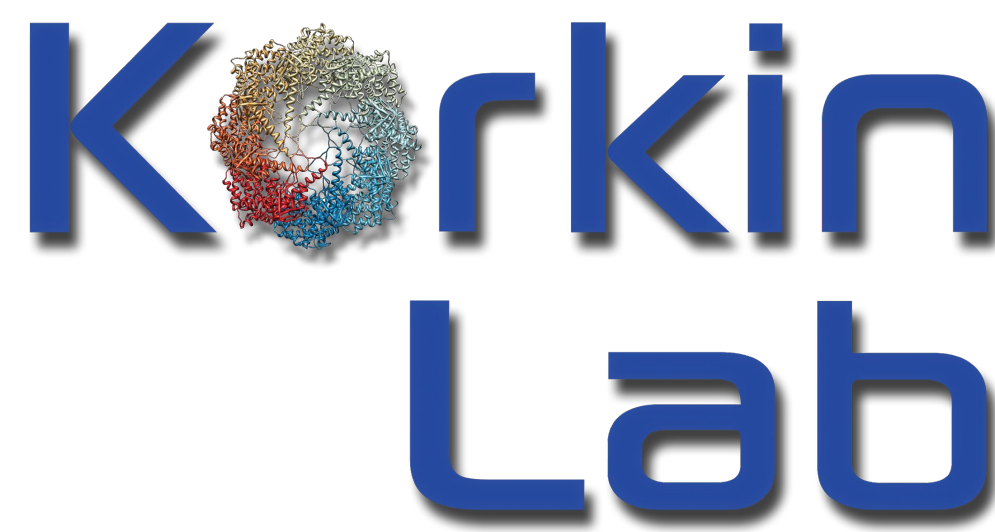




LIMEs in human population

LIME-ы в популяции человека



Chernova Elizaveta*, Katrukha Veronika*, Sverdlin Alexandr*, Bugrova Alisa, Gubler Arseniy, Demianchenko Oleg, Fedorenko Artem, Isakova Lada, Kuznetsova Anastasia, Lyovina Tatiana, Pelevin Arseniy, Rodionova Emma, Semenova Ariadna, Ivankov Dmitry, Kondrashov Fyodor, Lyulina Anastasia, Korkin Dmitry^{1,2,3} and Nuzhdina Ekaterina⁴

¹Bioinformatics and Computational Biology Program, ²Data Science Program, and ³Department of Computer Science Worcester Polytechnic Institute, MA, USA, ⁴BostonGene, Boston, USA

Summary

Problem:

Determine if LIMEs are variable in the human population and characterize the genetic variation

Определить вариативны ли LIME-ы в популяции человека и охарактеризовать найденные мутации

Results:

- We discovered that LIMEs are variable in the human population and mutations have different characteristics on their genomic position, frequency, and ethnic distribution
- We found sequence motifs shared between different functional groups of LIMEs

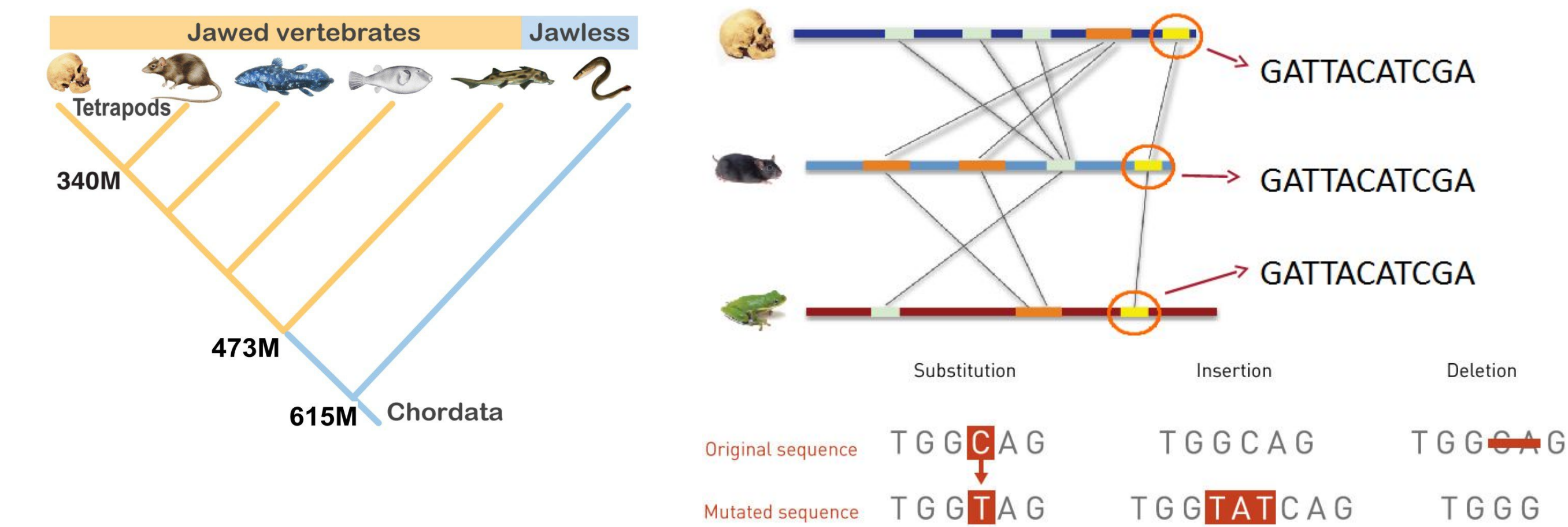
- Обнаружили, что LIME-ы вариабельны в человеческой популяции. Мутации отличаются локализацией, частотой встречаемости в этнических группах
- LIME-ы содержат мотивы, характерные лишь для определенных функциональных групп



Background

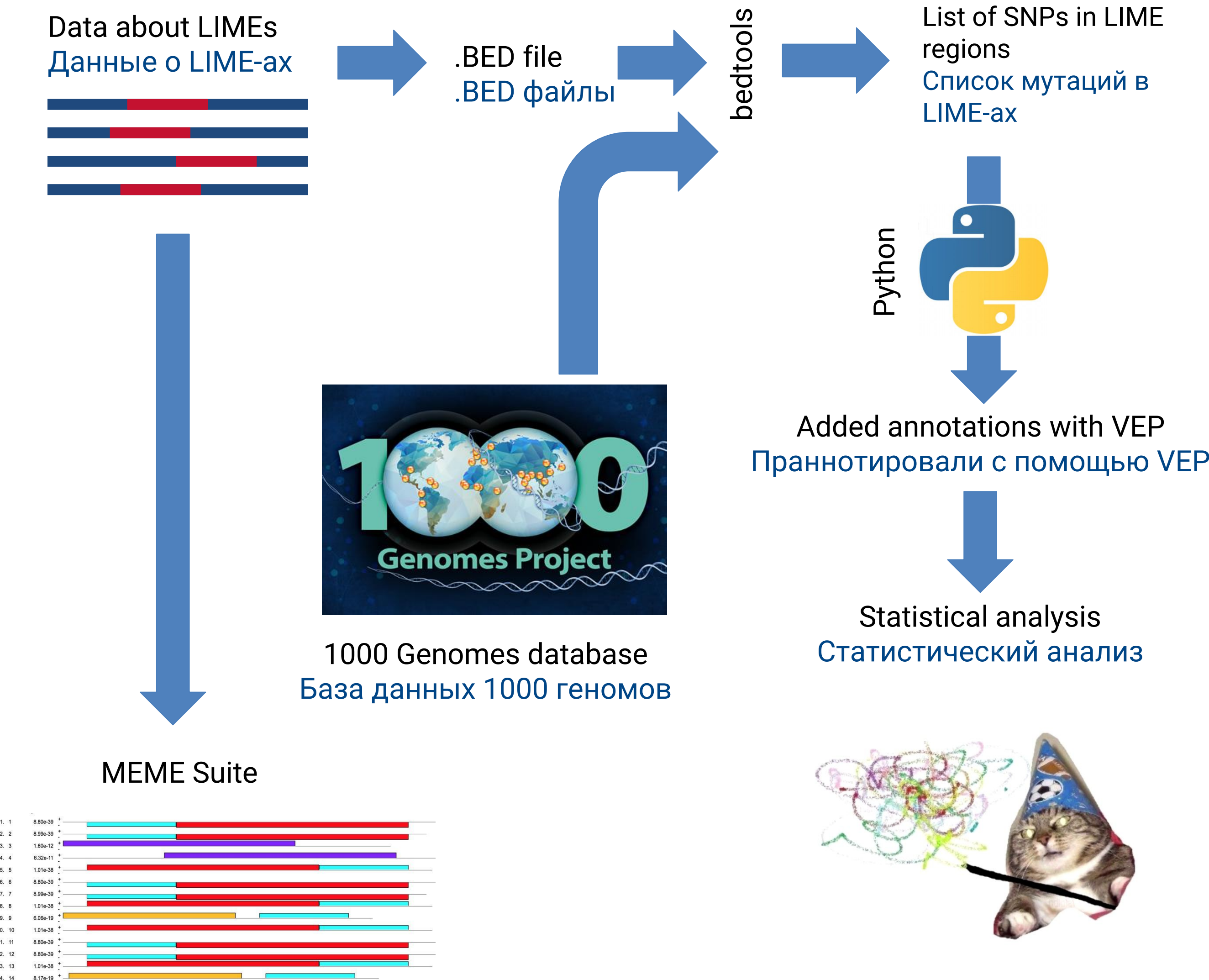
LIMEs (Long Identical Multispecies Elements) are extremely conservative sequences in vertebrates genomes, some of which have remained unchanged after a billion years of evolution. But do they differ across populations? To find the answer, we've analyzed mutations in LIMEs of different human populations, using the 1000 Genomes data

LIME-ы - экстремально консервативные участки генома позвоночных, оставшиеся неизменными на протяжении миллиарда лет эволюции. Но насколько они вариабельны в популяции? Мы задались этим вопросом и стали исследовать мутации в LIME-ах в геноме человека. Данные были взяты из проекта "1000 геномов", в котором есть геномы здоровых людей из разных популяций.



- 2,600 genomes sequenced
- 5 ethnic groups
- Healthy individuals
- 91.6 mln in/dels and substitutions

Methods



References

- 1.Korkin, D., Reneker, J. et al. PNAS, 109(19). 2012.
- 2.Craig Venter, J., Telenti, A. et al. Nature Genetics, 50. 2018.
- 3.The 1000 Genomes Project Consortium, Nature, 526. 2015.
- 4.Haussler, D., Bejerano, G., et al. Science, 304. 2004.

Results

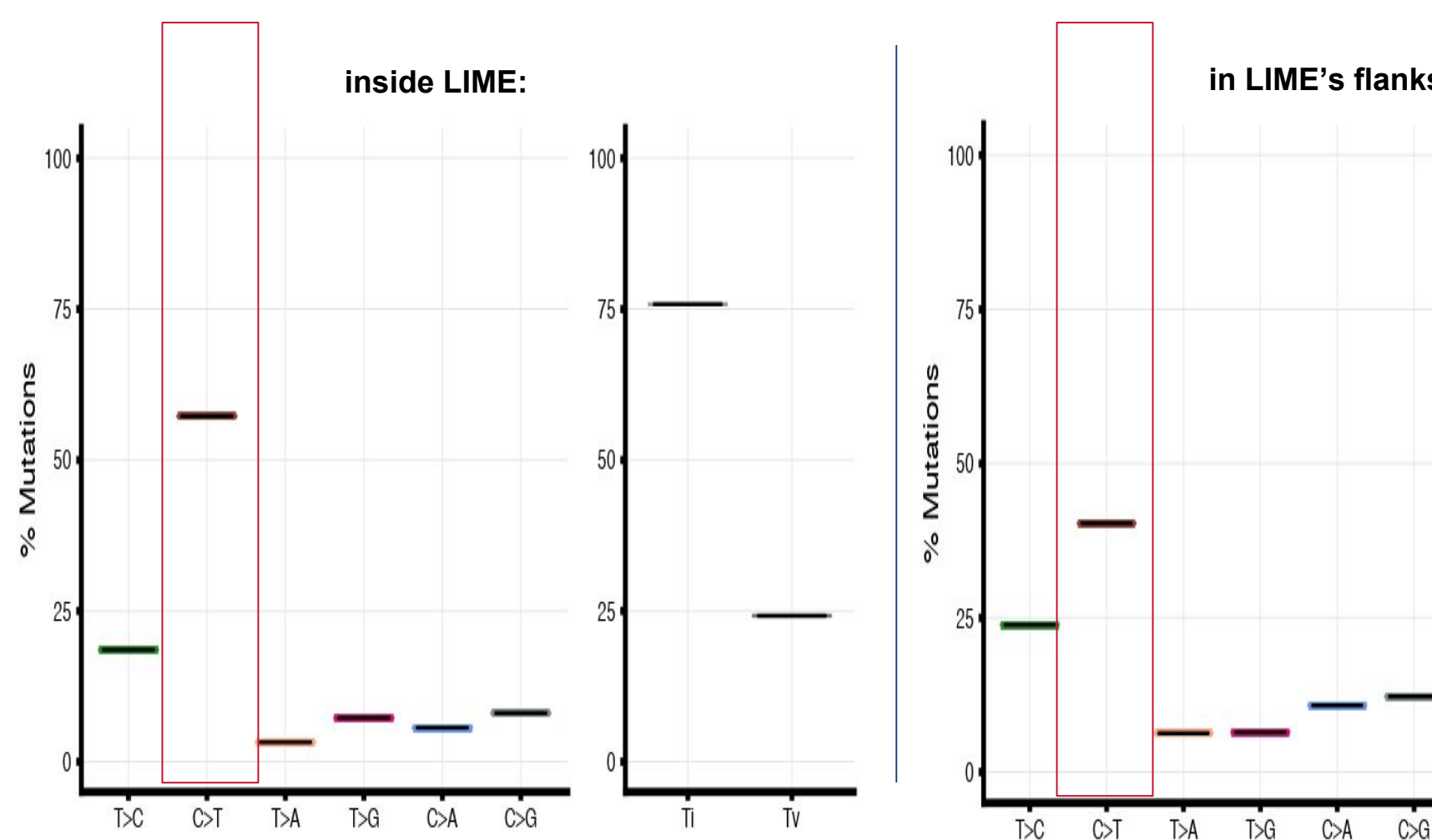
What was analyzed:

- The frequency of mutations in UTR, introns and non-genes in general, by RNA type and by chromosomes
- Types of mutations and their ratio in LIMEs and their flanking regions
- Motifs of LIMEs' sequences
- Frequencies of LIMEs' mutations in human populations

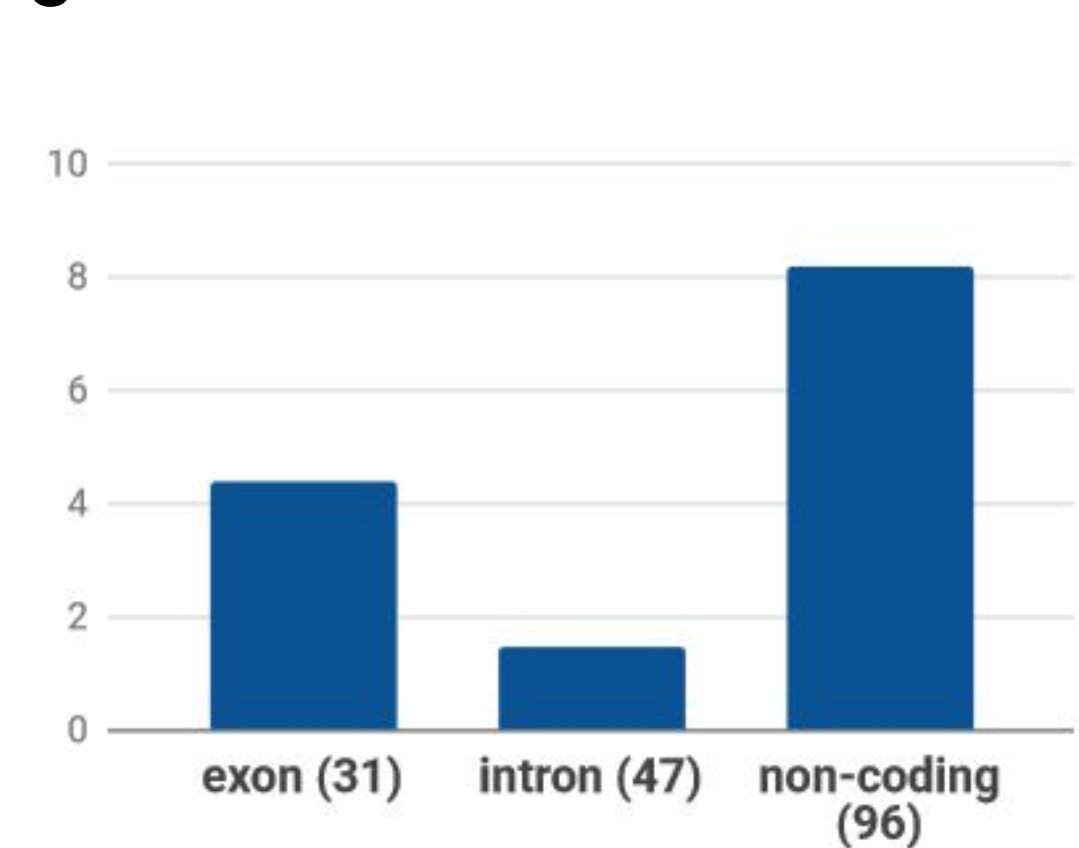
Что мы проанализировали:

- Частоту мутаций в нетранслируемых, интронных и межгенных LIME-ах в целом, по типу РНК и по хромосомам
- Типы мутаций и их соотношение в LIME-ах и их ближайшем окружении
- Мотивы последовательностей LIME-ов
- Частоту мутаций в LIME-ах по популяции людей

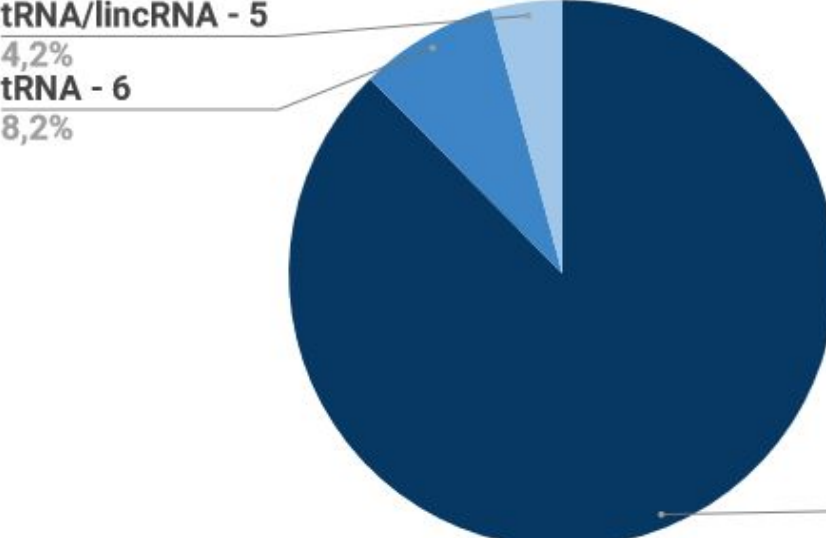
Percent of mutations in UTR LIMEs



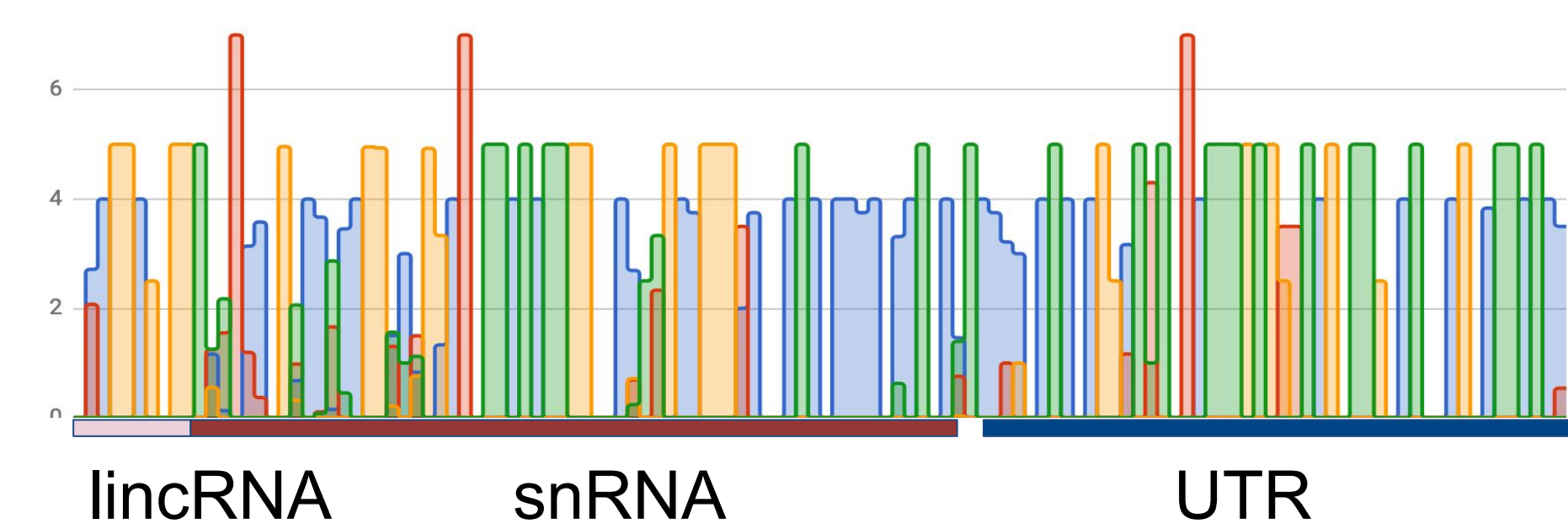
Average number of mutations in LIME



Distribution of mutations in different RNA types in UTR LIMEs



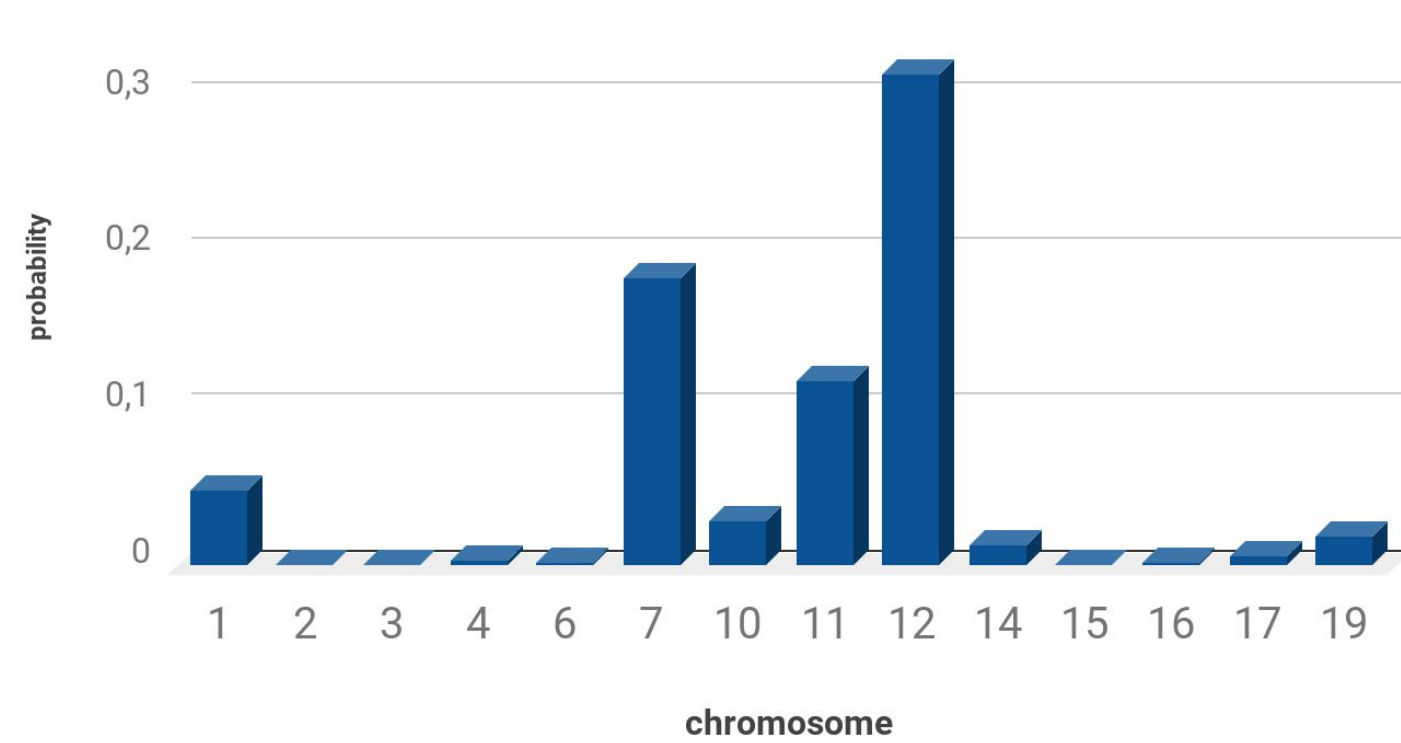
Mutation in human population



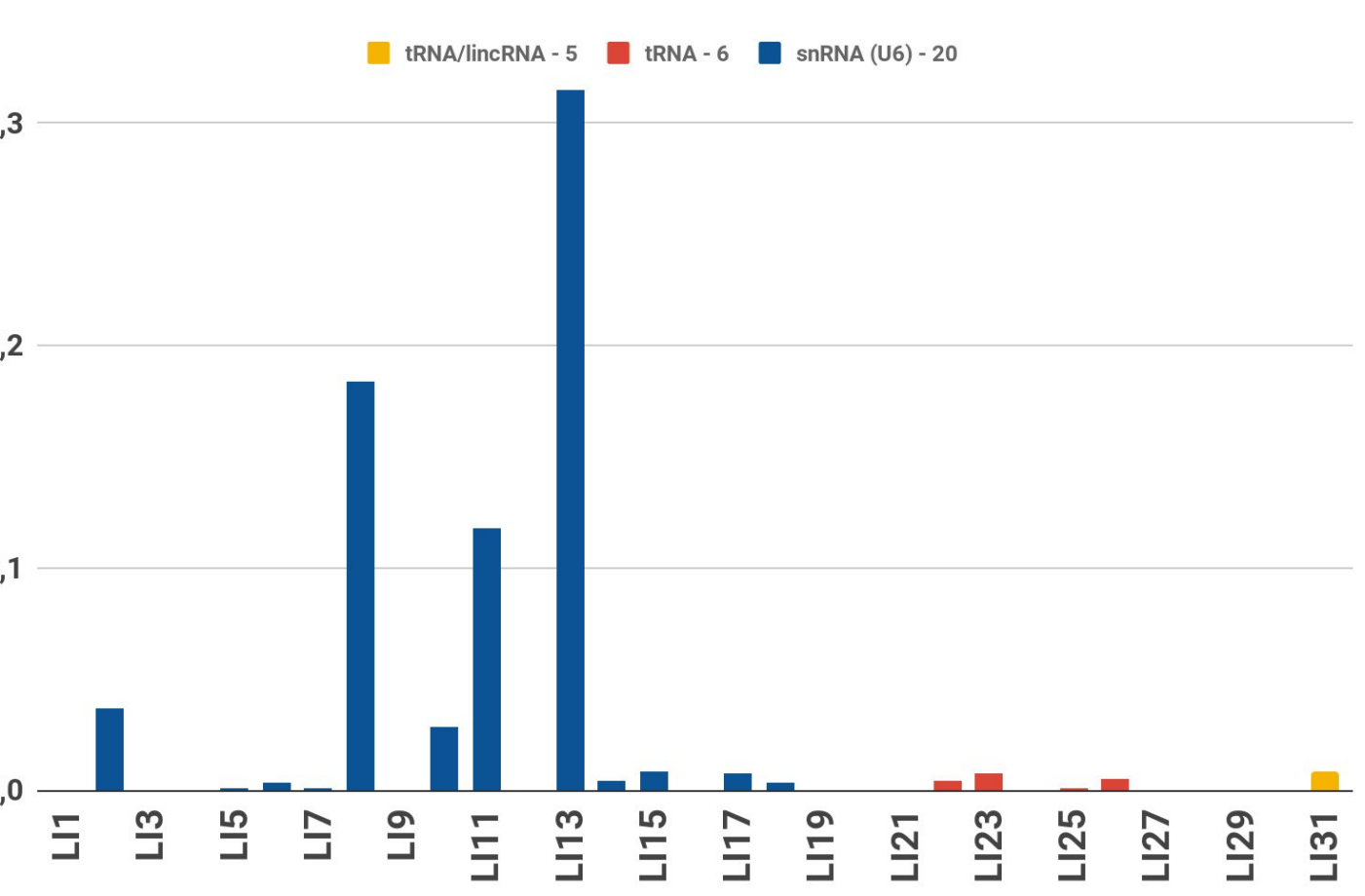
Motifs in UTR LIMEs



Probability of UTR LIMEs mutations by chromosomes



Probability of mutations in UTR LIMEs



Conclusions

- Different mutations are found in copies of the same LIME; the majority of them do not overlap.
- The frequency of mutations differs significantly across different LIMEs and chromosomes
- Intronic LIME have fewer mutations than UTR and non-genic LIMEs
- Unlike LIMEs found in tRNAs and lincRNAs, those found in snRNAs has motifs; LIME in snRNAs tend to have a much higher frequency of mutation occurrence
- The ratio of transversions to transitions of LIMEs is 1.5 times higher than in their flanking regions
- Найдены различные мутации в копиях одного и того же LIME, большинство из которых не пересекается.
- Частота мутаций существенно различается в разных LIME-ах и в хромосомах.
- Интронные LIME имеют меньше мутаций, чем UTR и некодирующие.
- В отличие от tРНК и lincRNA, snRNA, содержащая LIME-ы, имеет мотивы. Кроме того, snRNA имеет гораздо большую частоту мутаций.
- Отношение трансверсий к транзициям в LIME-ах больше в 1.5 раза, чем в их окружении

Future Directions

- Analyze the distribution of mutations in copies of the same LIME sequence or motif
- Investigate population variability of LIMEs within other species and compare to variability in humans
- Comparatively analyze mutations by RNA types in introns and non-coding genome

- Проанализировать распределение мутаций в копиях LIME и в найденных мотивах
- Изучить изменчивость LIME внутри других видов и сравнить с изменчивостью в человеческой популяции
- Анализ мутаций внутри разных групп по типам РНК в интронах и межгенных участках