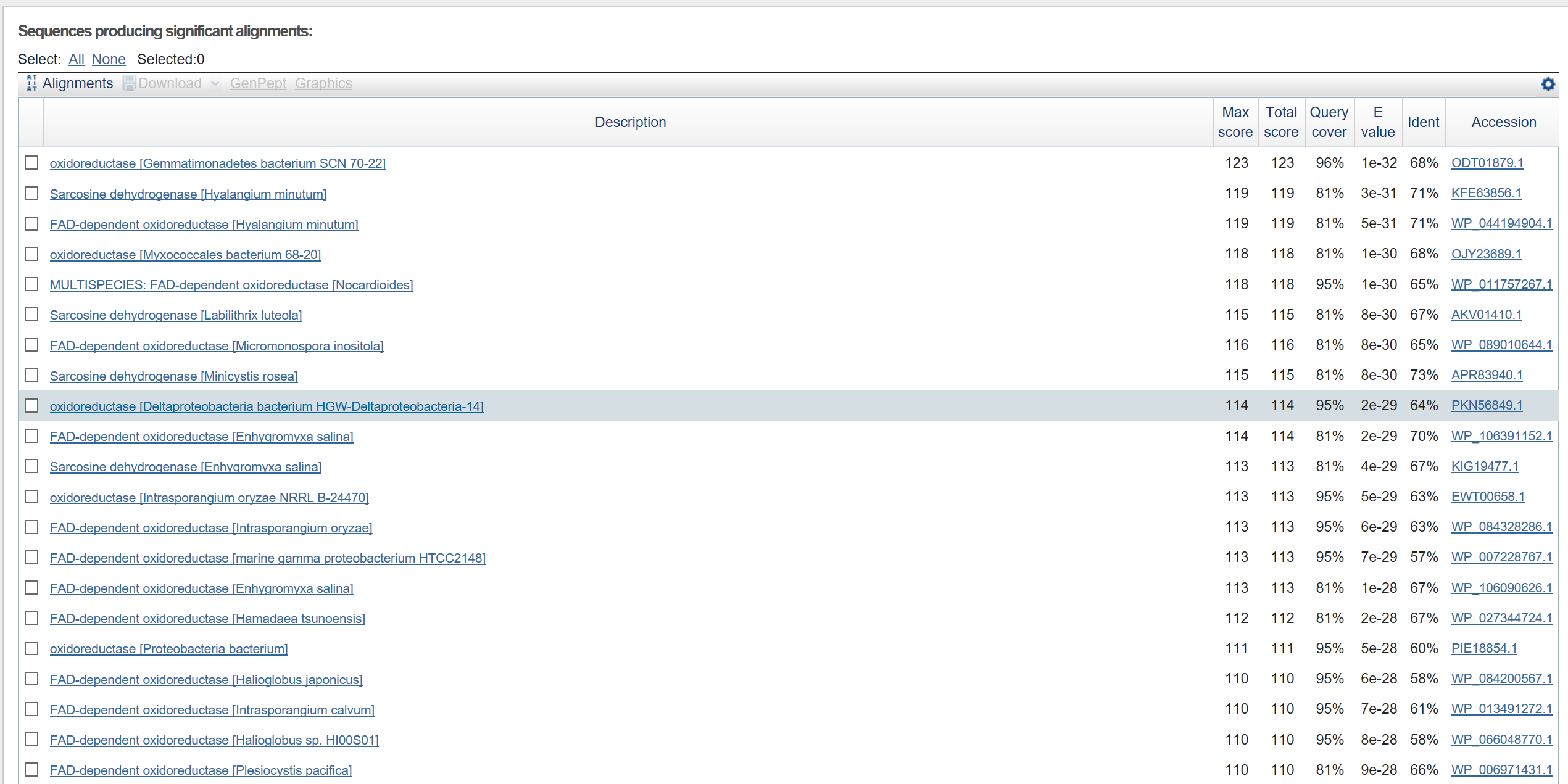
2630/2

2914 1



2914/12914/2

2914 2



**Hoe zeker ben je van het gevonden organisme? Ben je zeker op soort niveau, geslacht niveau, of op familie niveau?**

**Klopt het organisme wat je gevonden hebt?**

2630/1

bifunctional proline dehydrogenase/L-glutamate gamma-semialdehyde [dehydrogenase PutA](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376)

**Geslacht** [Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376) of [Aquamicrobium defluvii](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_737028353)

**Soort** van [Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376): [australicum](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376), [mediterraneum](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1239507221), [ciceri](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1017216850), [muleiense,](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1223561758) [plurifarium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1122250625)

2630/2

bifunctional proline dehydrogenase/L-glutamate gamma-semialdehyde [dehydrogenase PutA](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376)

**Geslacht** [Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376) en [Aquamicrobium defluvii](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_737028353)

**Soort** van [Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376): [**loti**](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_974595428), [temperatum,](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1239515622) [ciceri,](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_639163952) [plurifarium,](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_751298905) [australicum](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376)

2914/1

[**Sarcosine dehydrogenase**](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1127362537)

Minicystis rosea(sar) , Labilithrix luteola(sar)

[**FAD-dependent oxidoreductase**](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_494029291)

Kineosporia (FAD), [Desulfotignum balticum](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_636807783) (fad), [Plesiocystis pacifica](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_494029291) (fad), [Halioglobus japonicus](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1180715421) (fad), [Bradyrhizobium valentinum](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1177255627) (fad)

2914/2



[**FAD-dependent oxidoreductase**](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_763336575)

Hyalangium minutum

Micromonospora inositola

Hamadaea tsunoensis

Desulfotignum balticum

[**Sarcosine dehydrogenase**](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1127362537)

Minicystis rosea

Enhygromyxa salina

**Waar komt dit organisme normaliter voor? Wat voor eigenschappen heeft dit organisme?**

2630

Ik heb alleen gezocht op het geslacht [Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376) en het organisme [Aquamicrobium defluvii](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_737028353)

[Mesorhizobium](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376):

Mesorhizobium is een geslacht van de gram-negative soil bacteria. Soil bactieren zitten in de grond en zijn gunstig voor de plantengroei. Op z’n minst 1, de nitrogen fixing soort, mesorhizobium loti, kan een simbiose aangaan met de wortels van planten van het geslacht lotus.

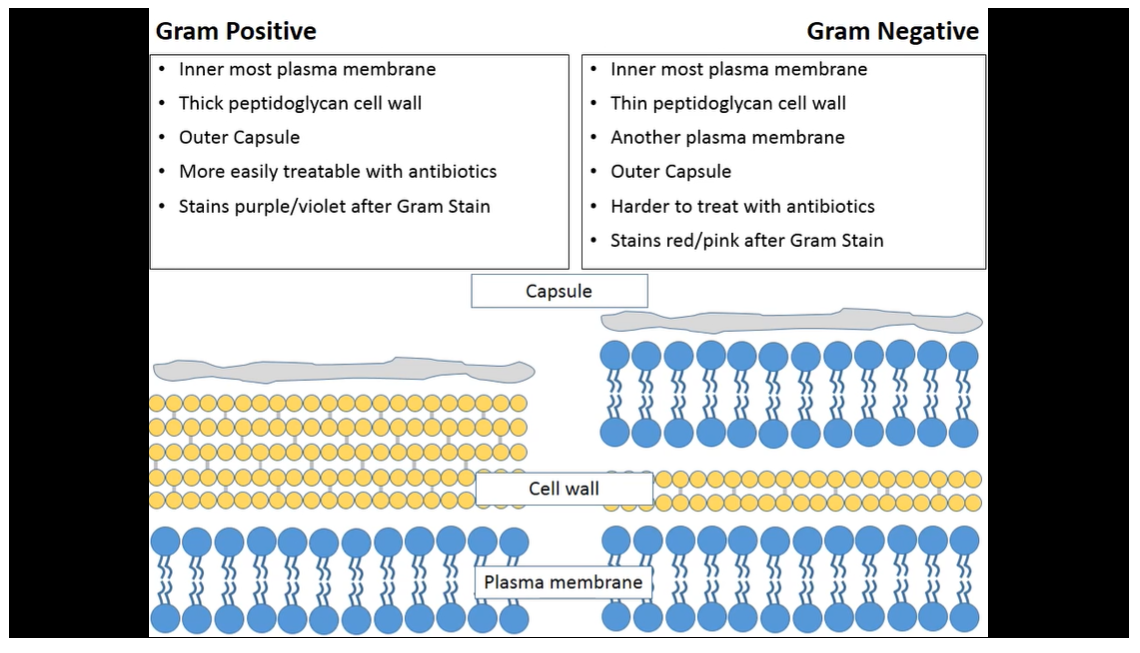
*Aquamicrobium defluvii:*

Aquamicrobium defluvii is een gram-negative, oxidase(enzym dat oxidatie en reductie reacties katalyseert)- en catalase-positive,(een enzym dat de decompositie van hydrogeen peroxide naar water en zuurstof katalyseert.) bacterie van het geslacht Aquamicrobium. Aquamicrobium defluvii gebruikt thiophene-2-carboxylate als enige bron van carbon.

Bifunctional proline dehydrogenase/L-glutamate gamma-semialdehyde [dehydrogenase PutA](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376)

Hits: in NCBI BioSystems:

* Alanine, aspartate and glutamate metabolism
* Arginine and proline metabolism



2914

[Sarcosine dehydrogenase](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1127362537) en [FAD-dependent oxidoreductase](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_494029291)

Ik heb alleen gezochtop het organisme Hyalangium minutum

Hyalangium minutum is a member of order Myxococcales, suborder Cystobacterineae and family Cystobacteraceae. It is a close relative of Cystobacter fuscus DSM 2262 and Stigmatella aurantiaca.

Wat is de functie van het eiwit gecodeerd door het geblaste fragment?

2630

Bifunctional proline dehydrogenase/L-glutamate gamma-semialdehyde [dehydrogenase PutA](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_505130376):

ACCESSION WP\_108701625

2914

Welke biologische domeinen heeft dit eiwit?

In welk biologisch proces wordt dit eiwit gebruikt?

Waar in het micro-organisme zit dit eiwit?

Zitten er transmembraan domeinen in dit eiwit? Hoe is de topologie van dit eiwit?

Welk(e) Blastalgoritme(s) ga je gebruiken?

BLASTX EN BLASTN, mochten hier geen resultaten uit komen dan zal tBLASTx gebruikt kunnen worden

Welke scorematrix ga je hiervoor gebruiken?

We gaan er vanuit dat de meeste sequenties verwant zijn, maar we gaan er ook vanuit dat er micro-organismen tussen zitten die niet verwant zijn. dus gebruiken we PAM250

Ga je nog andere parameters van Blast aanpassen?

Database 🡪 nr

Exclude 🡪 eukaryotes

Gap costs 🡪 existence: 11 extension: 1

Ga je gebruik maken van een cut-off voor het opslaan van hits? Doe je dat op score of op E-value?

Cut-off 🡪 op Evalue van 1\*10-10

Van hoeveel hits ga je de resultaten opslaan?

Per sequenctie willen we 10 hits opslaan. Eerst de slechte hits er uit filteren en daarna de 10 beste hits opslaan.

Welke resultaten ga je opslaan?

Alleen resultaten die een E-value hebben hoger dan 1\*10-10,

een identity hoger dan 80% (eiwitten hoger dan 25%),

positives hoger dan 75%,

coverage hoger dan 80%

Stel een ERD op voor de te ontwikkelen database.

Uit de database moeten per fragment minimaal de volgende gegevens kunnen worden opgehaald:

* De sequenties van Read 1 en Read 2
* Micro-organisme (familie, geslacht, soort)
* Van het eiwit die de meeste overeenkomst heeft met het eiwit gecodeerd door het fragment:
* De mate van overeenkomst met de sequentie data
* Het accessie nummer
* De functie(s) van het eiwit
* Het organisme.

Wat is de bio-informatica hoofdonderzoeksvraag van jouw onderzoek.

* **Welke micro-organismen zitten erin de compost?**
* **Welke eiwitten kun je identificeren aan de hand van de dataset?**

Welke subvragen horen daar bij?

* **Wat zijn de functies/ genen van de gevonden eiwitten?**
* **Welke domeinen bevatten de eiwitten? (active site, transmembaandomeinen)**
* **Wat is de locatie van de eiwitten in de cel? (mitochondrien, ribosomen)**
* **Topologie opbouw van een eiwit (uniprot 3d plaatje helixen, alfa en beta sheets)**
* **Locatie organismen? Waar komt het organismen voor?**
* **Eigenschappen organisme?**

Let op dat je alle vragen kunt beantwoorden met het onderzoek dat jij gaat uitvoeren?

Welke micro-organismen zitten er in compst?

Wat zijn de eiwitten die de micro-organismen produceren?

blastx , organisme en eiwit

blosum62, standaardmatrix

cut off op evalue: 1\*10^-10

10 hits aanhouden

**Welke eiwitten zitten er in de micro-or**

**Wat zijn de functies van de gevonden eiwitten**

**Soorten en of geslachten**

**Welke qualityscore is een goede grens**