

ВВЕДЕНИЕ В ИНФОРМАЦИОННУЮ БИОЛОГИЮ

Н.А. Колчанов, С.А. Лашин

Электронно-лекционный курс
разработан в рамках реализации Программы развития НИУ-НГУ
2012 год

Лекция 11
Биоинформатика в генетике растений

Модельные объекты — это организмы, используемые в качестве моделей для изучения тех или иных свойств, процессов или явлений живой природы.

Использование модельных организмов основано на том, что все живые организмы имеют общее происхождение и сохраняют много общего в механизмах хранения и реализации наследственной информации, метаболизме и др.

можно предположить что открытые при их изучении закономерности окажутся свойственны и другим более или менее похожим организмам.



Многообразие модельных организмов среди животных:



Nematostella vectensis,
(актиния)



Nematostella vectensis,
(моллюск)



Xenopus laevis,
(земноводное)



Danio rerio,
(рыба)



Mus musculus,
(млекопитающее)



Pan troglodytes,
(млекопитающее)

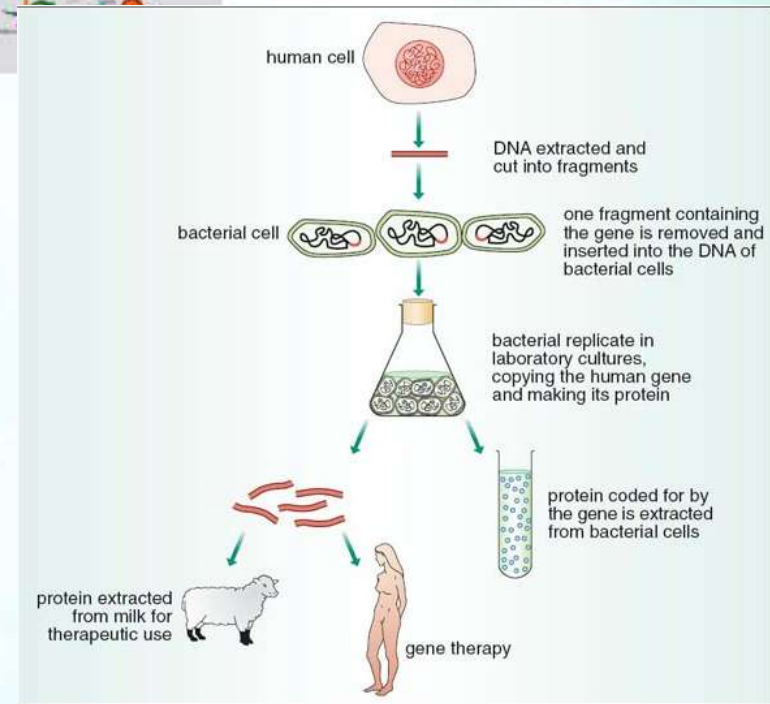
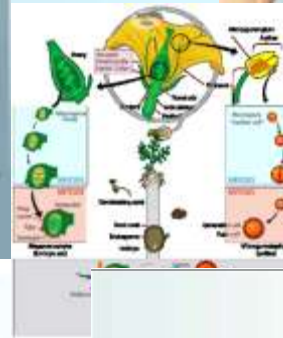
Характеристики, важные для модельного объекта:

1) легкость содержания
и разведения в лабораторных условиях

2) короткое время генерации
(быстрая смена поколений),

3) Пригодность для генетических исследований
и генетических манипуляций
(наличие инбредных линий,
в случае многоклеточных
возможность получения стволовых клеток,
разработанные методы
генетической трансформации).

4) положение на филогенетическом древе



Drosophila melanogaster - Чернотелая дрозофила



Используется для изучения генетики и биологии развития животных

Жизненный цикл - 10 - 20 дней

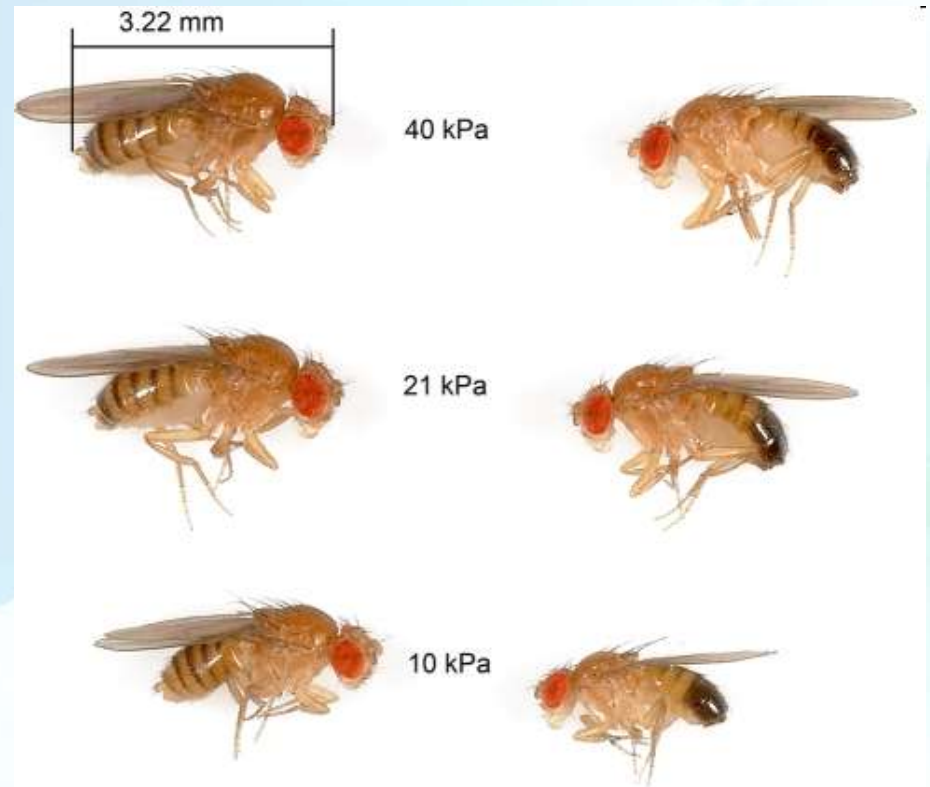
Количество потомков от 1 пары - до 400

Кариотип - $2n = 8$

Геном - 132 Млн.П.О.

приблизительно 14 000 генов.

секвенирован и аннотирован.



(Klok CJ, Harrison JF. 2009)

Arabidópsis thaliána - Резуховидка Таля



Используется для изучения генетики и биологии развития растений

Жизненный цикл – 21 - 32 дня

Количество потомков от 1 растения
– до 1000

Кариотип $2n = 10$

Геном состоит из порядка 157 миллионов пар оснований

приблизительно 27 000 генов.
секвенирован и аннотирован.

Drosophila melanogaster

Ж.ц. - 12 - 20 дней

Кол-во потомков - до 400

Кариотип - $2n = 8$

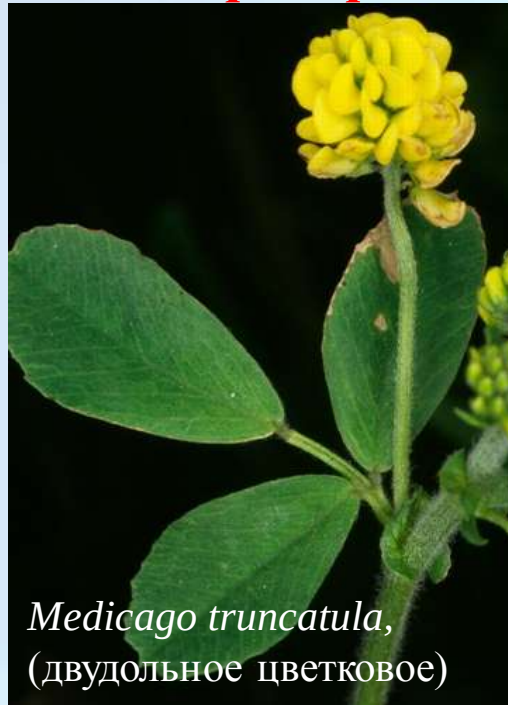
Геном - 132 Млн.П.О.

приблизительно 14 000 генов.
секвенирован и аннотирован.

Многообразие модельных объектов среди растений:



Physcomitrella patens,
(зелёный мох)



Medicago truncatula,
(двудольное цветковое)



Oryza sativa,
(Однодольное цветковое)



Zea mays L.,
(однодольное цветковое)



Selaginella moellendorffii,
(плаун)



Populus tremula,
(двудольное цветковое)

Triticum aestivum L.

Мягкая пшеница

Важен для человечества
как источник пищевых ресурсов

Стабильный геном большого размера
естественного происхождения,

Аллогексаплоид – 42 хромосомы
(по 7 пар А В D «геномов»)

Размер генома:

17 Gb = 17 000 000 000 п.н.

= 5 x Homo sapiens

= 108 x Arabidopsis thaliana

= 3 500 x Escherichia coli



Почему работают с растительными объектами?

Важность растений как пищевых р

Экономическая значимость.

Значимость для экосистемы.

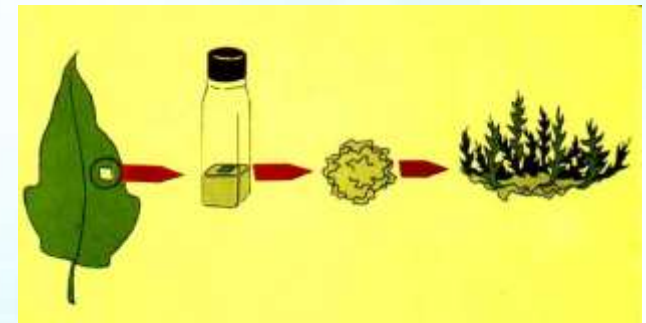


Удобство в культивировании.

- Неподвижны (относительно),
- позволяют в одну стадию восстановить целый организм из 1 клетки

Удобство для генетических манипуляций.

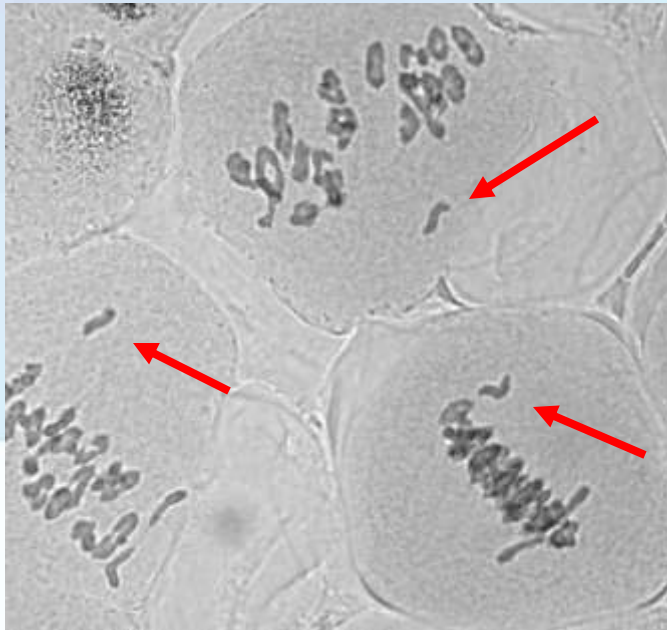
- Особи с изменённым числом хромосом часто жизнеспособны и стабильны
- Большое количество методов трансформации (например: бактериальная, баллистическая)



Классическая генетика и цитогенетика растений. методы

Скращивания

Манипуляции с хромосомами :



Мейоз моносомика *T.aestivum* по 7В
хромосоме (фото 2010 год)

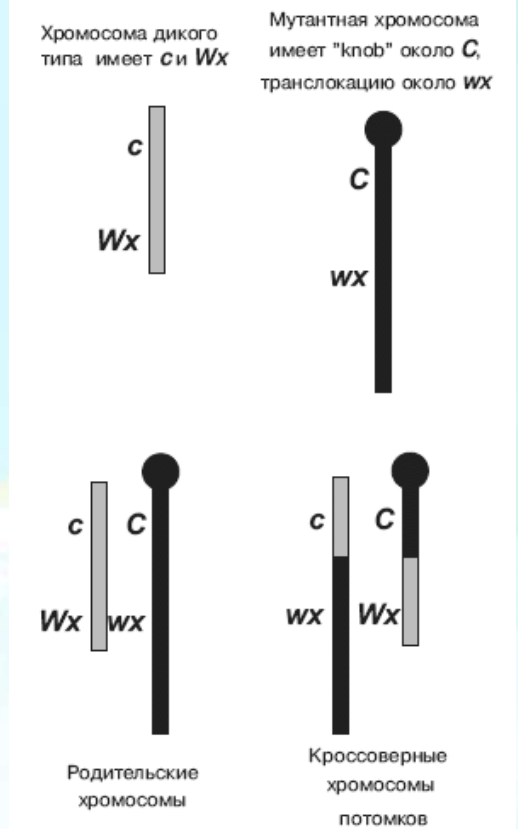


Рис 3.3. Схема кроссинговера в IX паре хромосом кукурузы в опытах Крейтона и МакКлинток (из: Lewin, 1994, р. 68)

Иллюстрация из книги
И.Ф. Жимулёва
«Общая и молекулярная генетика»

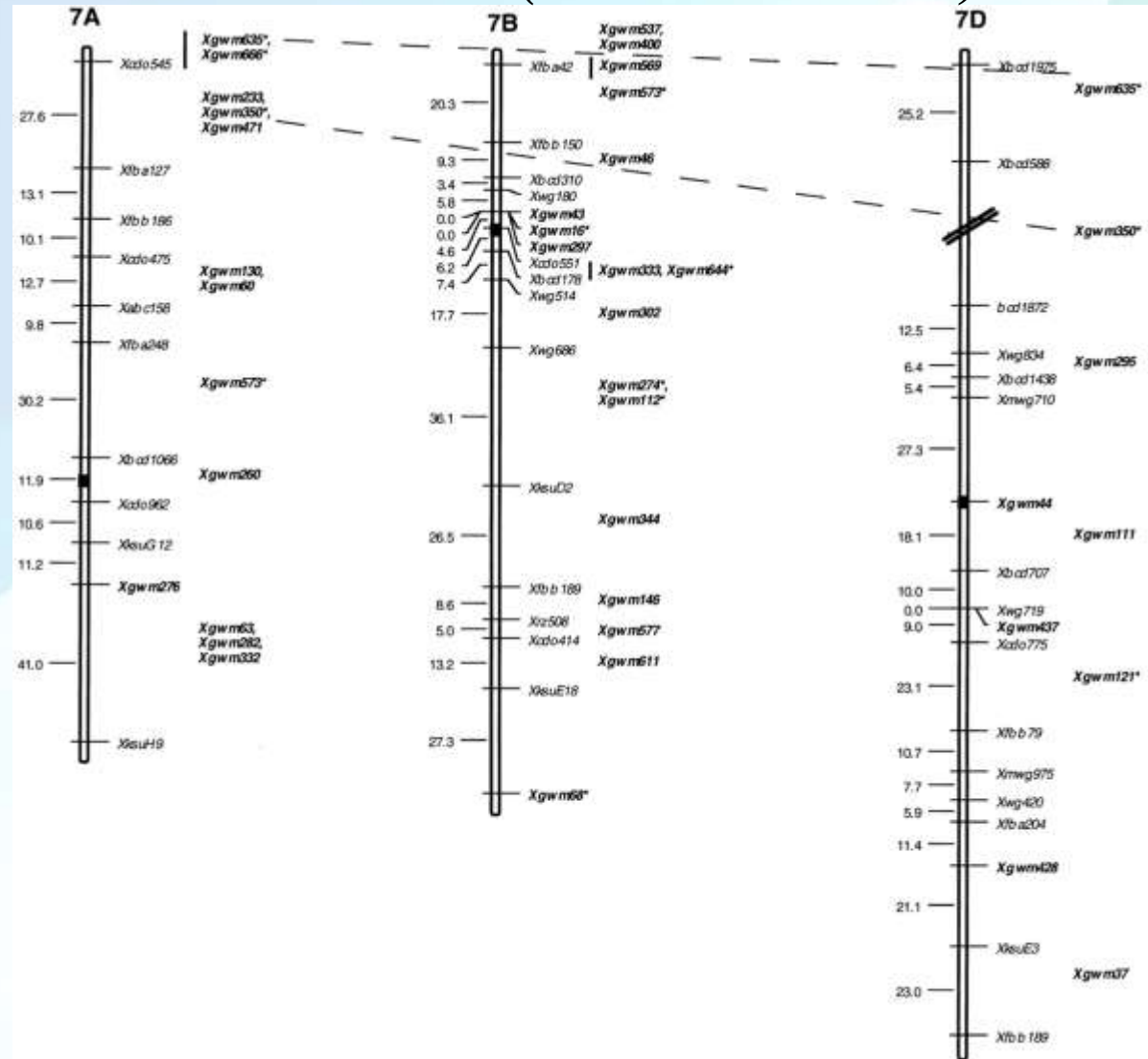
Молекулярная генетика и цитогенетика растений. методы

Работа с геномом

- картирование генов,
получение
последовательностей

- исследование экспрессии
генов

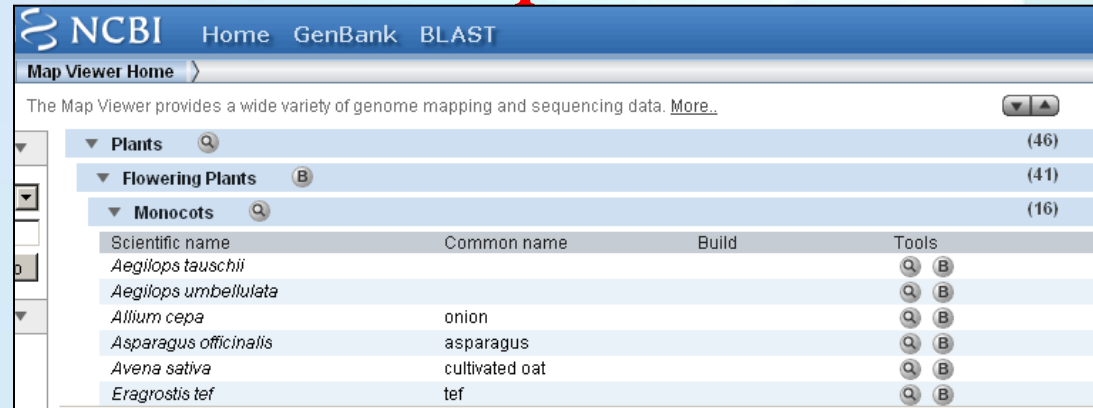
A Microsatellite Map of Wheat (Marion S. Röder et. al.)



Системная биология в генетике растений.

Базы данных

(нуклеотидные,
белковые,
фенотипические)



Scientific name	Common name	Build	Tools
<i>Aegilops tauschii</i>			Q B
<i>Aegilops umbellulata</i>			Q B
<i>Allium cepa</i>	onion		Q B
<i>Asparagus officinalis</i>	asparagus		Q B
<i>Avena sativa</i>	cultivated oat		Q B
<i>Eragrostis tef</i>	tef		Q B

Автоматическое фенотипирование

(цифровое описание признаков
(фенотипа) при помощи
современных инструментальных
методов)

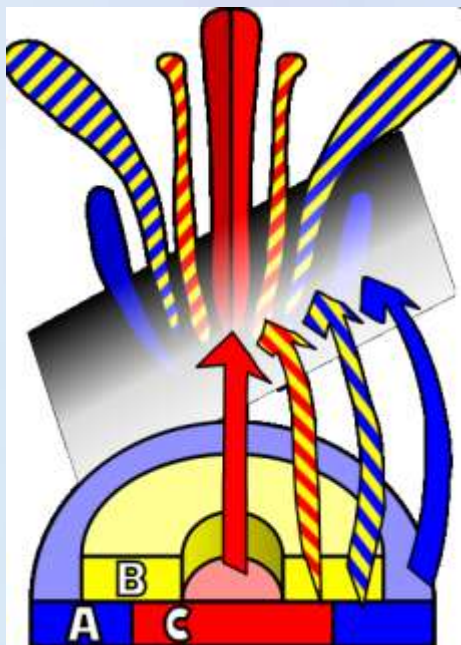


Моделирование процессов развития

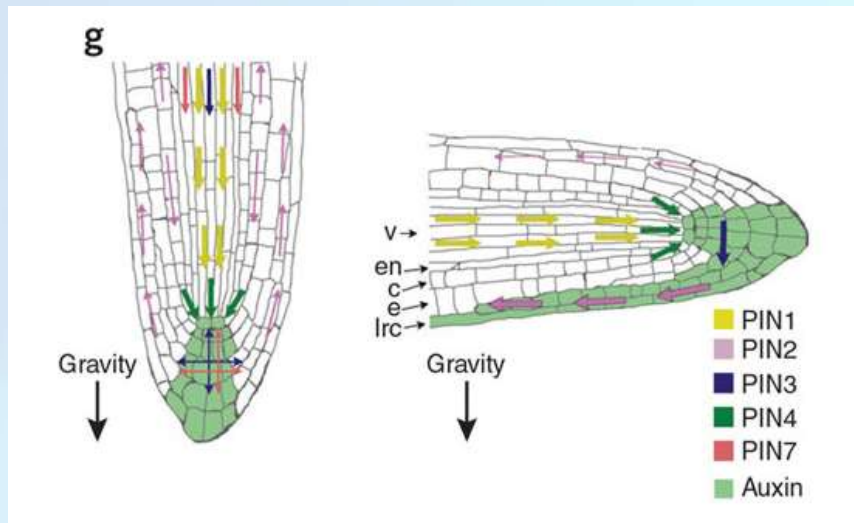
Генетика развития растений.

Исследование морфогенеза растительных органов

- выявление ключевых морфогенов и механизмов их работы



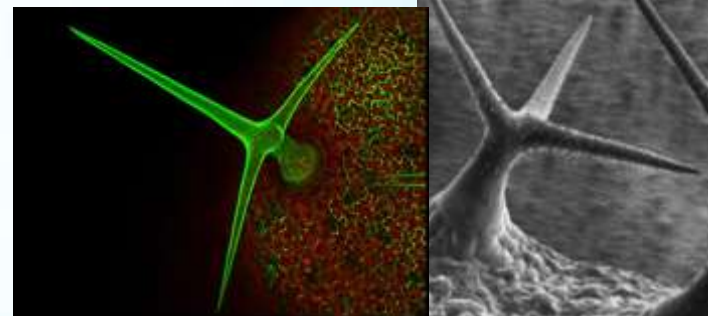
Модель развития цветка
(Enrico Coen and Elliot
Meyerowitz, 1991)



Модель реализации
положительного
геотропизма корня
(Hélène S Robert
& Jiří Friml 2009)



Трихомы cannabis sativa



Трихомы Arabidopsis thaliana

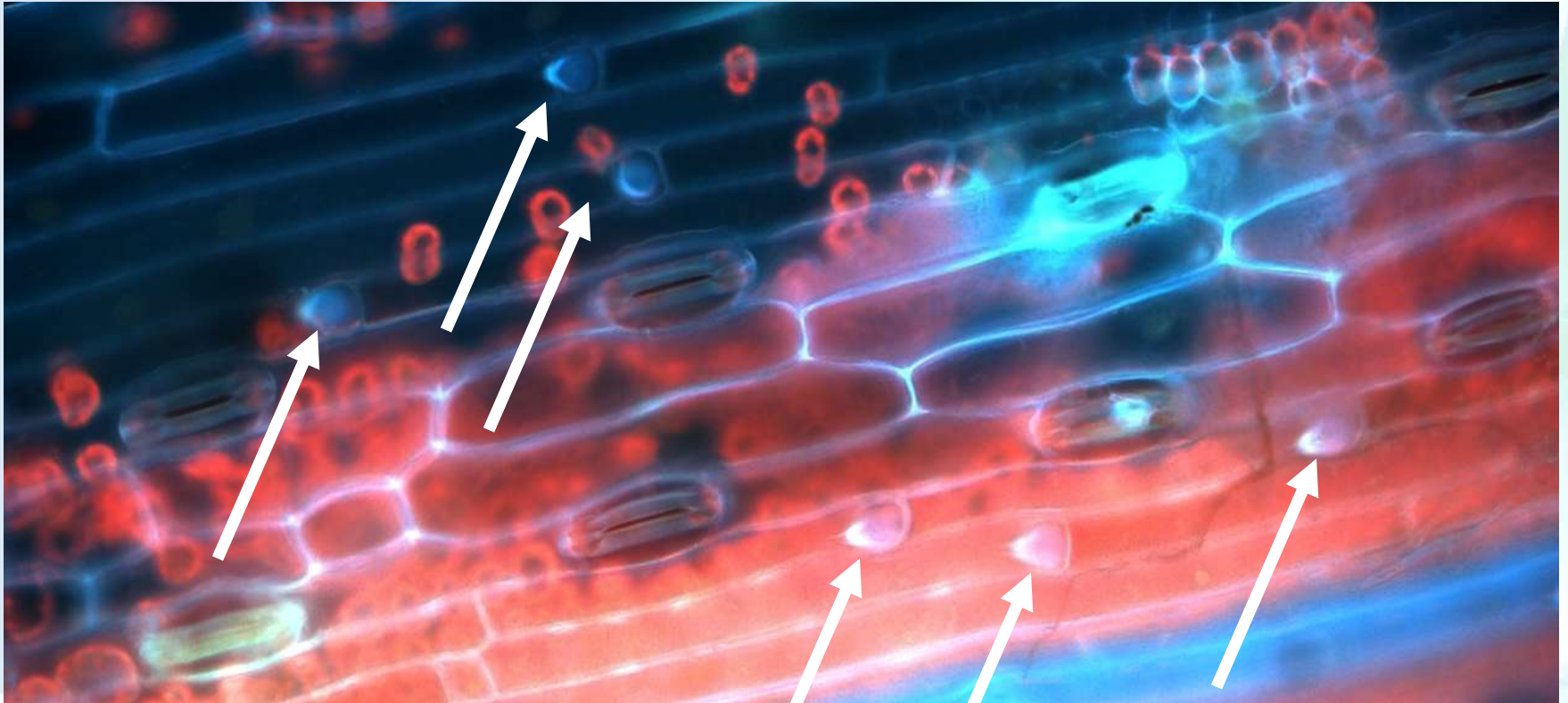
Опушение (органа растения) – совокупность поверхностных образований преимущественно эпидермального происхождения (трихом)



- Опушение листа у пшеницы имеет большое адаптационное значение.
- Сильное опушение характерно для ряда засухоустойчивых сортов, а для сортов, произрастающих во влажном климате, напротив, характерно очень слабое опушение
(Крупнов В.А., Цапайкин А.П. 1990)
- Сильно опушённые сорта пшеницы значительно более устойчивы к ряду вредителей, например:
жуку – листоеду *Oulema melanopus* L. (Schillinger J.A., Gallun R.L. 1968)
гессенской мухе *Mayetiola destructor* (Roberts J.J., Gallun R.L., Patterson F.L., Forster J.E. 1979)
- Опушение имеет также физиологическое значение. Предполагается, что клетки трихом, основной объём которых вынесен за плоскость эпидермы, расширяют возможности формирования локальных градиентов тургорных давлений эпидермиса, а также снимают нарастающее в онтогенезе напряжение покровной ткани, увеличивая таким образом эффективность работы устьичного аппарата (Журавлёва, 1992).

Анализ опушения листа:

Флуоресцентная микрофотография поверхности листа



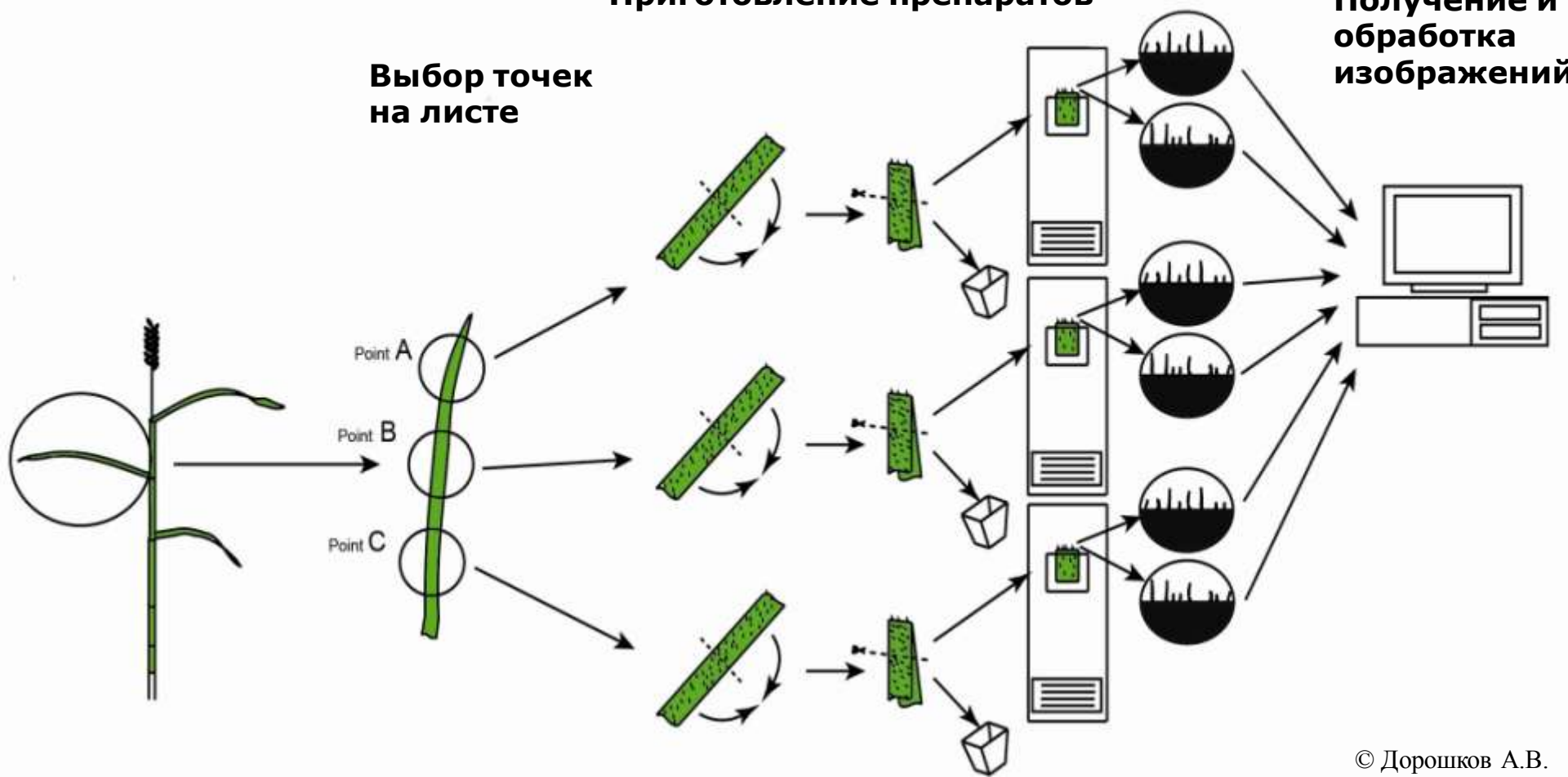
Разработка технологии высокопроизводительного фенотипирования для анализа опущения листа пшеницы.

Выбор листа

Выбор точек на листе

Приготовление препаратов

Получение и обработка изображений

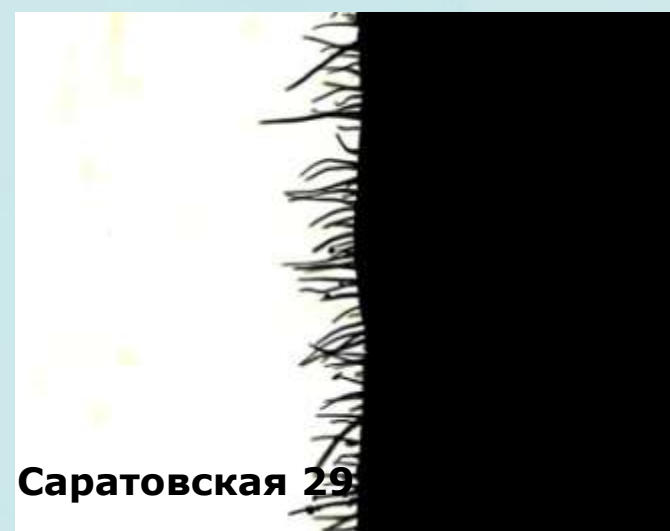


Типичные морфотипы опушения листа:

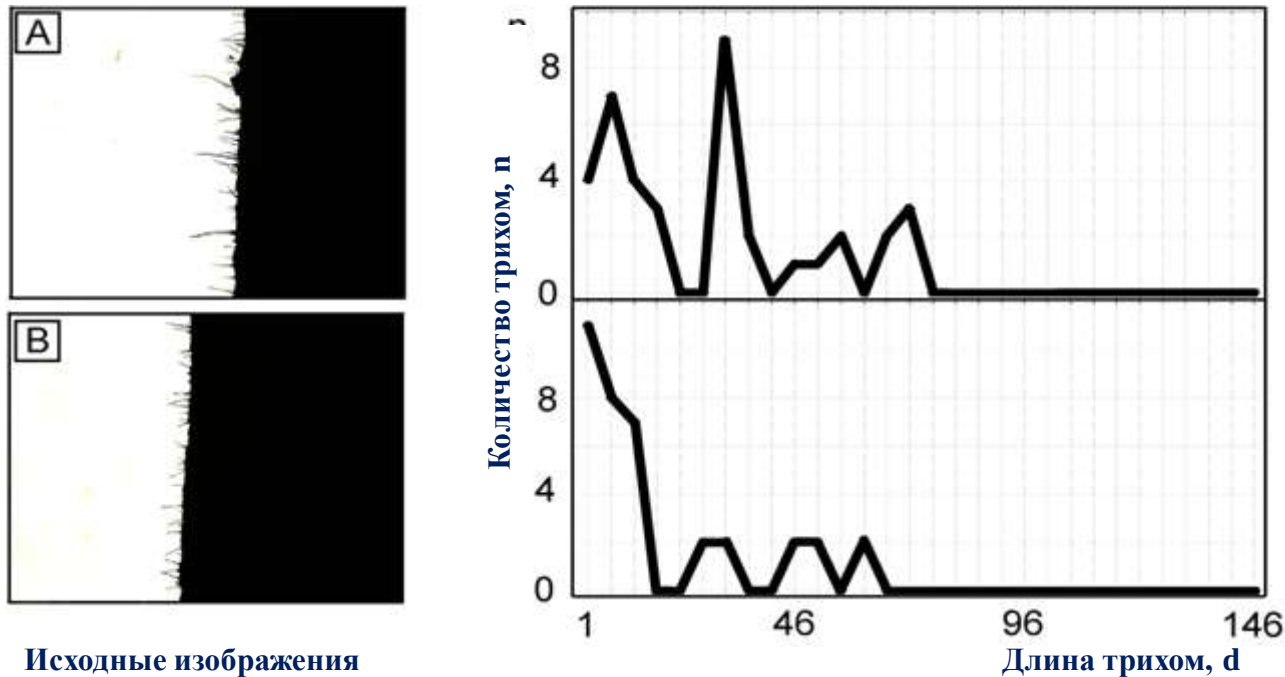
Сорта из
коллекции
Генбанка
(Гатерслебен,
Германия):



Сибирские
сорта и линии:



Результат компьютерного анализа изображений: Профили распределения трихом по длине



Исходные изображения

Применение разработанного метода фенотипирования позволяет получать детальное цифровое описание морфологии опушения листа в виде вектора чисел. Также могут быть рассчитаны интегральные параметры, такие как средняя длина трихом и общее количество трихом на сгибе.

$$\langle L \rangle = \left(\sum_{i=1}^k n_i d_i \right) / N$$

$$N = \sum_{i=1}^k n_i$$

Программа обработки изображений LNDetect интегрирована в базу данных.



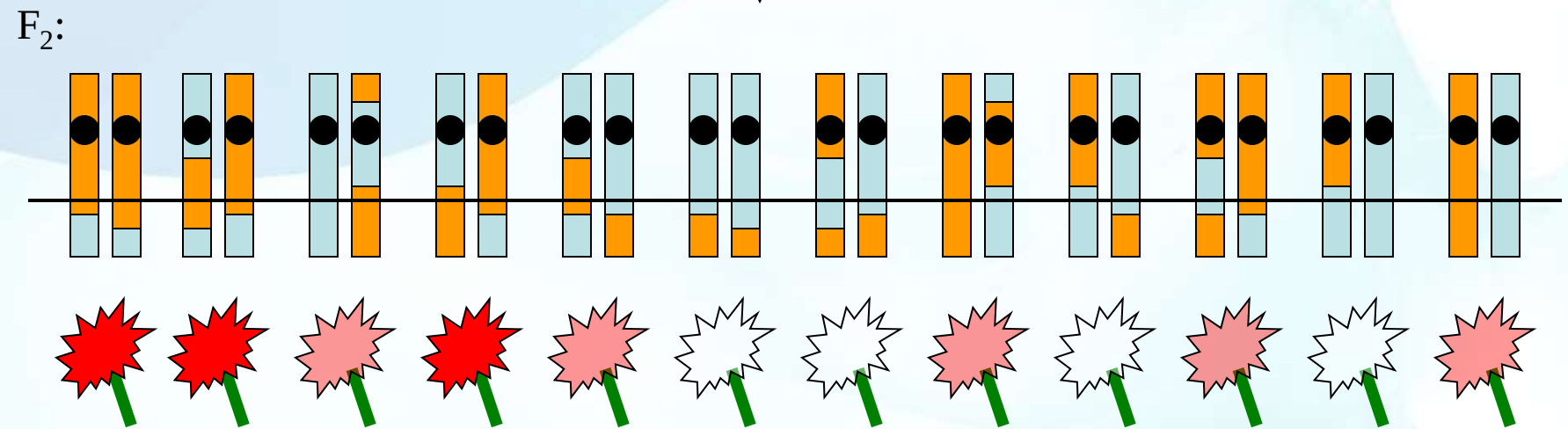
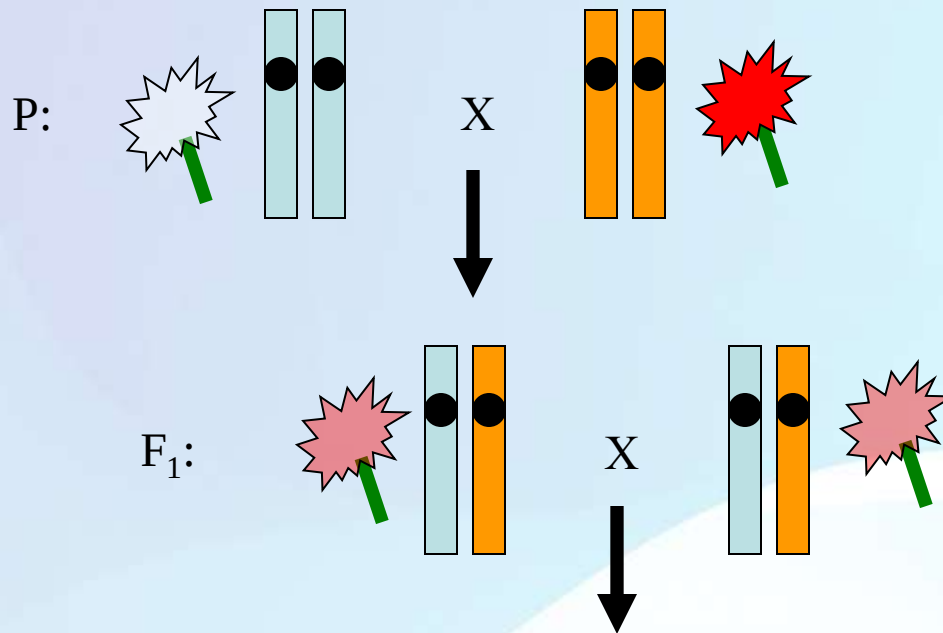
***Компьютерные методы
фенотипирования позволяют:***

- **Более детально описывать исследуемый признак**
- **Уменьшить время, необходимое для описания больших популяций**

Это необходимо для:

- **Картирования генов,**
- **Исследования влияния условий среды на проявление признака.**

*Картирование гена:
использование генетических и молекулярных маркеров*



«картирующая популяция»

Картирование гена: предсказание локализации гена методами системной биологии

Ресурсы:

- **KEGG**
- **PLAZA 2.5**
- **NCBI**

и др.

поиск целевых
генов в геномах
растений

анализ представленности
и консервативности
целевых генов среди
растительных организмов

Предсказание локализации при
помощи поиска по базе данных
последовательностей

Предсказание
локализации при
помощи
сравнительных
хромосомных карт

Поиск генов, влияющих на развитие трихом.

<http://www.ncbi.nlm.nih.gov/>

NCBI Resources How To

Gene Limits Advanced

Display Settings: ☒ Full Report Send to: ☒

GL2 homeobox-leucine zipper protein GLABRA 2 [*Arabidopsis thaliana*]

Gene ID: 844323, updated on 8-Apr-2012

Summary

Gene symbol: GL2

Gene description: homeobox-leucine zipper protein GLABRA 2

Primary source: [TAIR:AT1G79840](#)

Locus tag: AT1G79840

Gene type: protein coding

RefSeq status: REVIEWED

Organism: [Arabidopsis thaliana \(ecotype: Columbia\)](#)

Lineage: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; malvids; Brassicales; Brassicaceae; Camelineae; Arabidopsis

Also known as: F19K16.20; F19K16_20; GL2; GLABRA 2; HOMEODOMAIN PROTEIN GLABRA 2

Summary: Glabra 2, a homeodomain protein affects epidermal cell identity including trichomes, root hairs, and seed coat. It also down-regulates seed oil content. Expressed in atrichoblasts and required to suppress root hair development. Also expressed abundantly during early seed development. Directly regulated by WER.

Название: GLABRA 2

epidermal cell identity
including trichomes

Поиск генов, влияющих на развитие трихом. Взаимодействия.

The screenshot displays the AND Visio - Alpha version software interface. The main window shows a network diagram of gene interactions. The left sidebar lists categories: Protein (0), Gene (14), Metabolite (0), Pathway (0), Component (0), Disease (0), and Взаимодействия (8). The central canvas shows a network of genes represented by red DNA helix icons, including ZWI, ETC2, STI, FASS, ETC1, EGL3, TRY, GL3, GL2, TTG1, RHL1, RHL2, SPY, and KAK. The right sidebar contains two detail panels. The top panel, titled 'Свойства', shows 'Тип взаимодействия association' and 'Участники взаимодействия' with a table listing RHL1 and GL2 as genes with undetermined roles. The bottom panel, also titled 'Свойства', shows 'Тип взаимодействия association' and 'Участники взаимодействия' with a table listing TTG1 and GL2 as genes with undetermined roles. Below this table is a 'Предложения' section with a text snippet about trichome initiation. At the bottom of the right sidebar is a 'Ссылки' section with a link to PMID - 12084832.

AND Visio - Alpha version

Файл Редактировать Вид Раскладка Анализ Настройка Помощь

Graph 1 +

Поиск ...

100%

Protein (0)
Gene (14)
Metabolite (0)
Pathway (0)
Component (0)
Disease (0)
Взаимодействия (8)

ЗWI
ETC2
STI
FASS
ETC1
EGL3
TRY
GL3
GL2
TTG1
RHL1
RHL2
SPY
KAK

Весь граф Пути Циклы Поиск

Disconnected MySQL

Свойства

Тип взаимодействия association

Участники взаимодействия

Название	Тип	Роль
RHL1	Gene	undetermined
GL2	Gene	undetermined

Предложения

We have examined the genetic relationship between RHL1 and GL2, an inhibitor of root hair initiation non-hair cells. >>>

Свойства

Тип взаимодействия association

Участники взаимодействия

Название	Тип	Роль
TTG1	Gene	undetermined
GL2	Gene	undetermined

Предложения

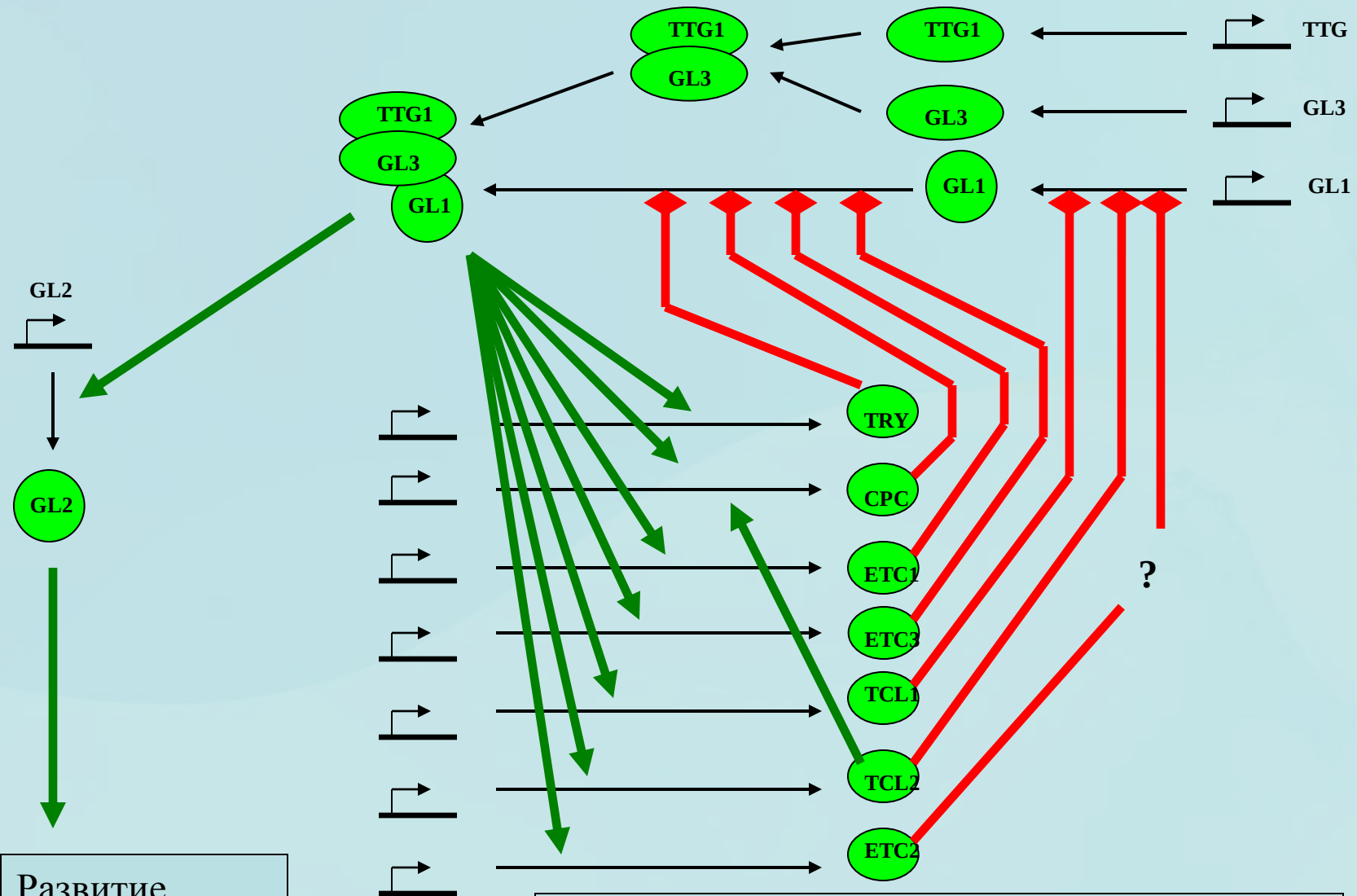
TTG2 acts downstream of the trichome initiation TTG1 and GLABROUS1, although trichome expansion of TTG2 continues to occur if they are inactivated

Ссылки

PMID - 12084832

Программа AND Visio позволяет извлекать информацию о взаимодействиях между генами из баз данных и публикаций.

Генная сеть инициации трихом *Arabidopsis thaliana*

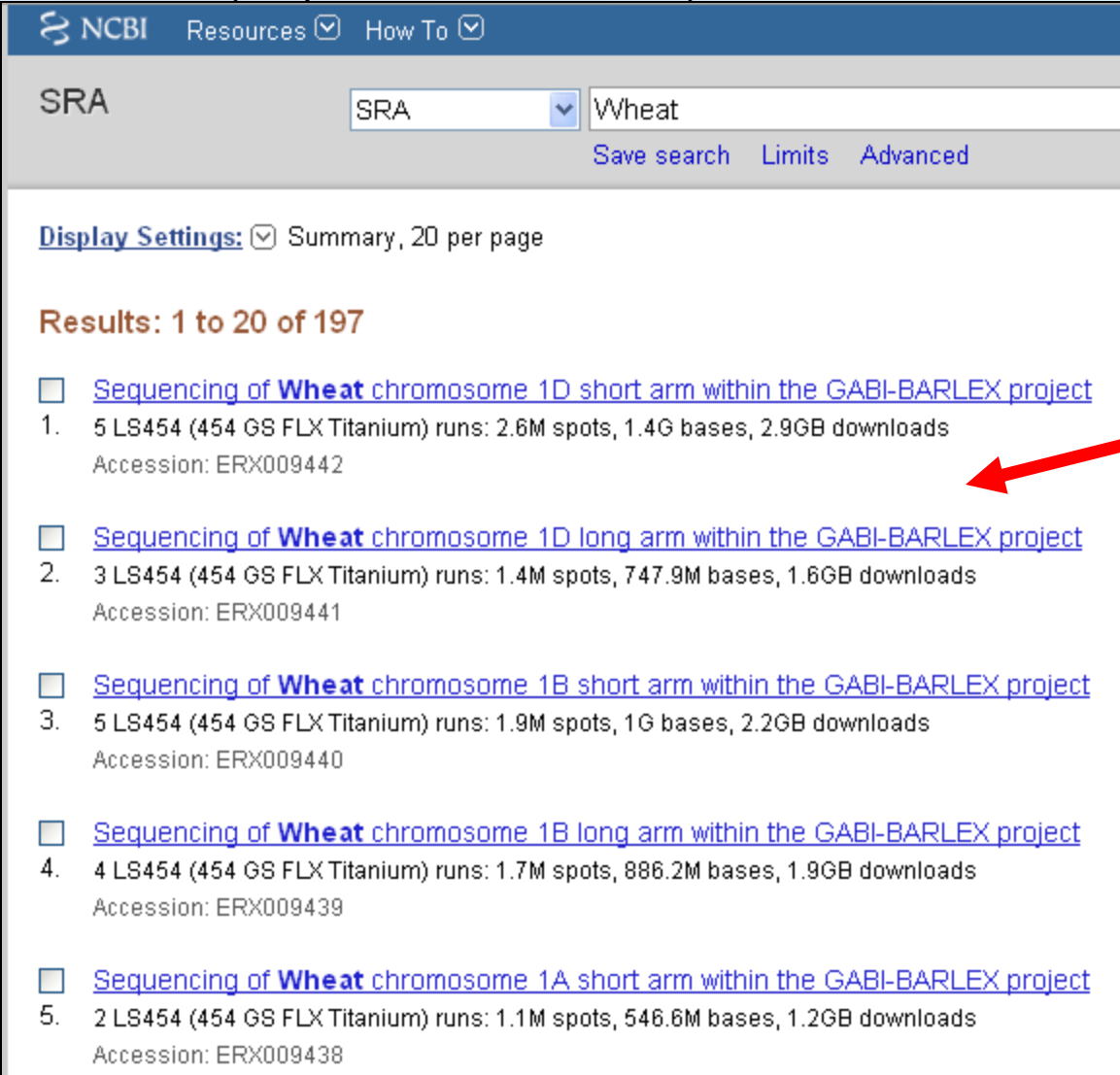


Развитие
трихомы

Каскад ингибиторов способен транспортироваться в соседние клетки, обеспечивая формирование равномерного паттерна распределения трихом

Предсказание локализации гена методами системной биологии: поиск последовательностей в геноме пшеницы

NCBI SRA (Sequence Read Archive)



The screenshot shows the NCBI SRA search interface. At the top, there's a navigation bar with 'NCBI', 'Resources', and 'How To'. Below it, the 'SRA' section has a dropdown menu set to 'SRA' and a search box containing 'Wheat'. Links for 'Save search', 'Limits', and 'Advanced' are visible. Under 'Display Settings', 'Summary' is selected with '20 per page'. The results section, titled 'Results: 1 to 20 of 197', lists five entries. Each entry has a checkbox, a project title, a brief description, and an accession number. A red arrow points from the Russian text on the right to the first result.

Checkbox	Project Title	Description	Accession
☐	Sequencing of Wheat chromosome 1D short arm within the GABI-BARLEX project	1. 5 LS454 (454 GS FLX Titanium) runs: 2.6M spots, 1.4G bases, 2.9GB downloads	ERX009442
☐	Sequencing of Wheat chromosome 1D long arm within the GABI-BARLEX project	2. 3 LS454 (454 GS FLX Titanium) runs: 1.4M spots, 747.9M bases, 1.6GB downloads	ERX009441
☐	Sequencing of Wheat chromosome 1B short arm within the GABI-BARLEX project	3. 5 LS454 (454 GS FLX Titanium) runs: 1.9M spots, 1G bases, 2.2GB downloads	ERX009440
☐	Sequencing of Wheat chromosome 1B long arm within the GABI-BARLEX project	4. 4 LS454 (454 GS FLX Titanium) runs: 1.7M spots, 886.2M bases, 1.9GB downloads	ERX009439
☐	Sequencing of Wheat chromosome 1A short arm within the GABI-BARLEX project	5. 2 LS454 (454 GS FLX Titanium) runs: 1.1M spots, 546.6M bases, 1.2GB downloads	ERX009438

Данные прочтения
последовательностей
отдельных хромосом
пшеницы

Картирование гена: предсказание локализации гена методами системной биологии

<http://www.genome.jp/kegg/genes.html>

KEGG Arabidopsis thaliana (thale cress): AT1G79840			
Help			
Entry	AT1G79840	CDS	T00041
Gene name	GL2		
Definition	homeobox-leucine zipper protein GLABRA 2		
Orthology	K09338 homeobox-leucine zipper protein		
Organism	ath Arabidopsis thaliana (thale cress)		
Class	BRITE hierarchy		
SSDB	Ortholog Paralog GFIT		
Motif	Pfam: START Homeobox Homeobox_KN HALZ PROSITE: HOMEBOX_1 START HOMEBOX_2 Motif		
Other DBs	NCBI-GI: 334184032 NCBI-GeneID: 844323		
Position	1		
AA seq	776		

Целевой ген: glabra 2
Отвечает за
формирование
опушения листа у A.
thaliana

Entry	KO	len	SW-score(margin)	bits	identity	overlap	best(all)
☒ rcu:RCOM_1493760 homeobox protein, putative		758	3241 (1304)	745	0.659	768	<-> 215
☒ pop:POPTR_817110 hypothetical protein		759	3215 (39)	739	0.655	768	<-> 297
☒ gmx:100809926 homeobox-leucine zipper protein GLABRA 2-		747	3212 (34)	738	0.668	748	<-> 509
☒ vvi:100256623 homeobox-leucine zipper protein GLABRA 2-		754	3203 (1280)	736	0.652	772	<-> 208
☒ mtr:MTR_2g101720 Homeobox-leucine zipper protein GLABRA		794	3136 (1264)	721	0.611	797	<-> 240
☒ bdi:100829793 homeobox-leucine zipper protein ROC9-like		758	2387 (595)	550	0.510	778	<-> 213
☒ smo:SELMODRAFT_450553 hypothetical protein		817	1912 (2)	442	0.429	777	-> 240
☒ zma:542468 outer cell layer1		784	1833 (58)	424	0.407	804	-> 136
☒ osa:4337074 Os04g0627000		784	1830 (51)	423	0.407	776	-> 192
☒ sbi:SORBI_06g029270 hypothetical protein		789	1827 (26)	422	0.398	789	-> 258
☐ aly:ARALYDRAFT_324546 hypothetical protein	K09338	811	1714 (1169)	397	0.394	761	-> 264
☐ ppp:PHYPADRAFT_183629 hypothetical protein		871	386 (10)	94	0.241	714	-> 191
☐ ssl:SS1G_00669 hypothetical protein		646	257 (108)	64	0.316	158	-> 133
☐ bfu:BC1G_06476 hypothetical protein		465	255 (86)	64	0.327	159	<-> 133
☐ nhe:NECHADRAFT_49051 hypothetical protein							
☐ mgr:MGG_07437 hypothetical protein							
☐ pan:PODANSg8303 hypothetical protein							

Список ортологов целевого гена

Картирование гена: предсказание локализации гена методами системной биологии

<http://bioinformatics.psb.ugent.be/plaza/>

Целевой ген: glabra 2
Отвечает за
формирование
опущения листа у *A.
thaliana*

Gene Overview

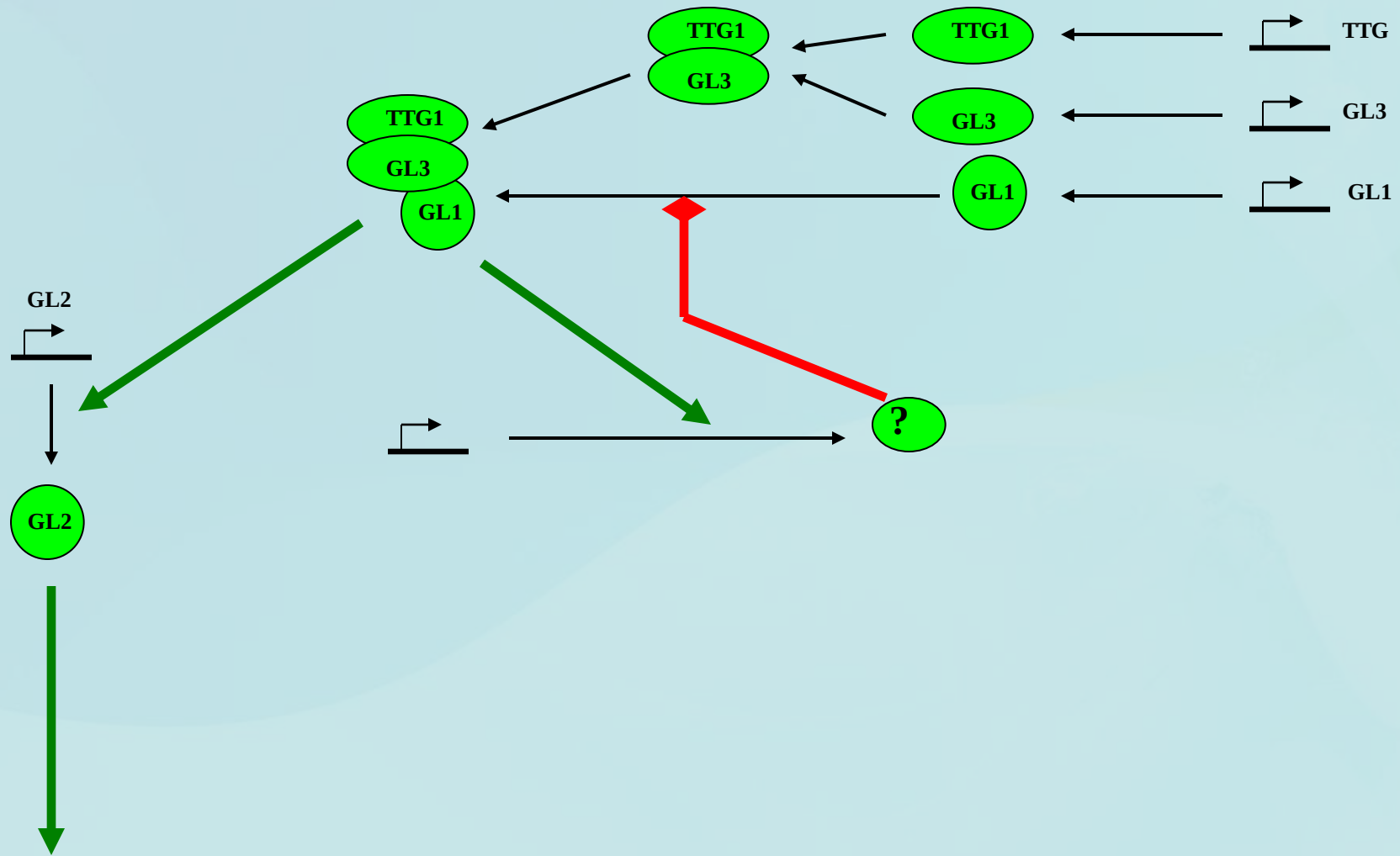
General information

Gene id	AT1G79840 	Chromosome	1
Species	Arabidopsis thaliana	Type	Coding gene
Transcript	AT1G79840.2	Sequence	Show sequence(s)
Coordinates	30037093 - 30041013	Gene family	HOM000230 338 genes in 20 species
Strand	positive	Sub family	ORTHO008985 14 genes in 15 species
Description	Glabra 2, a homeodomain protein affects epidermal cell identity including trichomes, root hairs, and seed oil content. Expressed in atrichoblasts and required to suppress root hair development. Also expressed in root development. Directly regulated by WER.;HD-ZIP IV family of homeobox-leucine zipper protein with lipid-binding domain.		
Comment	transcript	eq	
Gene structure			

Группа ортологов целевого гена
ORTHO008985

ath	Arabidopsis thaliana (1 gene)
bdi	Brachypodium distachyon (1 gene)
cpa	Carica papaya (1 gene)
fve	Fragaria vesca (1 gene)
gma	Glycine max (4 genes)
lja	Lotus japonicus (1 gene)
mdo	Malus domestica (2 genes)
mes	Manihot esculenta (2 genes)
mtr	Medicago truncatula (1 gene)
osa	Oryza sativa ssp. japonica (1 gene)
osaindica	Oryza sativa ssp. indica (1 gene)
ptr	Populus trichocarpa (2 genes)
rco	Ricinus communis (1 gene)
tca	Theobroma cacao (1 gene)
vvi	Vitis vinifera (1 gene)

Консервативная часть генной сети инициации трихом *Arabidopsis thaliana*



Развитие
трихомы

Каскад ингибиторов способен транспортироваться в соседние клетки, обеспечивая формирование равномерного паттерна распределения трихом

Картирование гена: предсказание локализации гена методами системной биологии

<http://bioinformatics.psb.ugent.be/plaza/>

Целевой ген: glabra 2
Отвечает за
формирование
опущения листа у *A.
thaliana*

Gene Overview

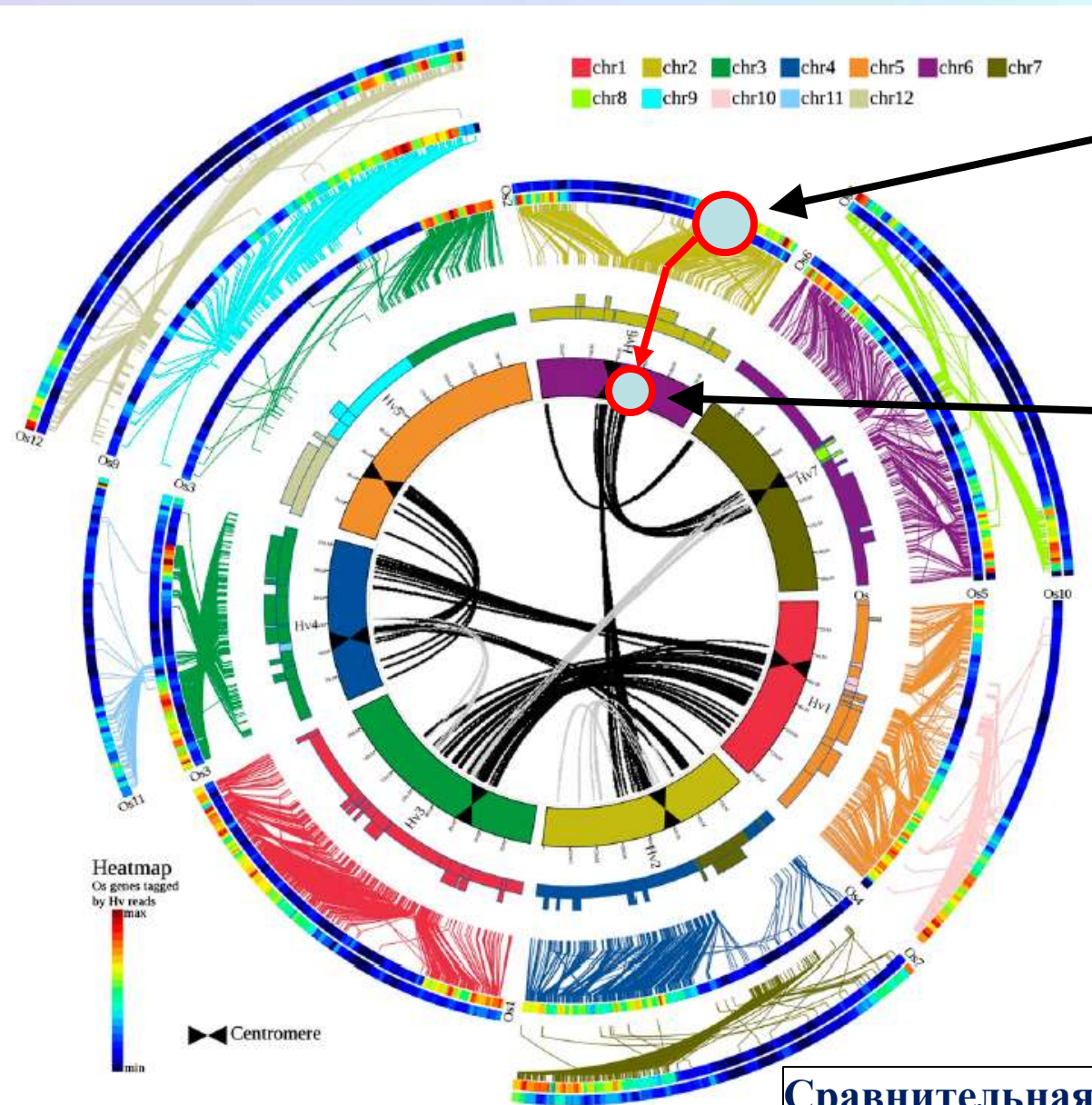
General information

Gene id	AT1G79840 	Chromosome	1
Species	Arabidopsis thaliana	Type	Coding gene
Transcript	AT1G79840.2	Sequence	Show sequence(s)
Coordinates	30037093 - 30041013	Gene family	HOM000230 338 genes in 20 species
Strand	positive	Sub family	<u>ORTHO008985</u> 14 genes in 15 species
Description	Glabra 2, a homeodomain protein affects epidermal cell identity including trichomes, root hairs, and seed oil content. Expressed in atrichoblasts and required to suppress root hair development. Also expressed in root development. Directly regulated by WER.; HD-ZIP IV family of homeobox-leucine zipper protein with lipid-binding domain. transcript eq		
Comment			
Gene structure			

Группа ортологов целевого гена
ORTHO008985

ath	Arabidopsis thaliana (1 gene)
bdi	Brachypodium distachyon (1 gene)
cpa	Carica papaya (1 gene)
fve	Fragaria vesca (1 gene)
gma	Glycine max (4 genes)
lja	Lotus japonicus (1 gene)
mdo	Malus domestica (2 genes)
mes	Manihot esculenta (2 genes)
mtr	Medicago truncatula (1 gene)
osa	<u>Oryza sativa ssp. japonica (1 gene)</u>
osaindica	<u>Oryza sativa ssp. indica (1 gene)</u>
ptr	Populus trichocarpa (2 genes)
rco	Ricinus communis (1 gene)
tca	Theobroma cacao (1 gene)
vvi	Vitis vinifera (1 gene)

Использование сравнительных хромосомных карт:

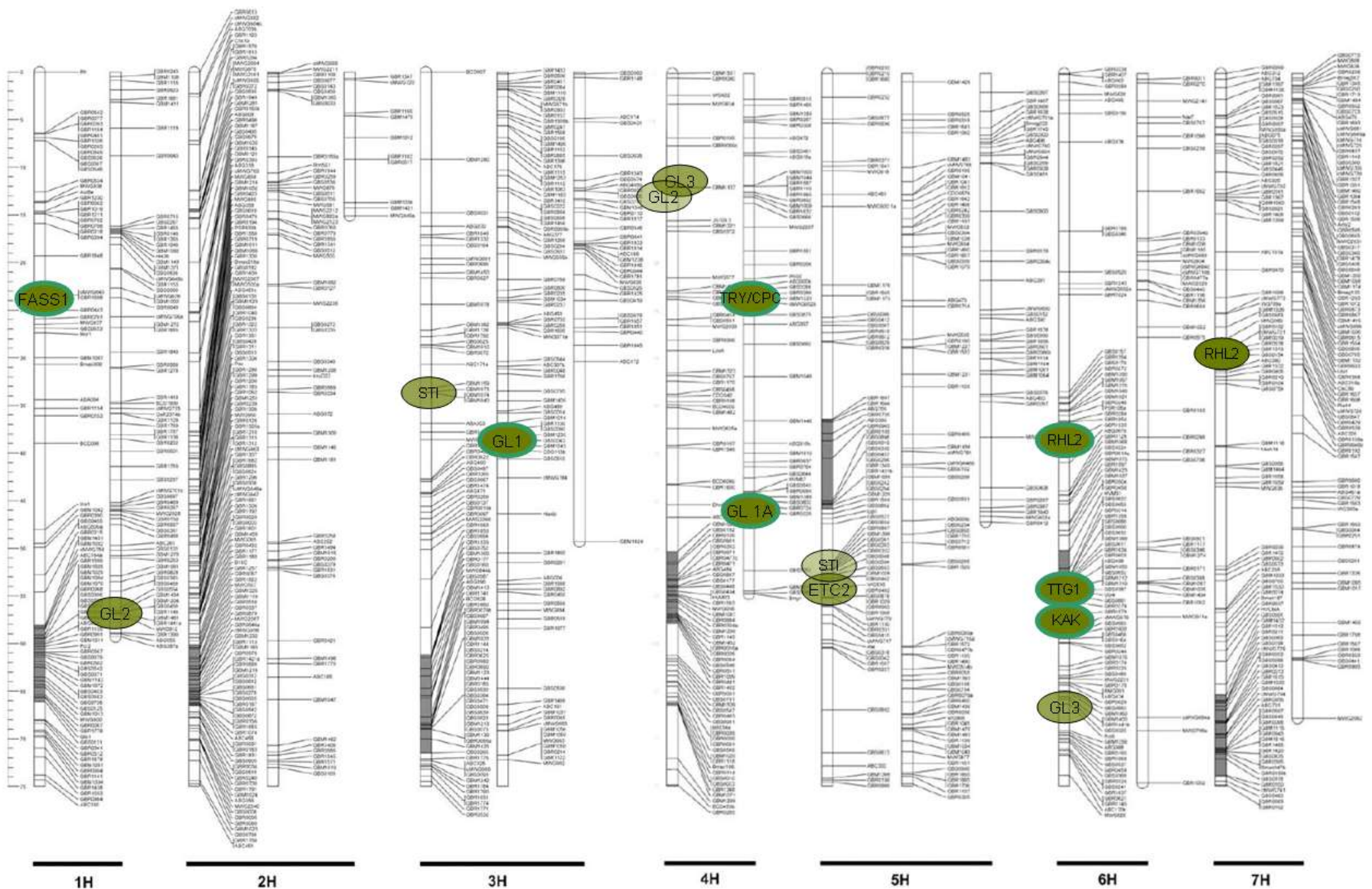


Положение целевого
гена на 2 хромосоме
O. sativa

Соответствующее
положение целевого
гена на 6 хромосоме
H. vulgare

Сравнительная хромосомная
карта *O. sativa* и *H. vulgare*

Предсказанное положение генов контроля опущения листа на хромосомах *N. vulgare*:



- Курс подготовлен с использованием лекционных материалов, прочитанных в 2008-2012 годах сотрудниками кафедры информационной биологии ФЕН НГУ и Института цитологии и генетики СО РАН:

- *Абердин И.Р.*
- *Алемасов Н.А.*
- *Афонников Д.А.*
- *Дорошков А.В.*
- *Иванисенко В.А.*
- *Иляскин А.В.*
- *Колчанов Н.А.*
- *Лашин С.А.*
- *Миронова В.В.*
- *Сормачева И.Д.*
- *Сорокина К.Н.*