

Эволюция генов рецепторов *LysM* Вавиловии прекрасной (*Vavilovia formosa*)

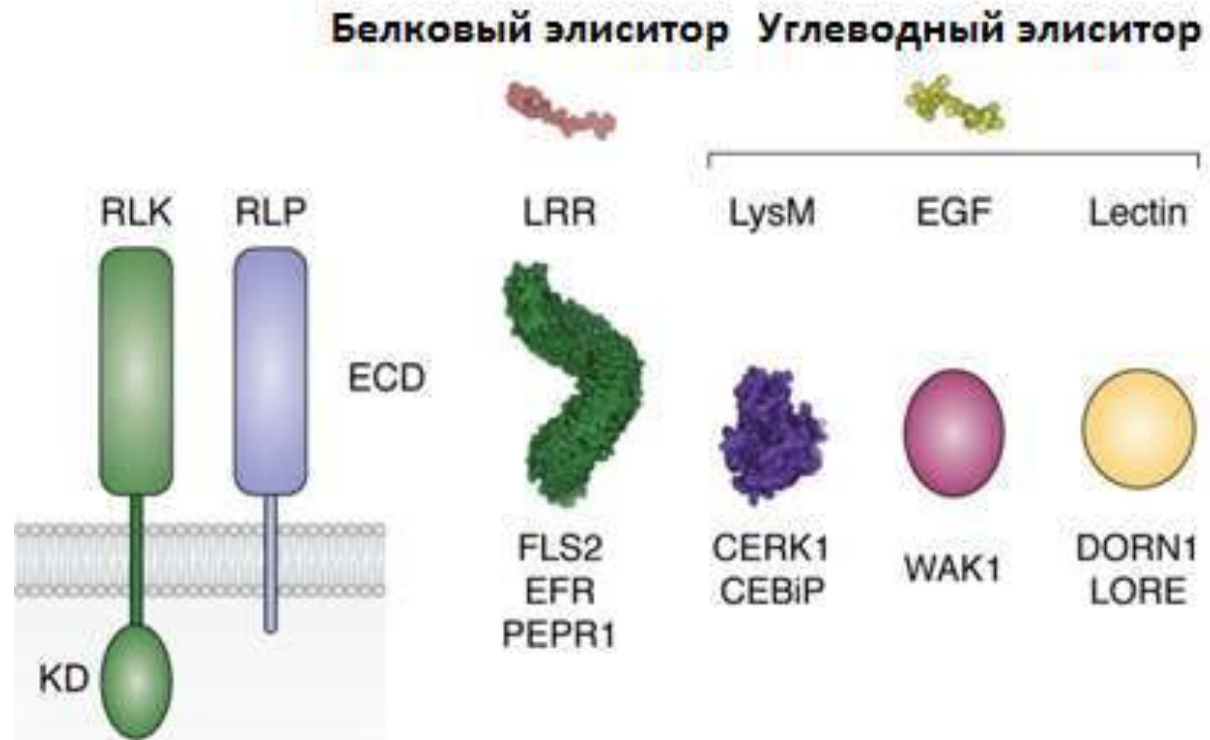
Новосибирский государственный университет
Институт Цитологии и Генетики СО РАН, Новосибирск

Выполнил: Каретников Дмитрий Игоревич, 4 курс ФЕН

Научный руководитель: Афонников Дмитрий Аркадьевич, канд. биол. наук, доцент,
заведующий лабораторией эволюционной биоинформатики и теоретической генетики
ИЦиГ СО РАН

Новосибирск, 2021 г.

Рецепторы неспецифического иммунитета

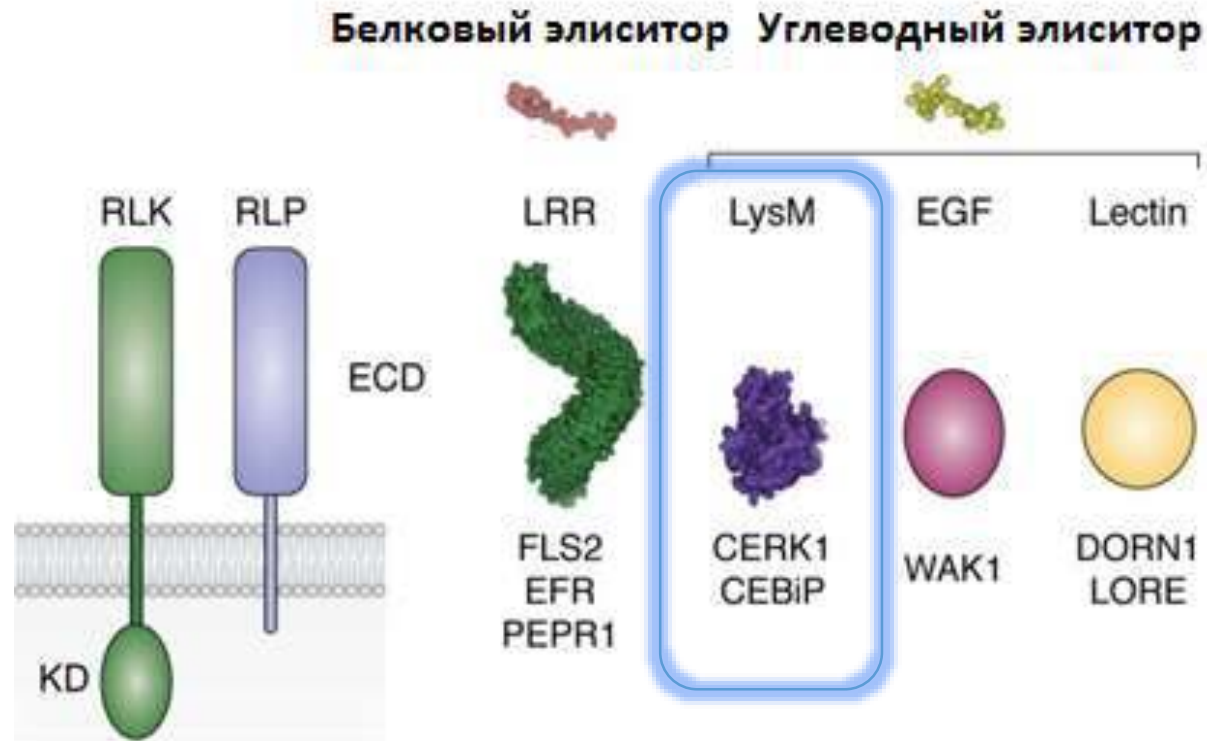


Основными участниками взаимодействия растений и микроорганизмов являются рецепторы растений и бактериальные лиганды.

Существует несколько типов рецепторов, каждый из которых взаимодействует со своими лигандами.

Все рецепторы подразделяются на RLK (receptor-like kinase) и RLP (receptor-like protein).

Рецепторы неспецифического иммунитета



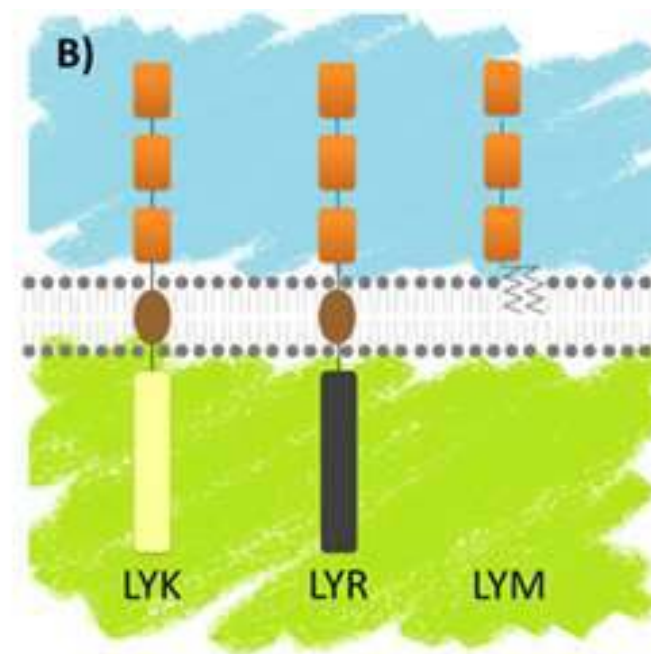
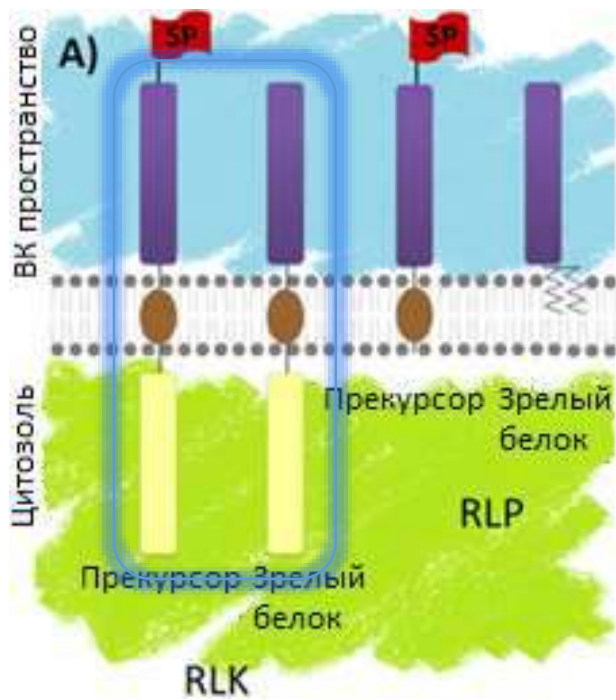
Основными участниками взаимодействия растений и микроорганизмов являются рецепторы растений и бактериальные лиганды.

Существует несколько типов рецепторов, каждый из которых взаимодействует со своими лигандами.

Все рецепторы подразделяются на RLK (receptor-like kinase) и RLP (receptor-like protein).

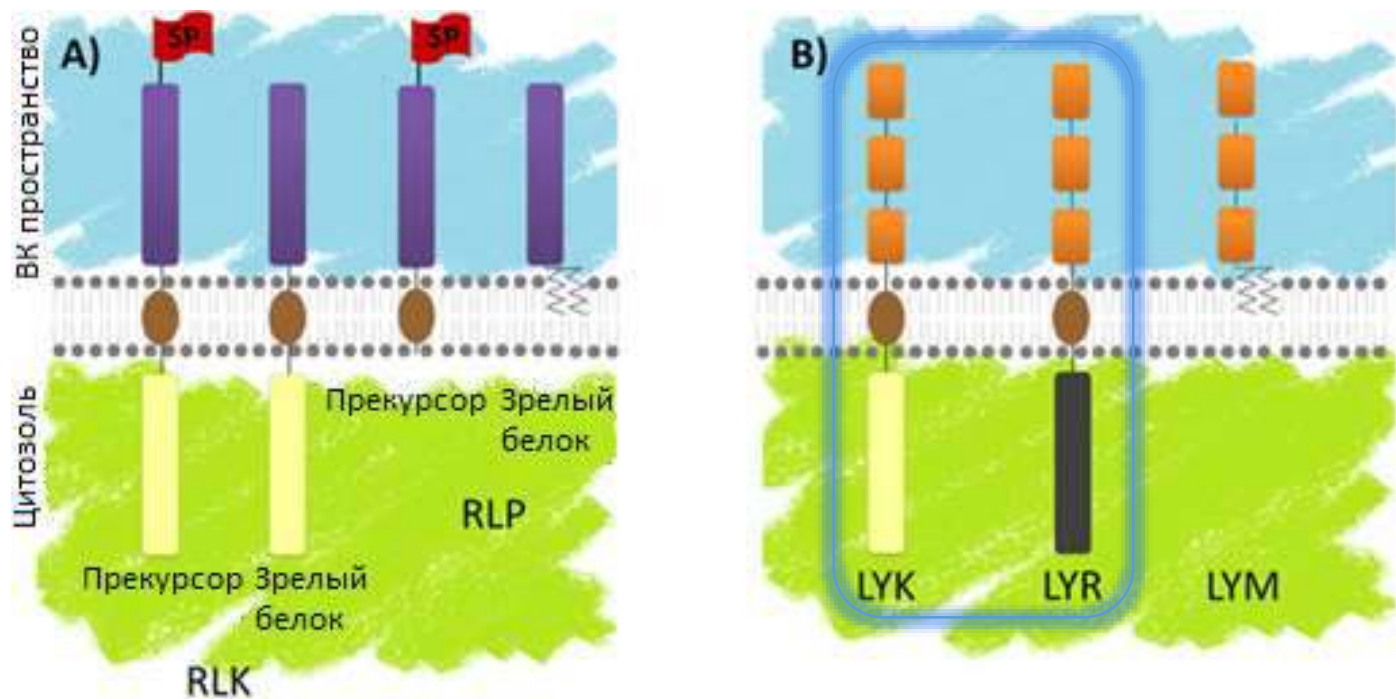
В нашем исследовании изучаются LysM рецепторы, функции которых выходят за рамки активации неспецифического иммунитета.

Классификация LysM рецепторов



LysM-RLK состоят из внеклеточного домена (LysM), внутриклеточного (киназный домен) и трансмембранного домена.

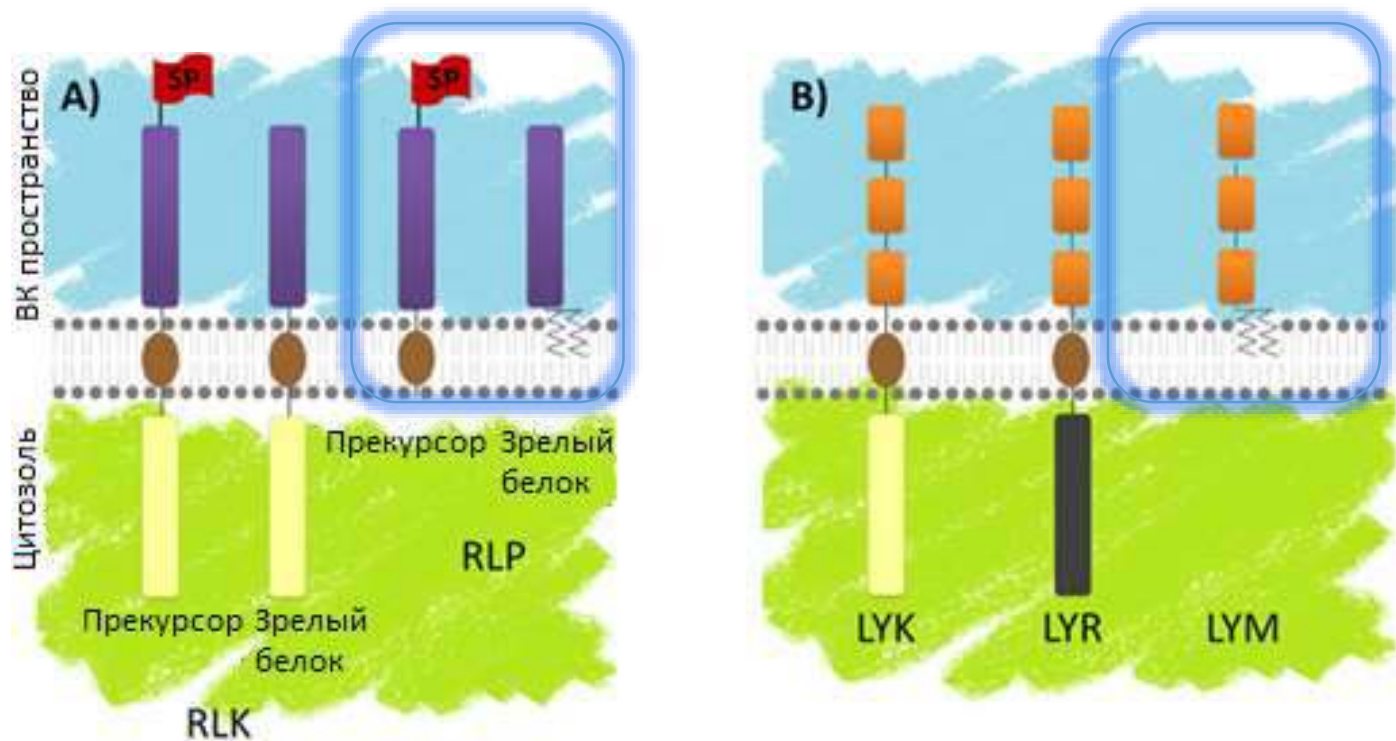
Классификация LysM рецепторов



LysM-RLK состоят из внеклеточного домена (LysM), внутриклеточного (киназный домен) и трансмембранного домена.

Существует два подтипа LysM-RLK: LYK и LYR.

Классификация LysM рецепторов

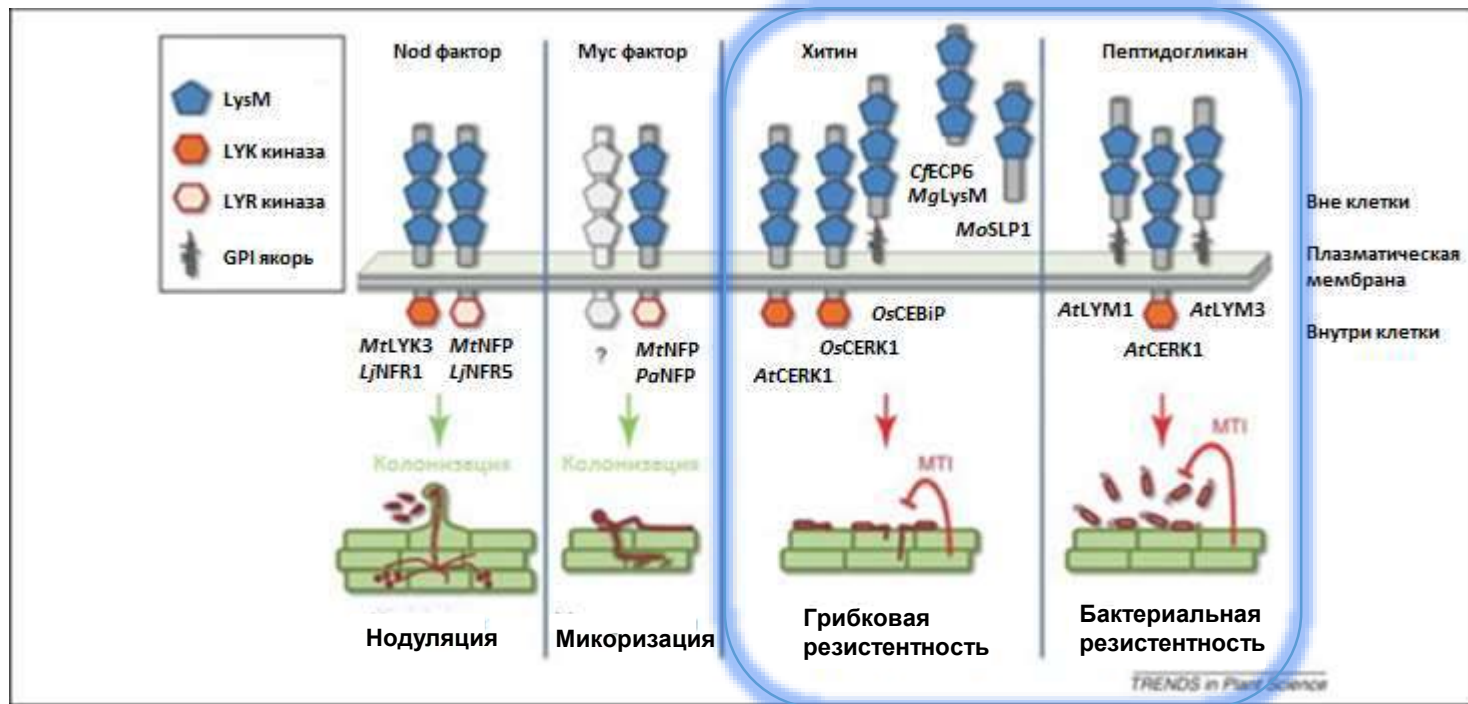


LysM-RLK состоят из внеклеточного домена (LysM), внутриклеточного (киназный домен) и трансмембранного домена.

Существует два подтипа LysM-RLK: LYK и LYR.

LysM-RLP имеют только внеклеточный LysM домен и якорь-гликолипид.

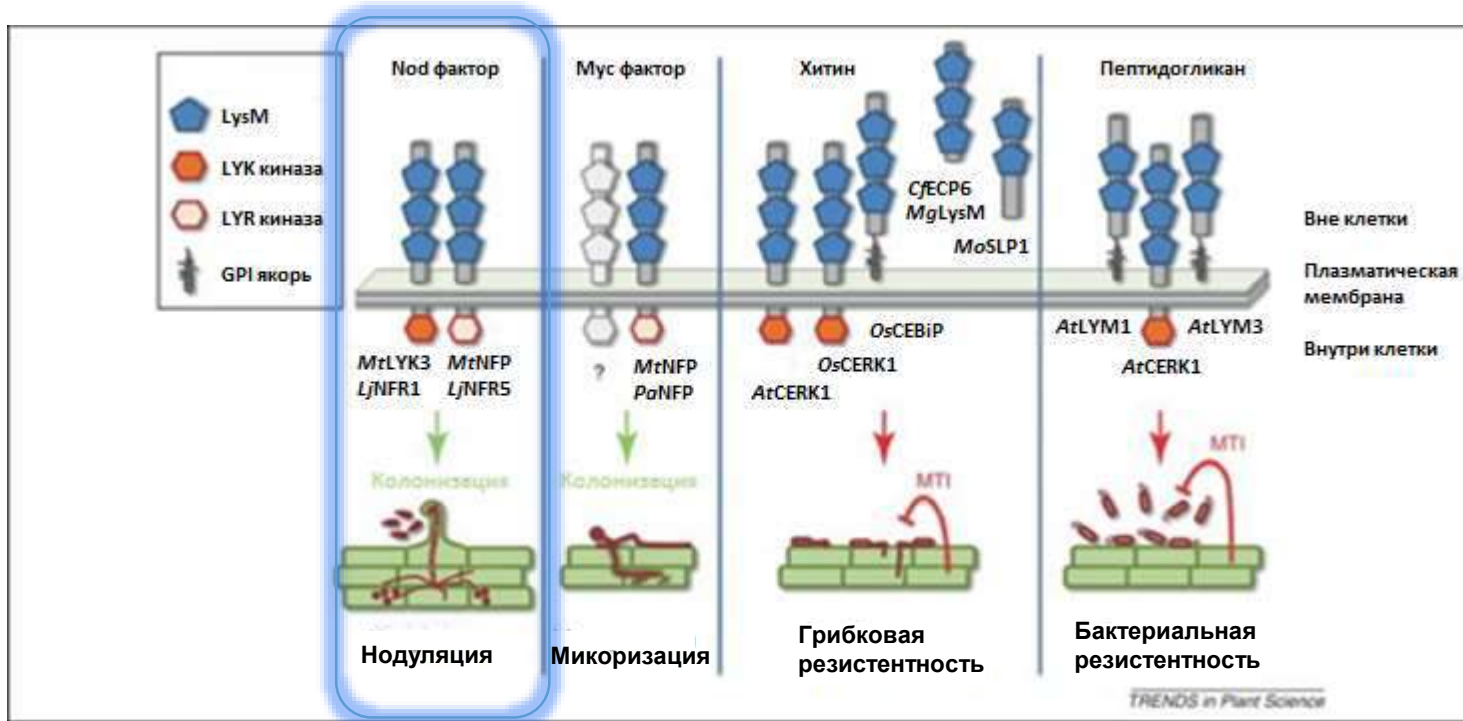
Функции LysM рецепторов



Функции LysM белков различны:

- Активация неспецифического иммунитета

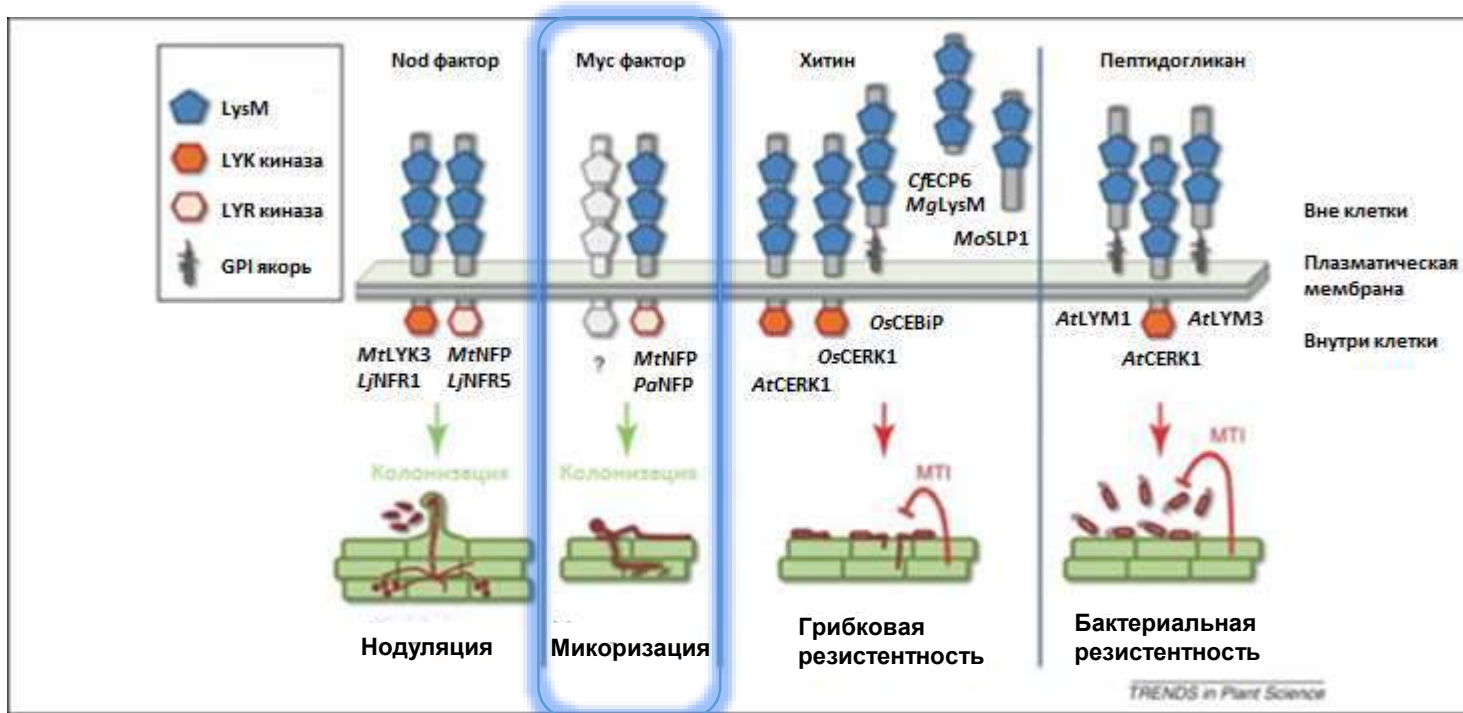
Функции LysM рецепторов



Функции LysM белков различны:

- Активация неспецифического иммунитета
- Симбиоз с бактериями, фиксирующими атмосферный азот (образование клубеньков)

Функции LysM рецепторов



Функции LysM белков различны:

- Активация неспецифического иммунитета
- Симбиоз с бактериями, фиксирующими атмосферный азот (образование клубеньков)
- Симбиоз с грибами Glomeromycota (образование арбускулярной микоризы)

Vavilovia formosa



Краткая характеристика:

Семейство: Fabaceae (бобовые)

Триба: Fabeae (бобовые)

Род: Vavilovia

Вид: Vavilovia formosa

- Эндемик Кавказа и Передней Азии
- Произрастает на высоте 2200 м. и более над уровнем моря
- Занесена в Красную книгу России
- Является стенобионтом (сложно вырастить при неподходящих условиях среды)
- Способна образовывать клубеньки

Цели и задачи

Цель настоящей работы заключается в идентификации и анализе структуры и эволюции генов, кодирующих LysM рецепторы, в геноме растительного организма семейства Fabaceae *V. formosa*. Для достижения поставленной цели были сформулированы следующие задачи:

1. Аннотация сборки генома *V. formosa*;
2. Определение репертуара LysM генов *V. formosa* на основе его геномной аннотации;
3. Классификация белковых последовательностей LysM у *V. formosa* с помощью филогенетического анализа и предсказание функций белков на основе гомологии; верификация классификации на основе анализа ортологии генов семи родственных организмов;
4. Реконструкция потенциального расположения генов LysM в геноме *V. formosa* на основе анализа ортологических групп и синтении геномов.

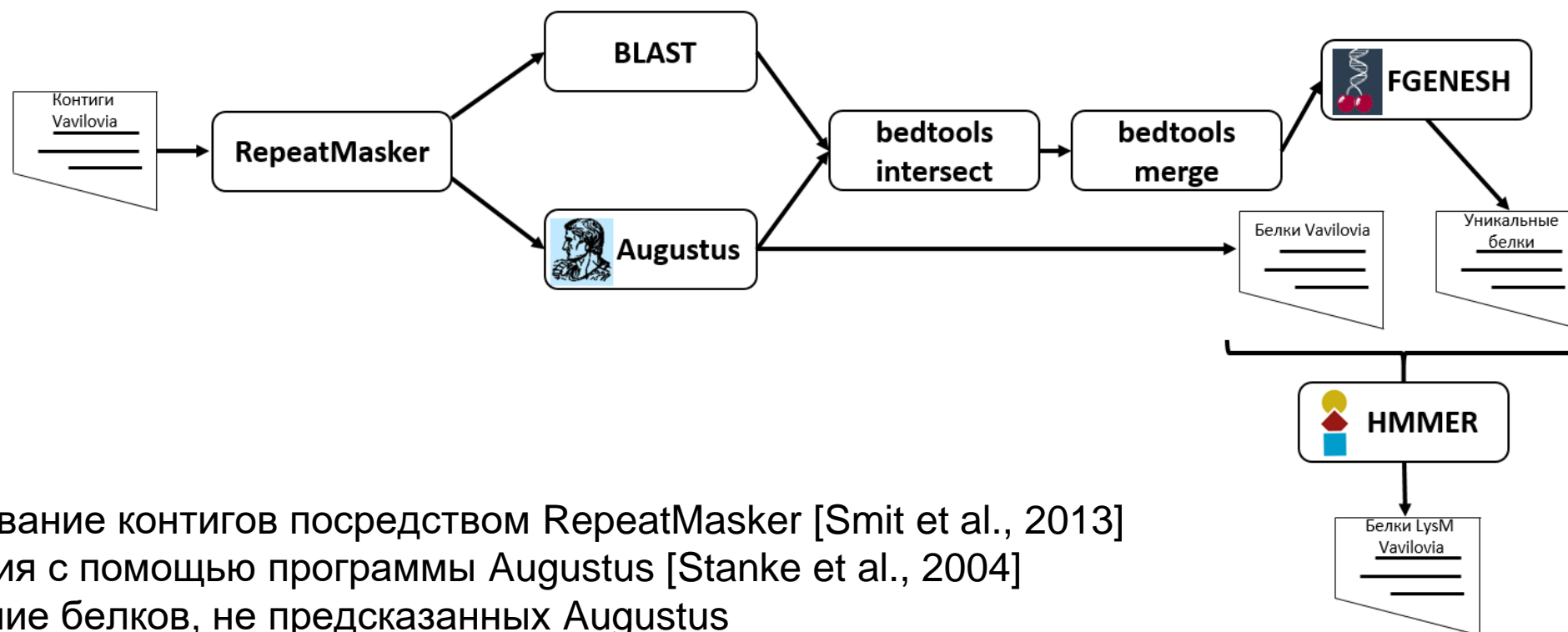
Материалы

В исследовании использовались:

1. Контиги генома *Vavilovia formosa*, собранные сотрудником ИЦиГ СО РАН Генаевым М.А. из прочтений технологий HiSeq и PacBio, предоставленных сотрудниками ФГБНУ “Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии”;
2. Хорошо аннотированные последовательности *LysM*, представленные в работах Buendia et al., 2018 и Zhang et al., 2009, и геномы *Pisum sativum* и *Trifolium pratense*;
3. Два белка водорослей, гомологичных хитиназе *Chlamydomonas*: KAF6263818.1 и KAF5843657.1.

Методы

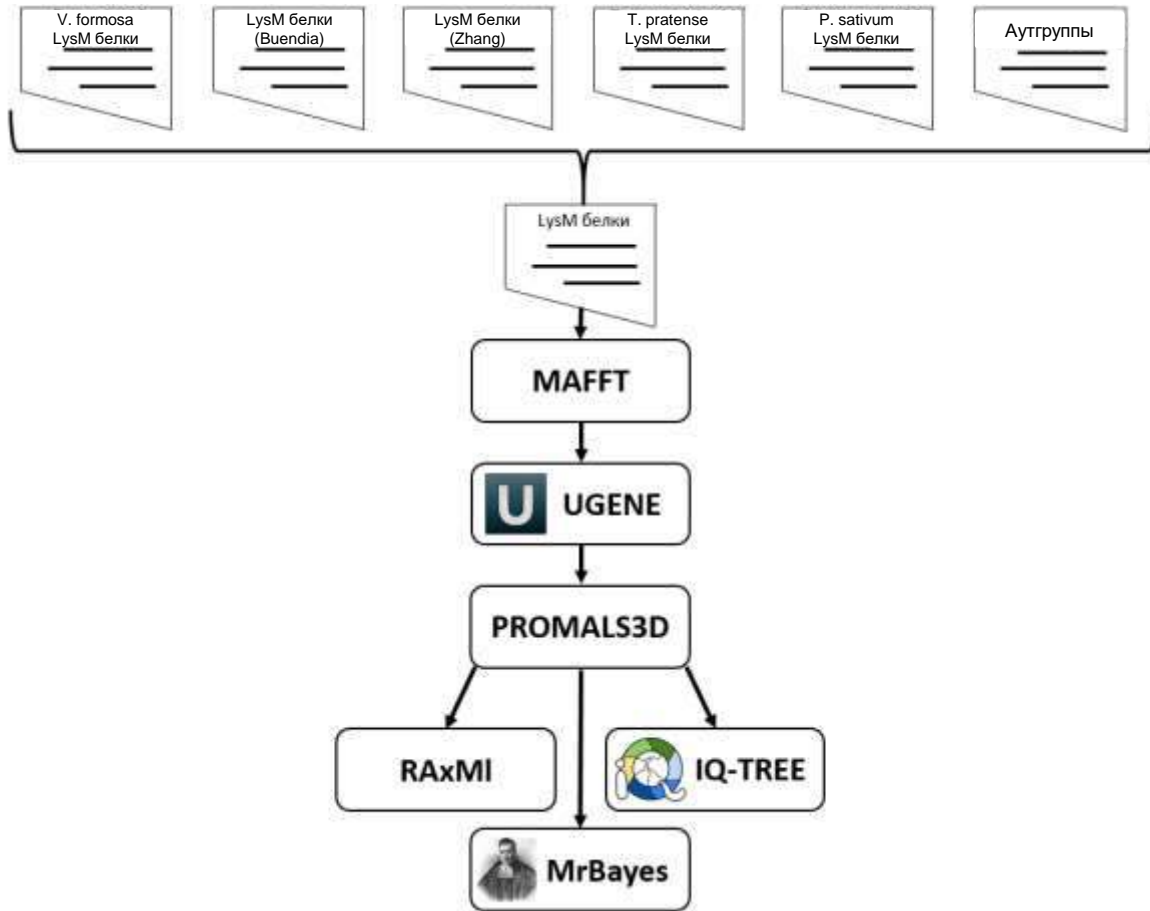
1. Идентификация гомологов LysM-рецепторов в геноме *Vavilovia formosa*



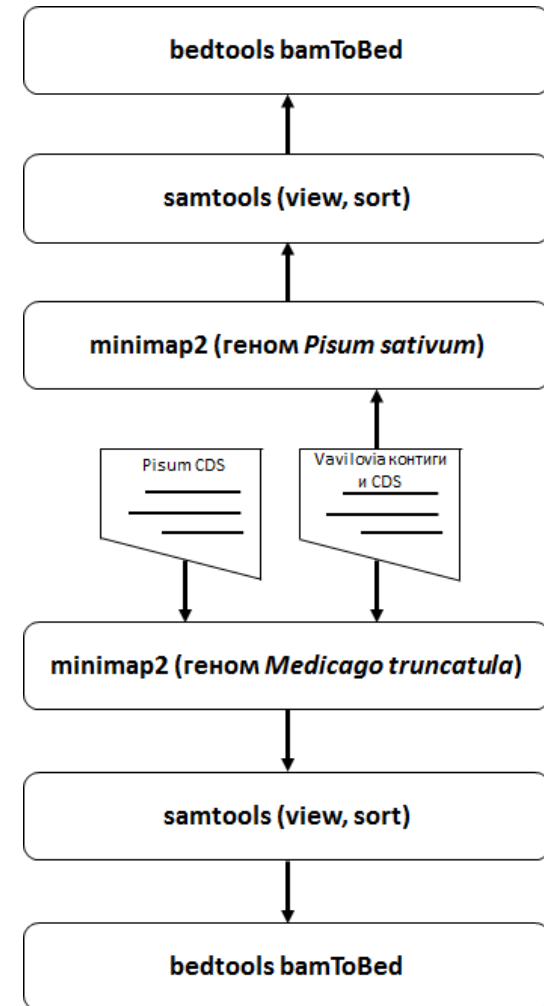
1. Маскирование контигов посредством RepeatMasker [Smit et al., 2013]
2. Аннотация с помощью программы Augustus [Stanke et al., 2004]
3. Выделение белков, не предсказанных Augustus
4. Поиск белков, содержащих LysM и/или домен протеинкиназы
5. Перепроверка предсказанных белков с помощью InterPro [Blum et al. 2021]

Методы

2. Филогенетический анализ LysM белков



3. Анализ картирования геномов



Предсказание генов в геноме *Vavilovia formosa* и сравнение с геномами гороха и клевера

Статистика по *Vavilovia formosa*

Количество контигов	52880
Количество белков после предсказания (Augustus)	35990
Количество белков после предсказания (FGENESH)	102
Количество полученных белков LysM	24
Количество белков LysM после перепроверки (InterPro)	18

Исследованные геномы оказались близки по количеству белков LysM: число последовательностей содержащих LysM домены варьирует в пределах от 24 до 39.

Статистика по *Pisum sativum*

Количество белков в протеоме	57835
Количество предсказанных белков LysM	32
Количество белков LysM после перепроверки (InterPro)	18

Количество последовательностей, у которых представлены все функциональные домены, составляет 18.

Статистика по *Trifolium pratense*

Количество белков в протеоме	41270
Количество предсказанных белков LysM	39
Количество белков LysM после перепроверки (InterPro)	18

Представленность белков LysM в расширенном наборе организмов

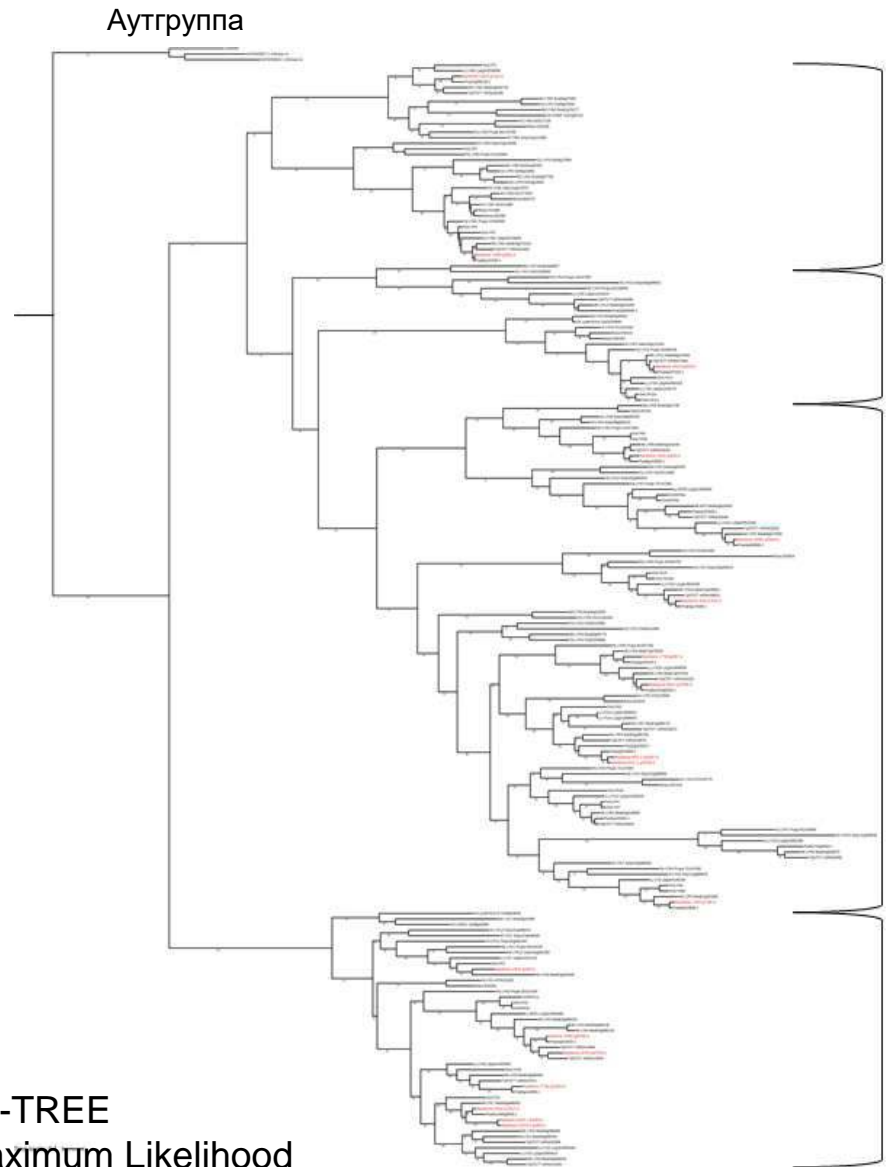
Растение	Число последовательностей
<i>Vavilovia formosa</i>	18
<i>Pisum sativum</i>	18
<i>Trifolium pratense</i>	18
<i>Medicago truncatula</i>	24
<i>Lotus japonicus</i>	19
<i>Prunus persica</i>	15
<i>Arabidopsis thaliana</i>	8
<i>Brassica rapa</i>	10
<i>Oryza sativa</i>	15
<i>Brachypodium distachyon</i>	11
<i>Solanum lycopersicum</i>	17
<i>Glycine max</i>	25
<i>Dunaliella salina</i>	1
<i>Scenedesmus sp</i>	1
<i>Chlamydomonas reinhardtii</i>	1

Всего в филогенетическом анализе использовалось 15 организмов. Из них 12 – высшие растения.

Общее количество последовательностей, использованных для реконструкции, составляет 201.

Данный пул последовательностей послужил нам хорошей основой для реконструкции филогении и классификации LysM белков *V. formosa*.

Филогенетическое дерево LysM белков



Выделяется четыре основных кластера:

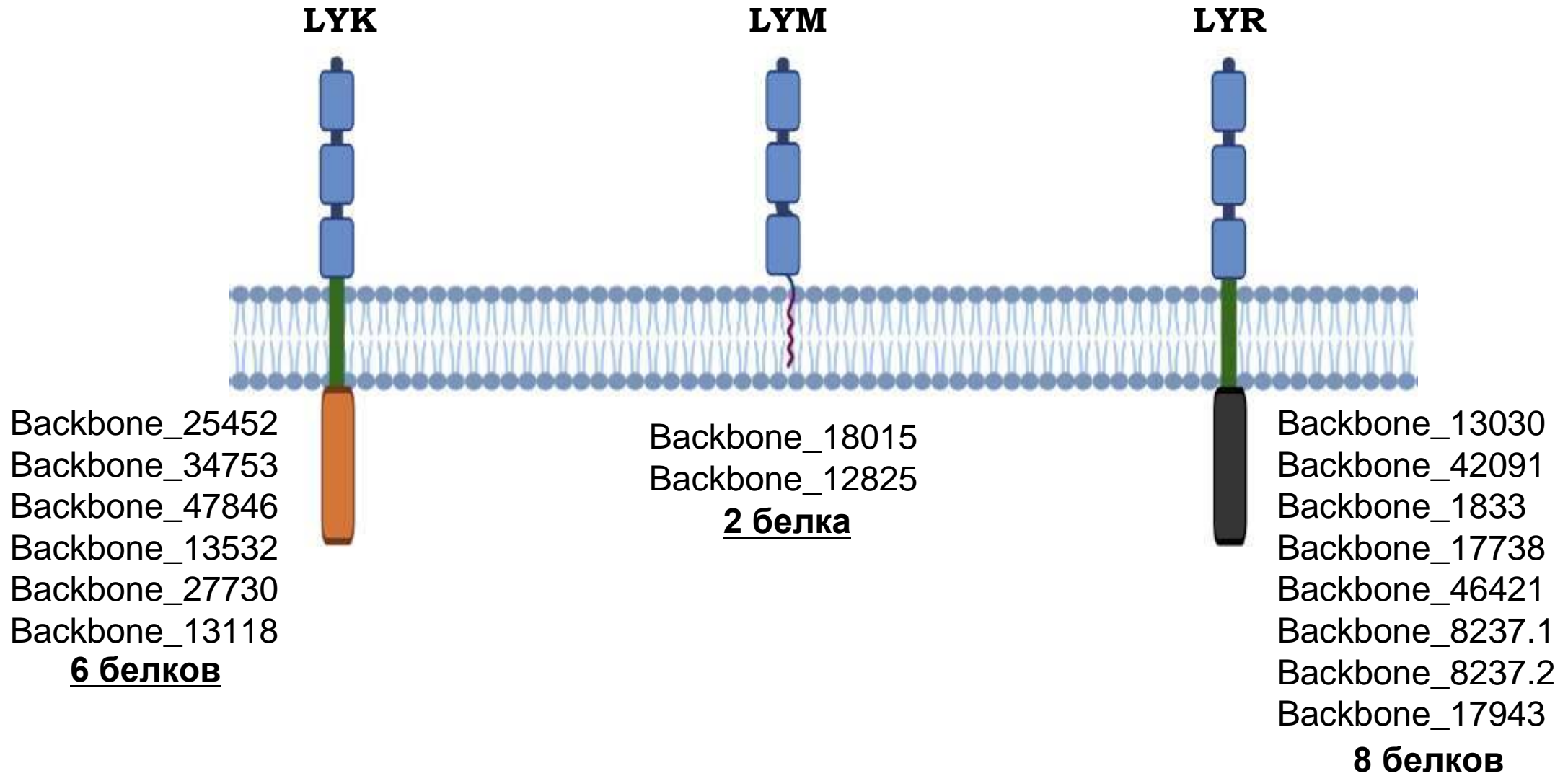
Кластер 1 – LYM белки

Кластер 2 – LYK белки

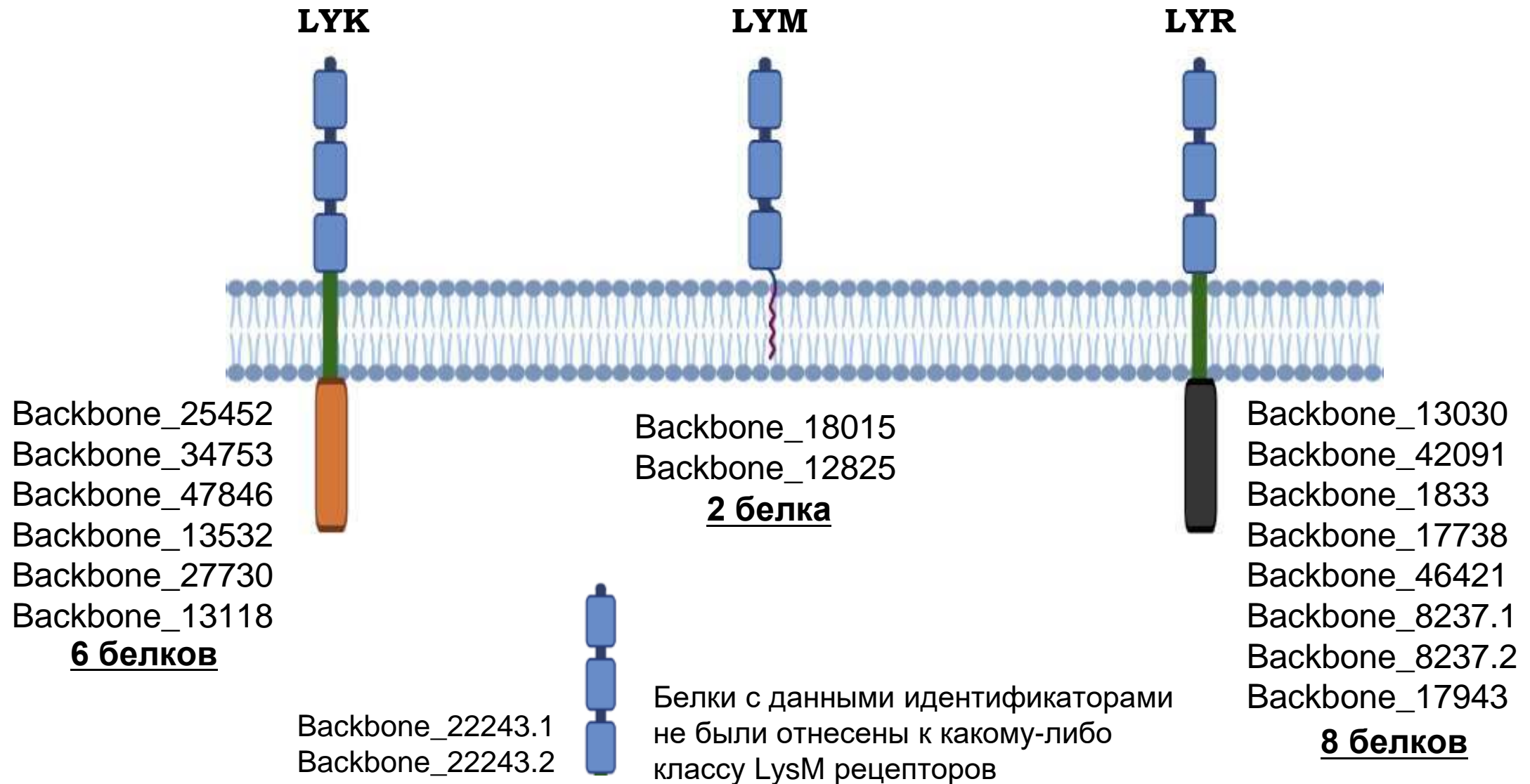
Кластер 3 – LYR белки

Кластер 4 – LYK белки

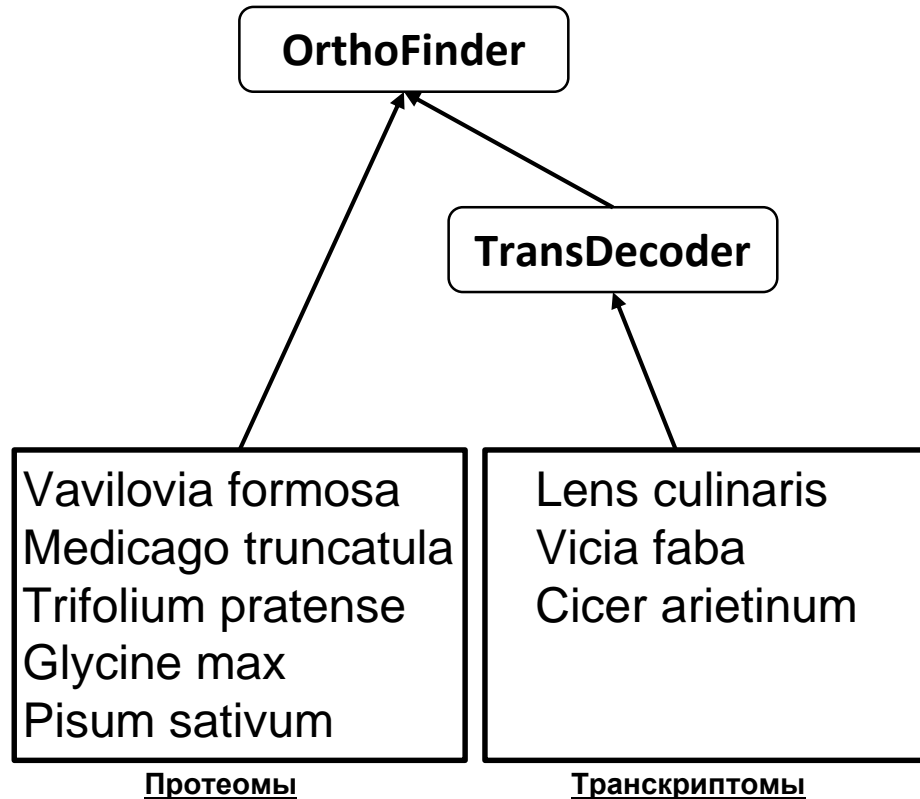
Классификация LysM белков *Vavilovia formosa*



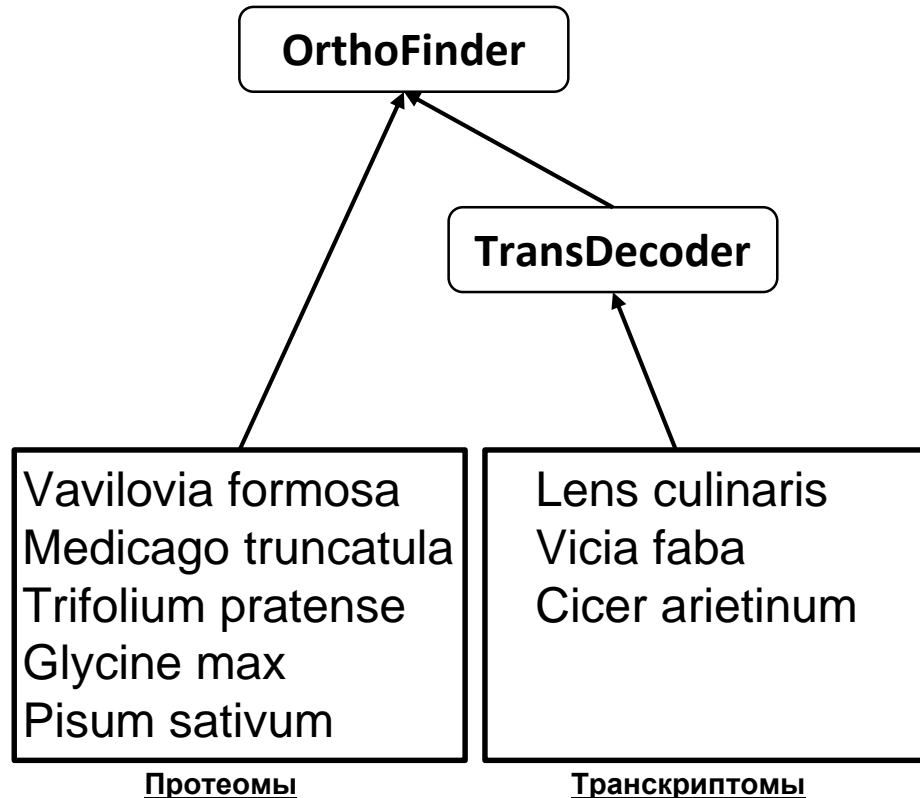
Классификация LysM белков *Vavilovia formosa*



Ортологические группы клеверных и бобовых



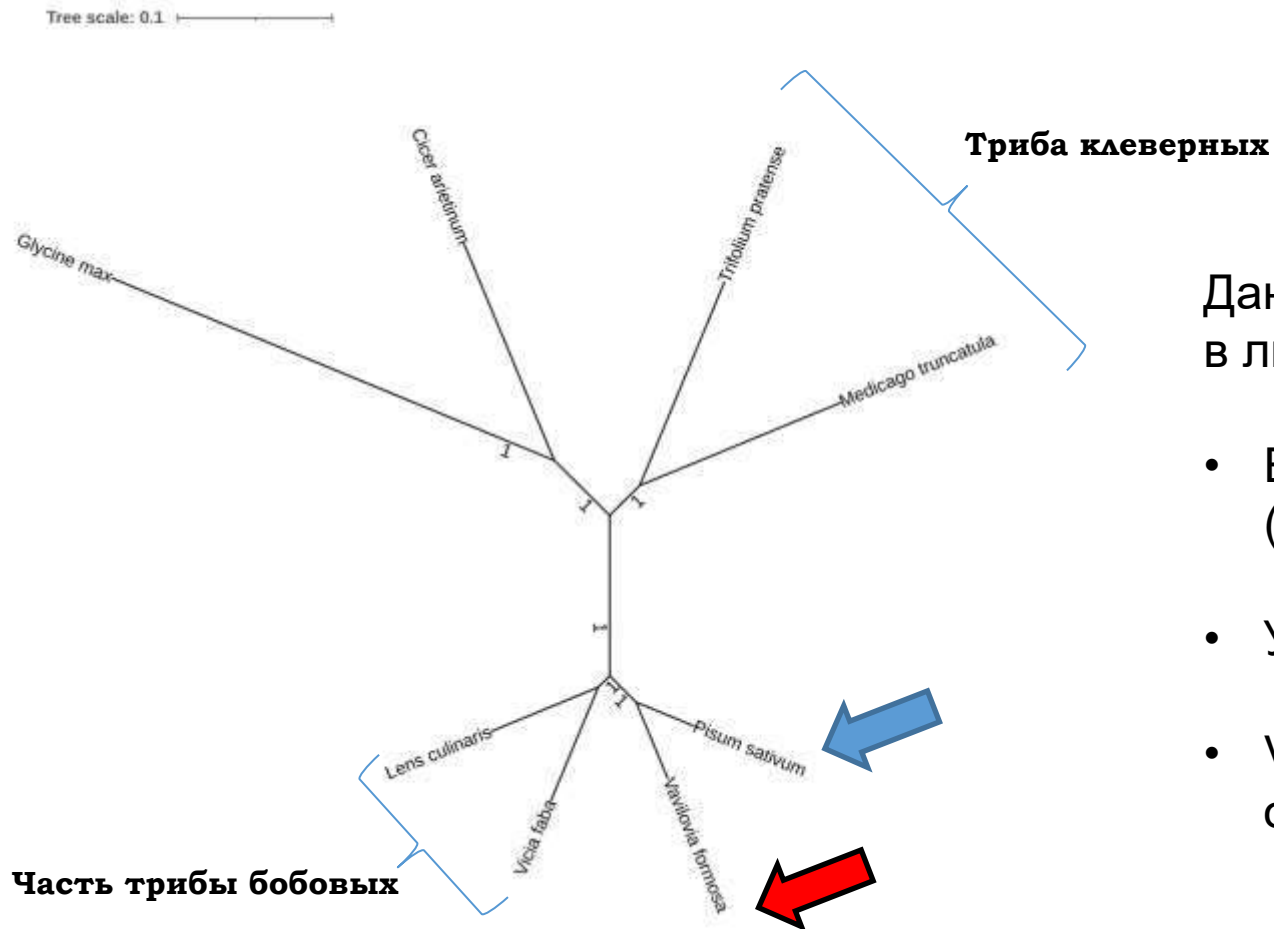
Ортологические группы клеверных и бобовых



Организм	Изначальные данные	Кол-во посл. до предсказания	Кол-во посл. после предсказания	Размер протеома, Мб	База данных
Medicago truncatula	Протеом	41939	41939	23.0	NCBI (MedtrA17_4.0)
Trifolium pratense	Протеом	41270	41270	22.6	Ensembl
Glycine max	Протеом	71219	71219	37.7	NCBI (Glycine_max_v2.1)
Cicer arietinum	Транскриптом	56176	59890	28.5	Pulse Crop Database
Vicia faba	Транскриптом	37378	43617	20.4	Pulse Crop Database
Pisum sativum	Протеом	57835	57835	20.4	URGI
Lens culinaris	Транскриптом	33613	38644	17.7	Pulse Crop Database
Vavilovia formosa	Геном	52880	35990	11.8	-

Статистика по протеомам и транскриптомам

Ортологические группы клеверных и бобовых

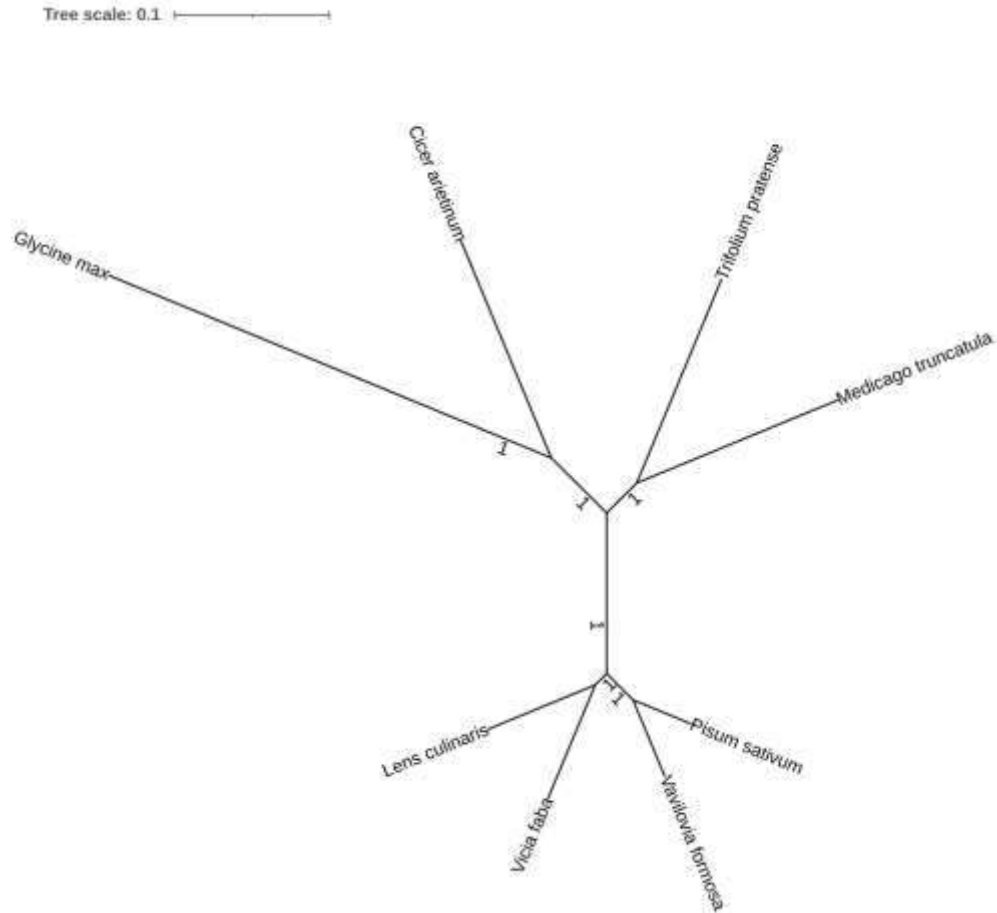


Данная топология соответствует принятой топологии в литературных данных, где:

- Бобовые разделяются на трибу клеверных (Trifolieae) и бобовых (Fabaeae)
- Укореняет две трибы *G. max* и *C. arietinum*
- *V. formosa* является сестринским организмом по отношению к *P. sativum*

Неукорененное дерево, построенное на основе 217 уникальных ортогрупп (-m MSA)

Ортологические группы клеверных и бобовых

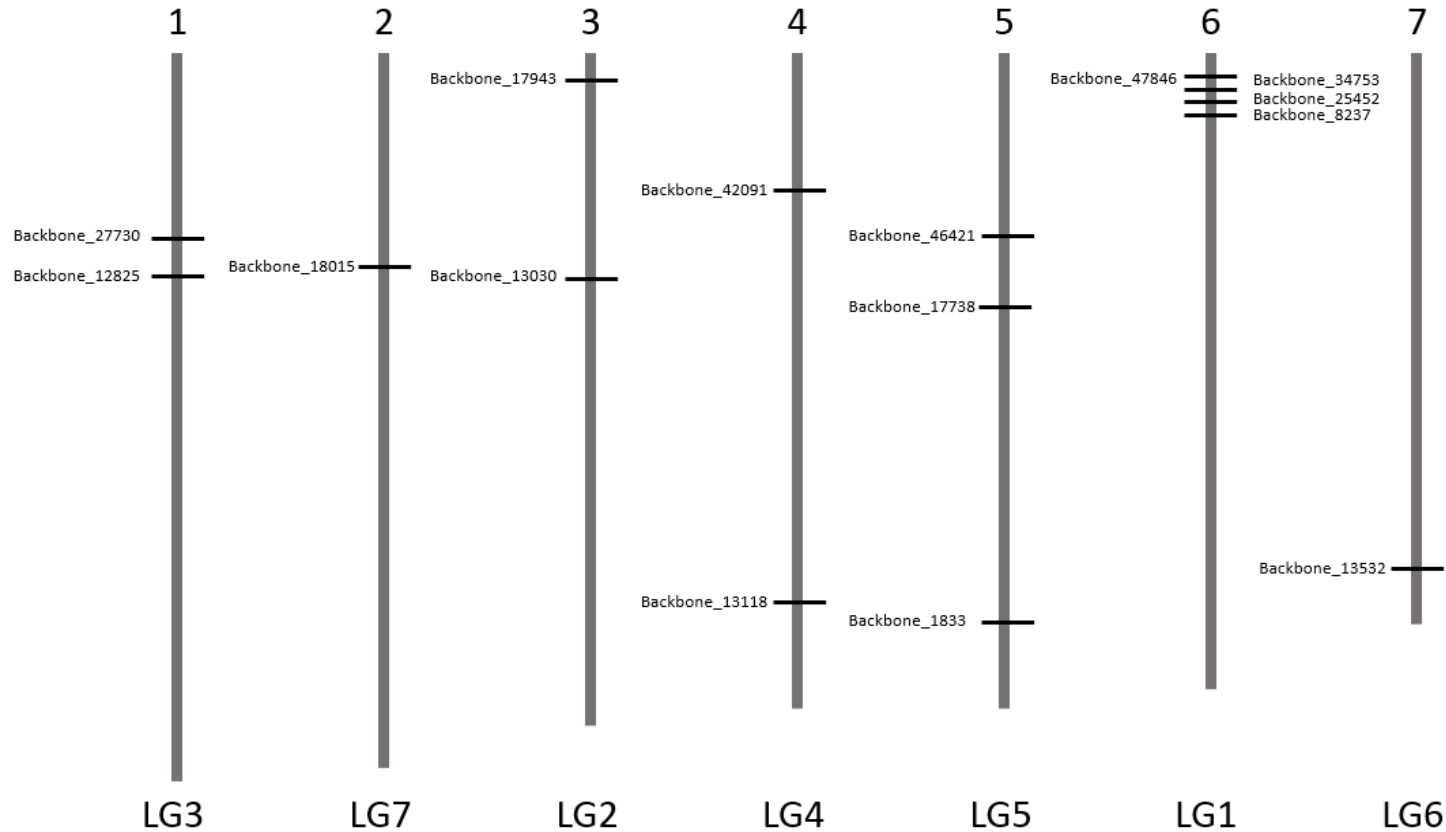


В таблице присутствуют все белки LysM *V. formosa* и *T. pratense*, определенные ранее

Ортогруппа	<i>Vavilovia formosa</i>	<i>Pisum sativum</i>	<i>Medicago truncatula</i>	Клада
OG0010100	Backbone_18015	Psat7g088120.1	MTR_4g094730	LYMII
OG0010030	Backbone_12825	Psat5g132320.1	MTR_3g072410	LYMI
OG0004500	Backbone_17738 Backbone_46421	Psat3g105440.1 Psat0s1633g0040.1	MTR_7g079320 MTR_7g079350	LYRIV
OG0008386	Backbone_42091	Psat4g090800.1 Psat2g162600.1	MTR_8g078300 MTR_5g019040	LYRI A
OG0019933	Backbone_1833	Psat3g176480.1	MTR_7g029650	LYRII A
OG0012903	Backbone_17943	Psat6g010960.1	MTR_1g021845	LYRIII B
OG0000465	Backbone_34753 Backbone_27730 Backbone_25452 Backbone_47846 Backbone_13532	Psat0s1360g0040.1 Psat1g177920.1 Psat2g024320.1 Psat5g112080.1	MTR_3g080050 MTR_5g086030 MTR_2g024290 MTR_5g086310	LYKI
OG0004453	Backbone_13118	Psat4g197320.1	MTR_8g014500	LYKIII
OG0008082	Backbone_22243.1 Backbone_22243.2		MTR_5g086660	LYKI
OG0010518	Backbone_8237.1 Backbone_8237.2	Psat2g024920.1 Psat2g024880.1	MTR_5g085790	LYRIII C
OG0017145	Backbone_13030	Psat6g102960.1	MTR_5g042440	LYRI B
OG0034034			MTR_3g080170	LYRIII C
OG0015210		Psat0s744g0040.1	MTR_4g058570	LYRII B
OG0013118		Psat2g162560.1	MTR_5g019050	LYRIII A
OG0013229	Backbone_47588	Psat2g094040.1	MTR_5g033490	LYKII
OG0023994			MTR_5g086040	LYKI
OG0017671			MTR_5g086540 MTR_5g086090 MTR_5g086110 MTR_5g086120 MTR_5g086130	LYKI

Ортологические группы, содержащие белки LysM, полученные с помощью OrthoFinder

Потенциальное расположение генов LysM в геноме *Vavilovia formosa*

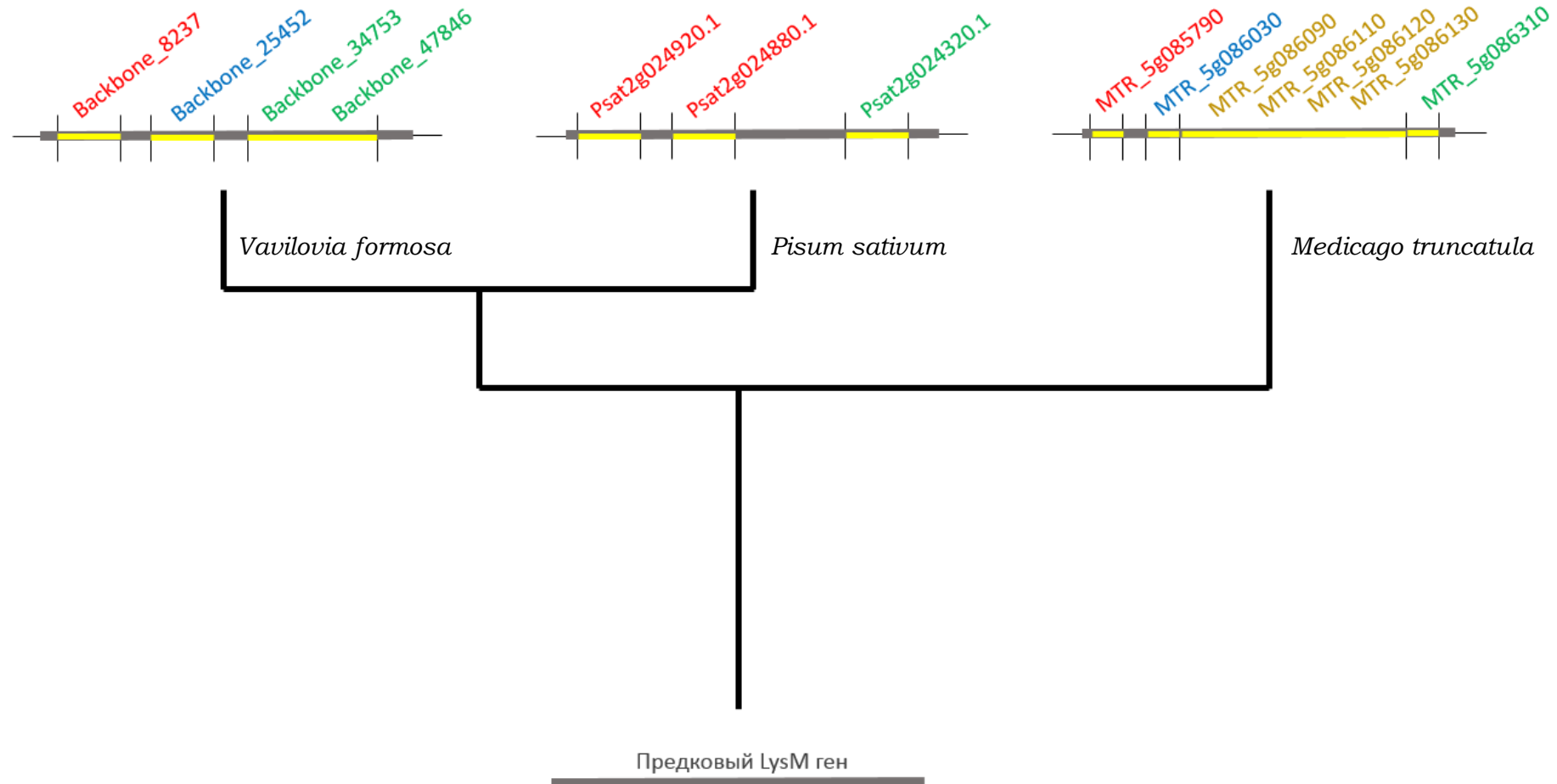


Расположение генов LysM на хромосомах *V. formosa*. Сверху показаны номера хромосом *V. formosa*, снизу – номера хромосом *P. sativum*

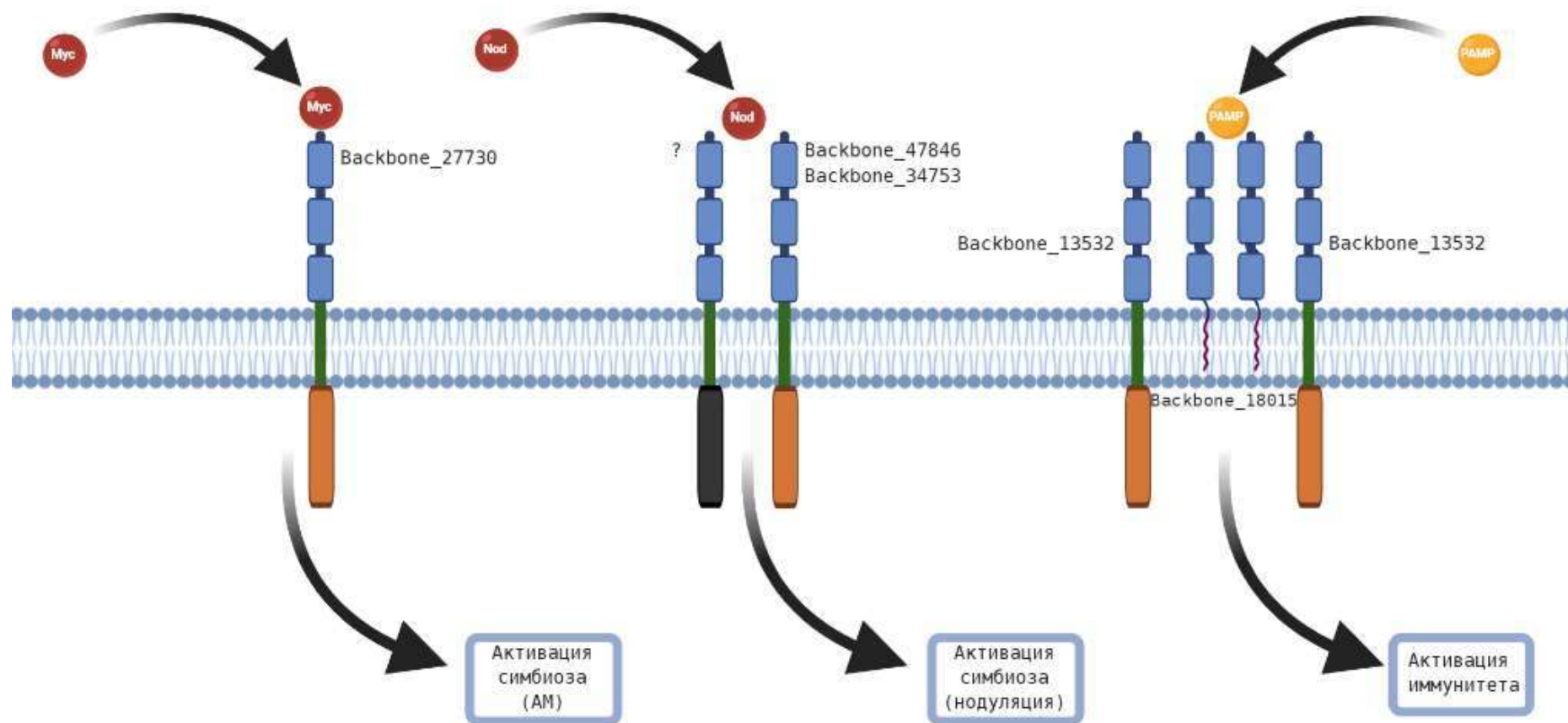
Результаты основаны на:

- Результаты анализа ортогрупп
- Результаты картирования CDS и контигов вавиловии и CDS гороха на референсный геном *Medicago truncatula*
- Результаты картирования CDS и контигов вавиловии на референсный геном *Pisum sativum*
- Литературных данных о количестве хромосом
- Литературных данных с метафазной пластинкой (для примерного определения соотношения размеров)

Эволюция синтенного блока LysM генов



Функции белков *Vavilovia formosa*, предсказанные на основе гомологии



Выводы

1. На основе компьютерной аннотации сборки генома *V. formosa* выявлено 35990 генов, кодирующих белки;
2. Впервые идентифицирован репертуар LysM генов *V. formosa*, включающий 18 последовательностей;
3. На основе филогенетического анализа проведена классификация белковых последовательностей LysM генов *V. formosa* и выявлено три их семейства: LYK (6 белков), LYR (8 белков), LYM (2 белка); данная классификация была подтверждена на основе поиска ортологических групп белков в протеомах и транскриптомах бобовых; отдельно были выявлены два белка *V. formosa* с доменами LysM, отнести которые к семействам LysM не представилось возможным;
4. На основе анализа ортологических групп и анализа синтении геномов была сформулирована гипотеза о расположении генов LysM на хромосомах *V. formosa*.

Благодарю за внимание!

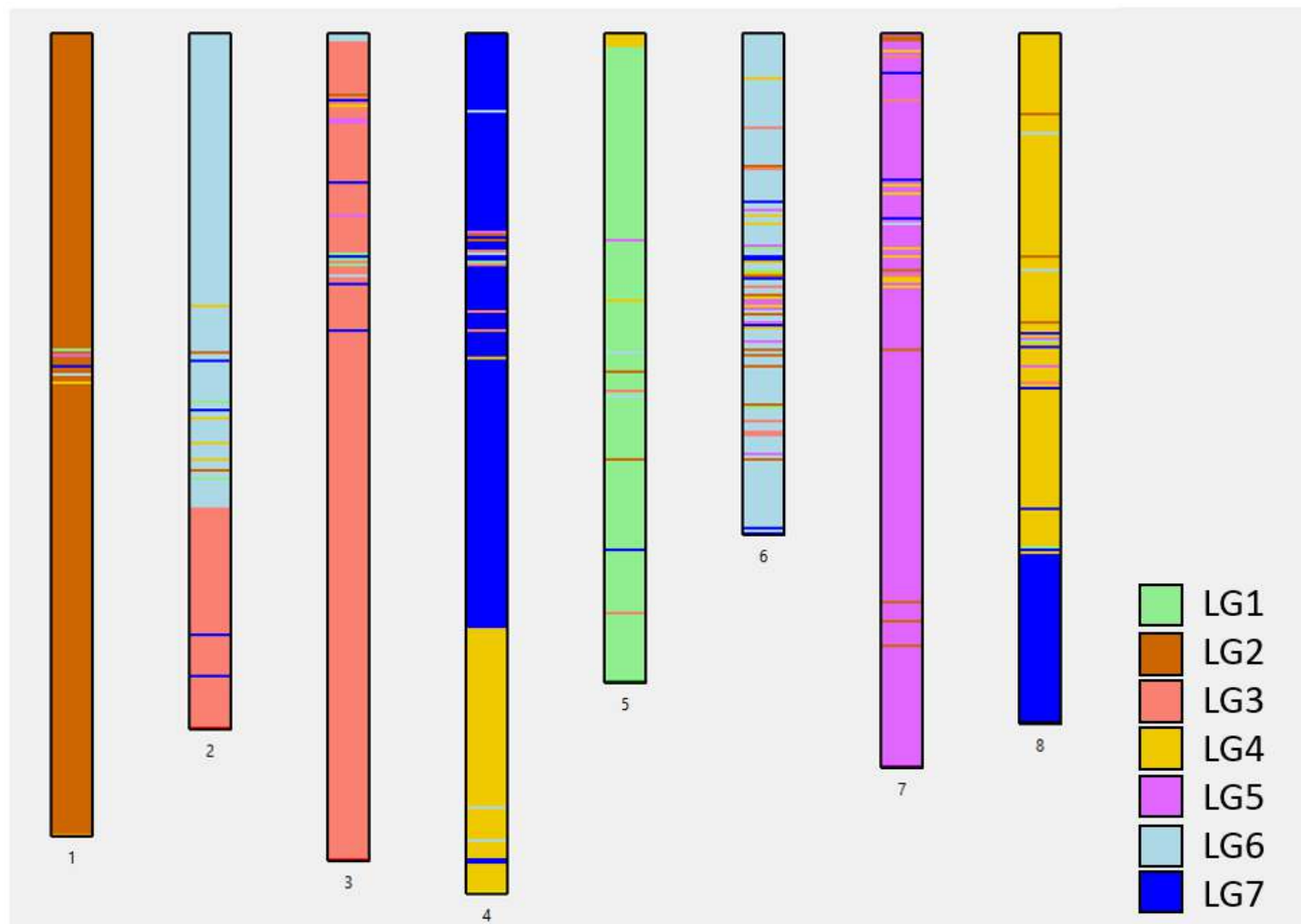


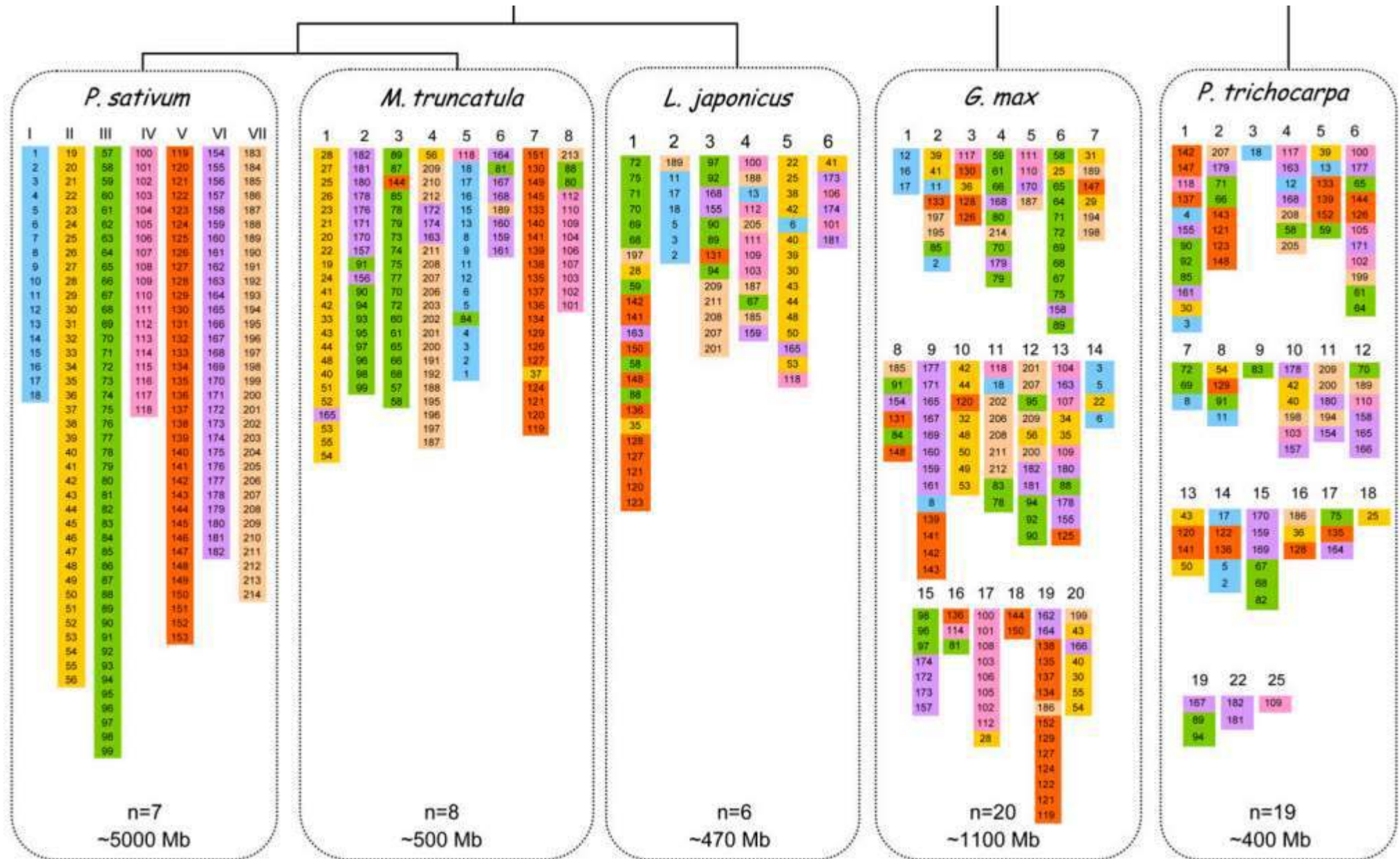
Контакты:

Каретников Дмитрий Игоревич

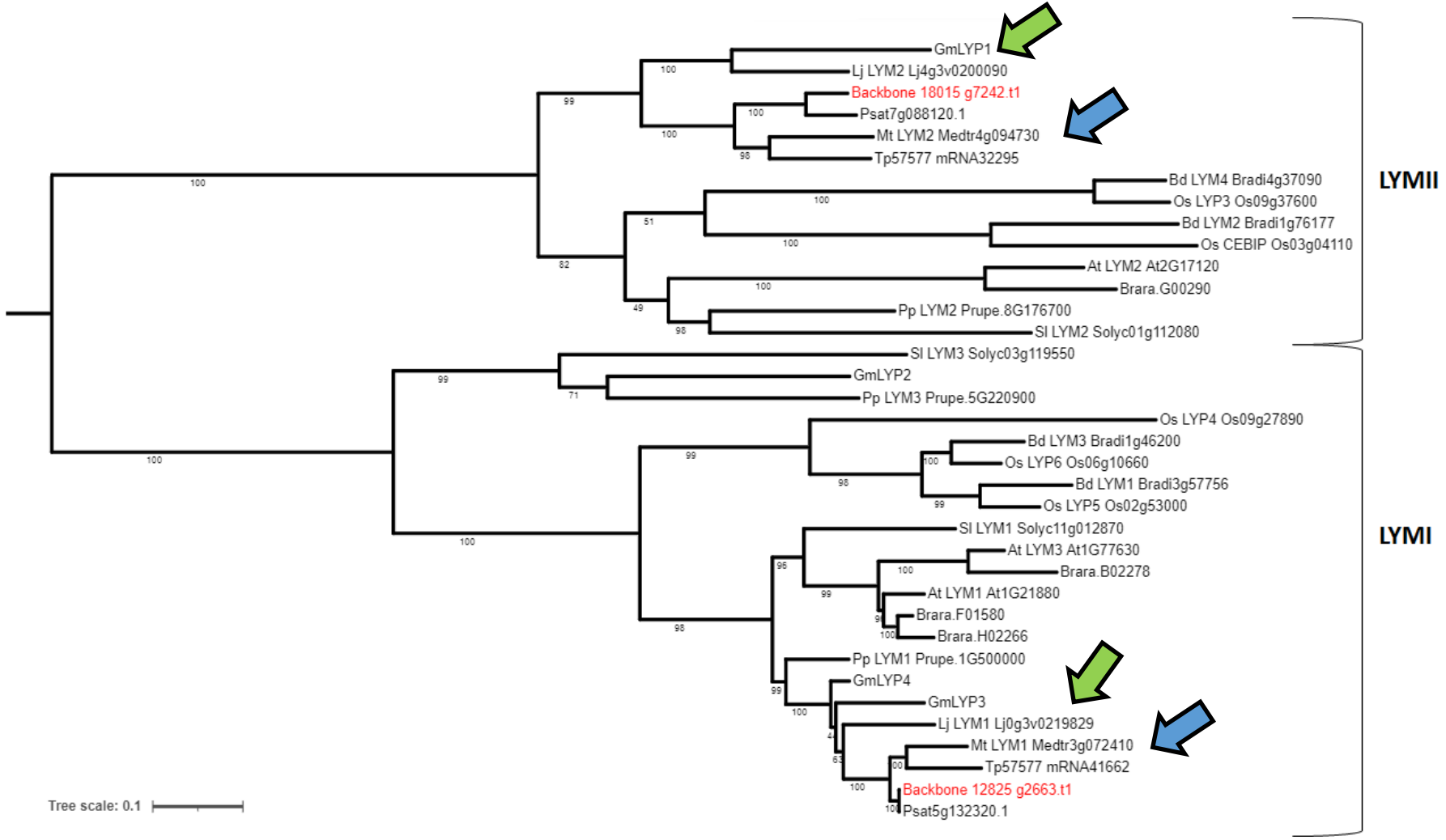
karetnikovmit@bionet.nsc.ru

Результаты картирования





Филогенетические отношения в белках семейств LYM



Филогенетическое описание белков семейства LYK

