

# Многообразие генов ОПИН-синтаз

Выполнили: Вервайн Е. Д., Волгин Д. А., Дускаева А. Г., Епифановская П. М.,  
Исмагилова Я. И., Сафонова Е. Е., Удалова К. Р., Чубаров Д. Д.  
Руководитель: Журбенко П.М.

# Что мы вообще делали всю неделю?

Искали гомологи генов синтеза опинов в  
базах данных



Строили филогенетические деревья



Искали случаи горизонтального  
переноса генов



# Границы исследования

Существует несколько известных генов синтеза опинов. Мы взяли некоторые из них:

- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>• <u>octopine synthase</u></li><li>• <u>vitopine synthase</u></li><li>• <u>nopaline synthase</u></li><li>• <u>cucumopine synthase</u></li><li>• <u>mikimopine synthase</u></li></ul> | <ul style="list-style-type: none"><li>• <u>succinamopine synthase</u></li><li>• <u>mannopine synthase</u></li><li>• <u>agropine synthase</u></li><li>• <u>chrysopine synthase</u></li><li>• <u>agrocinopine synthase</u></li></ul> |
|--|--|

- ❖ В процессе работы, мы увидели, что существует **множество гомологичных** генов с неизвестной функцией
- ❖ Оказалось сложно определить границы генов “синтеза опинов”
- ❖ Решили изучать **более крупно** и брать в работу все гомологичные белки

# Как работает psiBLAST?

- Основан на технике профилей
- Последовательно корректирует профиль для выравнивания новых найденных сиквенсов на базу данных
- Первая итерация – обычный белок-белковый бласт, далее – выравнивание по профилю
- Поиск считается сошедшимся, если после очередной итерации список лучших находок не изменился

профиль	1	2	3	4
AK1	0.5	0	1	0.6
AK2	0.25	0.3	0	0.23
AK3	0.25	0.7	0	0.17

вес AK

# Кто этот ваш HMMER?

- Алгоритм для выравнивания последовательностей (очевидно)
- Также основан на построении профилей, учитывает наиболее частые сайты смещения последовательности, использует только текущий профиль для получения результата
- Не способен выполнить несколько итераций



# Что это вообще за звери?

Это бласт

Standard Protein BLAST

BLASTP programs search protein databases using a protein sequence.

blastn blastp blastx tblastn tblastx

Enter Query Sequence

protein-blast

Enter accession number(s), gi(s), or FASTA sequence(s)

Query subrange  From  To

Or, upload file  Файл не выбран.

Job Title   
Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Databases  Standard databases (nr etc.)  Experimental databases [New](#) [For more info see What is clustered nr?](#)

Compare  Select to compare standard and experimental database

Standard

Database Non-redundant protein sequences (nr)

Organism Optional Enter organism name or id—completions will be suggested   exclude

Exclude Optional  Models (XM/XP)  Non-redundant RefSeq proteins (WP)  Uncultured/environmental sample sequences

Program Selection

Algorithm  Quick BLASTP (Accelerated protein-protein BLAST)  
 blastp (protein-protein BLAST)  
 PSI-BLAST (Position-Specific Iterated BLAST)  
 PHI-BLAST (Pattern Hit Initiated BLAST)  
 DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm

psi-алгоритм

# HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

Service Update: Improvements Underway for Enhanced Reliability

We understand that our service has experienced issues in recent months, and we sincerely apologize for any inconvenience caused. We want to assure you that we are actively working to address these challenges and enhance the overall reliability of our service.

Please bear with us as we make significant improvements to minimize disruptions. Your patience and understanding are greatly appreciated during this time.

## А это HMMER

Quick search

Paste in your sequence or use the example

Enter your sequence

Reference Proteomes  UniProtKB  SwissProt  Pfam

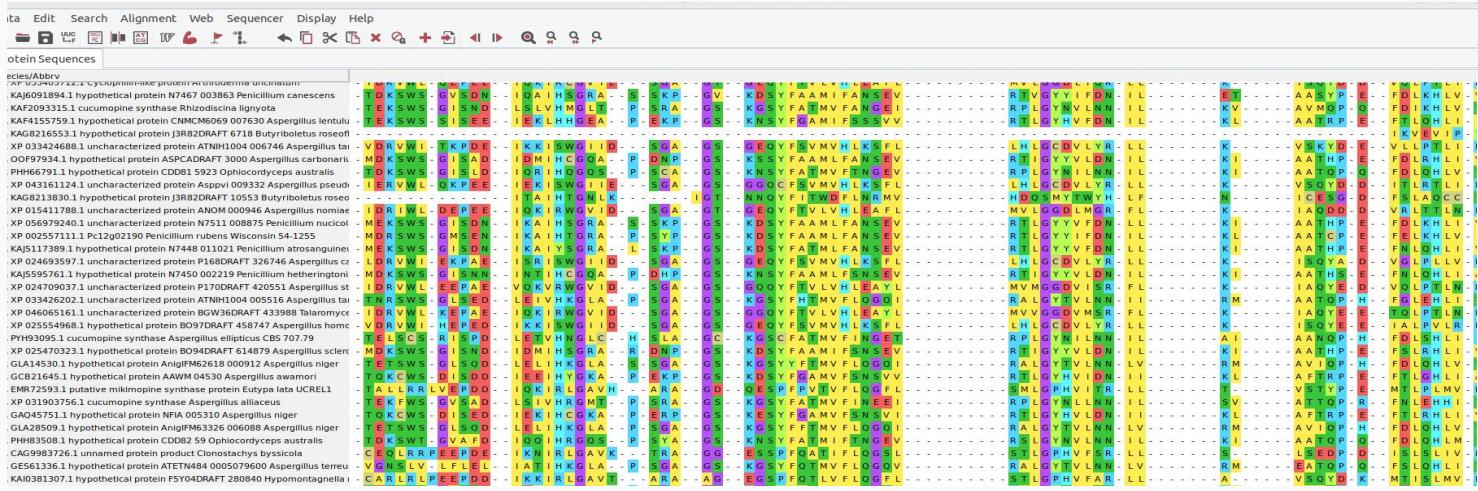
Quickstart tutorial

Online documentation

И с помощью них мы искали белковые последовательности

# Выравнивание последовательностей

- Использовали **MAFFT** (mafft --localpair --maxiterate 1000)
  - Визуализация в программе MEGA

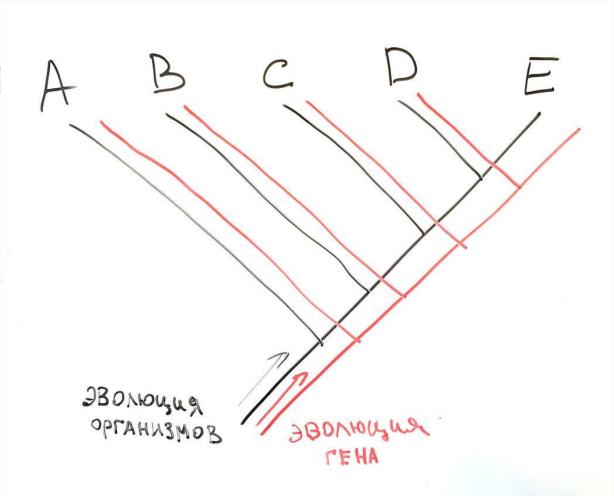


# Построение филогении

У нас было много последовательностей, поэтому мы выбрали **быстрый алгоритм** построения дерева - **neighbor joining** в программе FastTree.

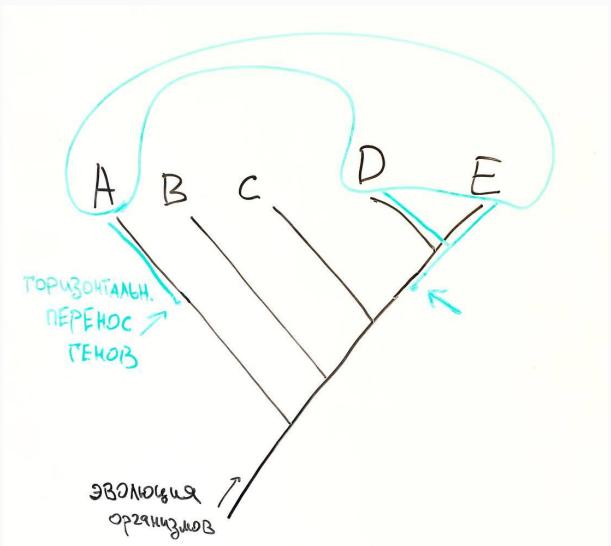
Визуализация деревьев проводилась на сайте iTOL.

# По каким критериям определяли горизонтальный перенос



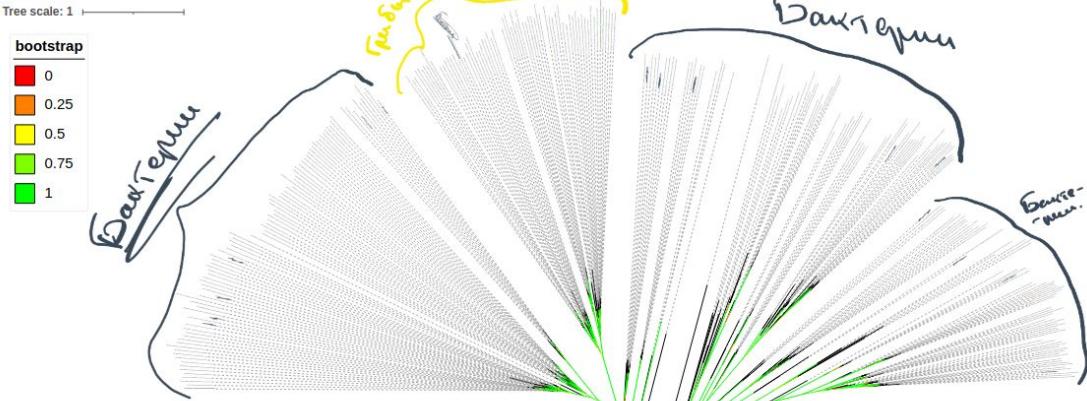
Вертикальный перенос.

Эволюция гена соответствует эволюции организмов

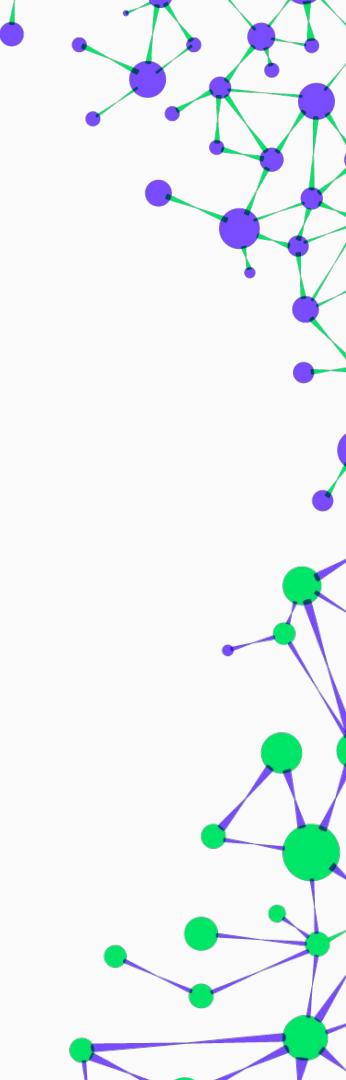
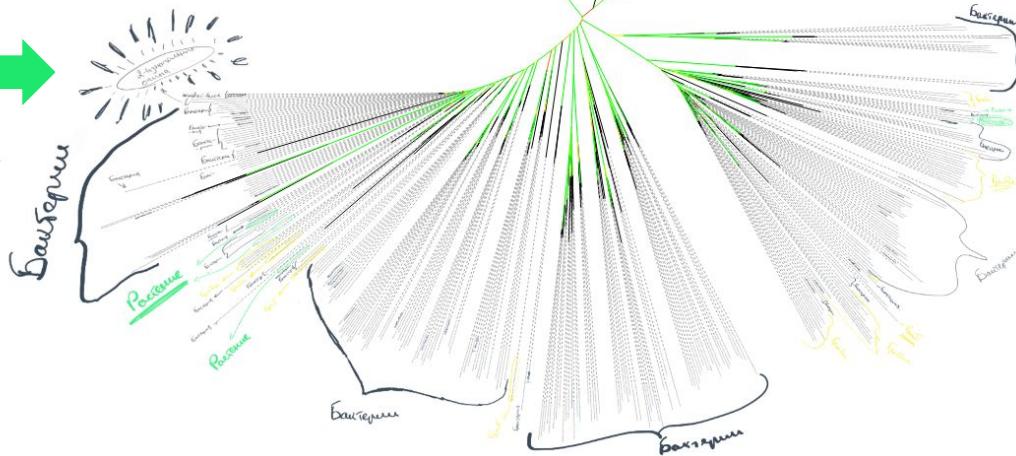


Множественный горизонтальный перенос гена. Отдаленные таксоны А и D/E окажутся близки на дереве, построенном по этому гену

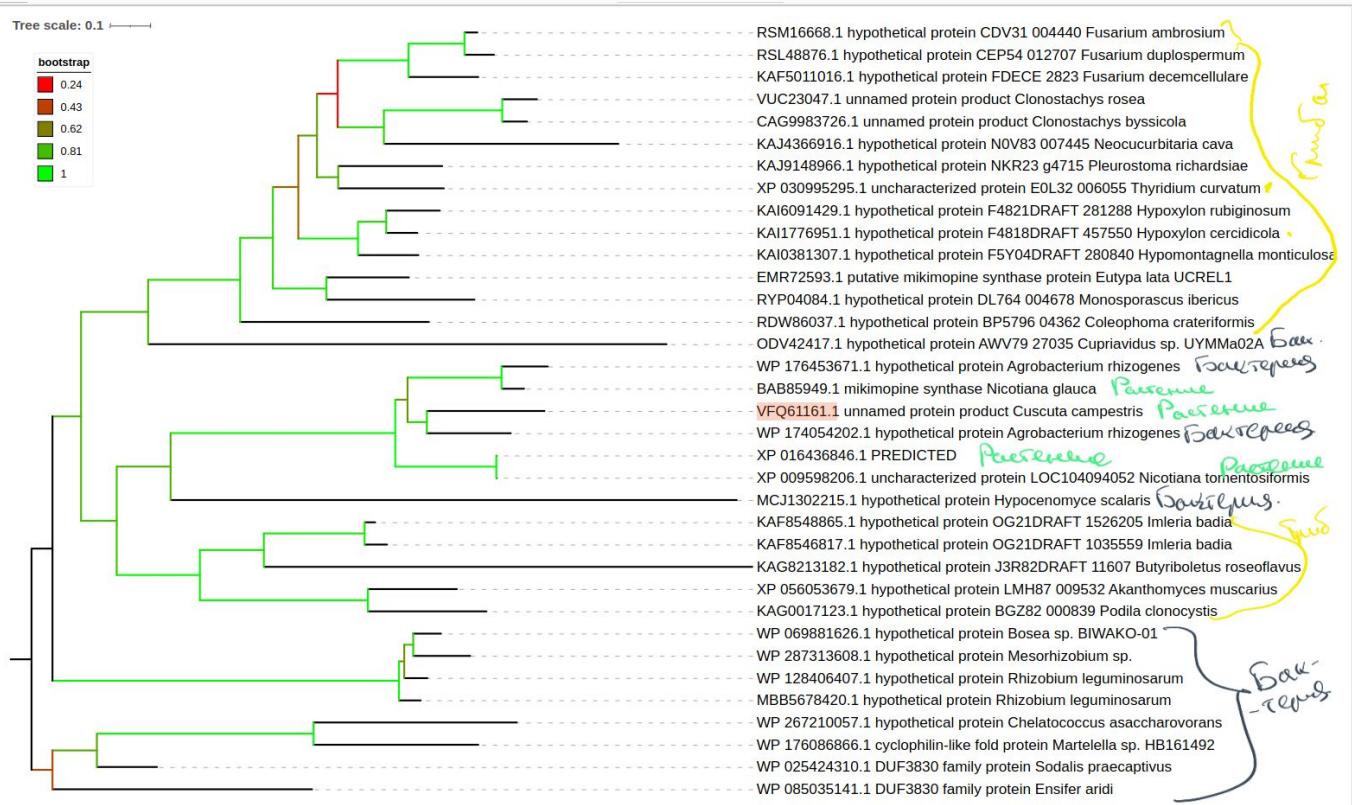
# Кукумопин (кукумопин-синтаза, Cus)



## Изначальный *cus*

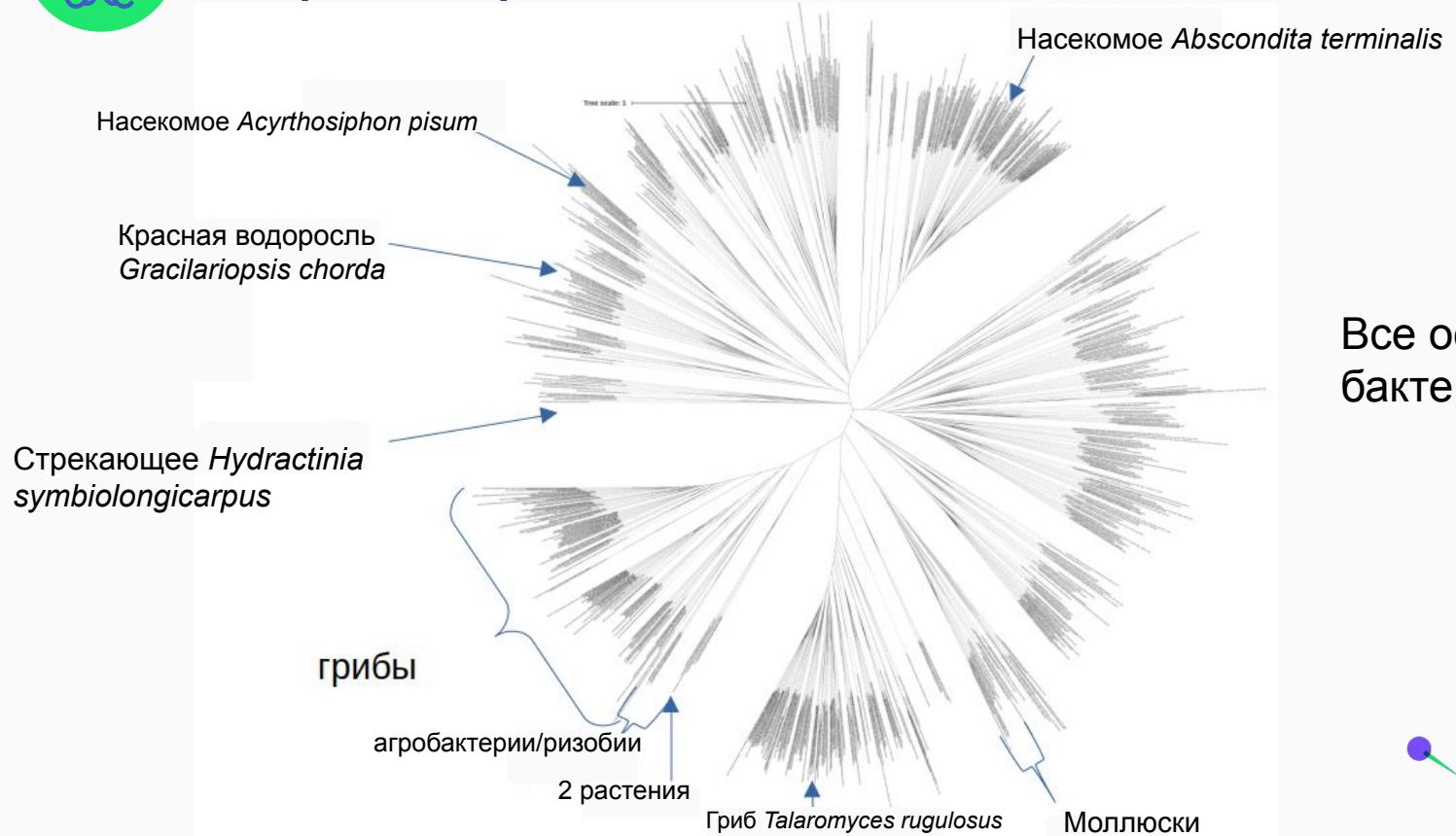


## Кукумопин (кукумопин-синтаза, Cus)

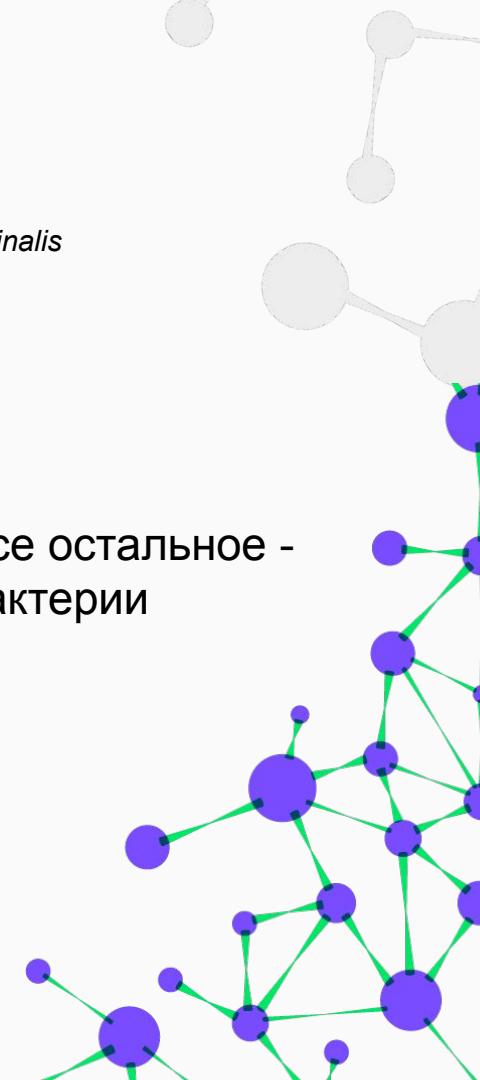




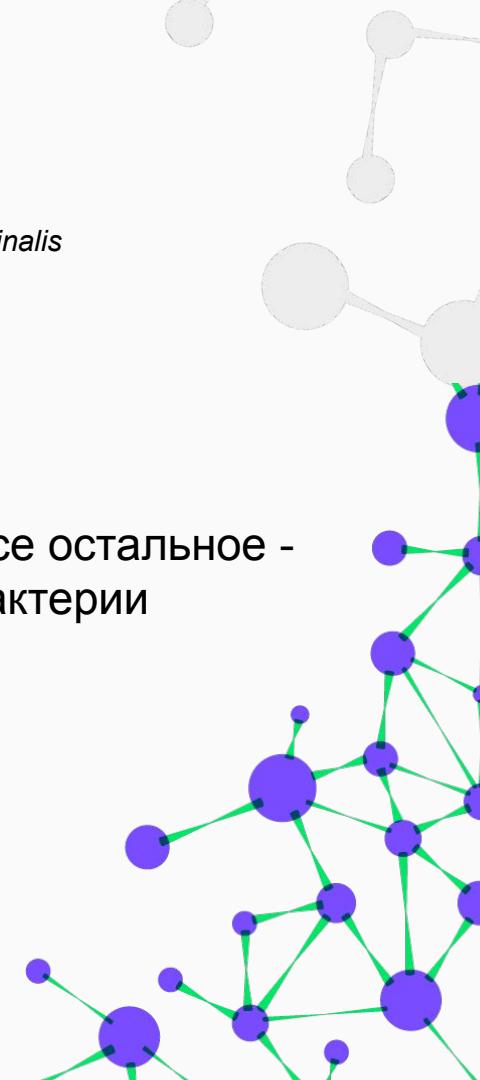
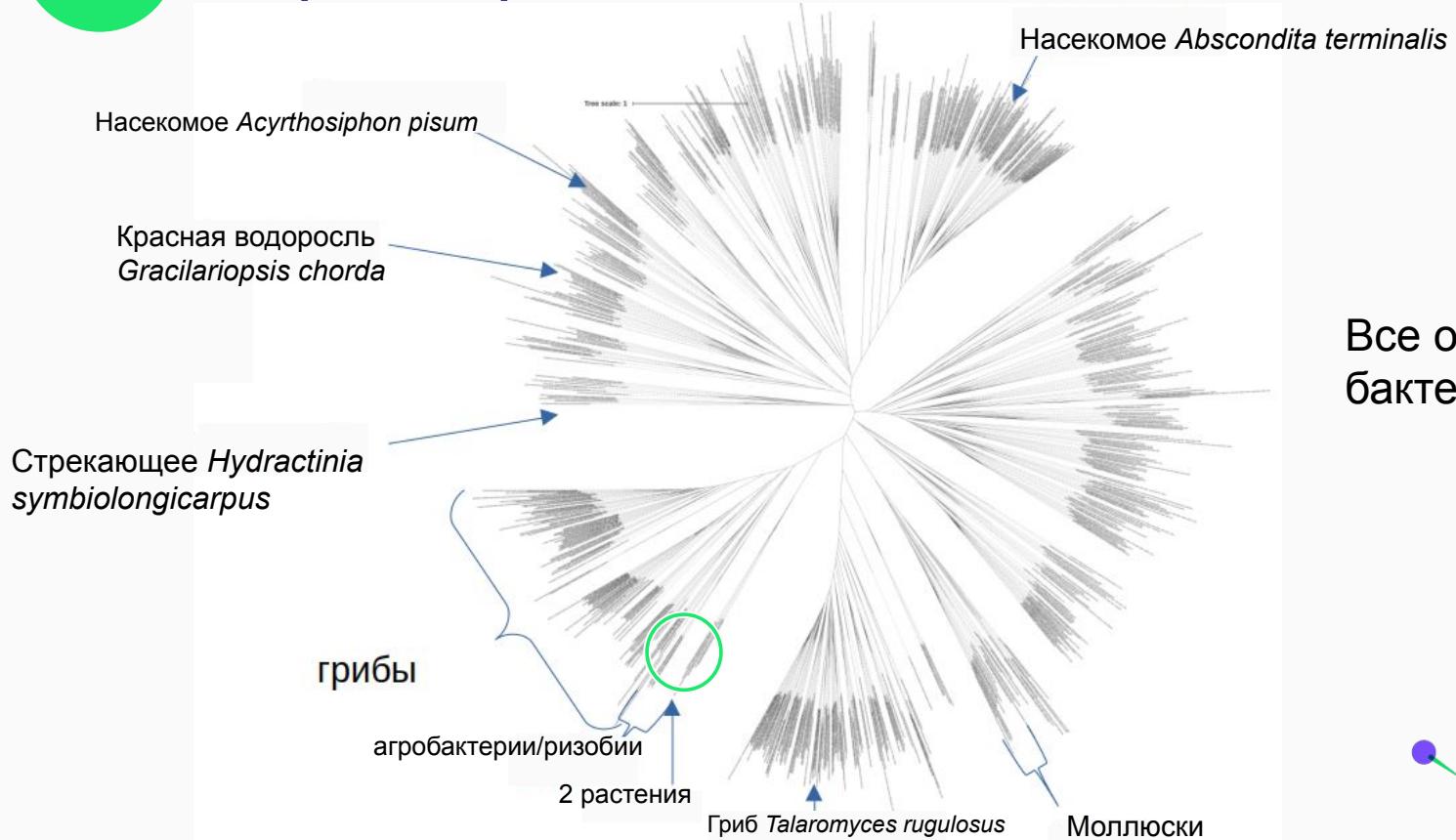
# Octopine synthase



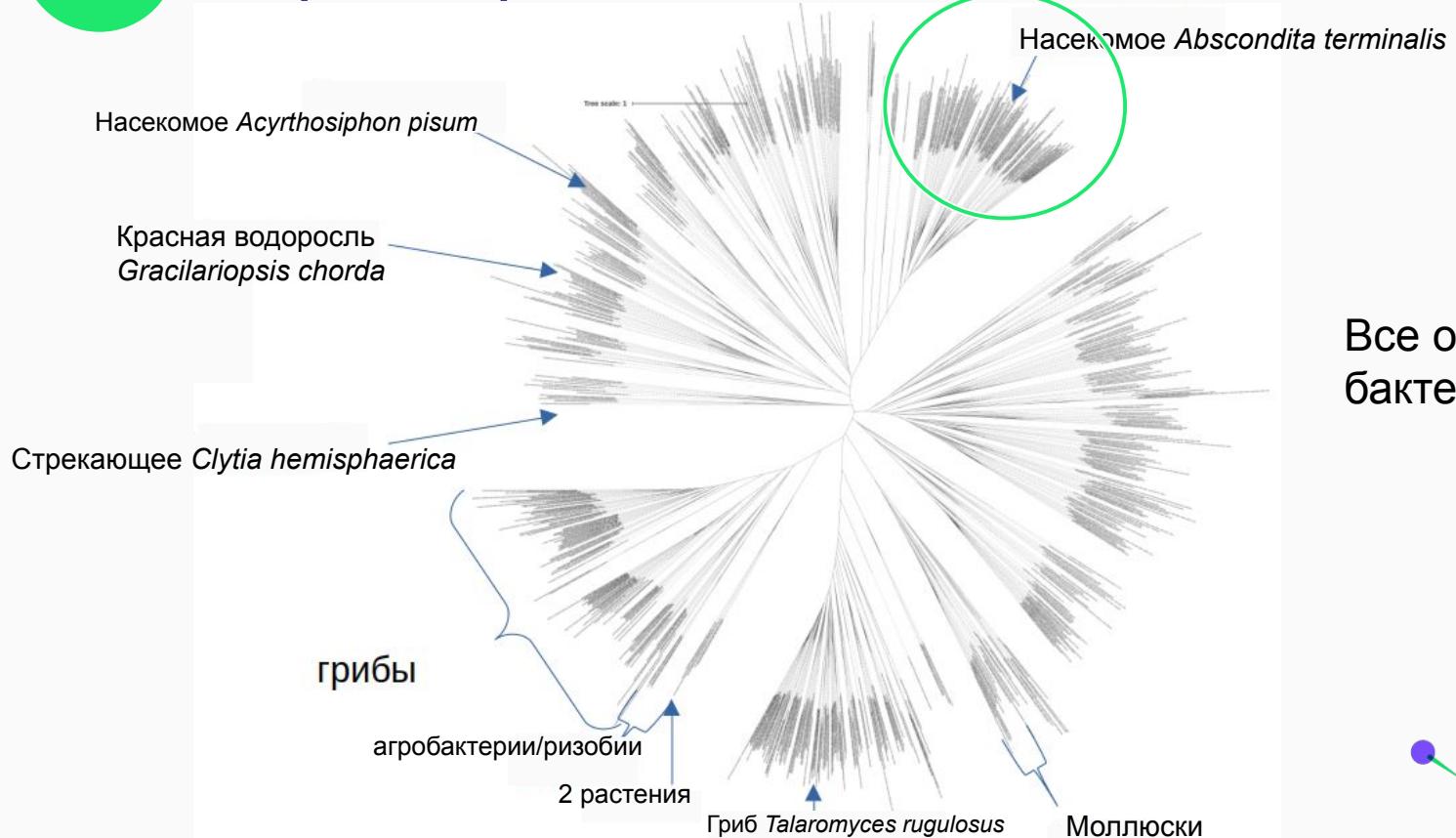
Все остальное -  
бактерии



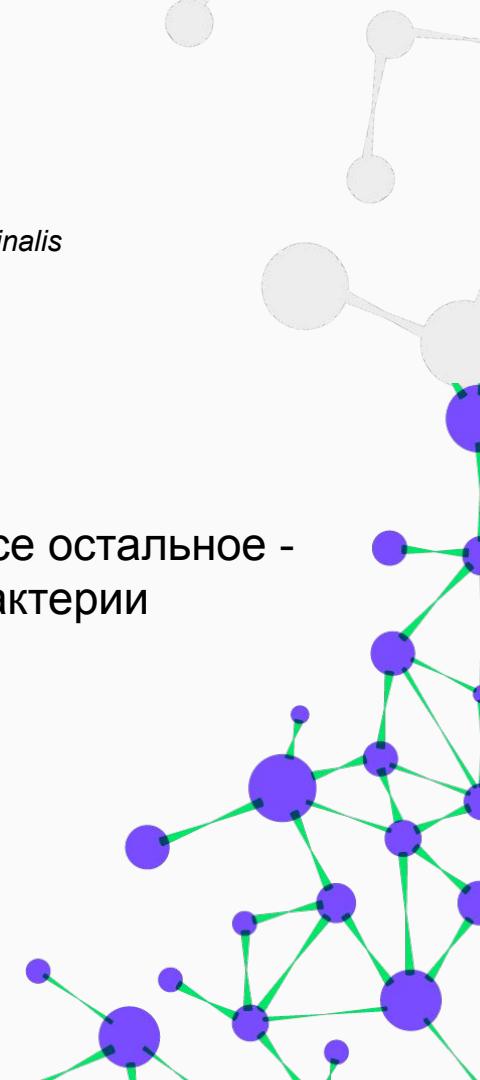
# Octopine synthase



# Octopine synthase

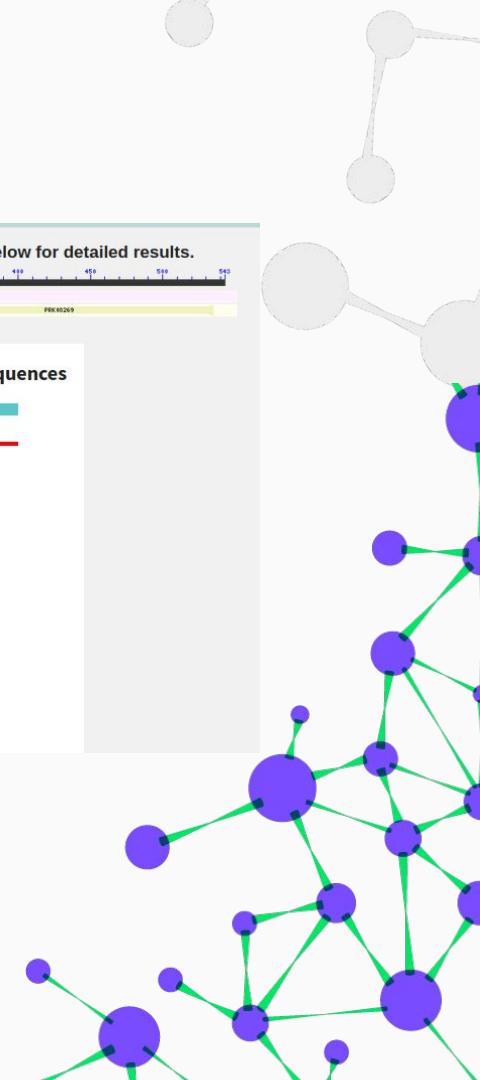
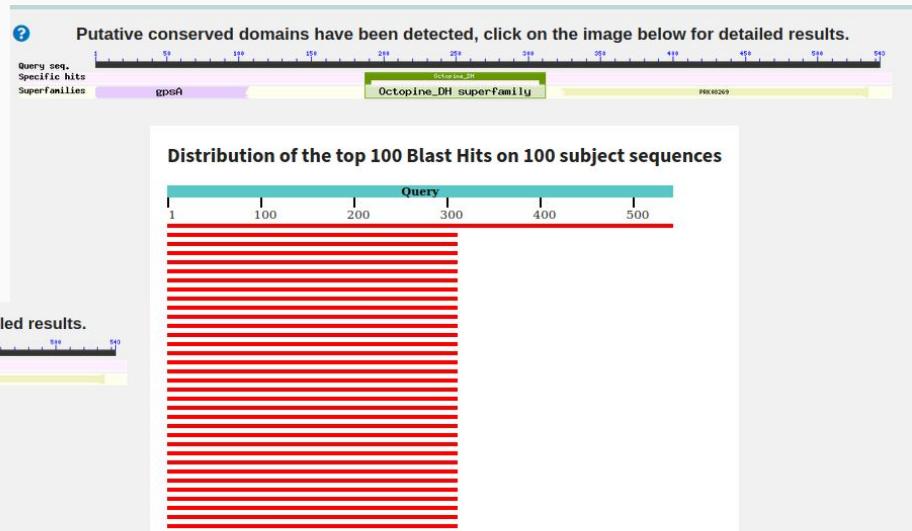
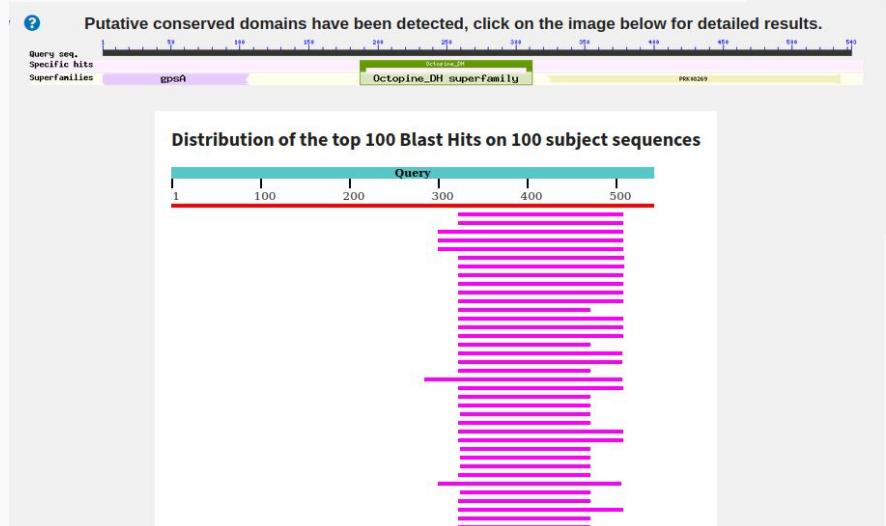


Все остальное -  
бактерии

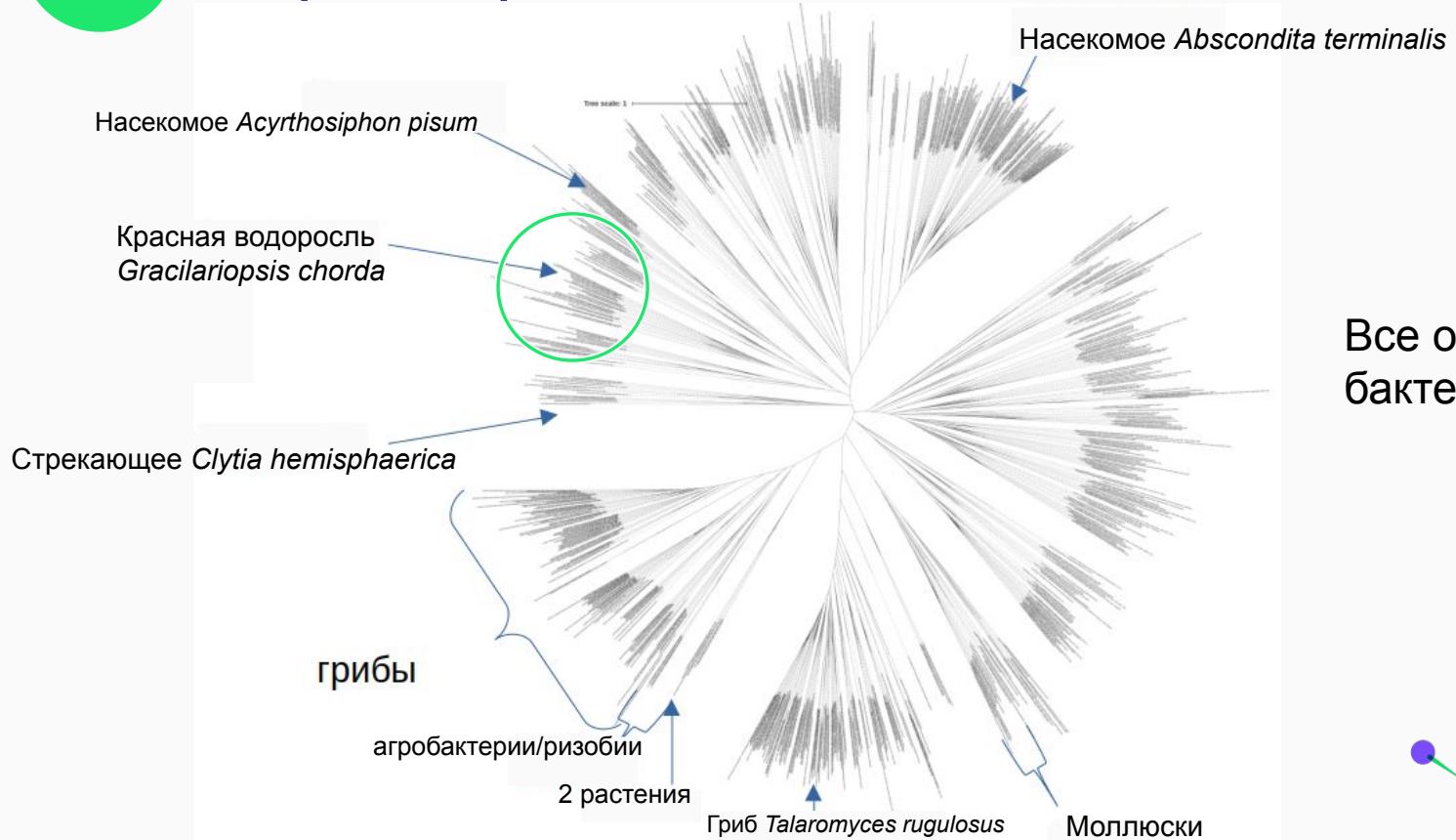




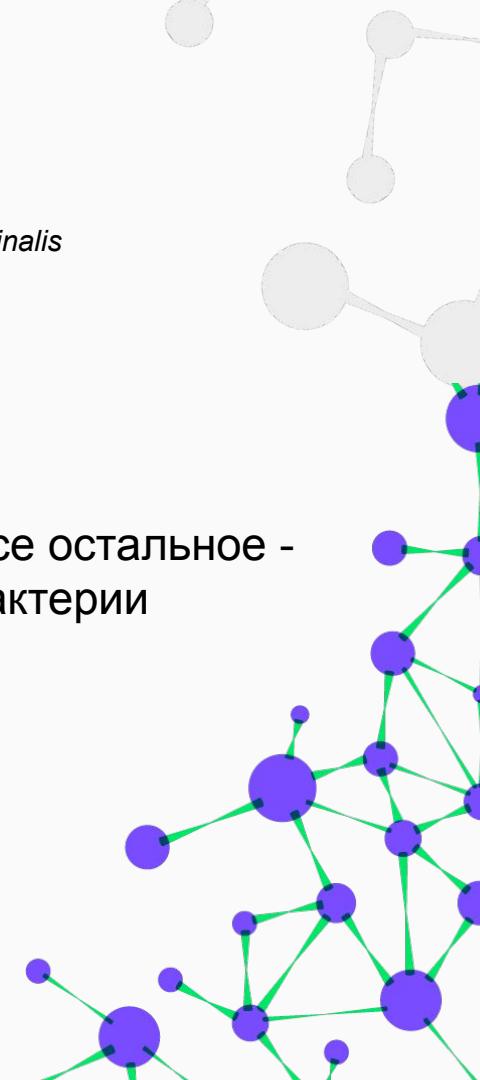
# Abscondita terminalis



# Octopine synthase



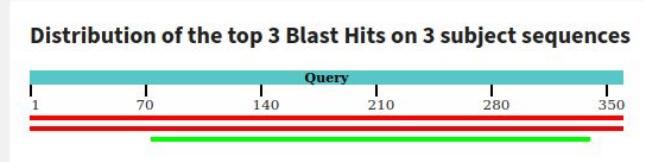
Все остальное -  
бактерии





# *Gracilaria*opsis *chorda*

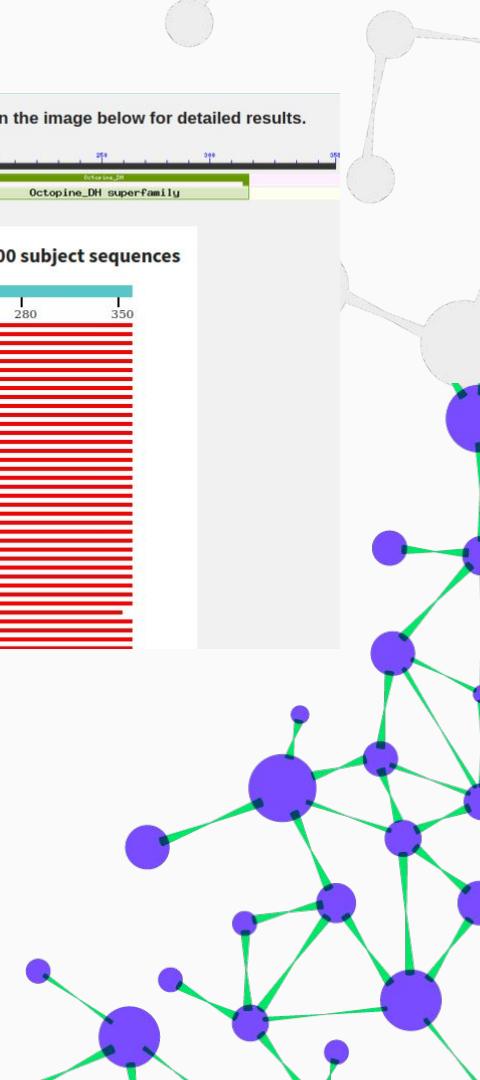
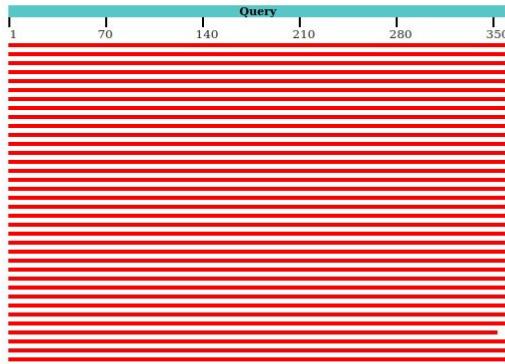
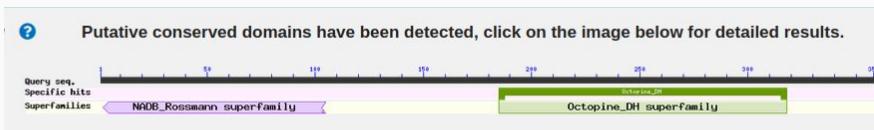
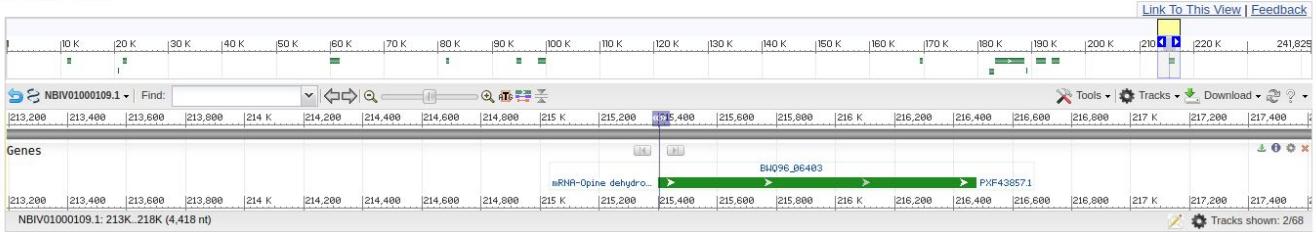
**Putative conserved domains have been detected, click on the image below for detailed results.**



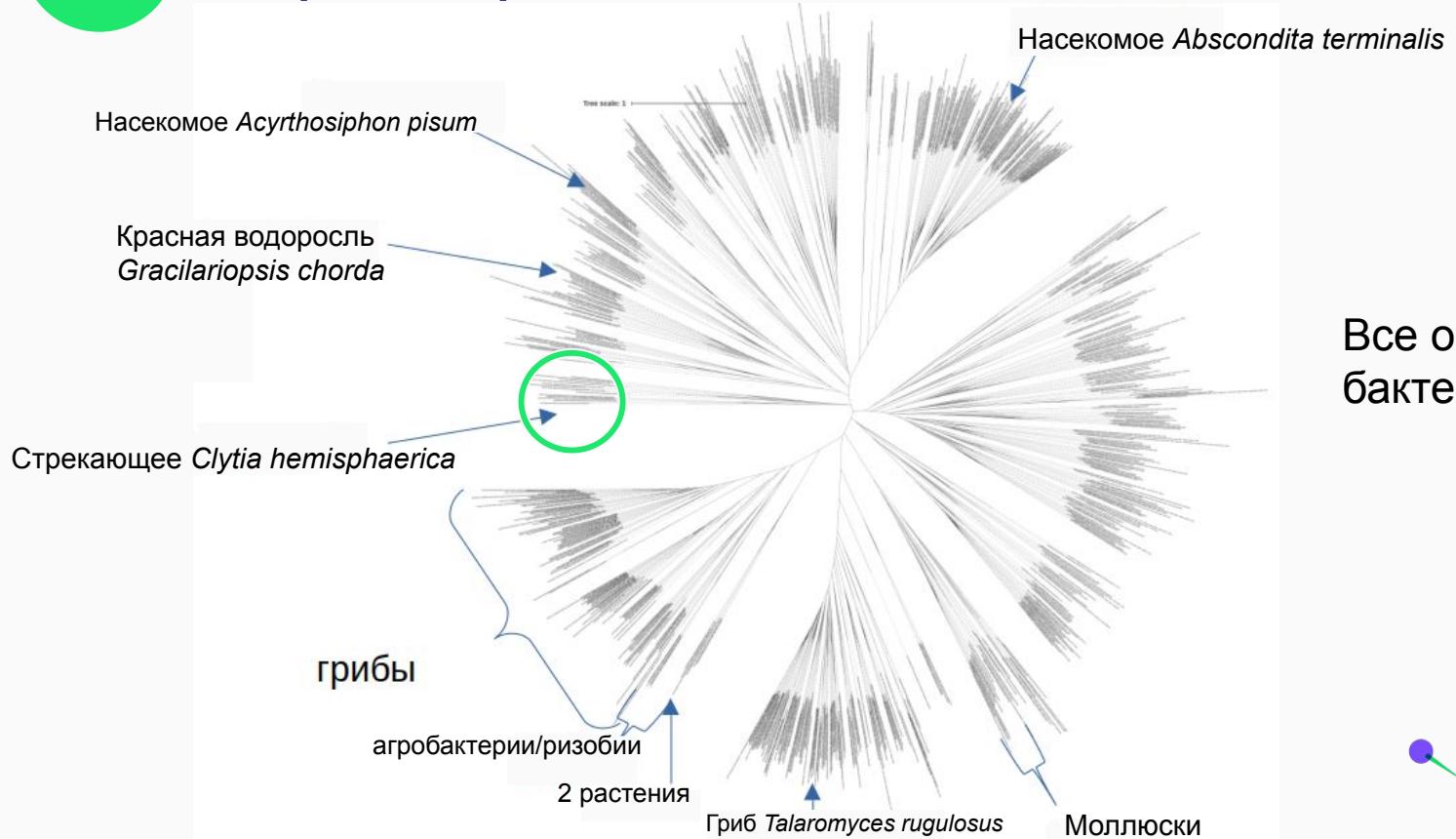
**Gracilariopsis chorda isolate SKKU-2015 GRC0111, whole genome shotgun sequence**

GenBank: NBIV01000109.1

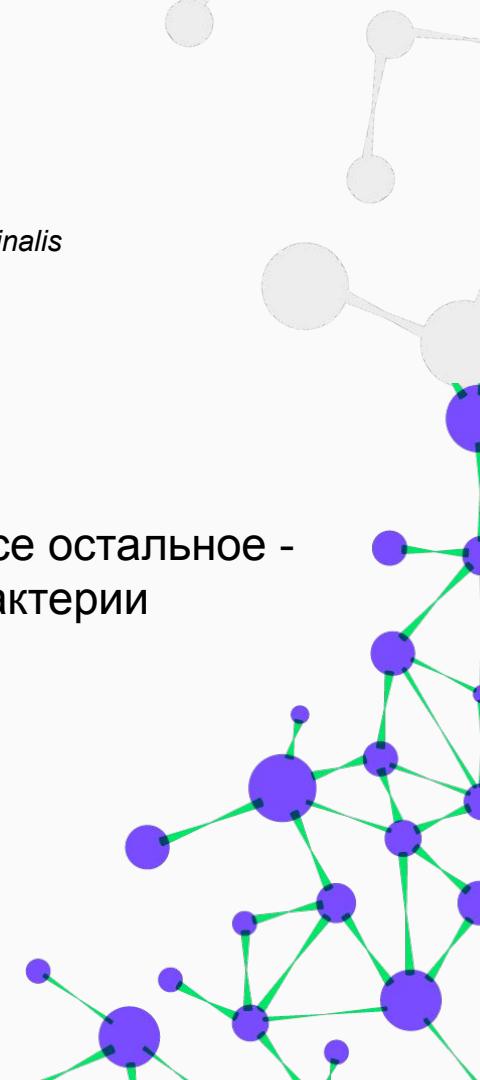
## GenBank FASTA

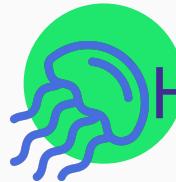


# Octopine synthase



Все остальное -  
бактерии



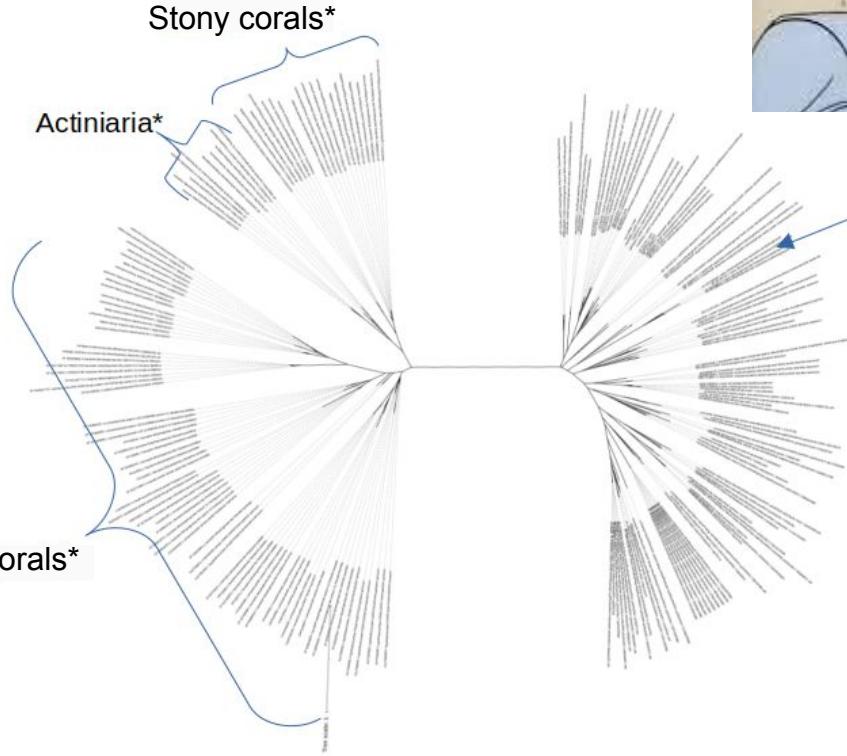


# Hydractinia symbiolongicarpus

## Стрекающие класс *Hexacorallia*

## Stony corals\*

\* порядок



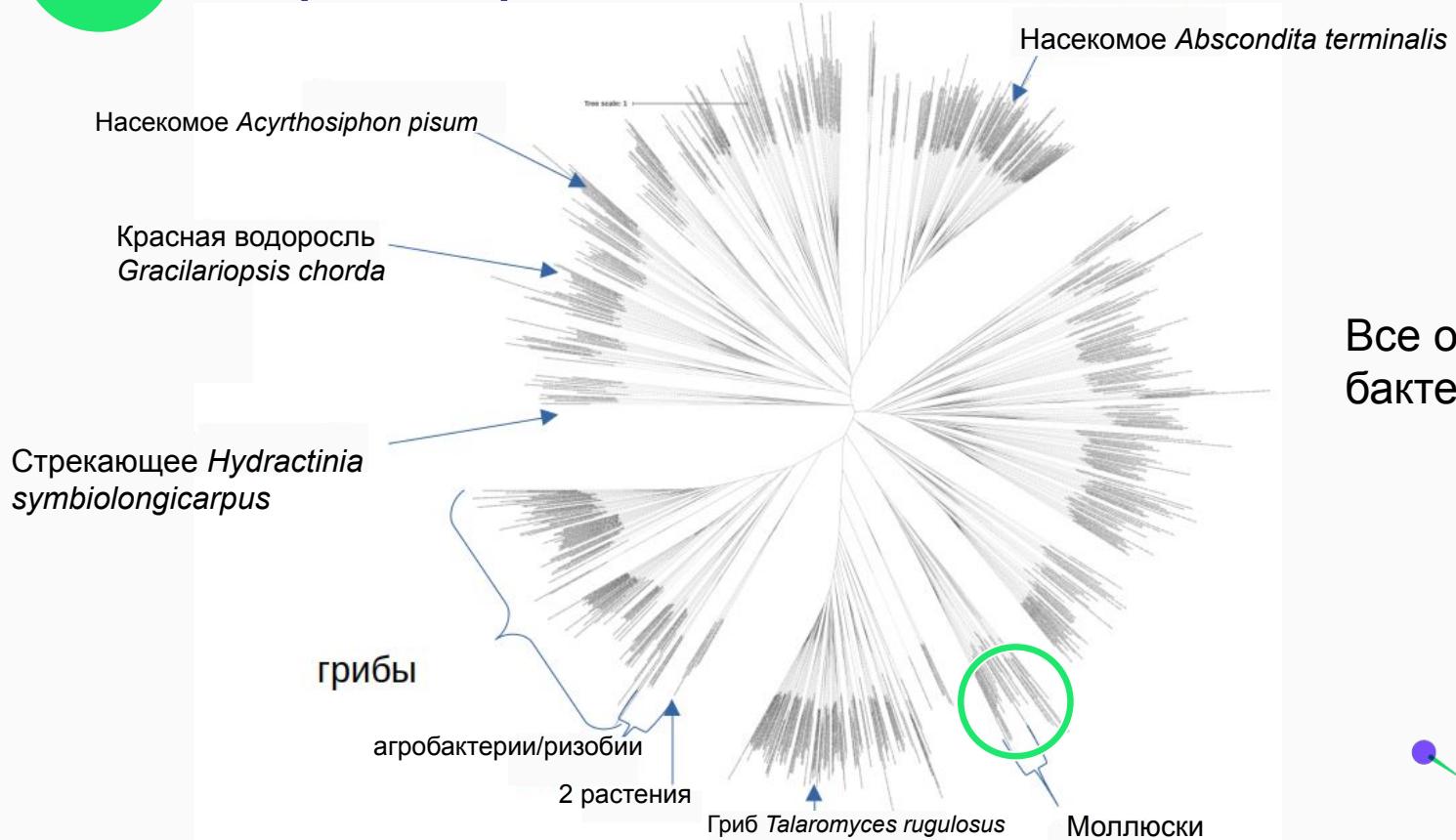
## Это бактерия?

## *Hydractinia symbiolongicarpus*

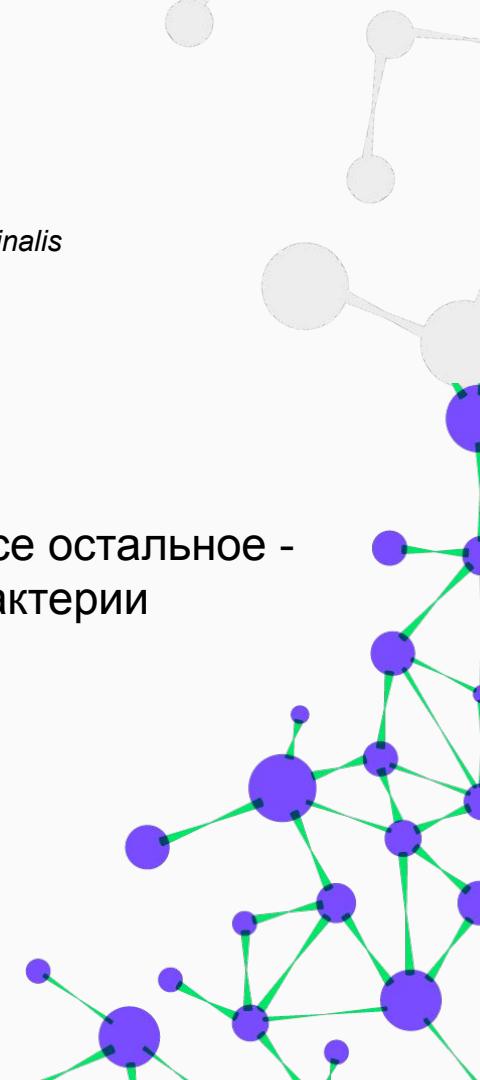
## Бактерии



# Octopine synthase

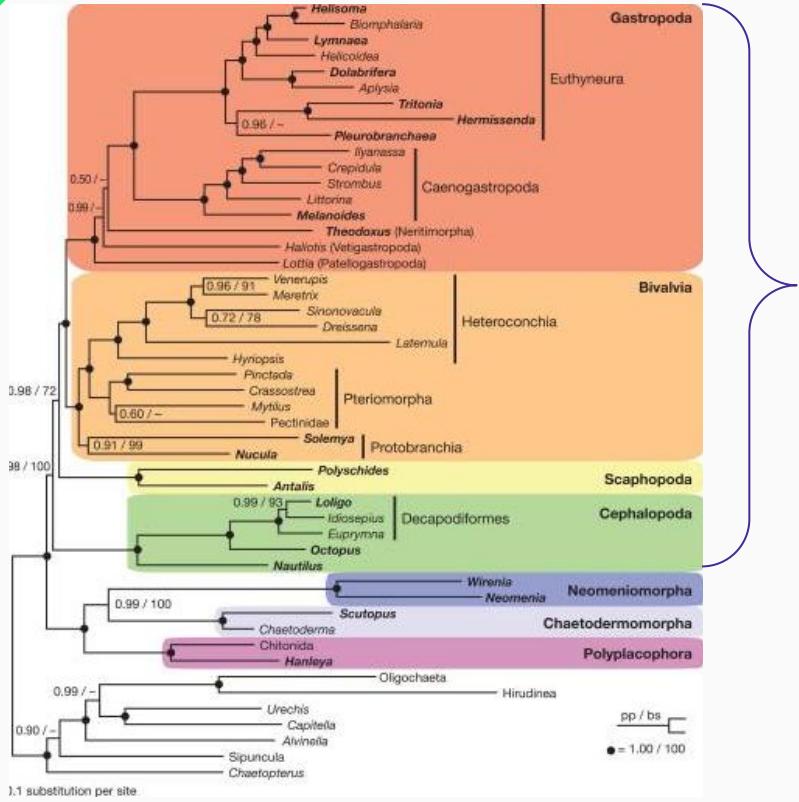


Все остальное -  
бактерии

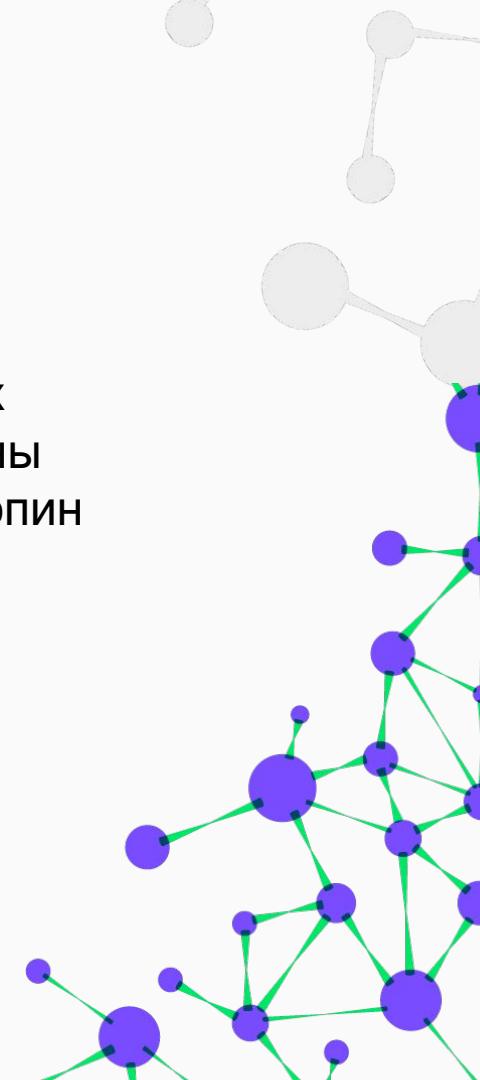




# Моллюски



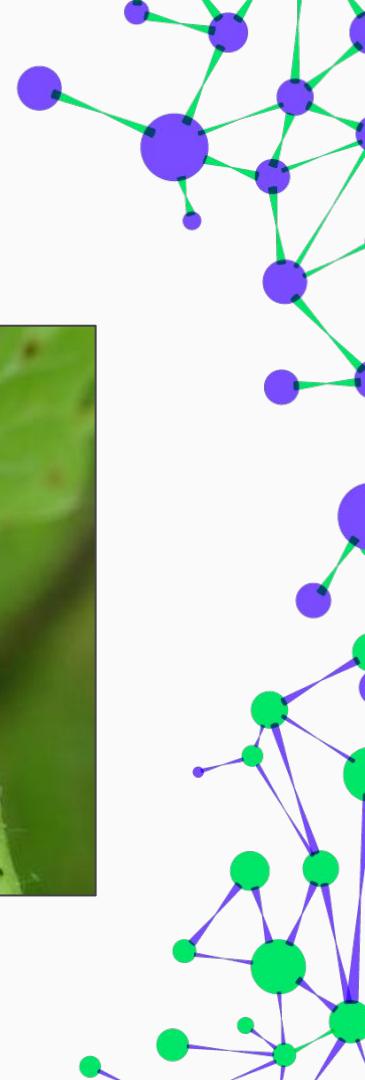
В этих группах  
встречаются гены  
родственные октопин  
синтазе





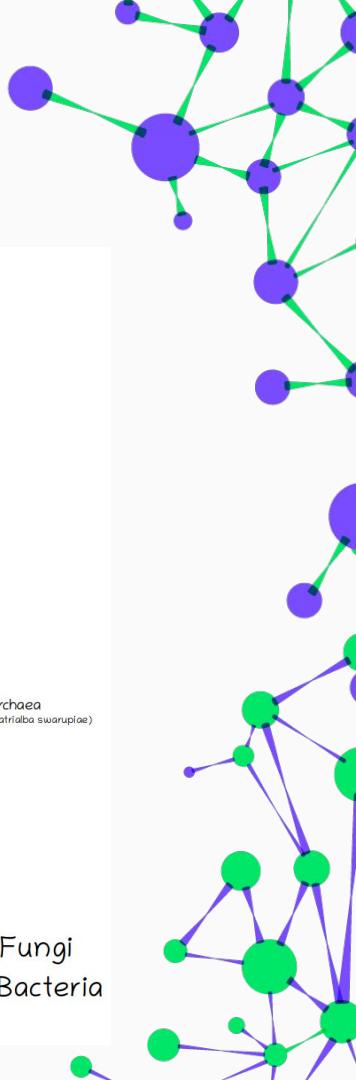
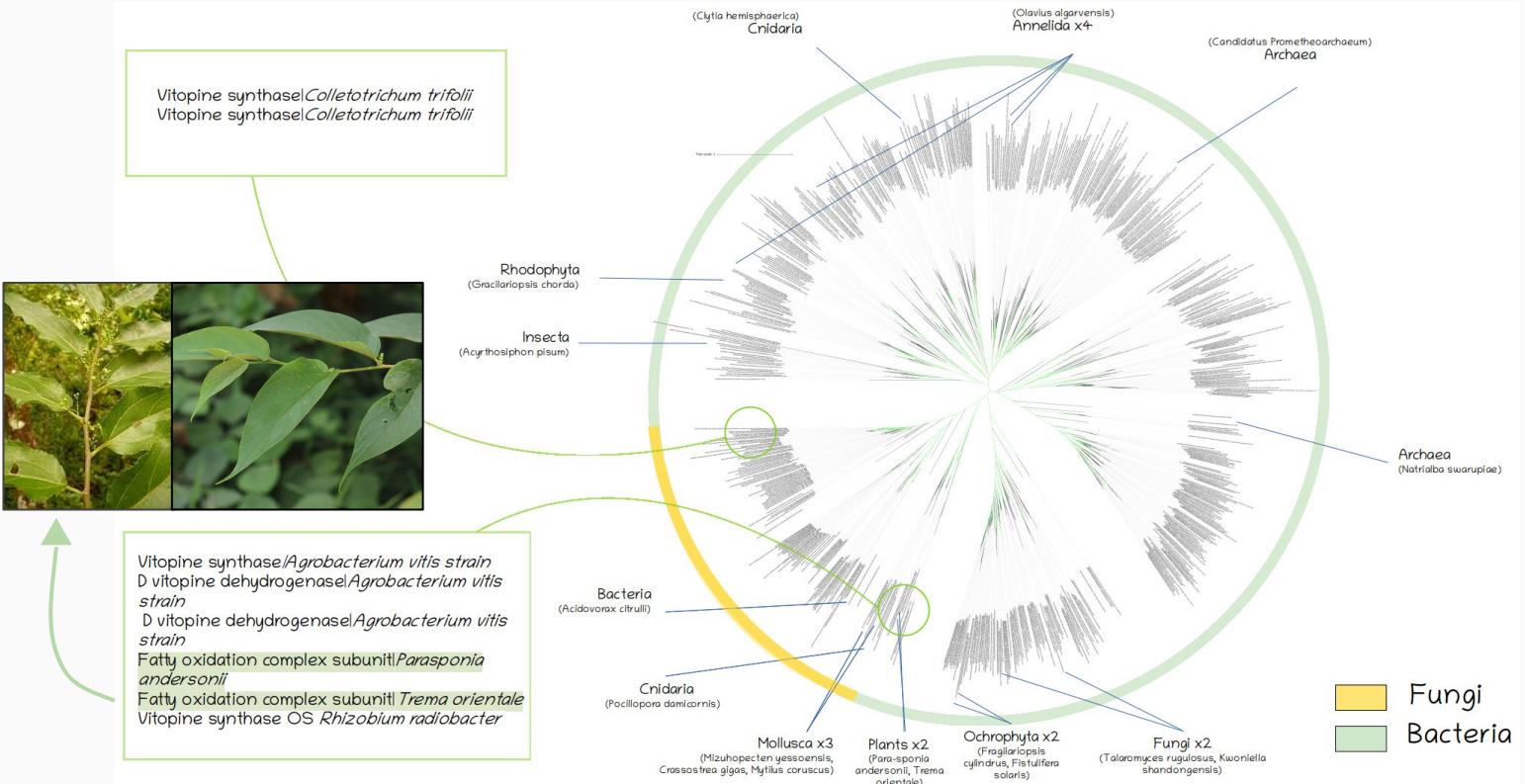
# Vitopine synthase (vis)

*Allorhizobium vitis*





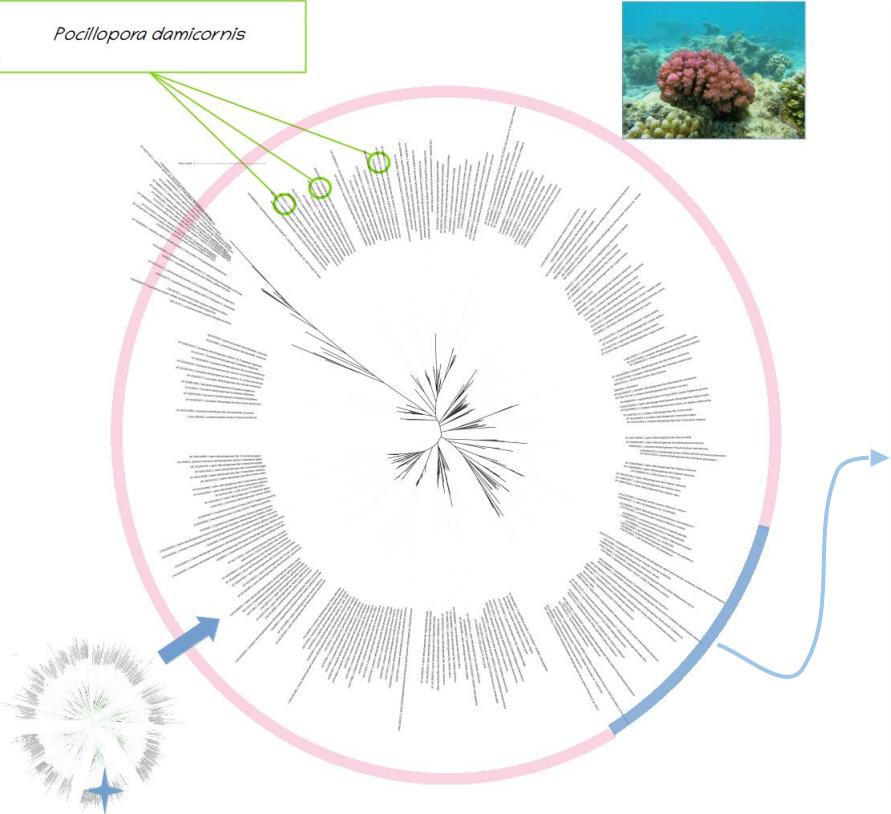
# Vitopine synthase (vis)





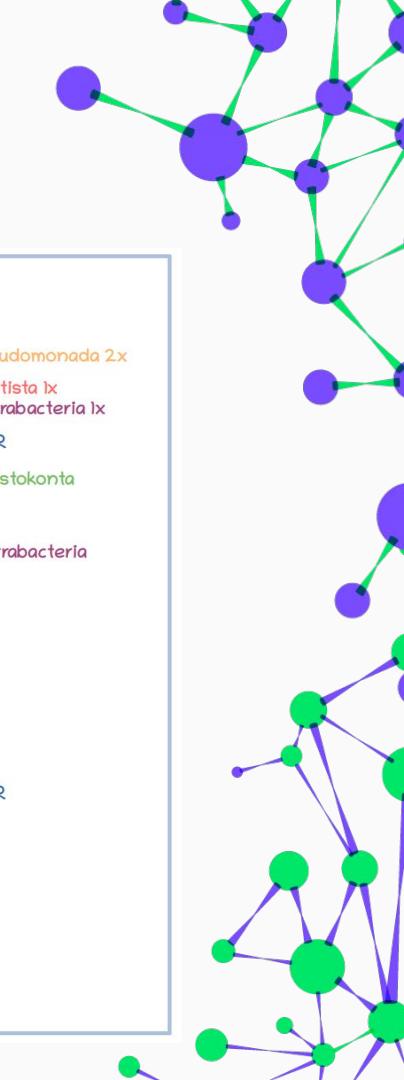
# Vitopine synthase (vis)

*Pocillopora damicornis*

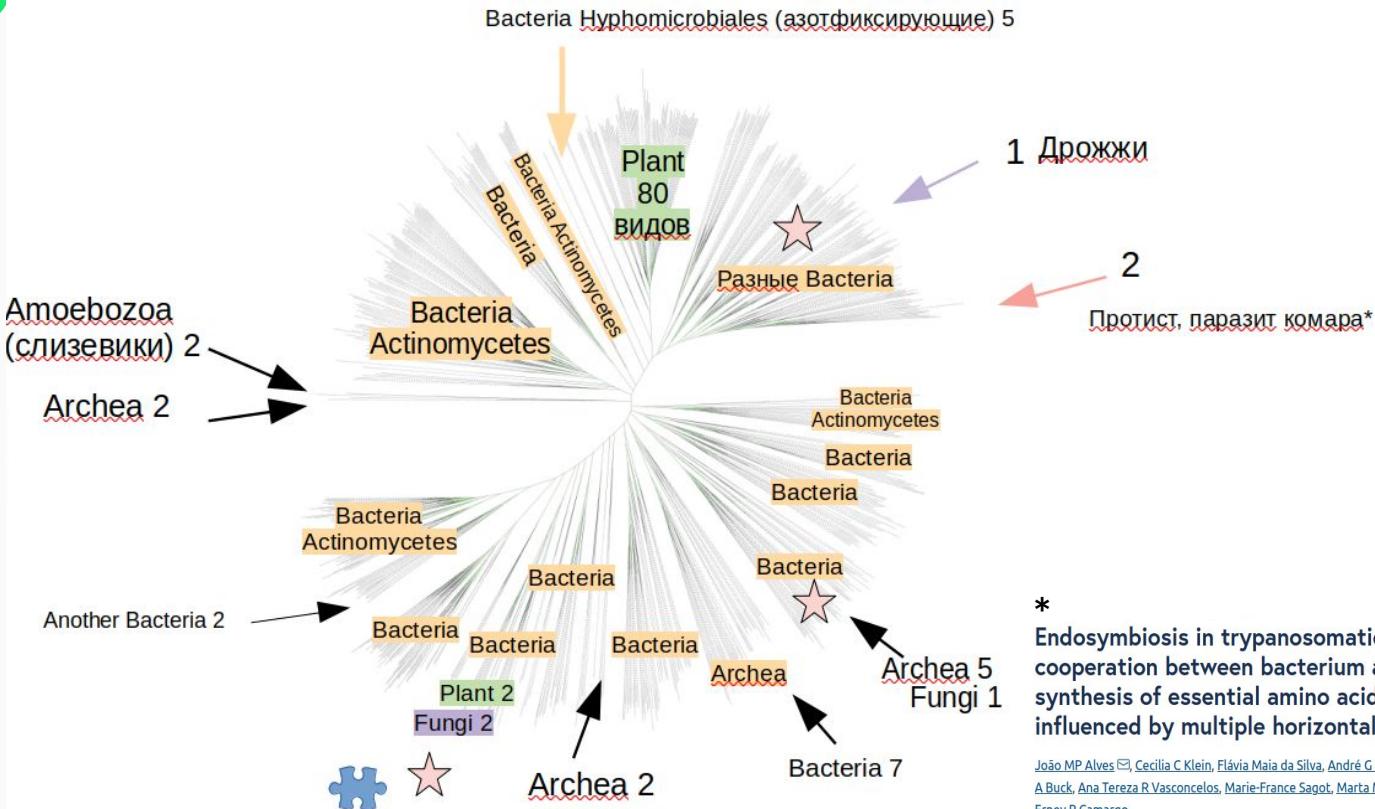


Porticoccus sp.  
Porticoccaceae bacterium  
Pavlovales sp.  
Anaerolineaceae bacterium  
Pythium insidiosum  
Pythium insidiosum  
Salpingoeca rosetta  
Monosiga brevicollis  
Anaerolineales bacterium  
Chloroflexota bacterium  
Chloroflexota bacterium  
Anaerolineales bacterium  
Caldilinea sp.  
Thalassiosira oceanica  
Skeletonema marinol  
Seminavis robusta  
nitzschia multistriata  
Nitzschia inconspicua  
Symbiodinium microadriaticum  
Symbiodinium microadriaticum  
Symbiodinium microadriaticum  
Cladocopium goreau  
Polarella glacialis  
Polarella glacialis  
Effrenium voratum  
Symbiodinium microadriaticum  
Symbiodinium sp.  
Effrenium voratum

Pseudomonada 2x  
Haptista Ix  
Terrabacteria Ix  
SAR  
Opistokonta  
Terrabacteria  
SAR



# The Imposter is SUS



\*

Endosymbiosis in trypanosomatids: the genomic cooperation between bacterium and host in the synthesis of essential amino acids is heavily influenced by multiple horizontal gene transfers

João MP Alves, Cecília C Klein, Flávia Maia da Silva, André G Costa-Martins, Myrna G Serrano, Gregory A Buck, Ana Tereza R Vasconcelos, Marie-France Sagot, Marta MG Teixeira, Maria Cristina M Motta & Erney P Camargo

# Юпитер это не планета

```
z = k
break
j = i[:z]+'_'+i[z:]
try:
    handle = Entrez.efetch(db="protein", id=j, retmode="xml")
    record = Entrez.read(handle)
    handle.close()
    print(str(j), record[0]["GBSeq_taxonomy"])
except Exception:
    print(str(j), "not found")
    c.append(str(j))
print("all not founds:")
for i in c:
```

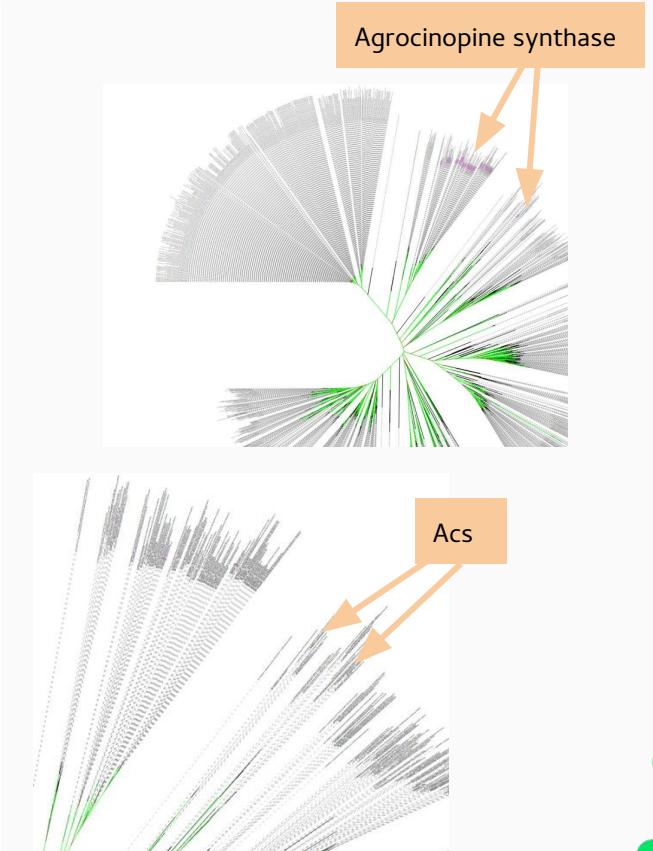
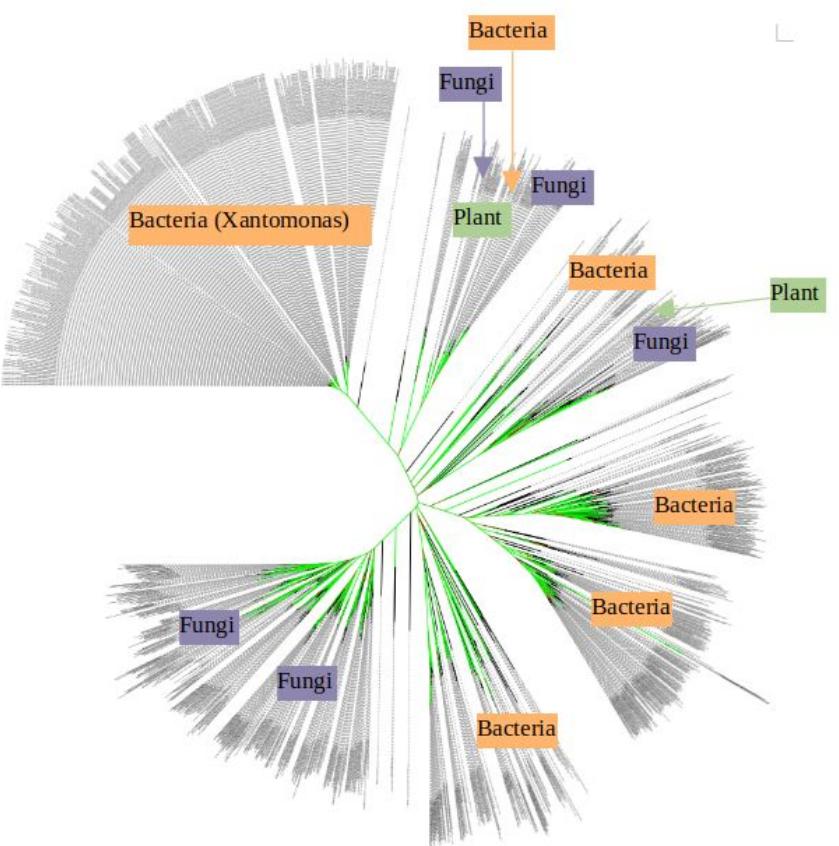
"The best way to learn a language is to speak to natives."

The guy learning python:

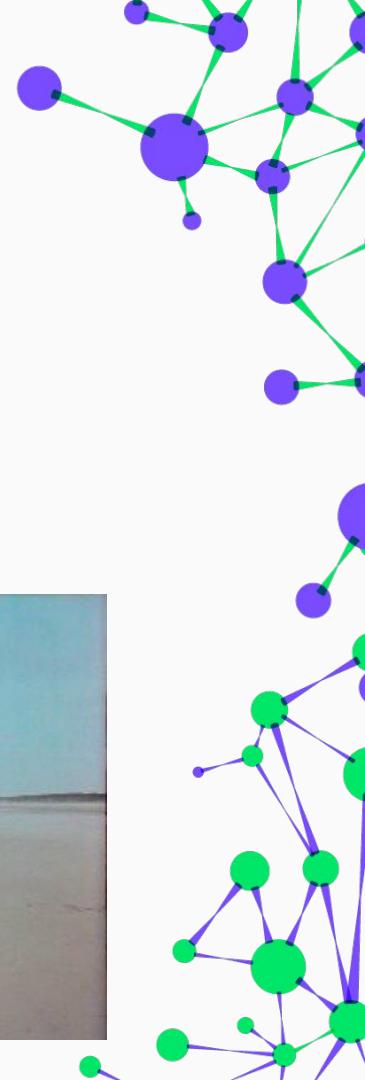
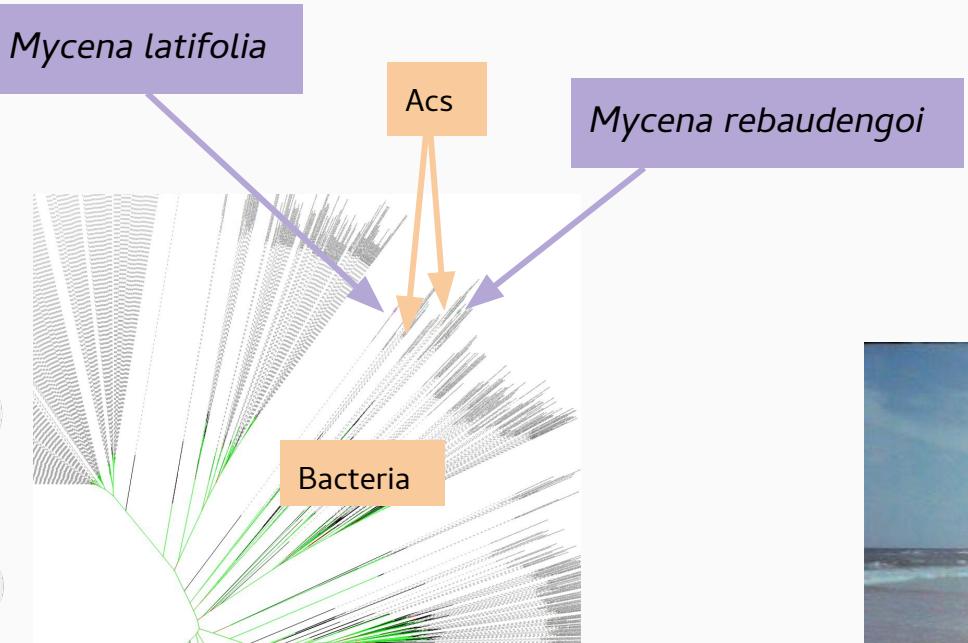


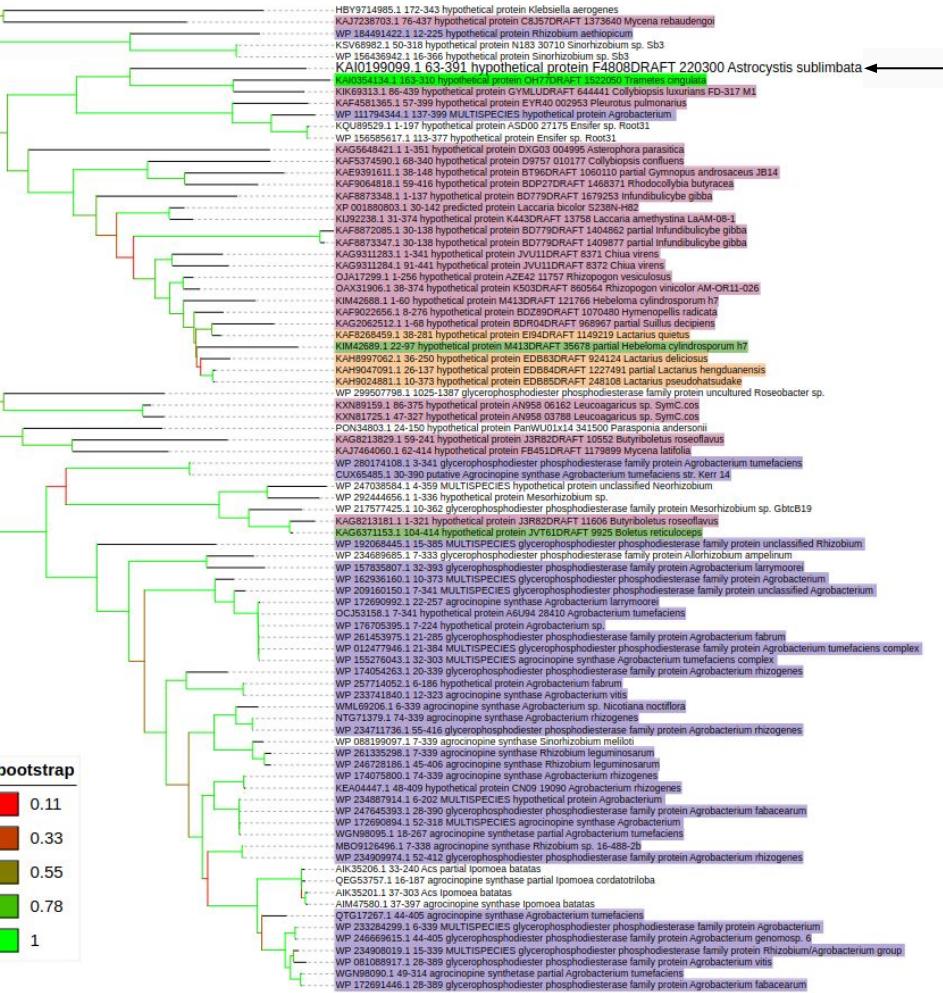
Код на python для автоматической  
таксономической аннотации по номеру  
NCBI

# Agrocinopine synthase (Acs)



# Откуда у грибов ген Acs?





*Ascomycota*

Дерево имеет мозаичную структуру



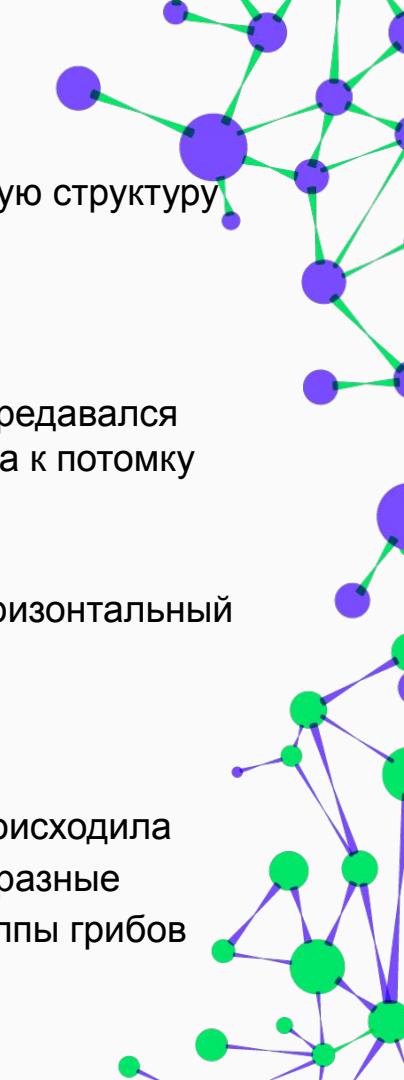
Значит ген Acs не передавался вертикально от предка к потомку



Вероятно, произошел горизонтальный перенос



Вставка гена Acs происходила множество раз в разные таксономические группы грибов



# Интроны у грибов

Ген



C8J57DRAFT\_1373640 [*Mycena rebaudengoii*]



C8J57DRAFT\_1376248 [*Mycena rebaudengoii*]



EDB85DRAFT\_248108 [*Lactarius pseudohatsudake*]



BDZ89DRAFT\_1070480 [*Hymenopellis radicata*]



DXG03\_004995 [*Asterophora parasitica*]

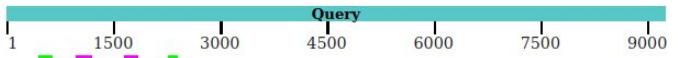
- Экзоны  
 - Интроны

# Место встраивания гена Acs у разных организмов

## *Gymnoporus androsaceus*

Description	Max Score	Total Score	Query Cover
<a href="#">Asterophora parasitica isolate AP01 scaffold_758, whole genome shotgun sequence</a>	103	336	15%

### Distribution of the top 4 Blast Hits on 1 subject sequences



Потому что

## *Chiua viren*

Description	Max Score	Total Score	Query Cover
<a href="#">Asterophora parasitica isolate AP01 scaffold_233, whole genome shotgun sequence</a>	193	193	3%
<a href="#">Asterophora parasitica isolate AP01 scaffold_19, whole genome shotgun sequence</a>	104	104	2%
<a href="#">Asterophora parasitica isolate AP01 scaffold_69, whole genome shotgun sequence</a>	68.9	68.9	3%

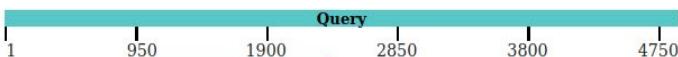
### Distribution of the top 3 Blast Hits on 3 subject sequences

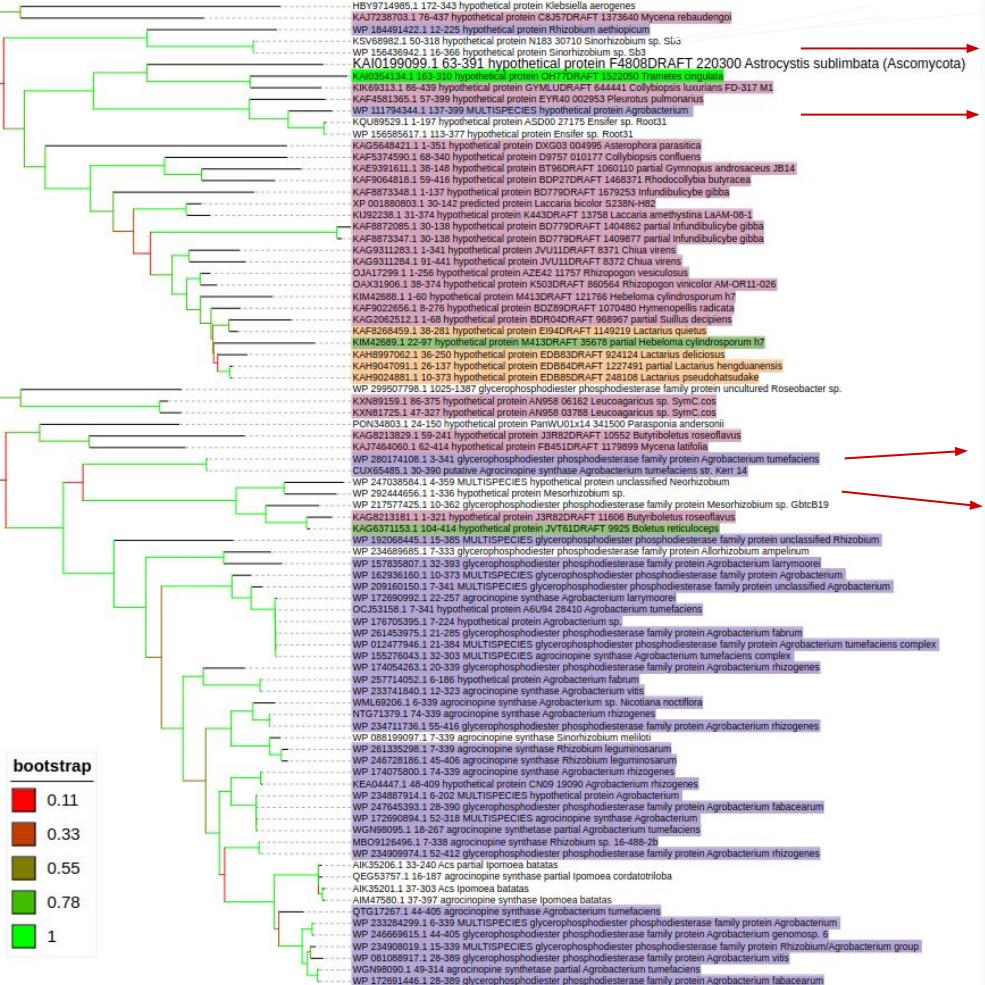


## *Laccaria amethystina*

Description	Max Score	Total Score	Query Cover
<a href="#">Asterophora parasitica isolate AP01 scaffold_20, whole genome shotgun sequence</a>	465	465	33%

### Distribution of the top 1 Blast Hits on 1 subject sequences





## Hypothetical protein *Sinorhizobium* sp.

## Hypothetical protein *Ensifer* sp.

## Agrocinopine synthase *Sinorhizobium meliloti*

## Glycerophosphodiester phosphodiesterase family protein *Mesorhizobium* sp.

## Agrobacterium

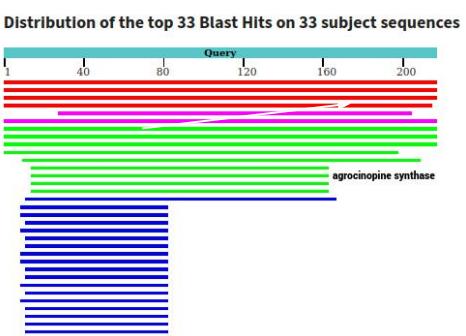
Hypothetical  
protein

Agocinopine  
synthase

Glycerophosphodiester  
phosphodiesterase  
family protein



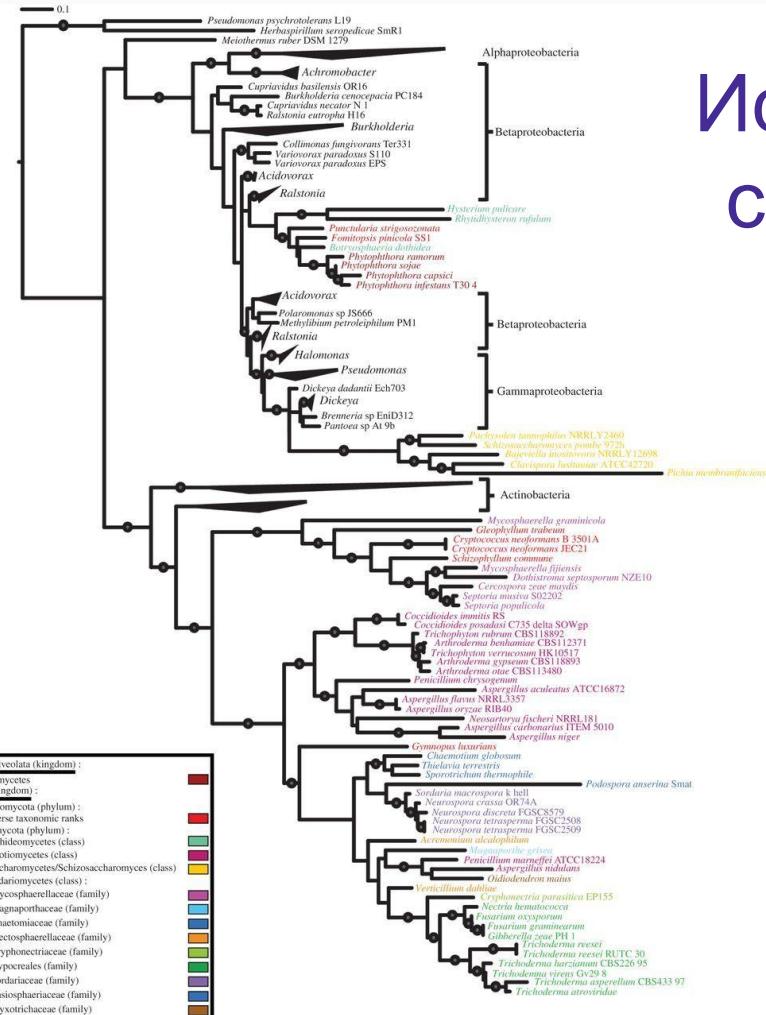
Hypothetical protein  
*Sinorhizobium sp.*



Hypothetical protein *Ensifer sp.*



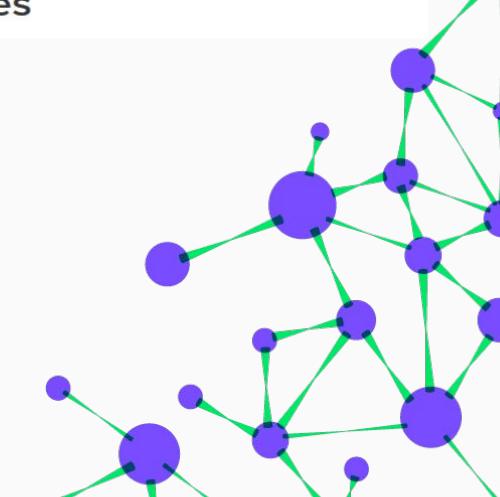
Glycerophosphodiester  
phosphodiesterase family protein  
*Mesorhizobium sp.*



# Исследование другого гена с похожими результатами

arch articles

quent, independent transfers of a metabolic gene from bacteria to contrasted filamentous eukaryotes



# Выводы и гипотезы

Мы обнаружили несколько случаев горизонтального переноса генов между **разными царствами** живых организмов

Существуют группы, включающие одновременно **грибы, бактерии и растения**, содержащие схожие последовательности генов

Некоторые группы без “мозаичной” таксономической структуры могли возникнуть в результате давнего переноса генов

