

Анализ эволюции генной сети развития трихом у растений

Константинов Д.К.

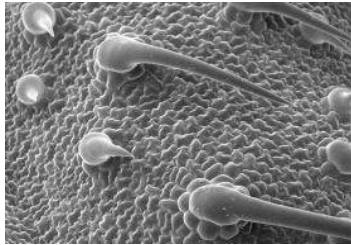
Новосибирский государственный университет
Институт цитологии и генетики СО РАН
Лаборатория эволюционной биоинформатики
и теоретической генетики

Научный руководитель: к.б.н. Дорошков А.В.

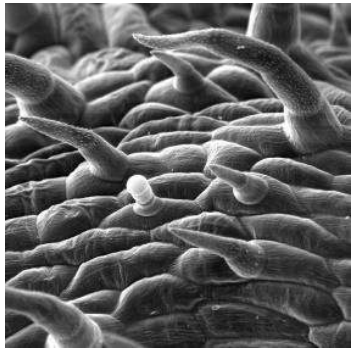
Новосибирск 2017

Трихомы

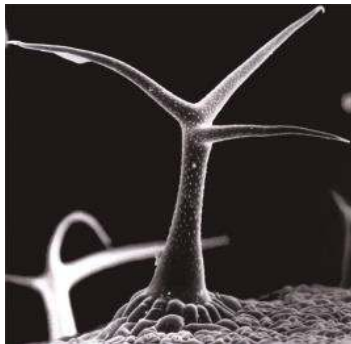
Трихомы – это клетки эпидермиса, образующие выросты на наземных органах растений



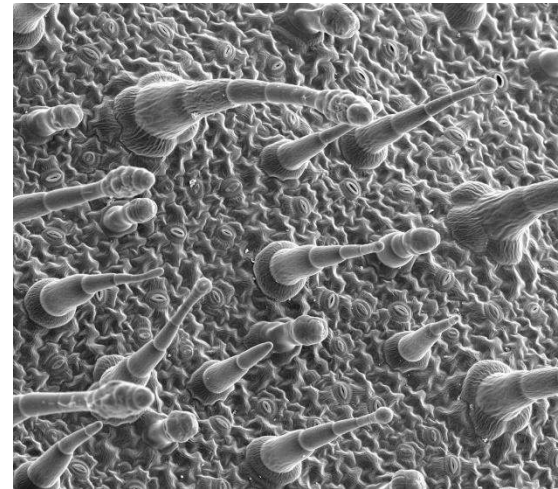
Одноклеточные



Многоклеточные



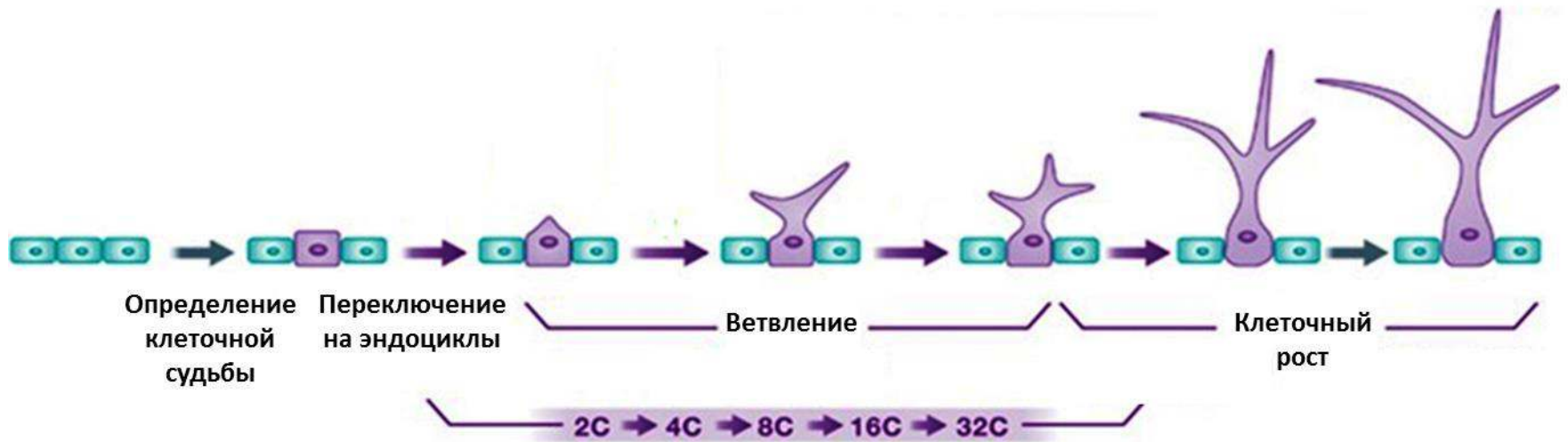
Ветвящиеся



Секретирующие

Трихомы как модель

Легко наблюдаемы
Факультативны
Удобны



E. Caro et.al. 2012

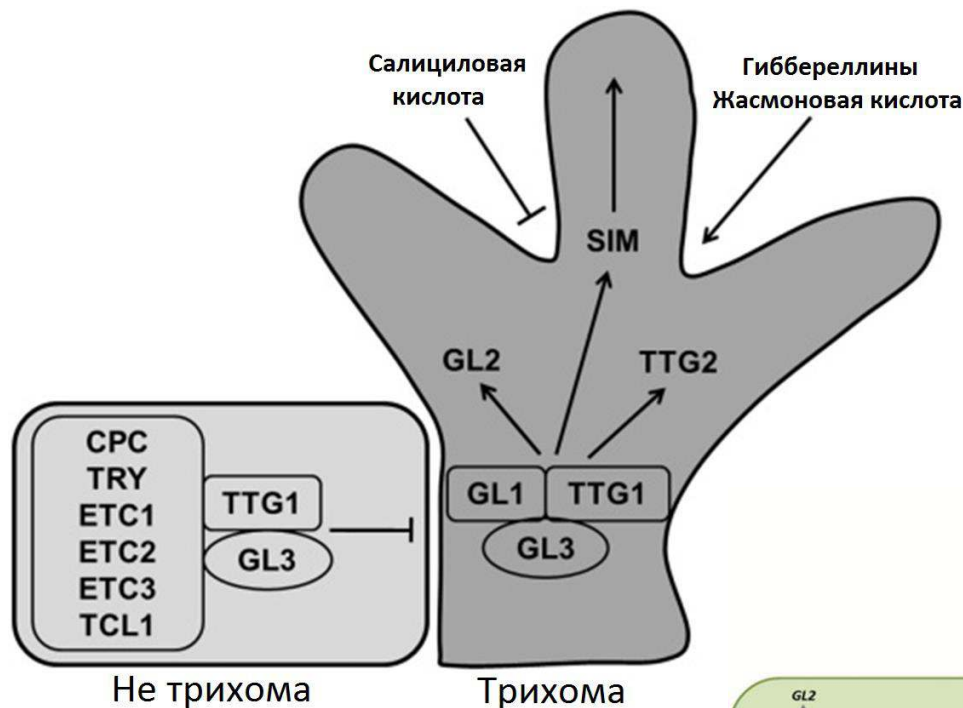
Цели и Задачи

Цель : проанализировать эволюцию компонентов и выявить консервативные элементы генной сети развития трихом у различных видов однодольных и двудольных растений

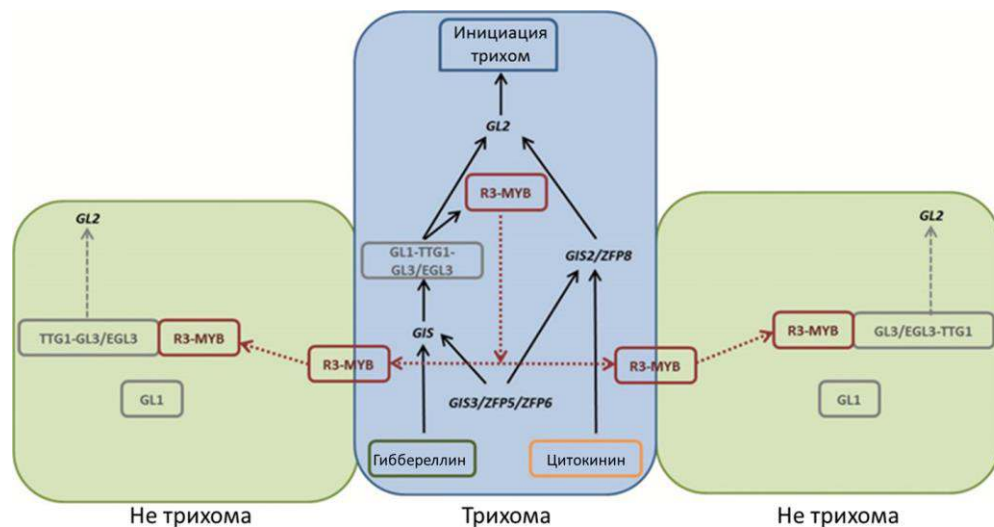
Задачи:

- Сформировать выборку генов, отвечающих за развитие трихом у *A. thaliana*.
- Реконструировать модель генной сети контроля развития трихом у *A. thaliana*.
- Провести анализ эволюции компонентов этой сети, выявить события дупликации, оценить количество ортологичных генов в геномах растений.
- Проанализировать доменную структуру белков и характер накопления замен в наиболее связных частях генной сети.
- Выявить момент появления элементов генной сети и этапы усложнения её топологии.

Модели генной сети контроля развития трихом у *A.thaliana*

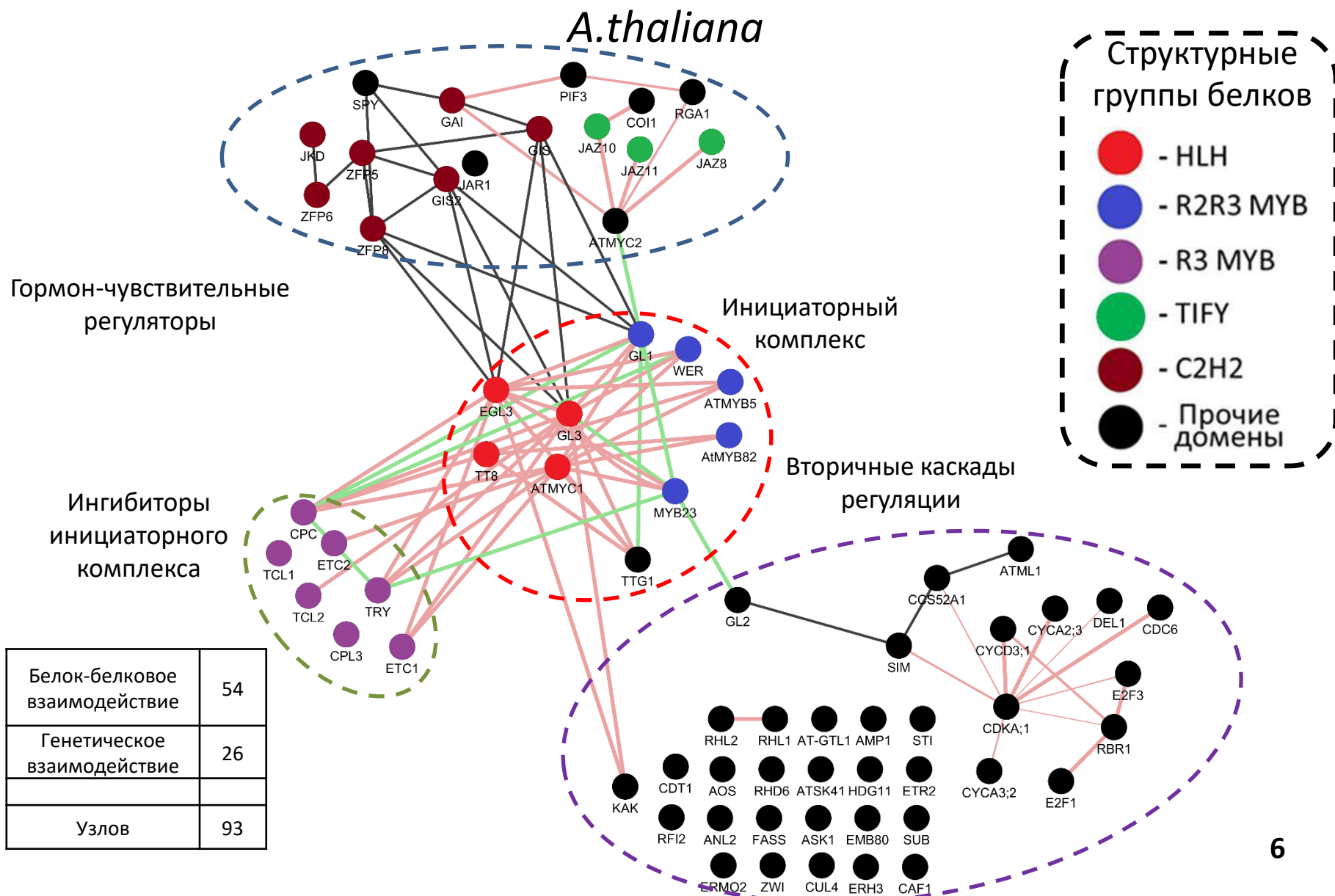


Shweta Kalve et. al. 2014



Luis Matías-Hernández et. al. 2015

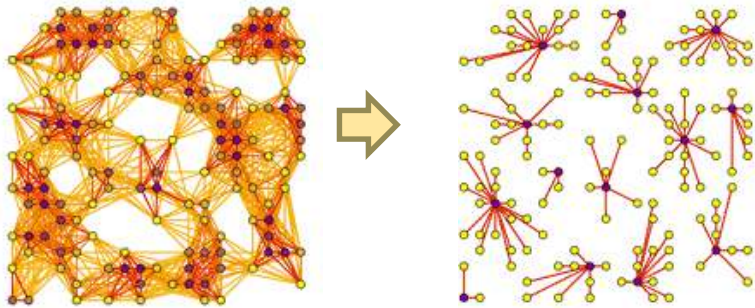
Реконструкция сети белок-белковых и регуляторных взаимодействий вовлеченных в контроль развития трихом у *A.thaliana*



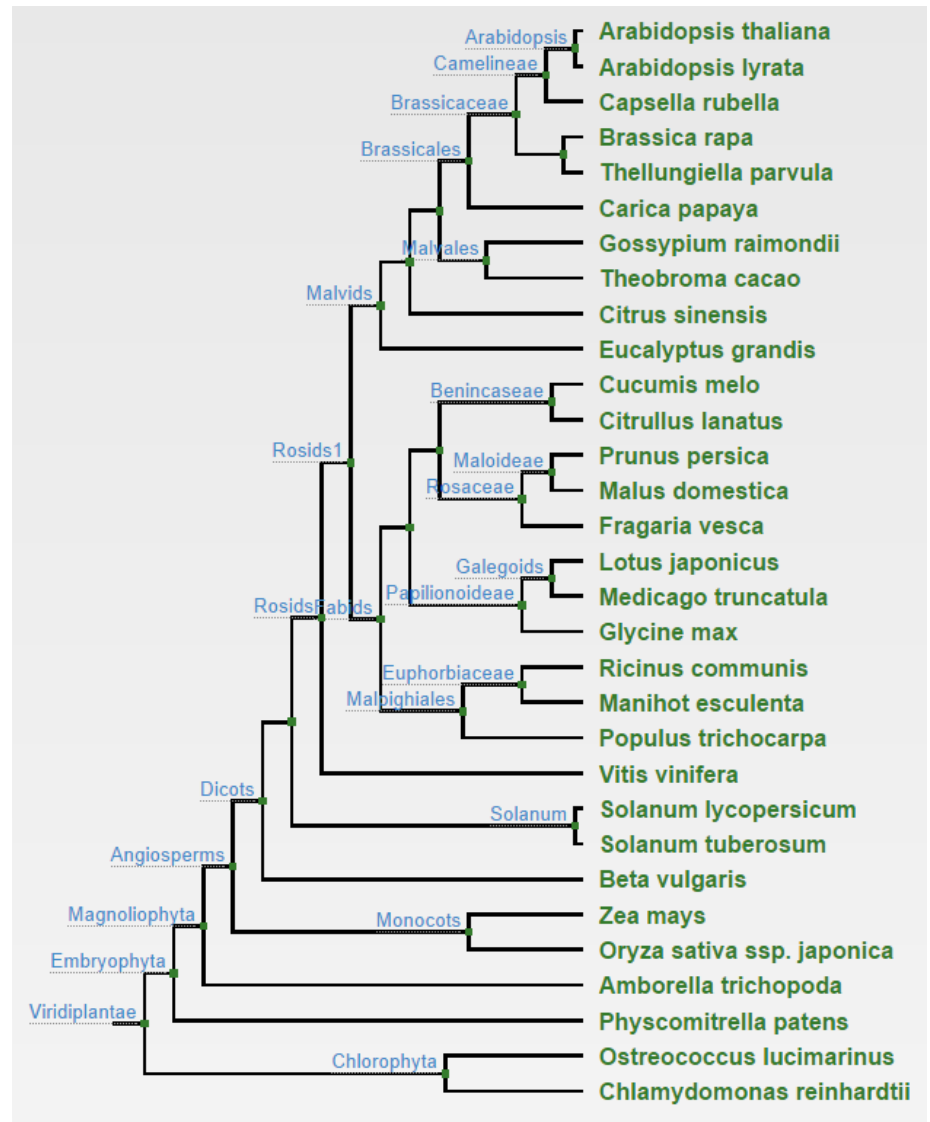
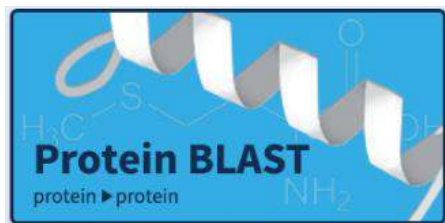
Поиск гомологов

В работе использовали
гомологичные группы из PLAZA

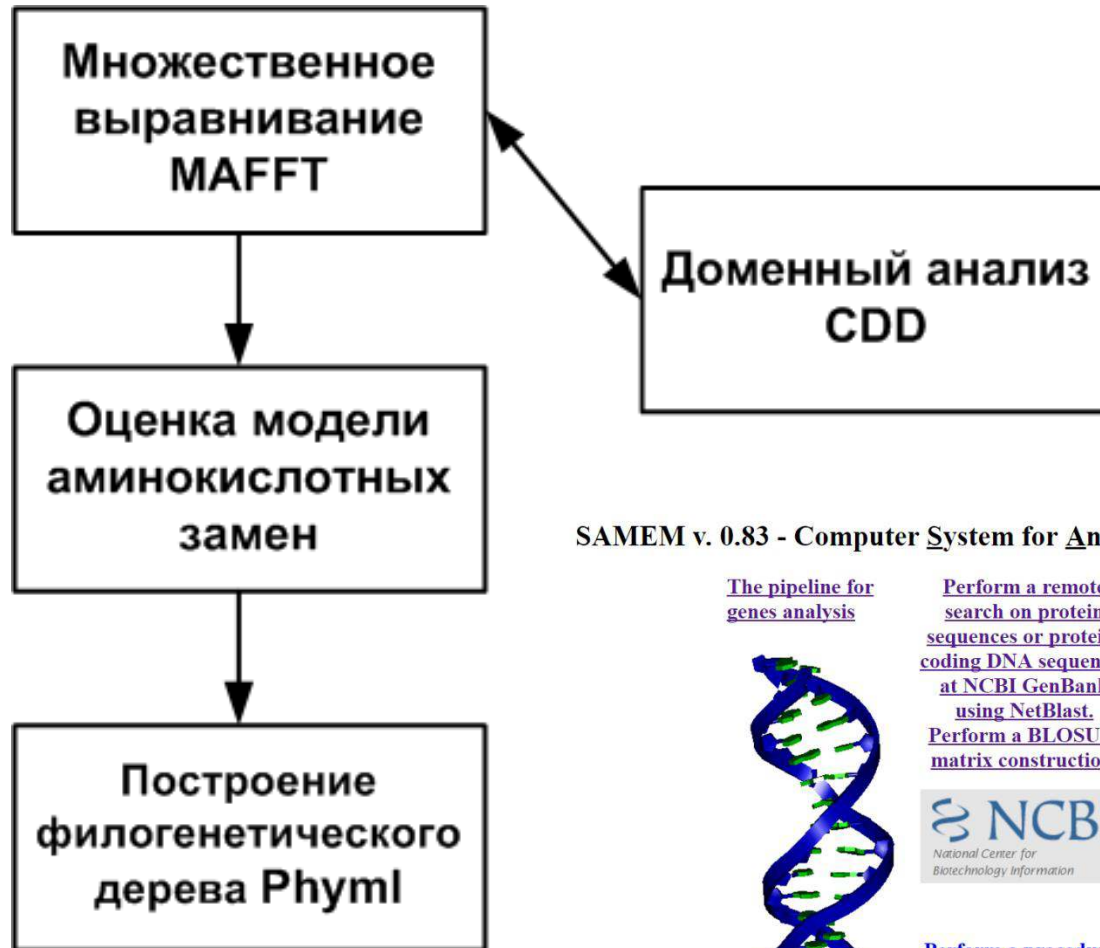
Марковский способ
кластеризации



Гомологичные группы были проверены
реципрокным BLAST поиском



Анализ гомологов



SAMEM v. 0.83 - Computer System for Analysis of Molecular Evolution Modes

[The pipeline for genes analysis](#)

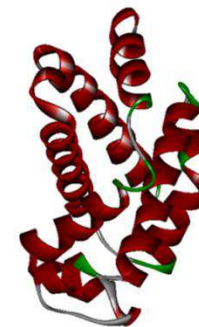


[Perform a remote search on protein sequences or protein-coding DNA sequences at NCBI GenBank using NetBlast. Perform a BLOSUM matrix construction.](#)



[Perform a procedures required for coding-sequence testing.](#)

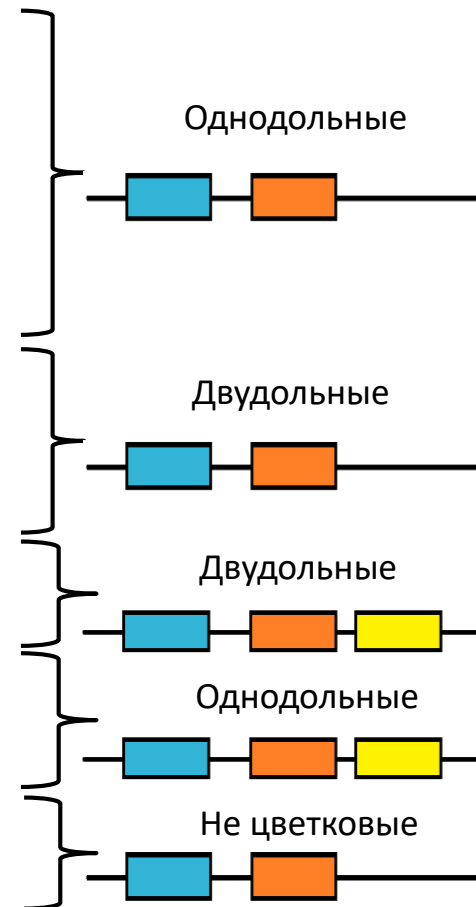
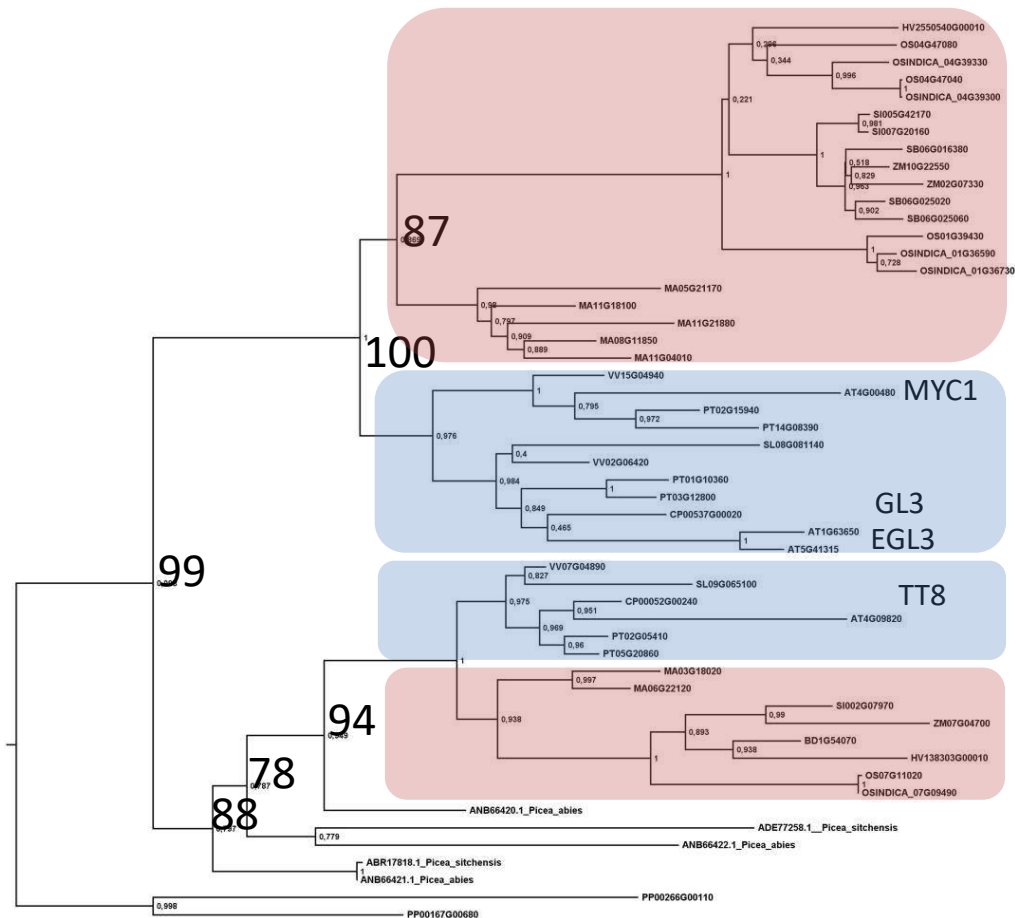
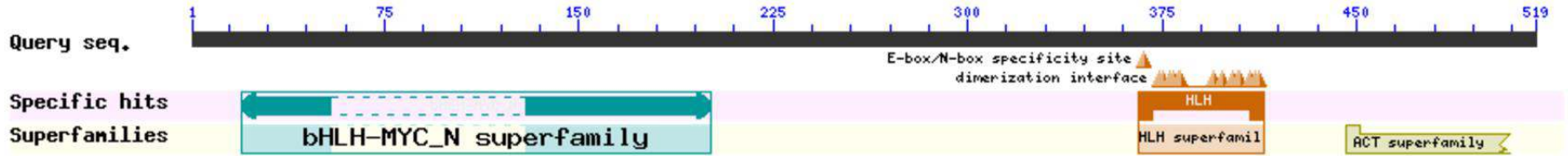
[The pipeline for proteins analysis](#)



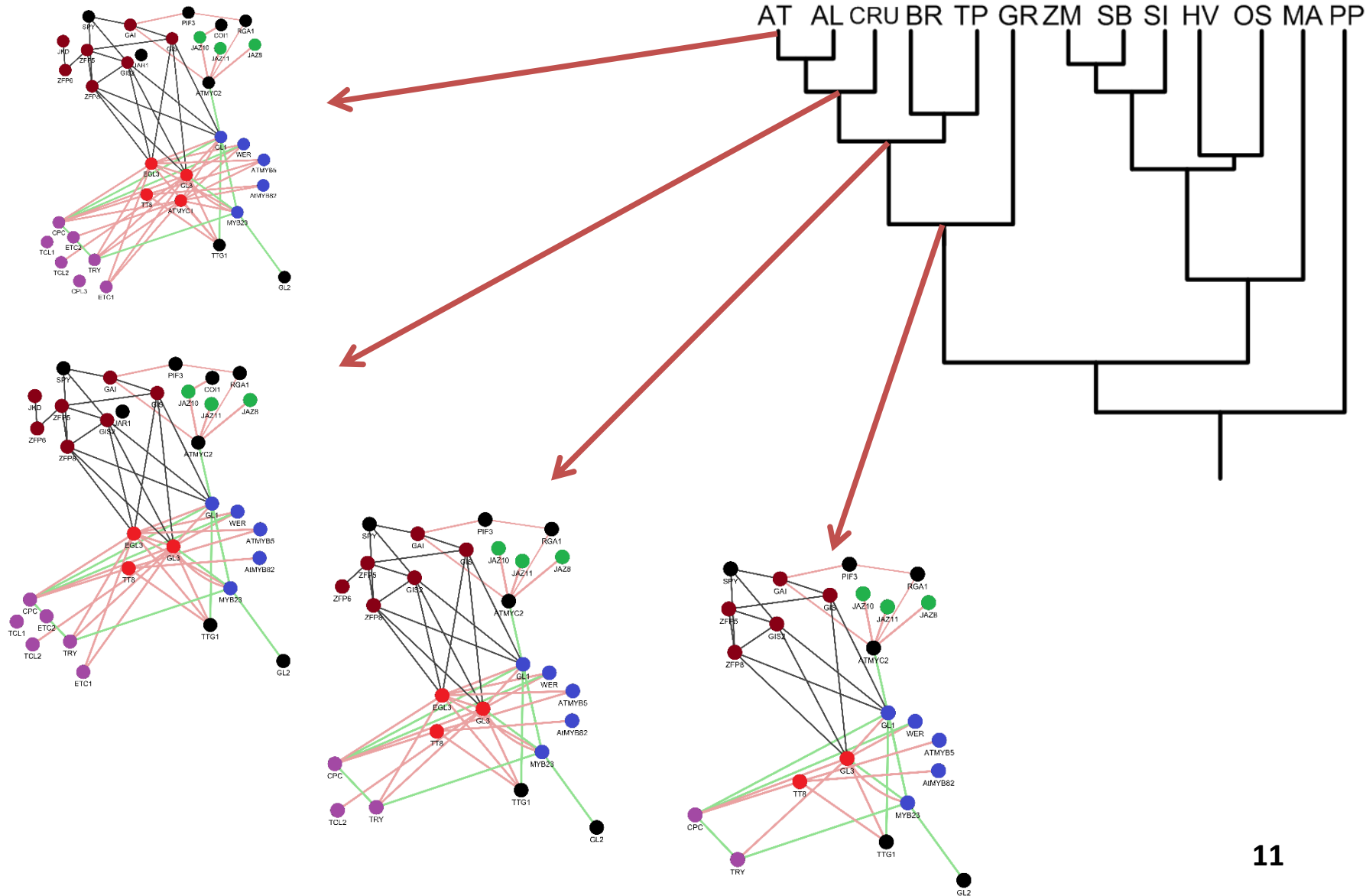
Доменный анализ и множественное выравнивание



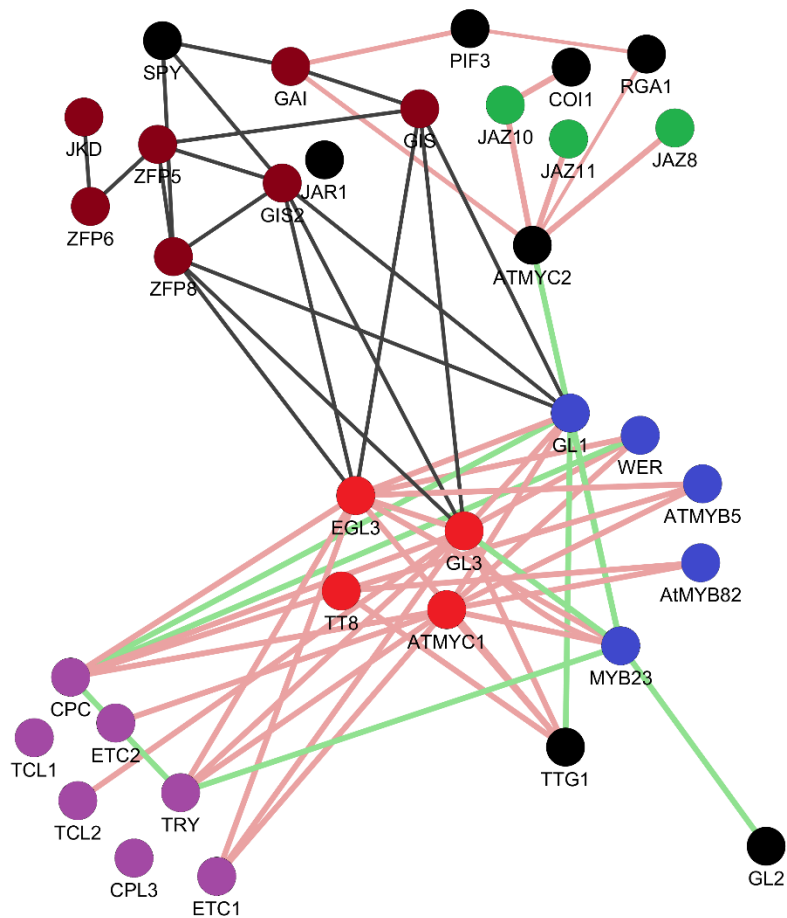
Филогенетический анализ



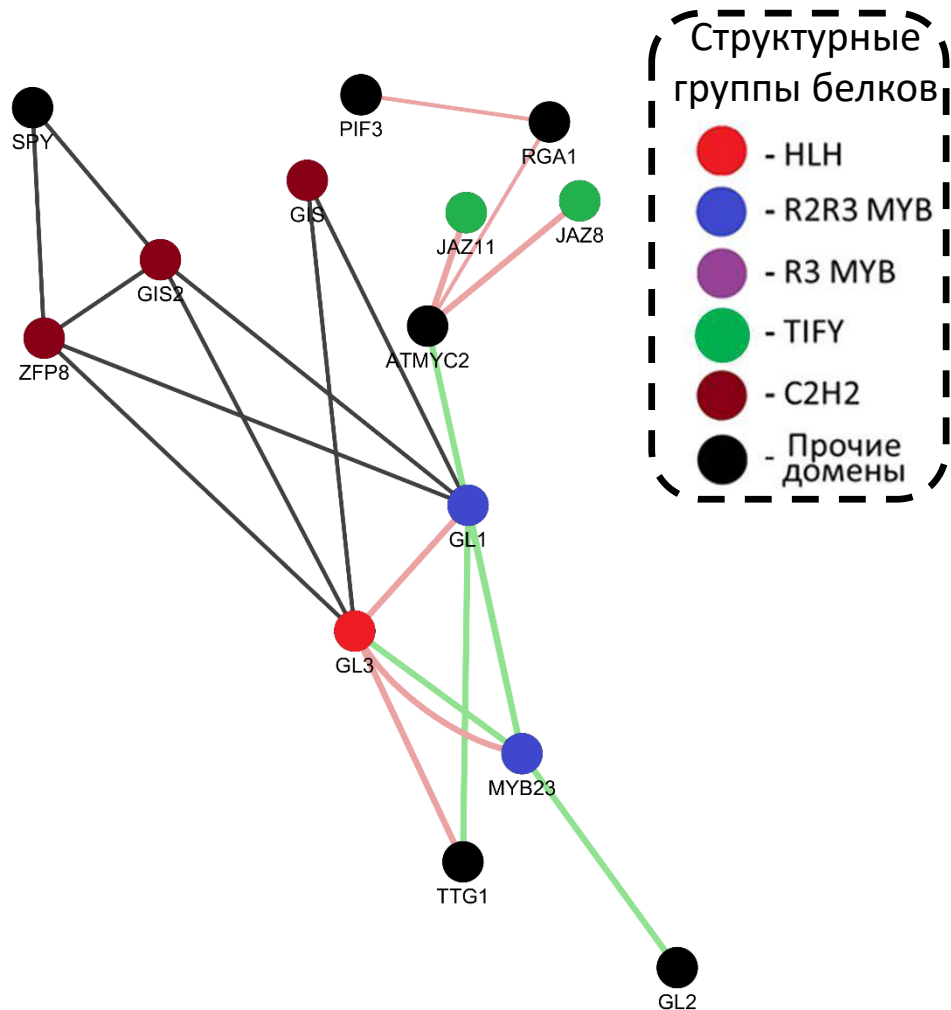
Эволюция геной сети у двудольных



Реконструкция предковой сети



A. thaliana

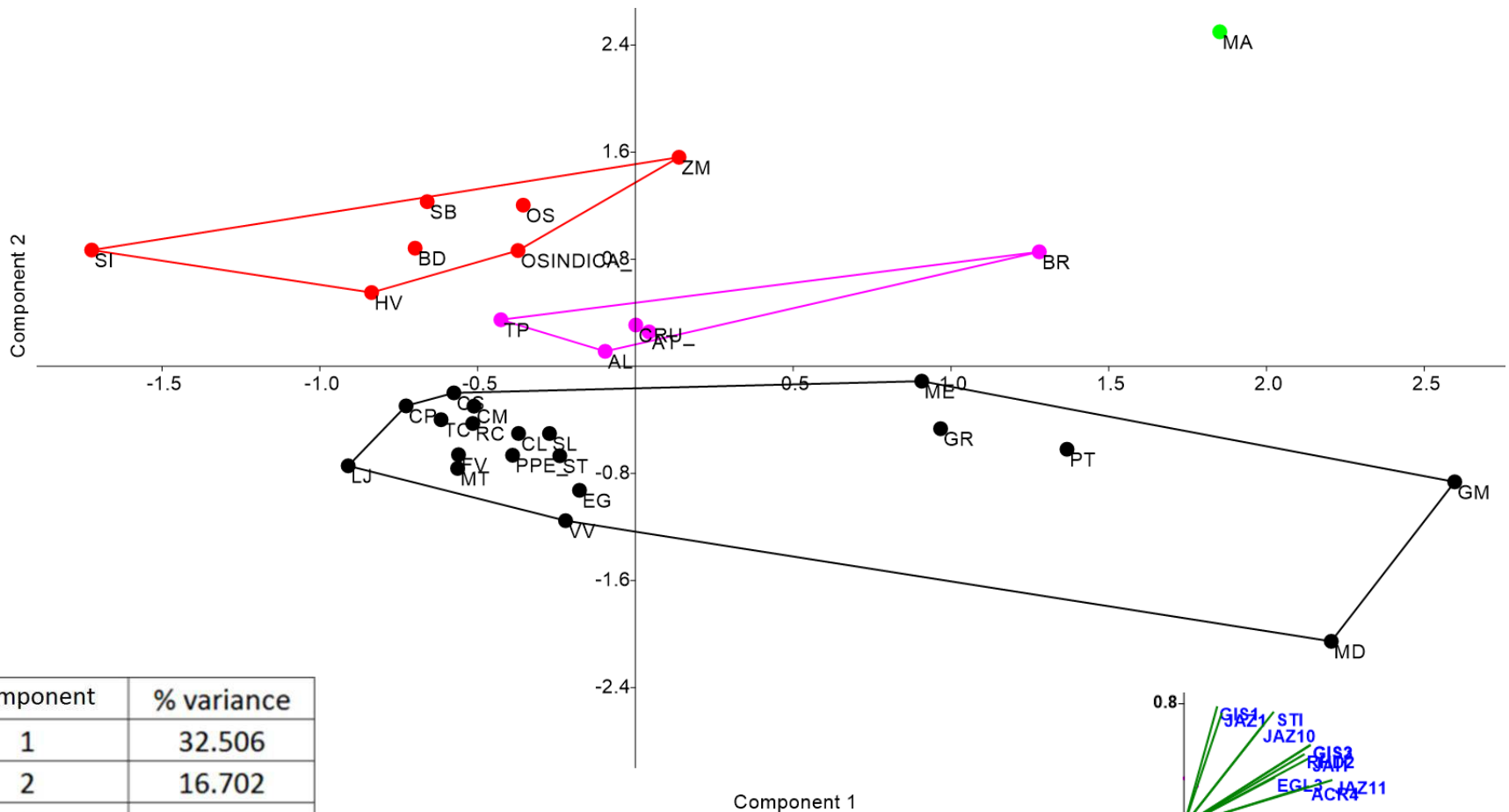


Предок сосудистых растений

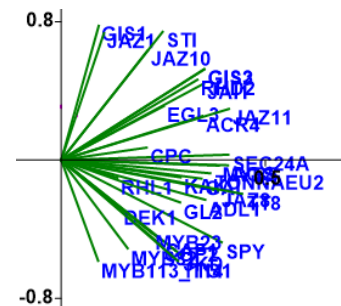
Структурные группы белков

- - HLH
- - R2R3 MYB
- - R3 MYB
- - TIFY
- - C2H2
- - Прочие домены

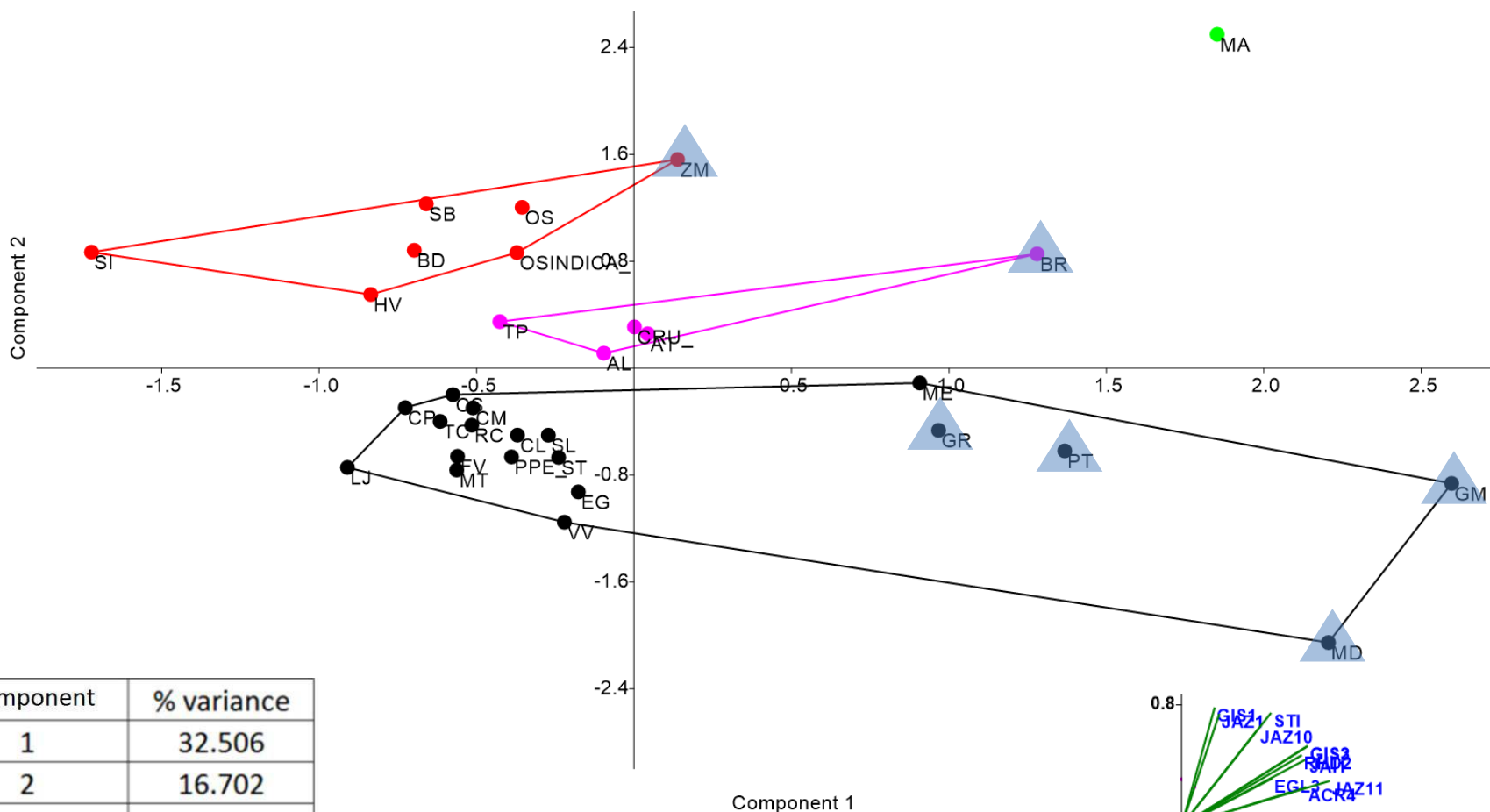
Количество ортологических генов в разных эволюционных ветвях: Анализ главных компонент



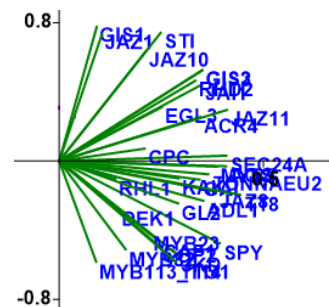
Component	% variance
1	32.506
2	16.702
3	10.134
4	6.2454
5	5.5939



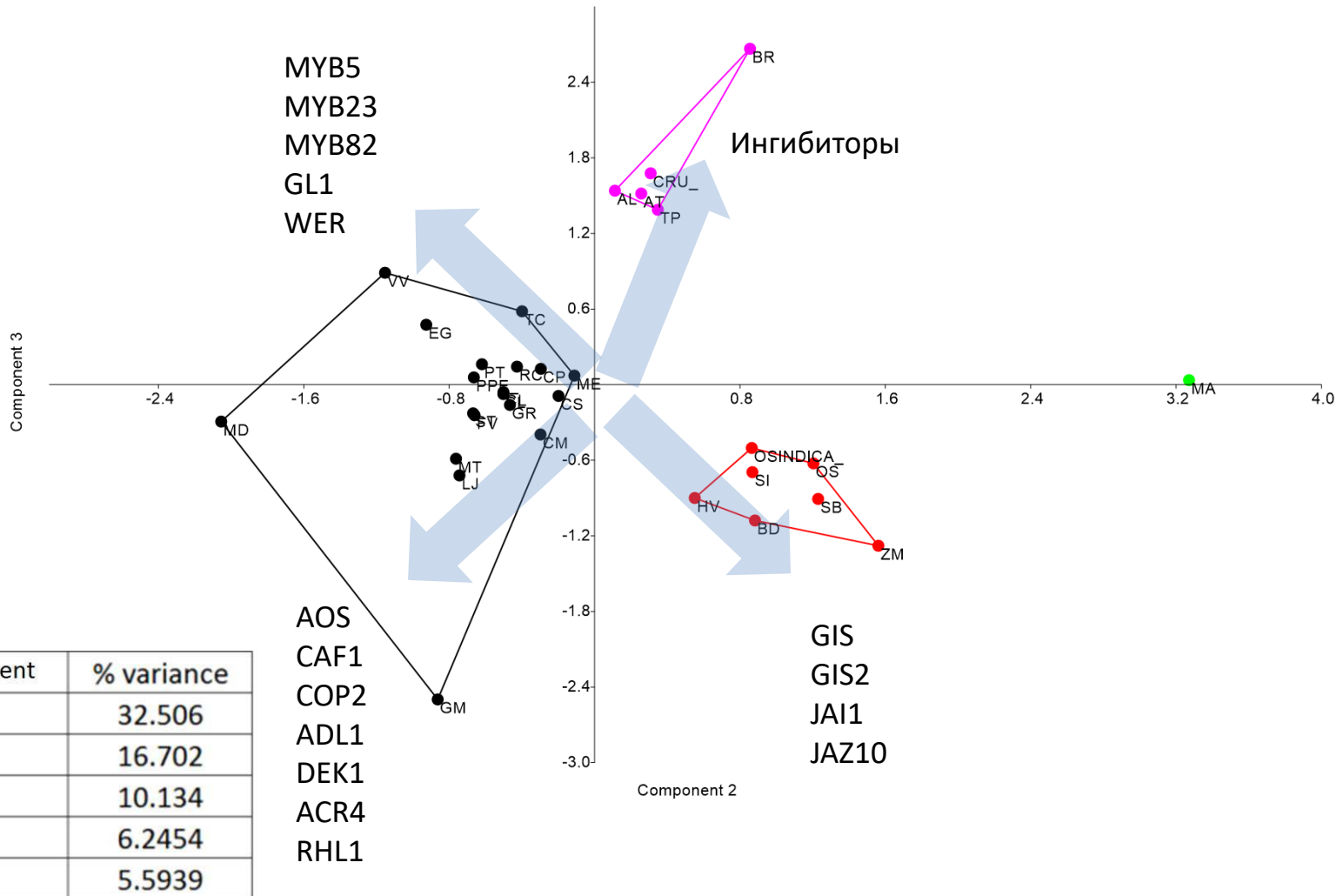
Количество ортологичных генов в разных эволюционных ветвях: Анализ главных компонент



Component	% variance
1	32.506
2	16.702
3	10.134
4	6.2454
5	5.5939



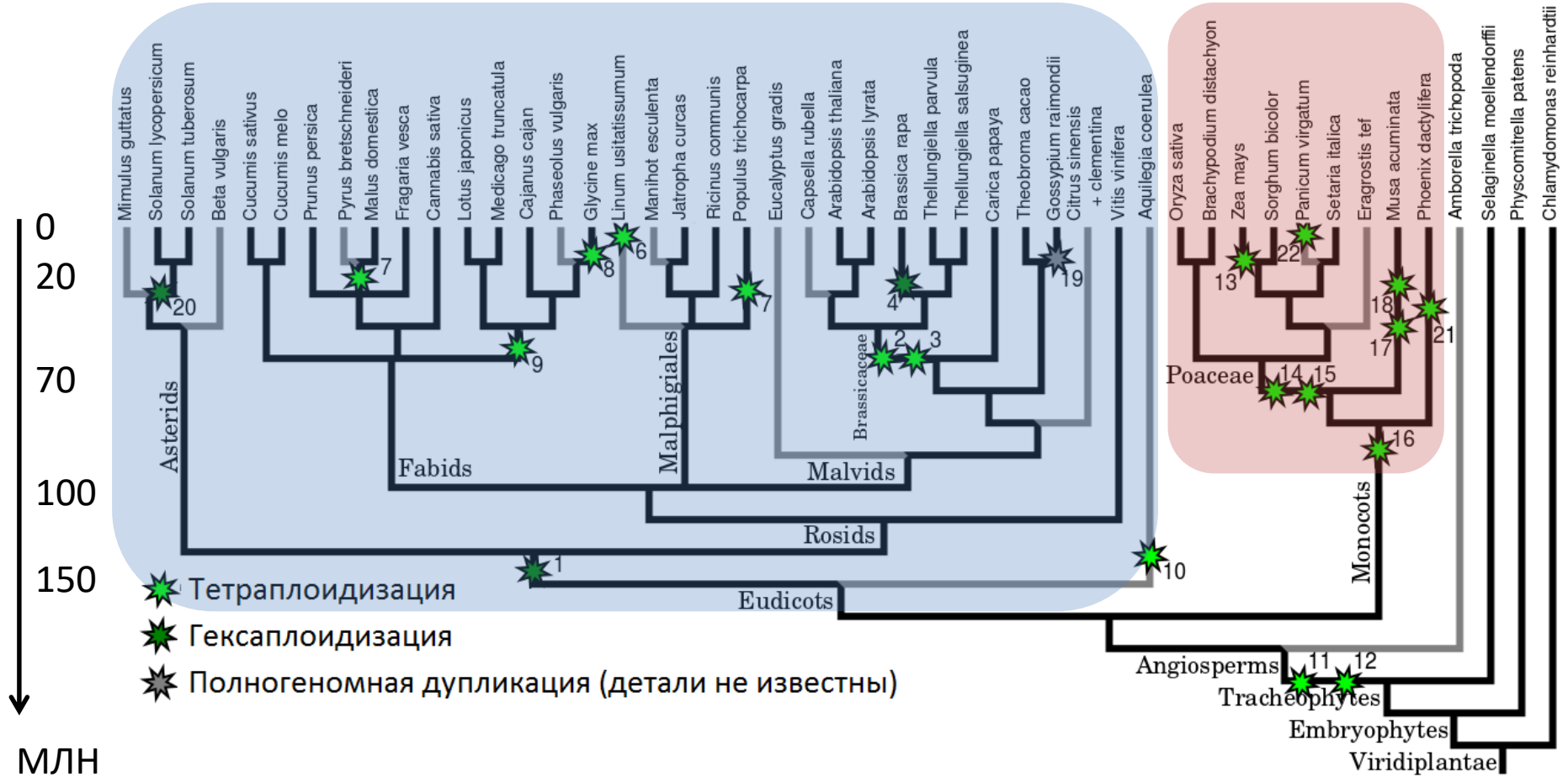
Количество ортологических генов в разных эволюционных ветвях: Анализ главных компонент



Выводы

- Генная сеть развития трихом у *A. thaliana* содержит не менее 69 генов и 323 различных связей (из которых 54 белок-белковых и 26 генетических взаимодействий). Она включает четыре фрагмента белков, участвующих на различных стадиях морфогенеза: “инициаторный комплекс”, гормон чувствительные регуляторы комплекса, его ингибиторы, и гены функционирующие после инициации трихом. Функции генов, входящих в один фрагмент генной сети вырождены.
- Эволюция структуры генной сети у одно и двудольных выявляет различные тенденции. У двудольных растений выявлено усложнение топологии “инициаторного комплекса” (R2R3-MYB факторы и белков с доменом HLH). У однодольных растений выявлено усложнение топологии подсети гормон-чувствительной регуляции “инициаторного комплекса” (ортологи GIS, ZFP6). В целом генная сеть морфогенеза трихом топологически проще у однодольных, что соответствует меньшему морфологическому разнообразию их трихом.
- Генная сеть развития трихом по нашим оценкам, присутствовала у предка высших растений, который существовал около 450 млн. лет назад. Она содержала белки со структурными доменами HLH, WD40 и R2R3-MYB, которые у современных видов необходимы для образования паттерна, однако не содержала ингибиторов, обеспечивающих формирование равномерных, упорядоченных паттернов. Эти белки со структурными доменами R3- MYB появились позже и присутствуют у общего предка, цветковых растений, который существовал около 150 млн. лет назад.

Филогения Растений



Основные базы данных используемые в работе



Tair – БД
специализирующаяся на
A. thaliana



NCBI - глобальная БД по
биологии и медицине



PLAZA - БД сравнительной
геномики растений

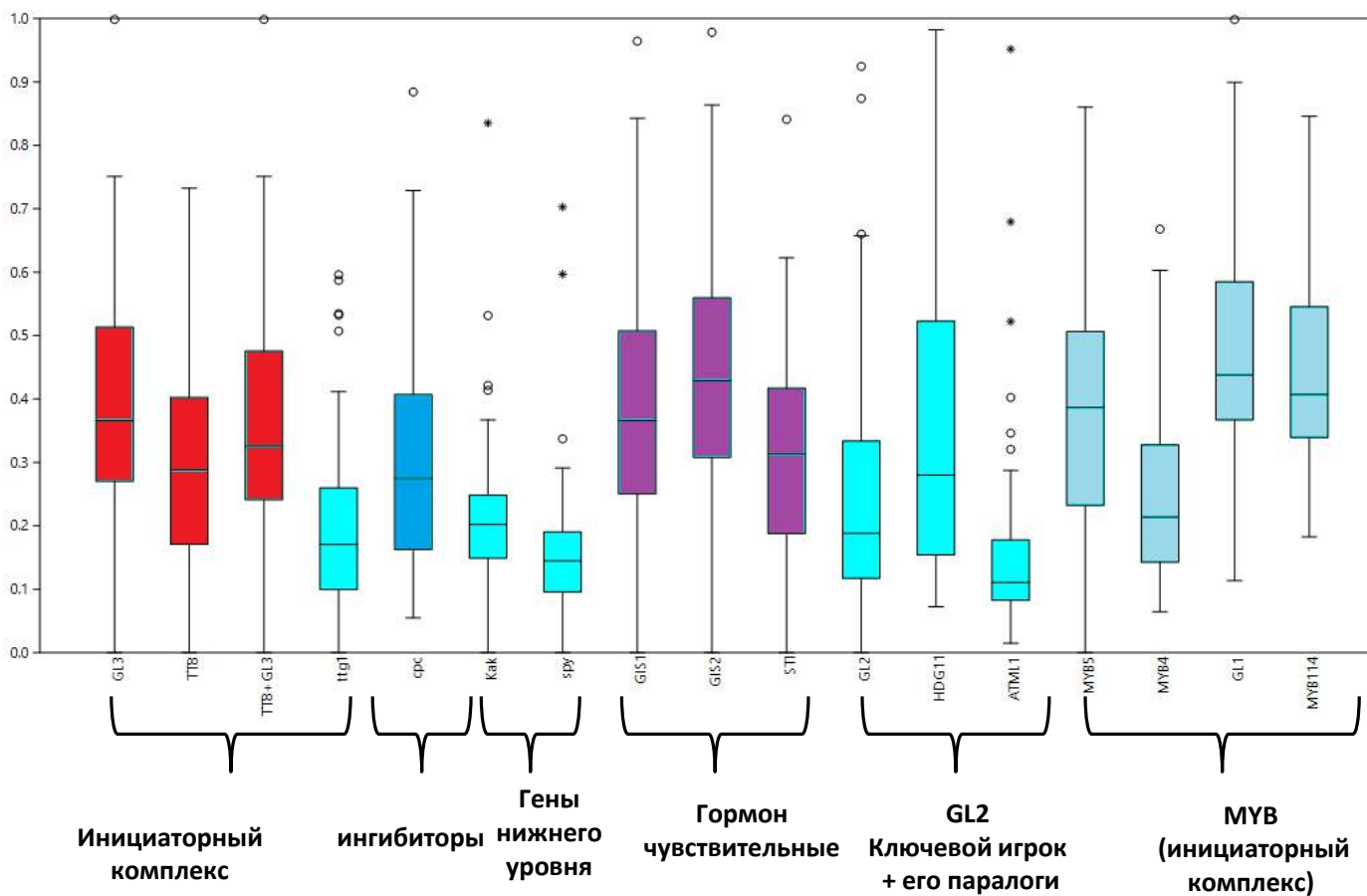
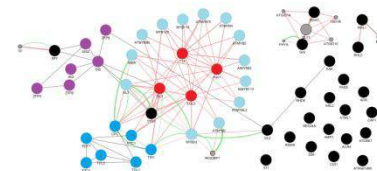


Gene Ontology - база
ассоциированных
терминов

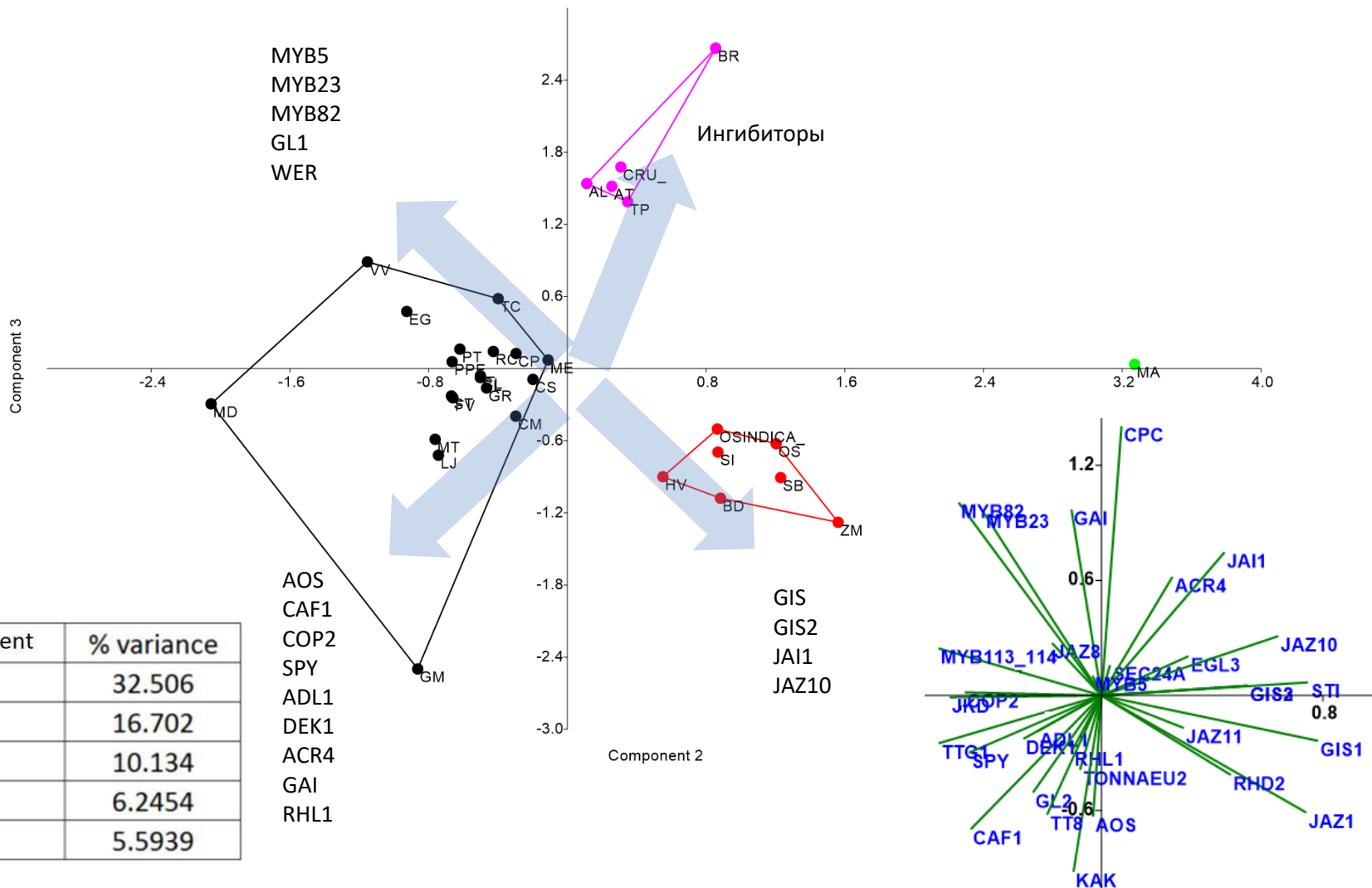
Анализ микрочипов

actin 7	AT5G09810	#Н/Д	AT2G18900
ORP4C	AT5G57240	#Н/Д	AT3G52550
MILDEW RESISTANCE LOCUS O 5	AT2G33670	ATPUP14	AT1G19770
MMS ZWEI homologue 2	AT1G70660	#Н/Д	AT5G14700
mitochondrial processing peptidase alpha subunit	AT3G16480	Q9FHQ1	AT5G37690
#Н/Д	AT1G15350	#Н/Д	AT5G51950
EXORDIUM like 2	AT5G64260	HAT14	AT5G06710
#Н/Д	AT2G39280	#Н/Д	AT3G62590
BGLU7	AT3G62740	CAM1	AT5G37780
CLMP1	AT1G62390	#Н/Д	AT1G51580
DA1-related protein 7	AT5G66610	ralf-like 33	AT4G15800
Q9LZV9	AT5G01870	FZR3	AT5G13840
TX1	AT2G04040	Q9LPJ7	AT1G32660
#Н/Д	AT3G11620	TPLATE	AT3G01780
ARABIDOPSIS THALIANA HOMEODOMAIN PROTEIN 14	AT2G34710	ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 13	AT2G44840
ATGSTF12	AT5G17220	ATTPSA	AT1G06410
RAP2.2	AT3G14230	#Н/Д	AT3G47836
#Н/Д	AT3G16680	#Н/Д	AT2G33350
Q9ZQ60	AT2G14290	#Н/Д	AT5G18460
synaptotagmin 3	AT5G04220	ETP2	AT3G18910
#Н/Д	AT1G06340	cytochrome P450, family 81, subfamily D, polypeptide 11	AT3G28740
SDG19	AT1G73100	#Н/Д	AT3G52760
#Н/Д	AT3G06180	#Н/Д	AT5G46070
#Н/Д	AT4G26620	RADIAL SWELLING 9	AT5G42080
#Н/Д	AT3G54080		

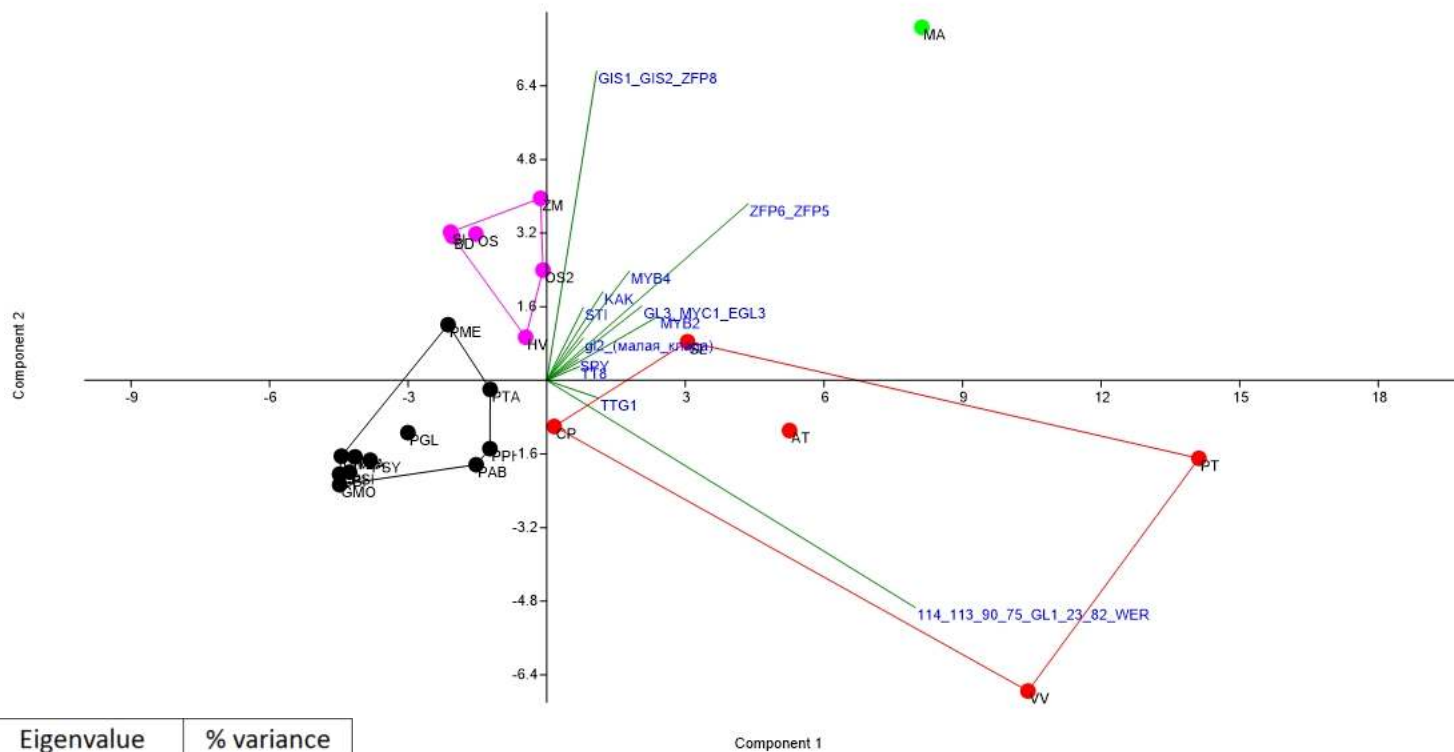
Анализ Ka/Ks у разных генов



Количество ортологических генов в разных эволюционных ветвях: Анализ главных компонент



Количество ортологических генов в разных эволюционных ветвях: Анализ главных компонент



PC	Eigenvalue	% variance
1	0.335546	32.506
2	0.172405	16.702
3	0.104609	10.134
4	0.064469	6.2454
5	0.0577434	5.5939