

## Summary

### Aim Цель

- Create a functional annotation of plant LIMEs and test the hypothesis about dependence of LIMEs copies on ploidy.

Создать функциональную аннотацию растительных ЛАЙМ-ов и протестировать гипотезу о зависимости количества копий лаймов от плоидности растений.

### Results Результаты

- The number of LIMEs copies does not depend on ploidy.  
Количество копий лаймов не зависит от плоидности.
- LIMEs are found in variable types of plants RNA, intergenic regions and protein-coding sequences, including exons.  
ЛАЙМ-ы обнаружены в различных видах растительных РНК, межгенных участках и белок-кодирующих последовательностях, включая экзоны.

ЛАЙМ-ы обнаружены в различных видах растительных РНК, межгенных участках и белок-кодирующих последовательностях, включая экзоны.

## Background

Our lab explored genomic regions of extreme sequence conservation - LIMEs (Long Identical Multispecies Elements). These parts of the genome are sequences with 100% identity and minimum length of 20 nucleotides in various species. The reason for this extreme conservation remains a mystery. This elements has been studied in tetrapods but nothing is known about plants.

Наша лаборатория исследовала геномные области с экстремальной консервативностью последовательностей - ЛАЙМЫ (Длинные идентичные мультивидовые элементы). Эти части генома представляют собой последовательности со 100% идентичностью и минимальной длиной 20 нуклеотидов у различных видов. Причина такой консервативности до сих пор остается загадкой. Эти элементы были изучены у четвероногих, но о расположении и функциях их в растениях ничего не было известно.

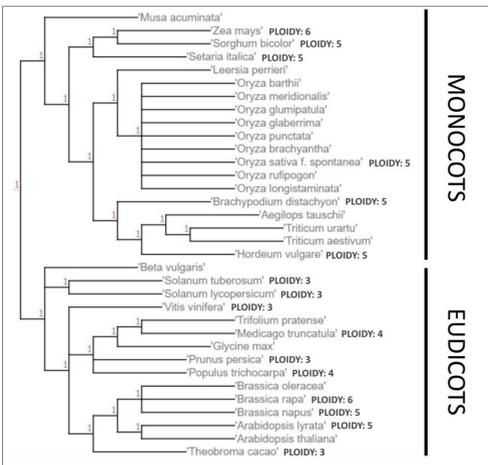


Fig 1. Phylogenetic tree of plant species used in the analysis.

Рис 1. Филогенетическое дерево видов растений, используемых в анализе.

## Methods

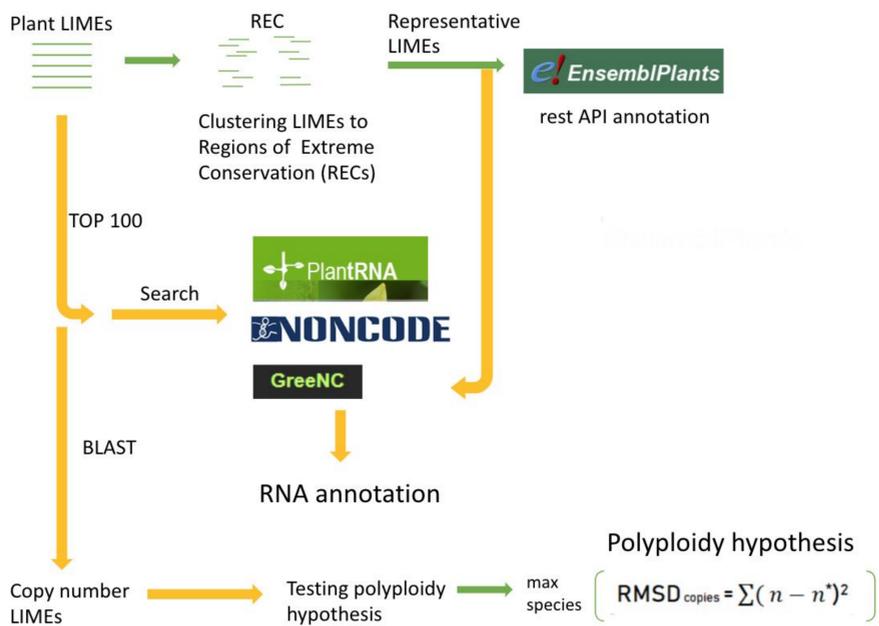


Fig 2. The scheme reflects main stages of our work. We clustered our LIMEs to create the annotation using rest API in Ensembl. Then we took 100 longest LIMEs to search them in RNA databases and created RNA annotation. Finally, we counted the number of LIMEs copies to test polyploidy hypothesis.

Рис 2. Схема отображает основные этапы нашей работы. Мы кластеризовали наши лаймы для аннотации при помощи Ensembl rest API. После этого произвели поиск 100-а самых длинных ЛАЙМ-ов в различных базах данных РНК и составили аннотацию. В конце, мы нашли количество копий ЛАЙМ-ов в разных видах растений для проверки гипотезы о зависимости количества копий от плоидности.

## Results

### Summary

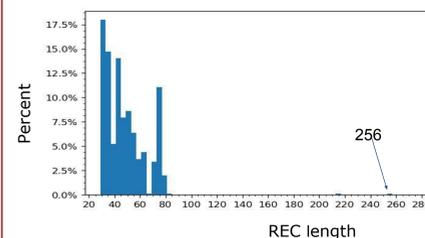
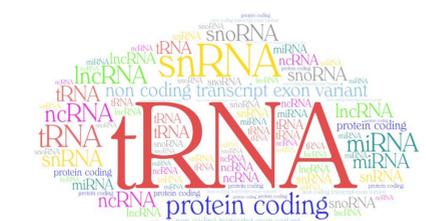


Fig 3. The distribution of lengths of RECs in genome of Arabidopsis thaliana, where "x" is a length and "y" is a percentage of RECs with that length.

Рис 3. Распределение длины REC-ов в геноме Резуховидки Таля, где по оси x – длина, а по оси y – процент REC-ов с такой длиной.



tRNA	529
snoRNA	1
lncRNA	19
snRNA	34
protein_coding	65
ng	1
ncRNA	1
miRNA	1
other	56

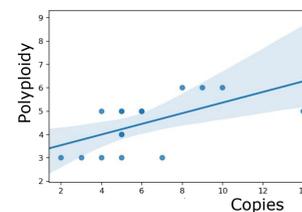


Fig 4. Correlation between polyploidy of plants and quantity of LIME's copies in different species.

Рис 4. Корреляция между полиплоидией растений и количеством копий ЛАЙМ-ов в разных видах.

### tRNA top 100 LIMEs

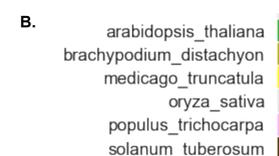
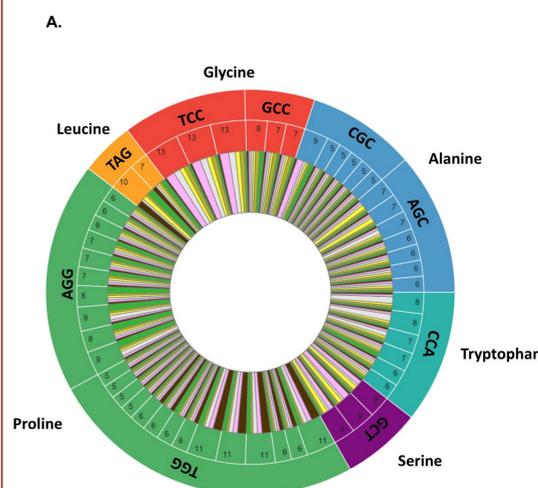


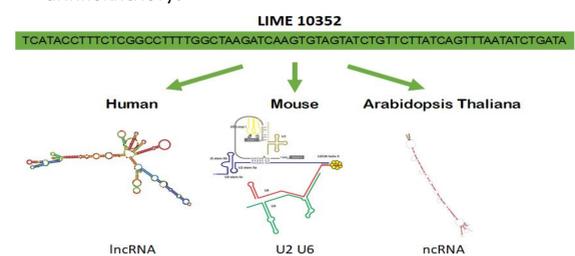
Fig 5. This pie chart reflects the frequency of finding various LIMEs in different tRNA types in diverse plant species. The inner layer stands for species (colours are on the right), the middle one for number of copies of each LIME, and the inside layer is for different anticodons (marked by letters). The colour of the middle and the inside layer reflects the type of amino acid.

Рис 5. Круговая диаграмма частоты встречаемости ЛИМЕ-ов в разных типах тРНК для нескольких видов растений. Внутренний слой показывает распределение видов (цветовая схема В.), средний – количество копий каждого ЛИМЕ-а, внешний слой – различные антикодоны для соответствующих аминокислот. Цвет внешнего и среднего слоя обозначает одну и ту же аминокислоту.

### LIME in different species

Fig 6. This figure reflects the organisms in which LIME 10352 can be found. In human it is located in lncRNA, in mouse - U2 U6 RNA and in Arabidopsis - ncRNA

Рис 6. Схема расположения ЛАЙМа 10352 в разных организмах. В человеке этот ЛАЙМ находится в длинной некодирующей РНК человека (lncRNA), в мыши - U2 U6 РНК, а в Резуховидке в некодирующей РНК (ncRNA).



## Conclusions

- The quantity of copies of different LIMEs presented in researched plant species does not correlate with their ploidy.

Количество копий различных ЛАЙМ-ов в исследуемых видах растений не коррелирует с их плоидностью.

- We found LIMEs in different RNA types and one LIME was found in human's lncRNA, mouse's U2 U6 RNA and Arabidopsis' ncRNA.

ЛАЙМы находятся в разных типах РНК, белок-кодирующих областях, межгенных участках.

- Different RNA types with various LIMEs prevail in diverse species.

В разных организмах преобладают разные типы РНК, в которых встречаются ЛАЙМы.

- Median length of REC - 44 bp, median LIME number in REC - 26.

Медиана длины REC - 44 нуклеотида, а медиана количества ЛАЙМ-ов, входящих в REC - 26.

## Future Directions

- Come up with the explanation of number of LIMEs copies.  
Исследовать различие количества копий ЛАЙМ-ов в разных организмах.
- Compare LIMEs in plant tRNA to those in other species.  
Сравнить ЛАЙМ-ы в растительных РНК с подобными в других видах.
- Analyse LIMEs in protein-coding regions of other organisms.  
Проанализировать ЛАЙМ-ы в белок-кодирующих последовательностях других организмов.

## References

- Nicholas Panchy, Melissa Lehti-Shiu, Shin-Han Shiu. *Plant Physiology*. 2016.
- Korkin, D., Reneker, J. et al. *PNAS*, 109(19). 2012.