

# Распространение и эволюция Tat-LTR-ретротранспозонов растений

Новосибирский государственный университет,  
Институт цитологии и генетики СО РАН

Устьянцев К. В.

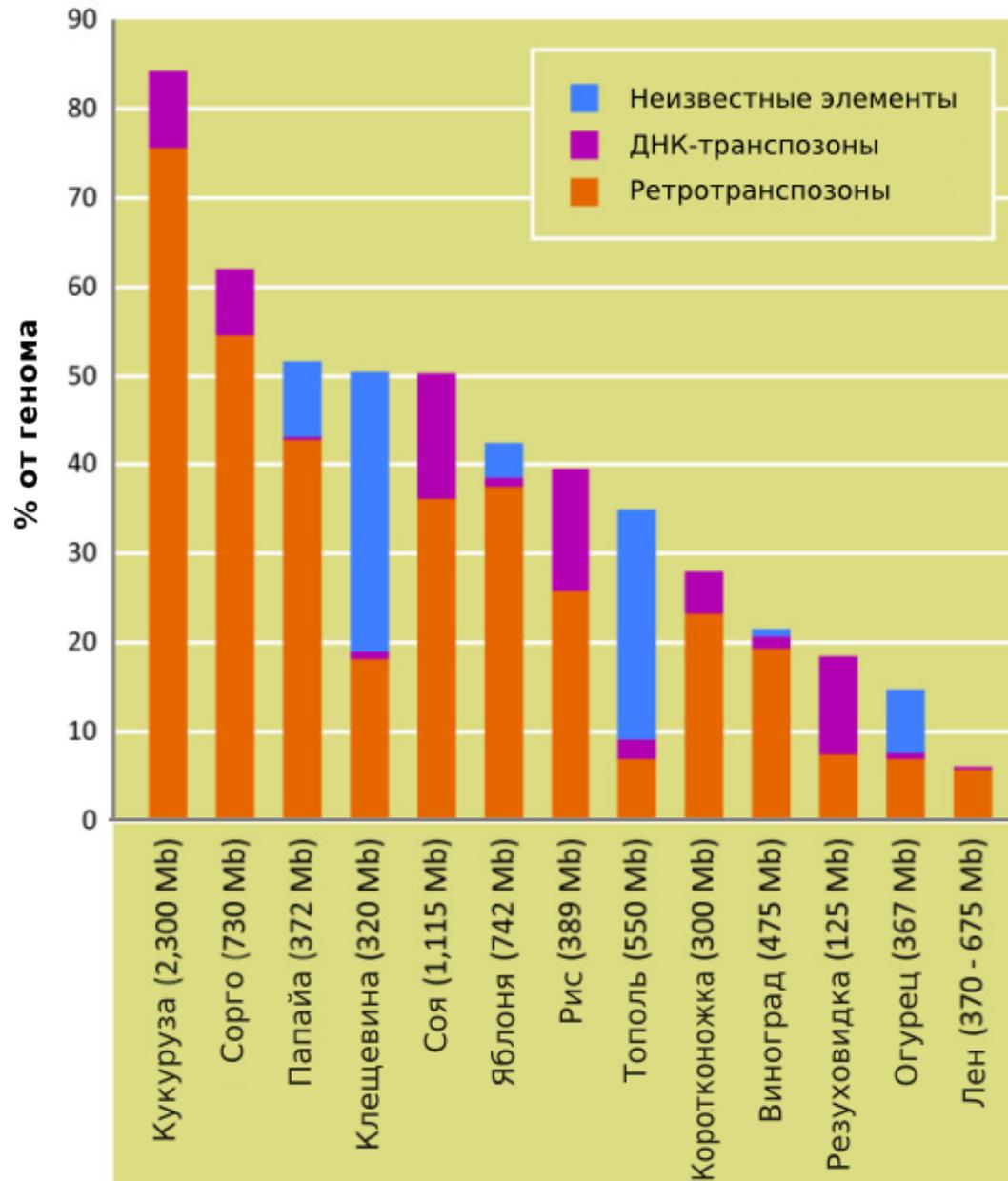
научный руководитель: к.б.н. Блинов А. Г.

# Мобильные генетические элементы (МГЭ)

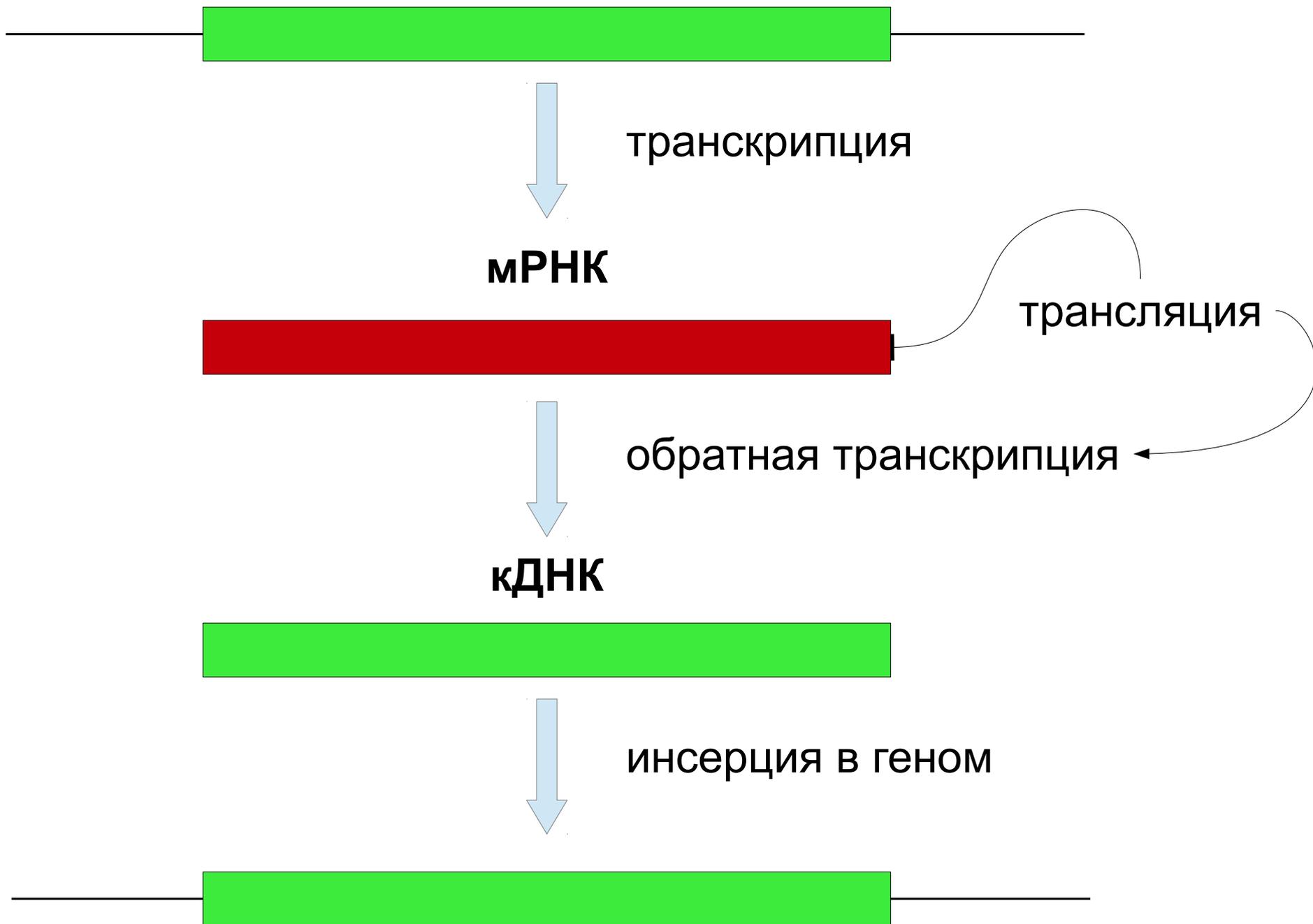


**МГЭ** – участки ДНК, способные перемещаться внутри генома.

# Содержание мобильных элементов в геномах растений

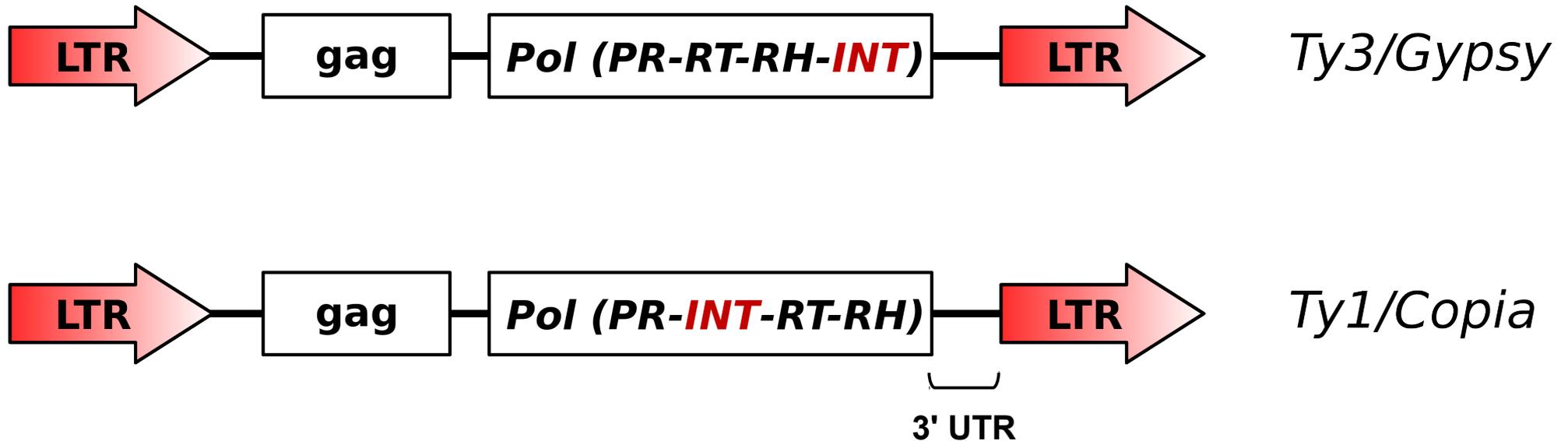


# Ретротранспозон





# LTR-ретротранспозоны



**LTR** – длинный концевой повтор (*Long Terminal Repeat*)

**PR** – протеаза (*Protease*)

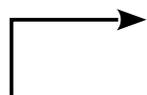
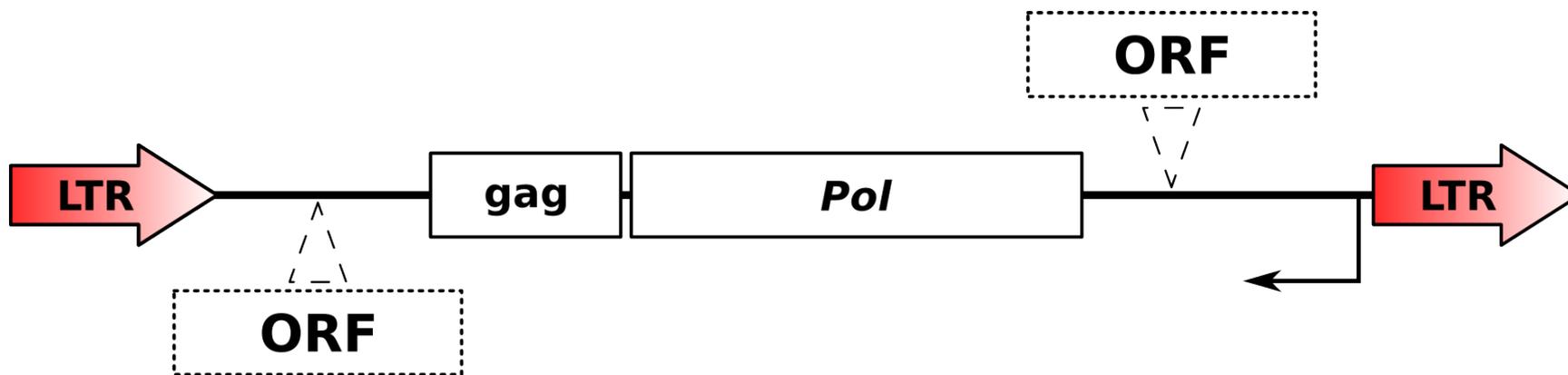
**RT** – обратная транскриптаза (*Reverse Transcriptase*)

**RH** – РНКазы H (*RNase H, RNH*)

**INT** – интегразы (*Integrase*)

**UTR** – нетранслируемая область (*Untranslated Region*)

# Tat-LTR-ретротранспозоны (Tat-LTR)



– положение старта транскрипции согласно Gomez-Orte et al.



– дополнительные рамки считывания (sense и antisense)

**Цель: изучение разнообразия, распространения и эволюции Tat-LTR-ретротранспозонов растений.**

**Цель: изучение разнообразия, распространения и эволюции Tat-LTR-ретротранспозонов растений.**

**1. Поиск нуклеотидных последовательностей Tat-LTR, содержащих домен архейной рибонуклеазы H, с помощью биоинформатических методов в последовательностях прочитанных геномов растений, а также экспериментальный поиск последовательностей обратной транскриптазы у представителей тех групп растений, для которых еще не было полностью прочитано геномов.**

# **Цель: изучение разнообразия, распространения и эволюции Tat-LTR-ретротранспозонов растений.**

- 1. Поиск нуклеотидных последовательностей Tat-LTR, содержащих домен архейной рибонуклеазы H, с помощью биоинформатических методов в последовательностях прочитанных геномов растений, а также экспериментальный поиск последовательностей обратной транскриптазы у представителей тех групп растений, для которых еще не было полностью прочитано геномов.**
- 2. Выявление основных филогенетических групп среди найденных Tat-LTR-ретротранспозонов и определение основных структурных характеристик элементов для каждой отдельной группы: наличие регулярных доменов и дополнительных рамок считывания.**

# **Цель: изучение разнообразия, распространения и эволюции Tat-LTR-ретротранспозонов растений.**

1. Поиск нуклеотидных последовательностей Tat-LTR, содержащих домен архейной рибонуклеазы H, с помощью биоинформатических методов в последовательностях прочитанных геномов растений, а также экспериментальный поиск последовательностей обратной транскриптазы у представителей тех групп растений, для которых еще не было полностью прочитано геномов.
2. Выявление основных филогенетических групп среди найденных Tat-LTR-ретротранспозонов и определение основных структурных характеристик элементов для каждой отдельной группы: наличие регулярных доменов и дополнительных рамок считывания.
3. **Установление основных событий в эволюции Tat-LTR-ретротранспозонов**

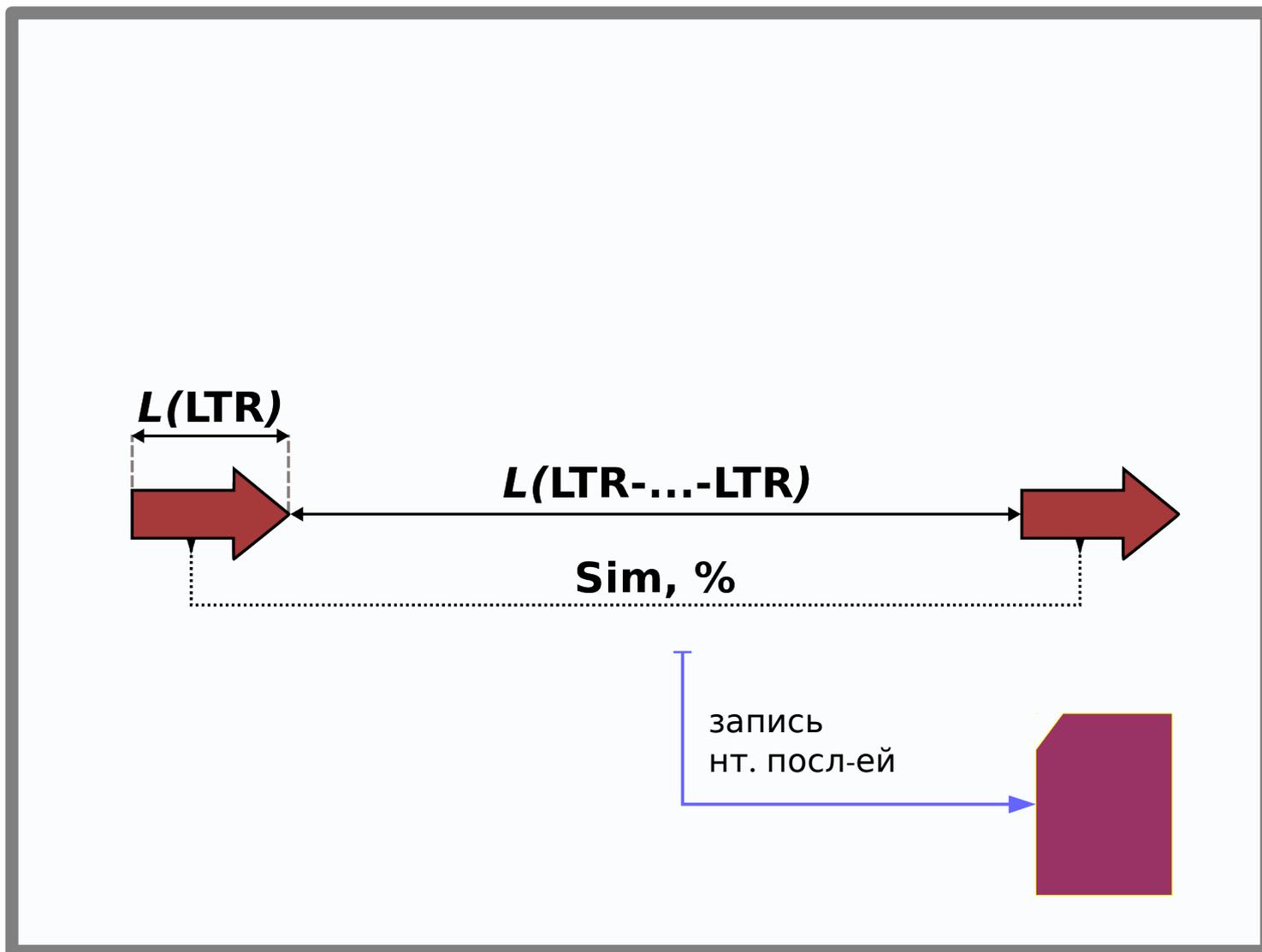
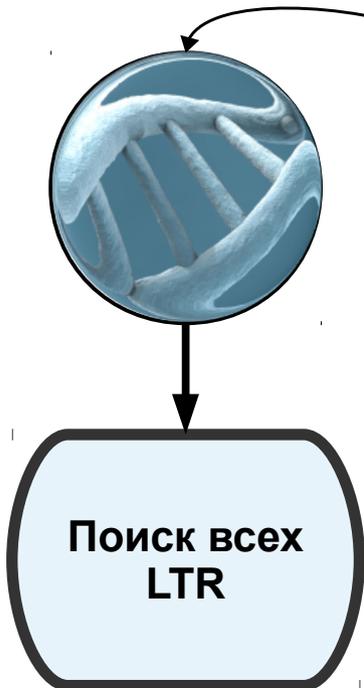
## Источник геномов

phytozome

[18 Mar 2013] Phytozome v9.1, containing new gene families calculated at 20 internal nodes, is now available.

[<http://phytozome.net/>]

# Алгоритм поиска



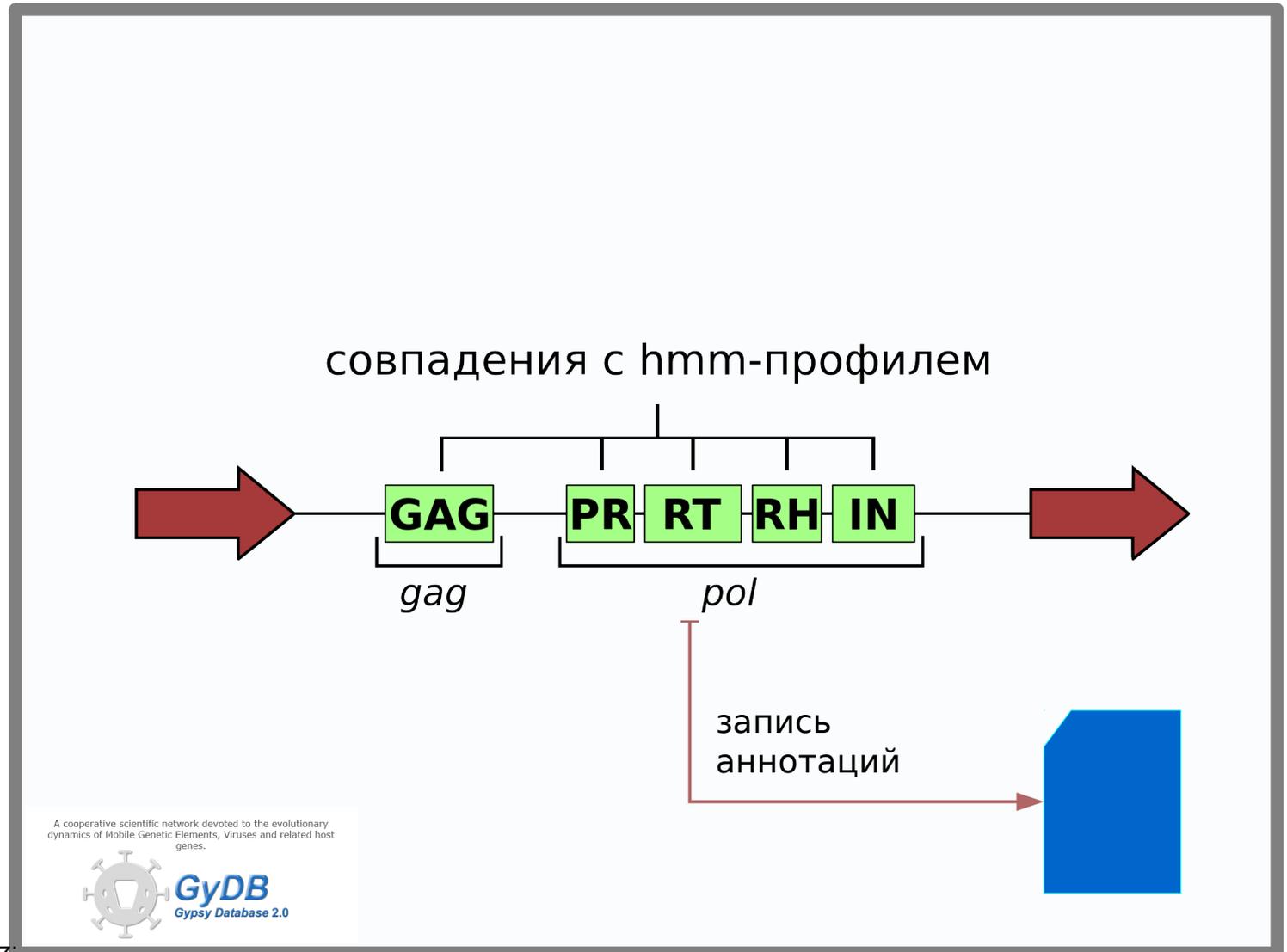
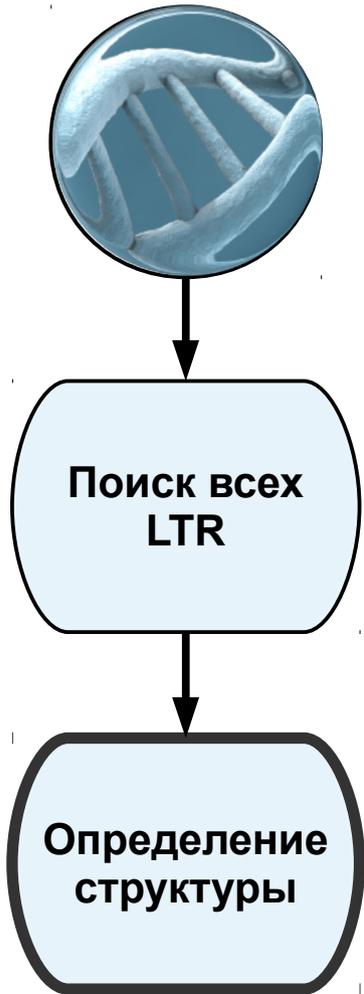
LTRharvest

Center for Bioinformatics  
ZBH



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

# Алгоритм поиска



LTRdigest

Center for Bioinformatics

ZBH



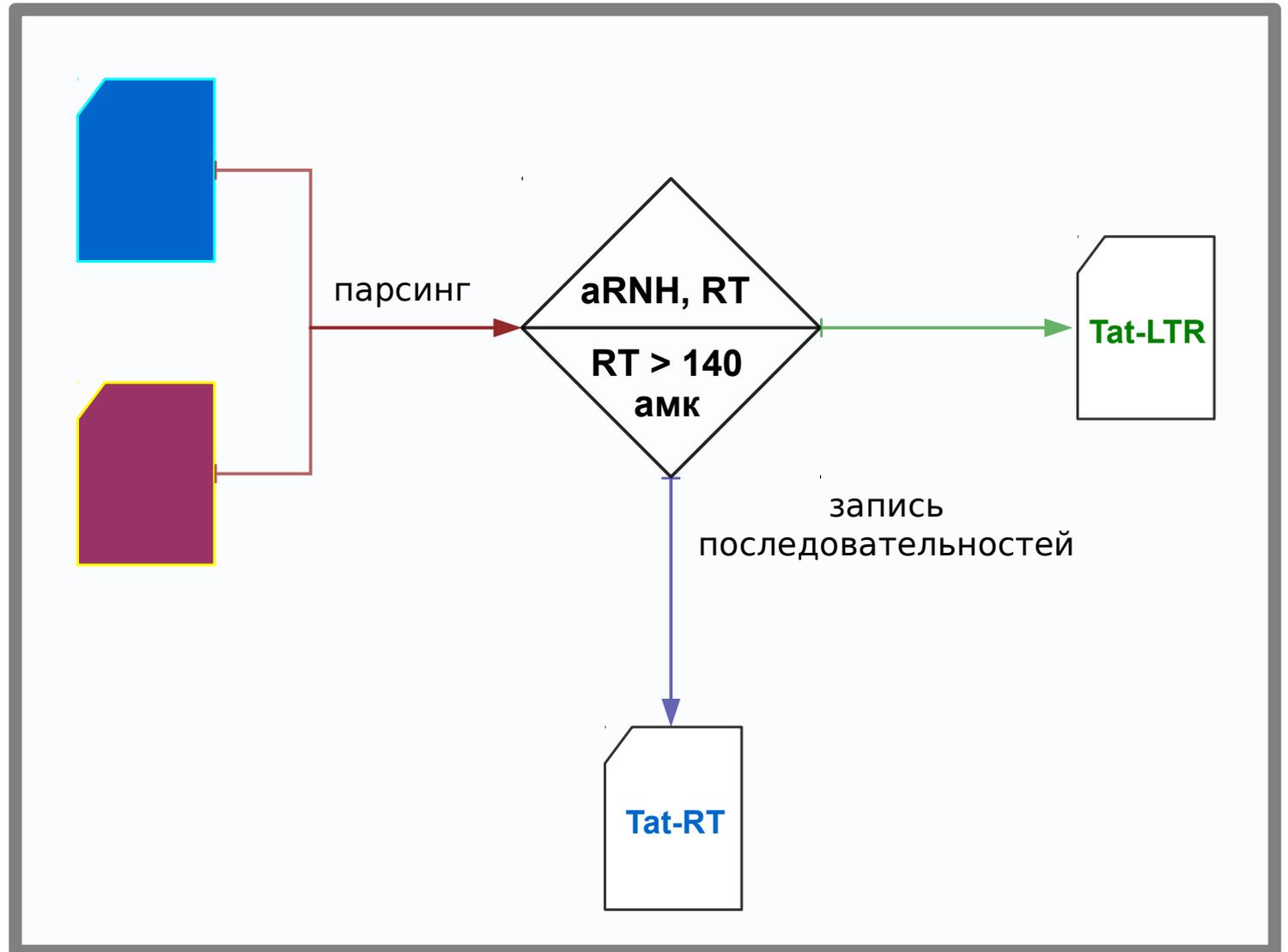
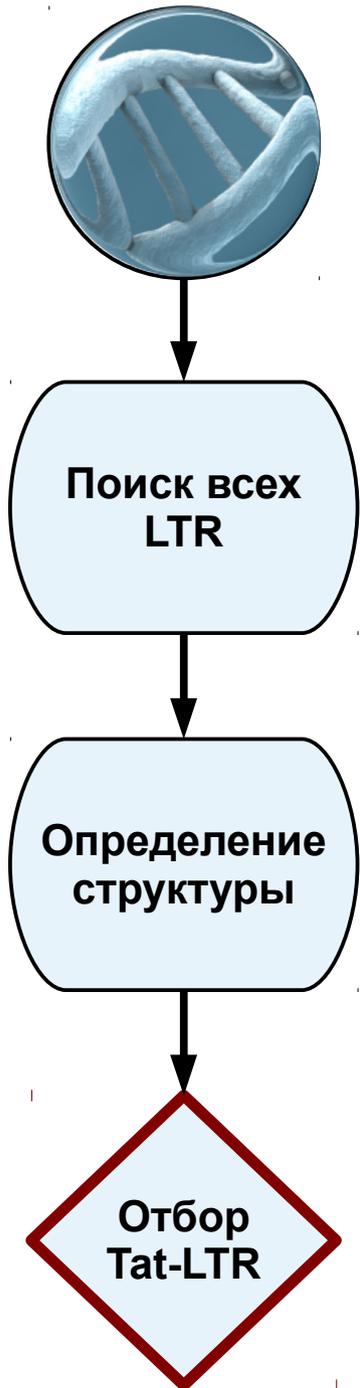
Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

A cooperative scientific network devoted to the evolutionary dynamics of Mobile Genetic Elements, Viruses and related host genes.

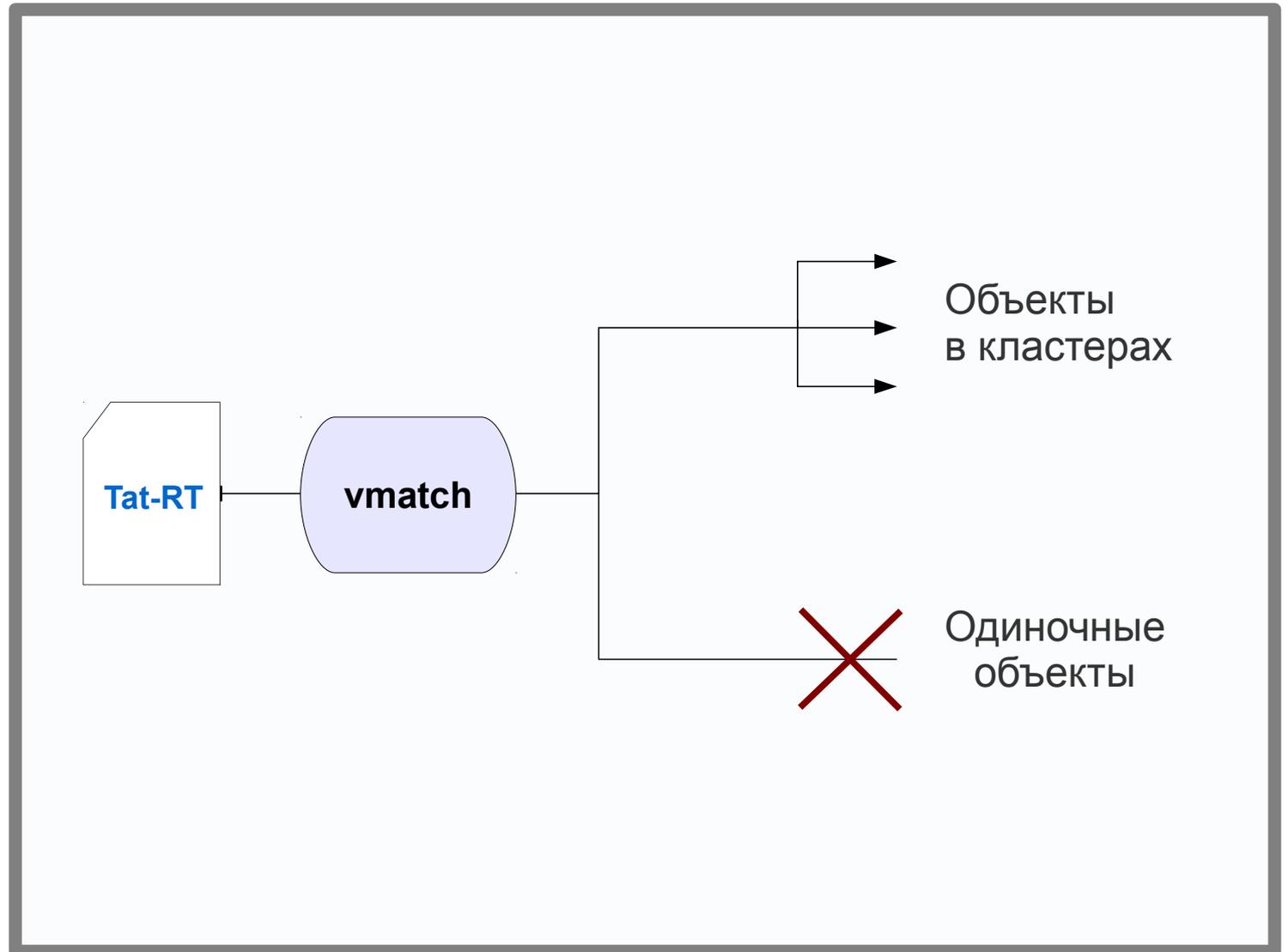
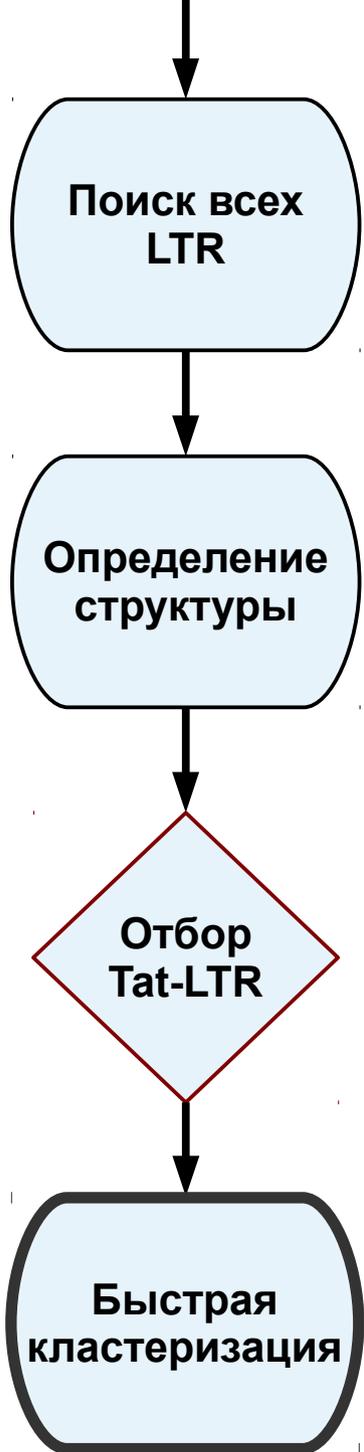


GyDB  
Gypsy Database 2.0

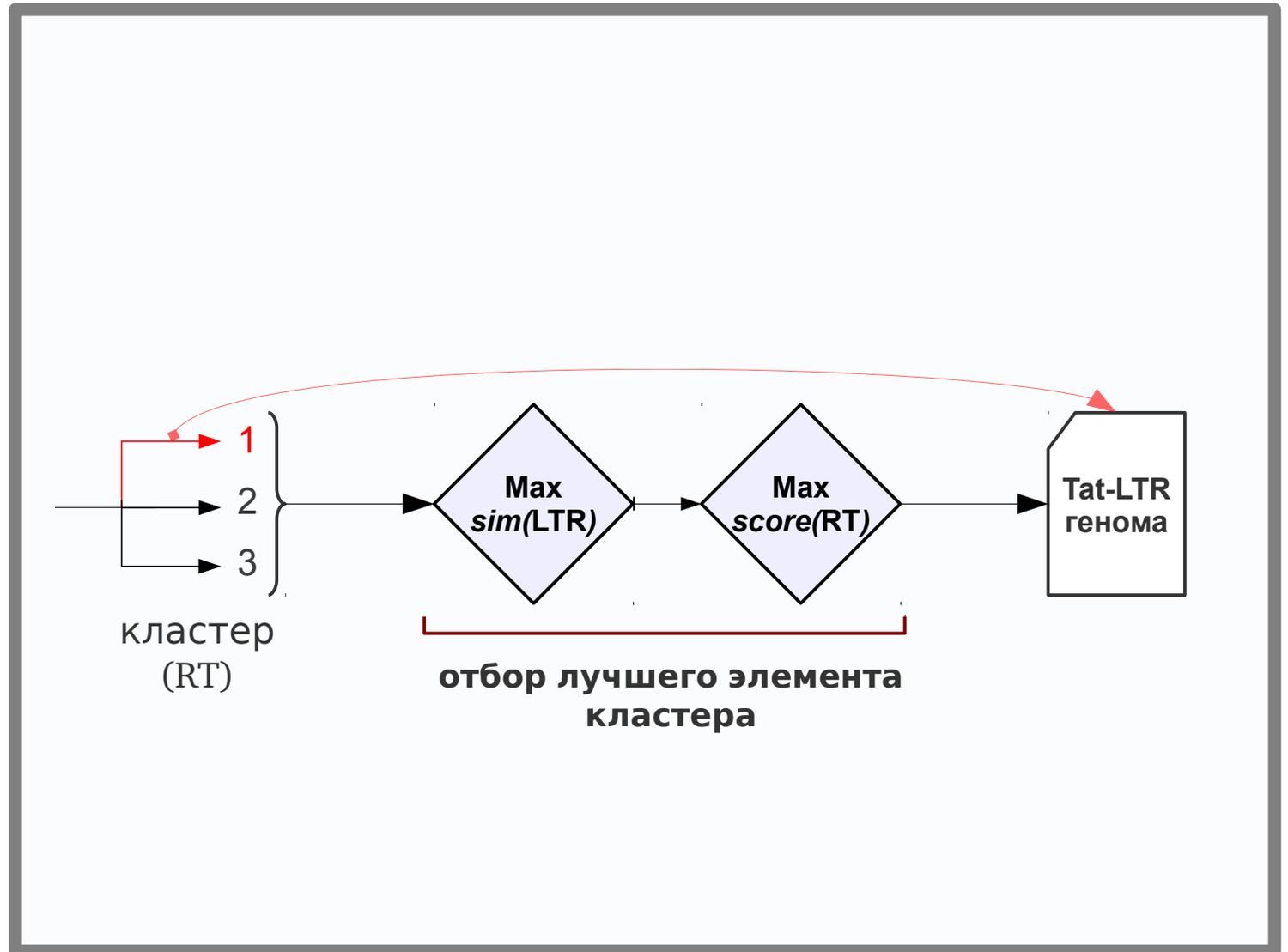
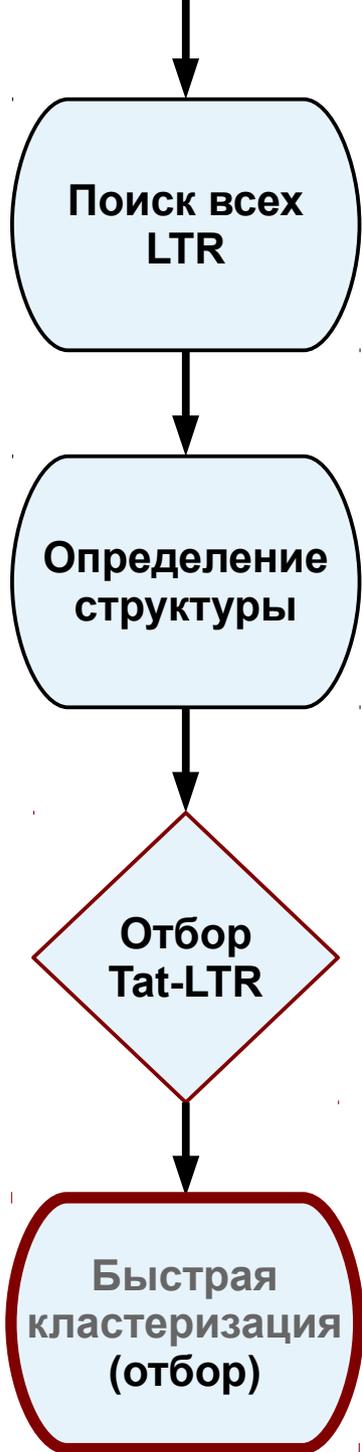
# Алгоритм поиска



# Алгоритм поиска



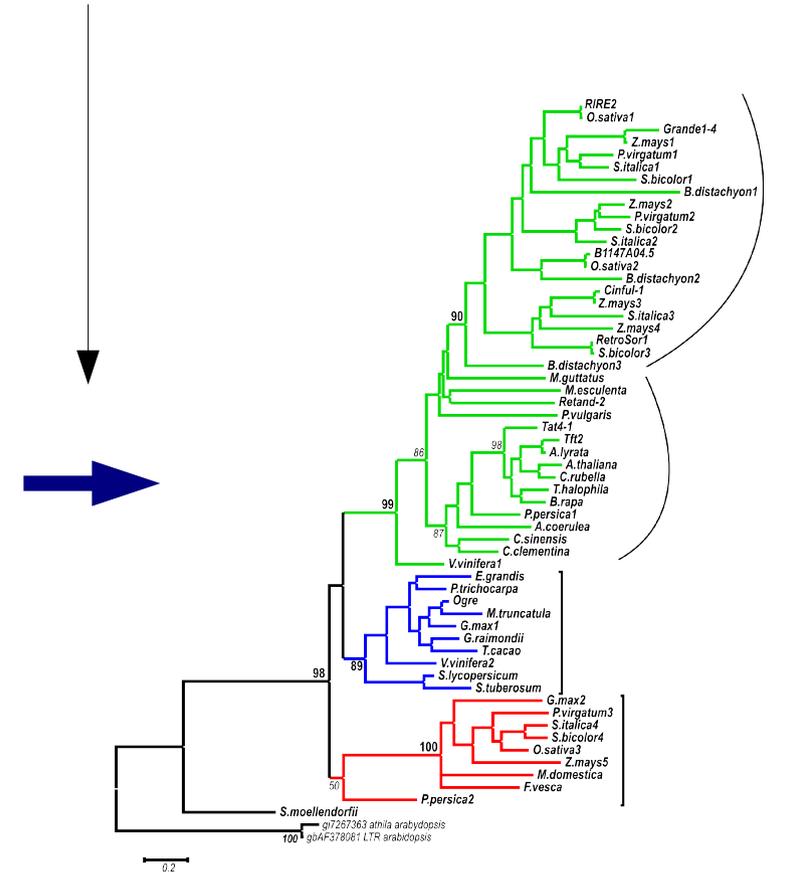
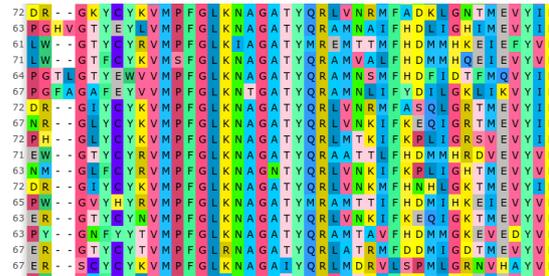
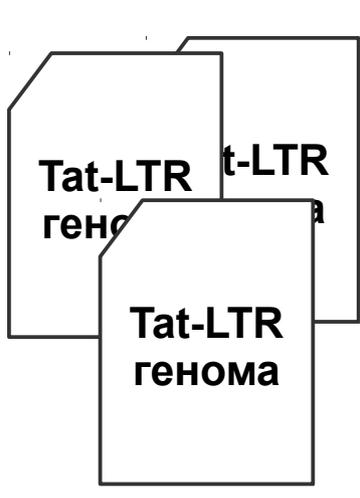
# Алгоритм поиска



# Построение филогении на основе RT-домена

UGENE <sup>1</sup>

PhyML 3.0 <sup>2</sup>



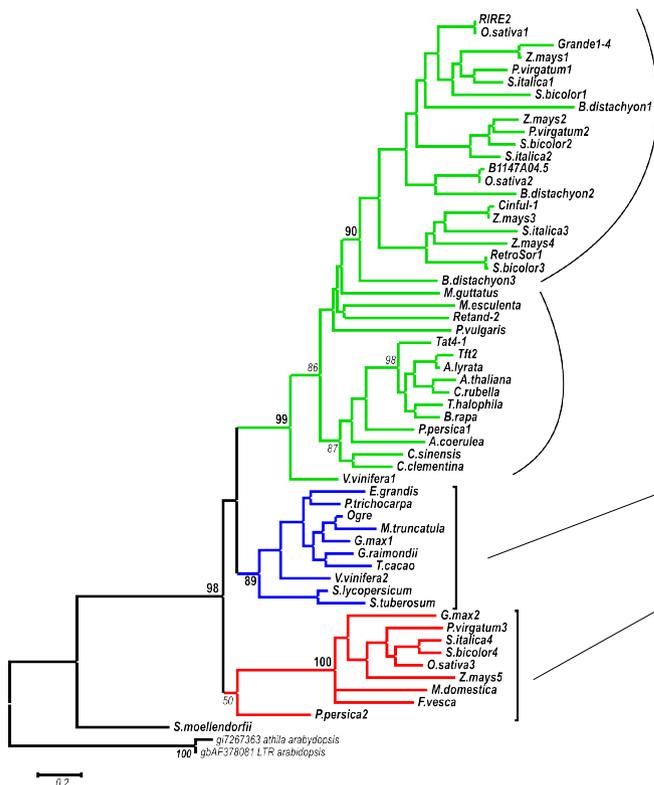
1. "Unipro UGENE: a unified bioinformatics toolkit". Konstantin Okonechnikov; Olga Golosova; Mikhail Fursov; the UGENE team.

Bioinformatics 2012 28: 1166-1167

2. "New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0."

Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. Systematic Biology, 59(3):307-21, 2010.

# Анализ структуры найденных Tat-LTR-ретротранспозонов



BLAST, NCBI <sup>1</sup> (поиск гомологии)

ORFfinder, NCBI <sup>2</sup> (поиск ORF)

HHpred <sup>3</sup> (поиск гомологии и пред. структуры)

TSSP <sup>4</sup> (предсказание растительных промоторов)

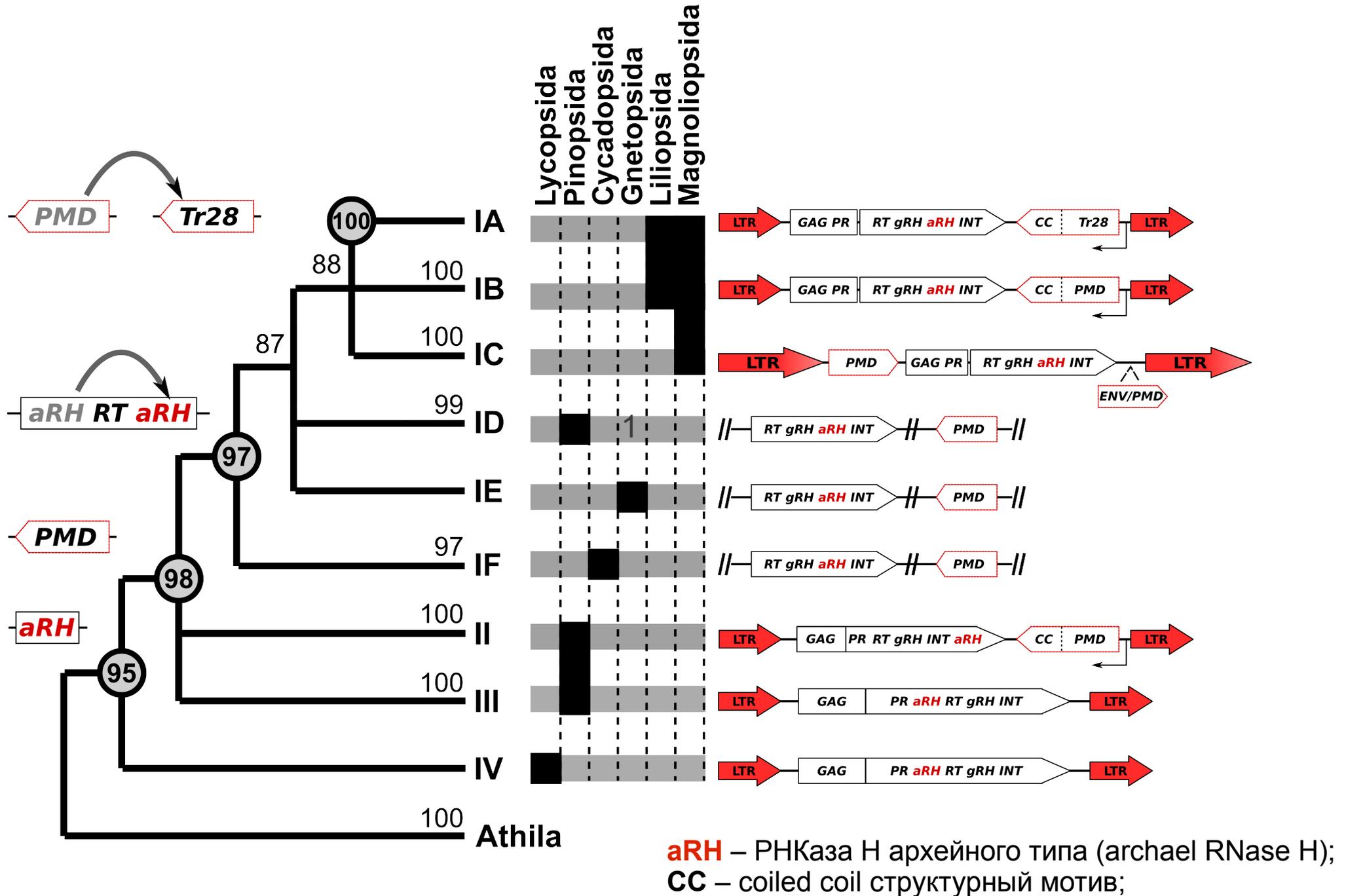
1. National Center for Biotechnology Information. <http://blast.ncbi.nlm.nih.gov/>;
2. National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>;
3. Max Planck Institute for Developmental Biology. <http://toolkit.tuebingen.mpg.de/hhpred>;
4. Softberry. <http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter>.

# Экспериментальная часть

## (поиск Tat-LTR-ретротранспозонов)

1. Выделение растительной тотальной ДНК;
2. ПЦР на дегенеративных праймерах, специфичных к последовательностям RT Tat-LTR-ретротранспозонов;
3. Выделение и очистка наработанного фрагмента;
4. Клонирование выделенных фрагментов в плазмидном T-векторе, трансформированном в компетентные клетки *E. coli*;
5. Отбор и анализ клонов;
6. Установление нуклеотидных последовательностей, встроенных в плазмиду и наработанных в *E. coli*;
7. Анализ нуклеотидных последовательностей.

# Результаты. Схема филогенетических взаимоотношений Tat-LTR и общая структура элементов одного кластера



# Результаты. Кластер I

- Присутствуют почти во всех таксонах семенных растений;

- Единая структура гена *pol*:

*RT gRH aRH INT*

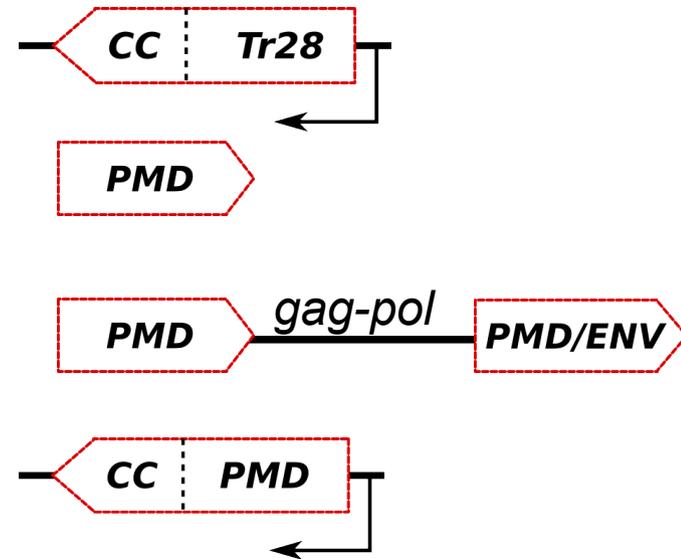
- Дополнительные рамки считывания:

\* 3' asPMD + as промотор

\* 5' sPMD

\* 5' sPMD + 3's(PMD/ENV)

\* 3' asTr28 + as промотор



**as** – antisense

**s** – sense

**PMD** – Plant Mobile Domain

**Tr28** – Transposase 28

**CC** – coiled coil структурный мотив;

**ENV** – белок вирусной оболочки (envelope)

# Результаты. Кластер II

- Присутствуют только у хвойных растений;
- Единая доменная структура для всех элементов:



- Впервые возникает дополнительная рамка считывания:  
3' asPMD + as промотор.

as – antisense

PMD – Plant Mobile Domain

CC – coiled coil структурный мотив;

# Результаты. Кластер III

- Присутствуют только у хвойных растений;
- Единая доменная структура для всех элементов:



- Отсутствуют дополнительные рамки считывания

## Кластер IV (*Selaginella moellendorffii*)

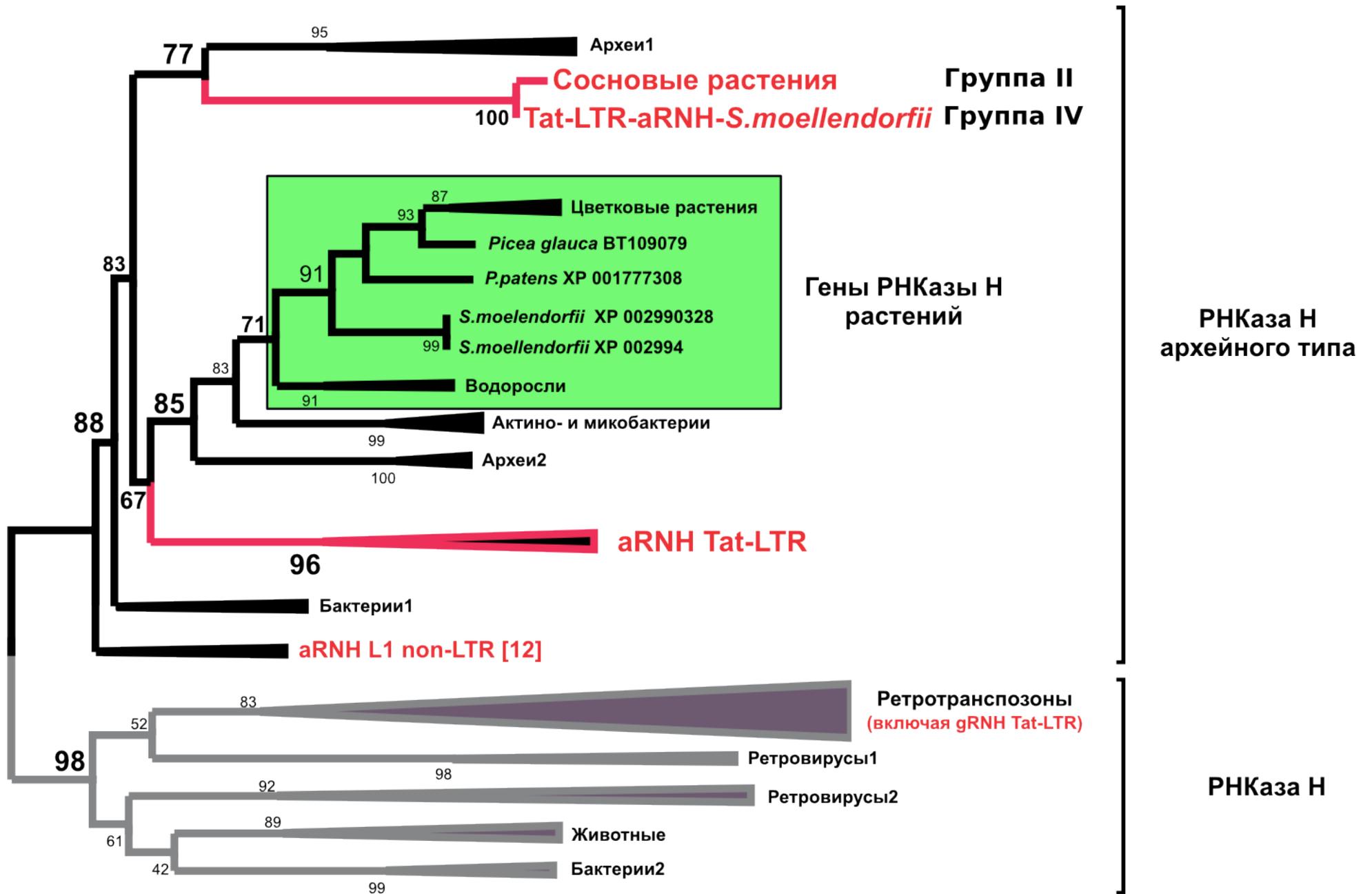
- Всего 5 элементов! Аналогичная структура.
- Самые древние Tat-LTR (~ 420 млн. лет назад)

as – antisense

PMD – Plant Mobile Domain

CC – coiled coil структурный мотив;

# Результаты. Эволюция домена RNH



# Выводы

- В результате биоинформатического и экспериментального анализа LTR-ретротранспозонов, содержащих домен архейной рибонуклеазы H (aRNH) было показано, что они относятся к одному семейству – Tat-LTR-ретротранспозонов и широко распространены во многих таксонах растений, начиная с плауновидных (Lycophyta).
- В рамках филогенетического анализа Tat-LTR ретротранспозонов было выявлено 4 крупных филогенетических кластера и 9 групп Tat-LTR-ретротранспозонов растений, внутри которых элементы характеризуются уникальным набором структурных характеристик.
- В филогенетический анализе генов RNH эукариот было обнаружено, что найденный домен aRNH Tat-LTR является наиболее близким к генам рибонуклеазы H растений, бактерий и архей. Данный факт говорит в пользу гипотезы “захвата” aRNH предковым ретротранспозоном на ранних этапах эволюции.

**Спасибо за внимание!**

