

# Оценка разнообразия и степени дивергенции повторов в различных эукариотических геномах

ПАВЕЛ ДОБРЫНИН, НИКИТА КАРТАШОВ, МАРИЯ КЛЮЕВА,  
ВАРВАРА КОНОВА, АЛЕНА МАЯКОВСКАЯ, НАДЕЖДА ПОТАПОВА

# Задачи

- ▶ Попытаться найти области консервативности в повторах двух типов LINE/L1 и SINE/tRNA.
- ▶ Определить насколько сильно отличаются повторы разных видов внутри семейства, используя дивергенцию от исходного генома как метрику.

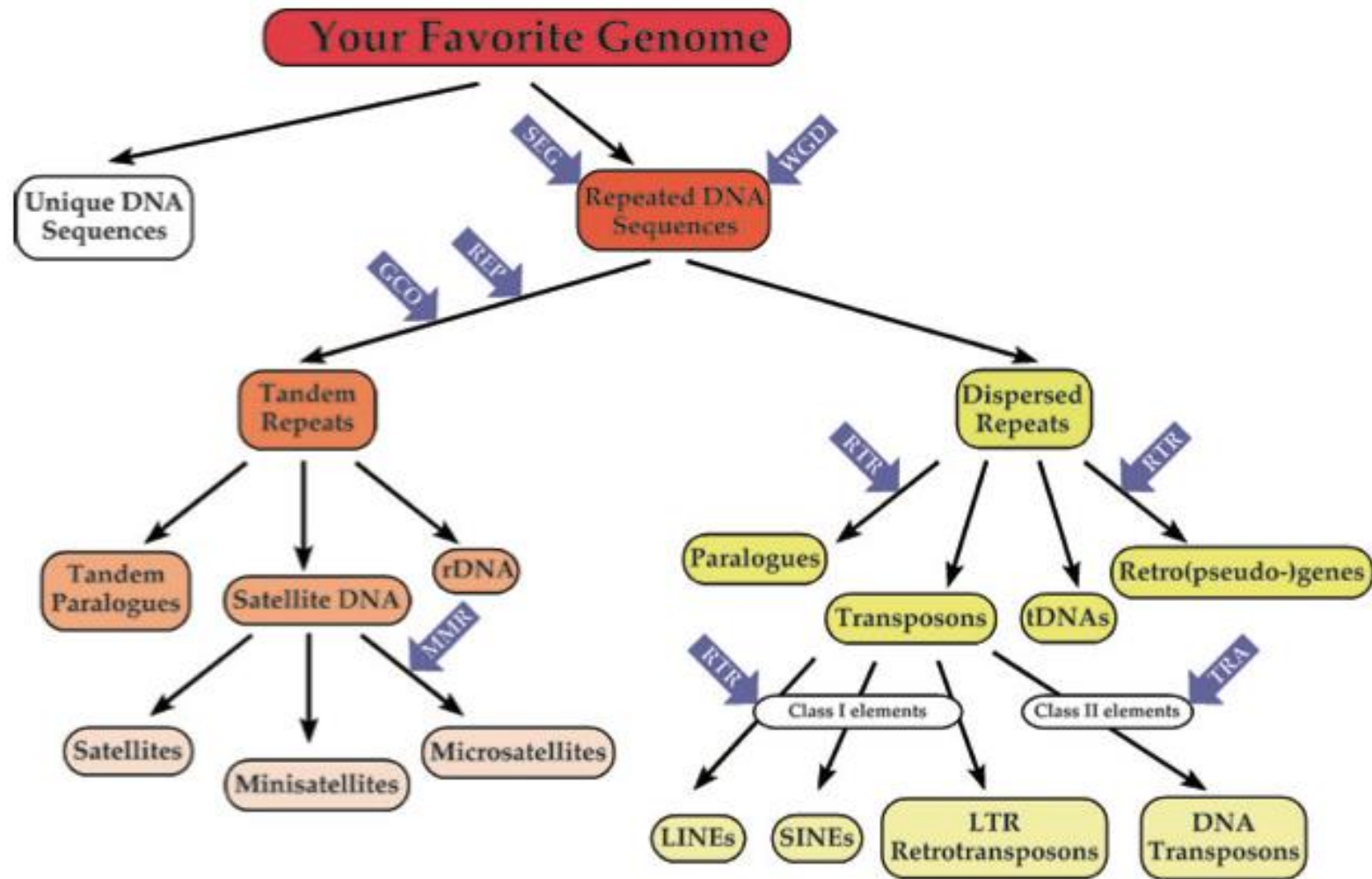
# Что?

Повторы - дублирование цепочек нуклеотидов в геномной последовательности.

Повторы: тандемные и диспергированные.

Тандемные повторы не были рассмотрены в результате работы.

# Диспергированные повторы

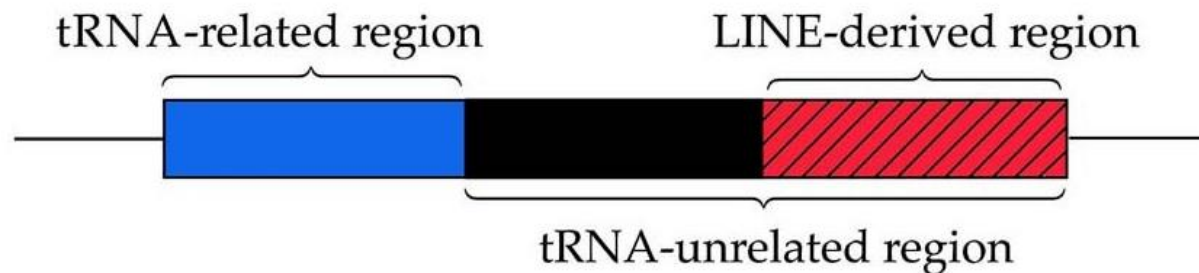


# LINE/L1

- ▶ Длина примерно 6-8 kbp
- ▶ Является автономным ретроэлементом
- ▶ Составляют до 20% генома человека/мыши

# SINE/tRNA

- ▶ Длина примерно 100-300 bp
- ▶ Самый частый элемент избыточности в геноме млекопитающих



# LTR

- ▶ Длину примерно 640 bp
- ▶ Рассмотренный представитель - MaLR
- ▶ Паразитируют на SINE

# Stats

- ▶ Входные данные - геном гепарда *acinonyx jubatus*
- ▶ Обработан: TRF, dust, RM для поиска повторов, выравнивание выполнялось в *Clustal Omega* и *MUSCLE*
- ▶ Размер входного генома: 2 375 874 546 bp
- ▶ Всего найдено повторов: 3 558 537
- ▶ из полностью совпавших: 107 761
- ▶ Всего повторами замаскировано около 40% генома

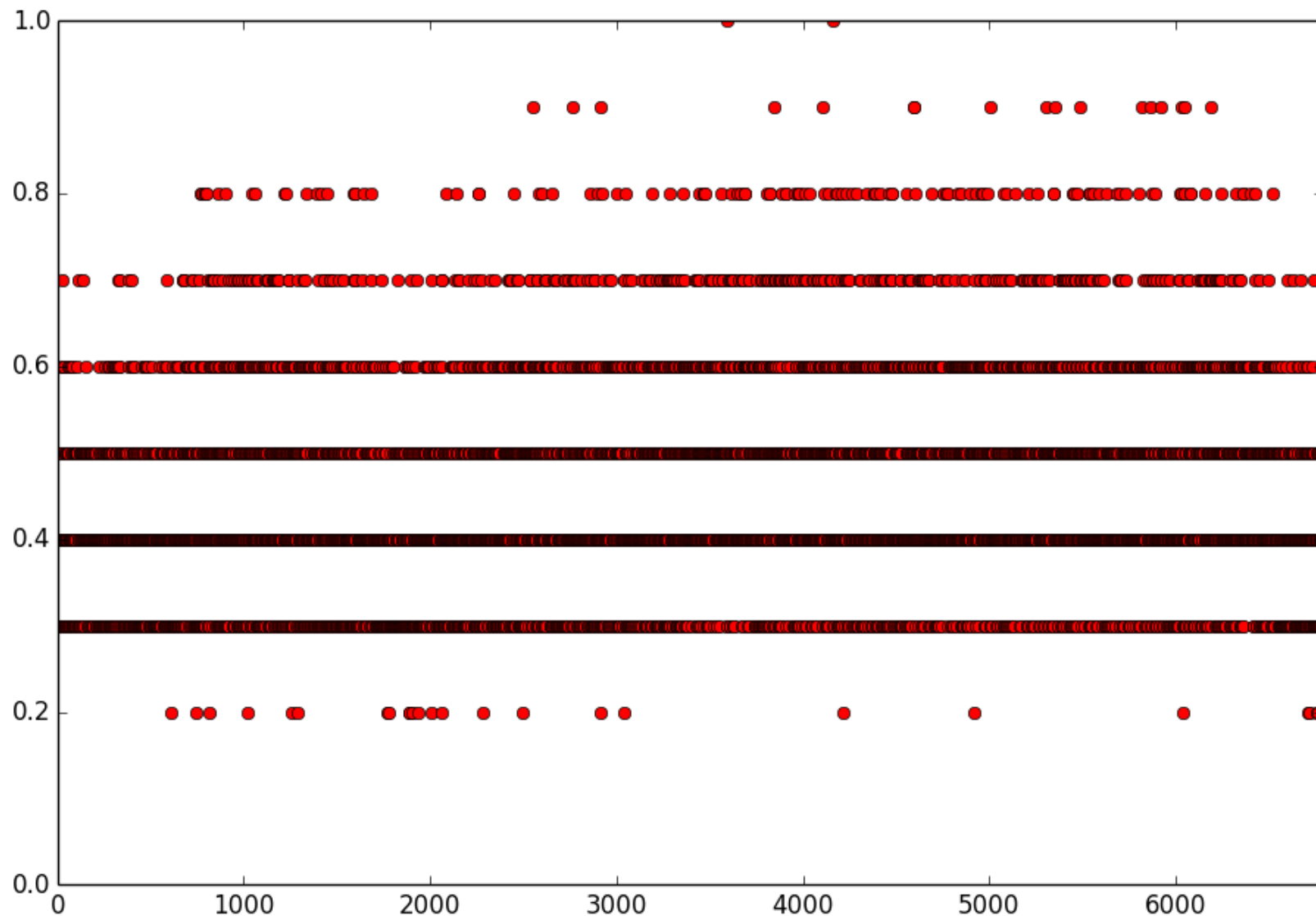


## Stats contd.

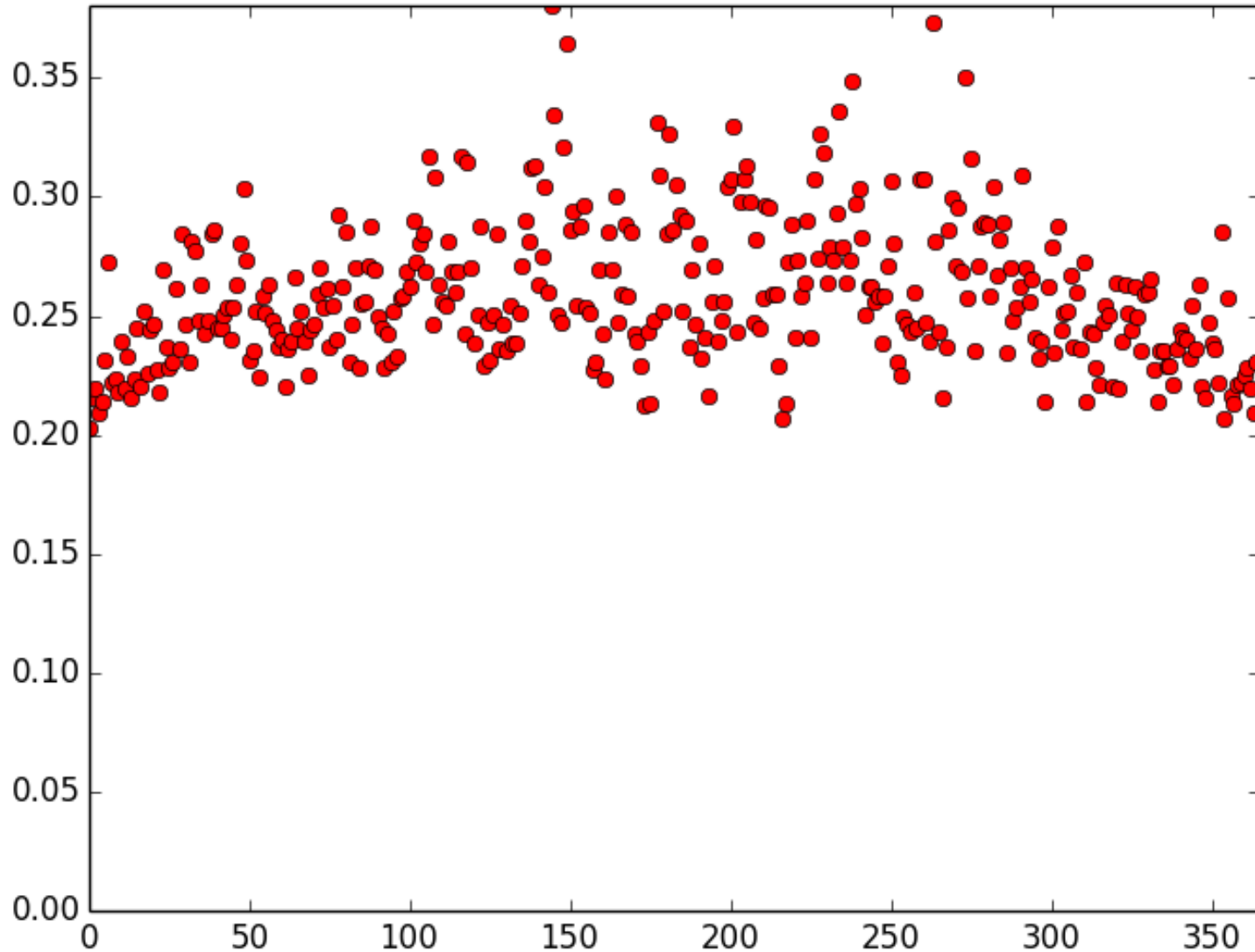
- ▶ LINE/L1: выбранные представители (количество)  
L1\_Carn1 (11)
- ▶ SINE/tRNA: SINEC\_Fc (7632), SINEC\_Fc2 (33292),  
SINEC\_Fc3 (39905)
- ▶ LTR/ERV1: MLT1A0 (1351), MLT1B (1271), MLT1D (1171),  
MLT1C (950)

# Консервативность LINE повторов

По горизонтали:  
номер нуклеотида в  
последовательности  
По вертикали:  
консервативность

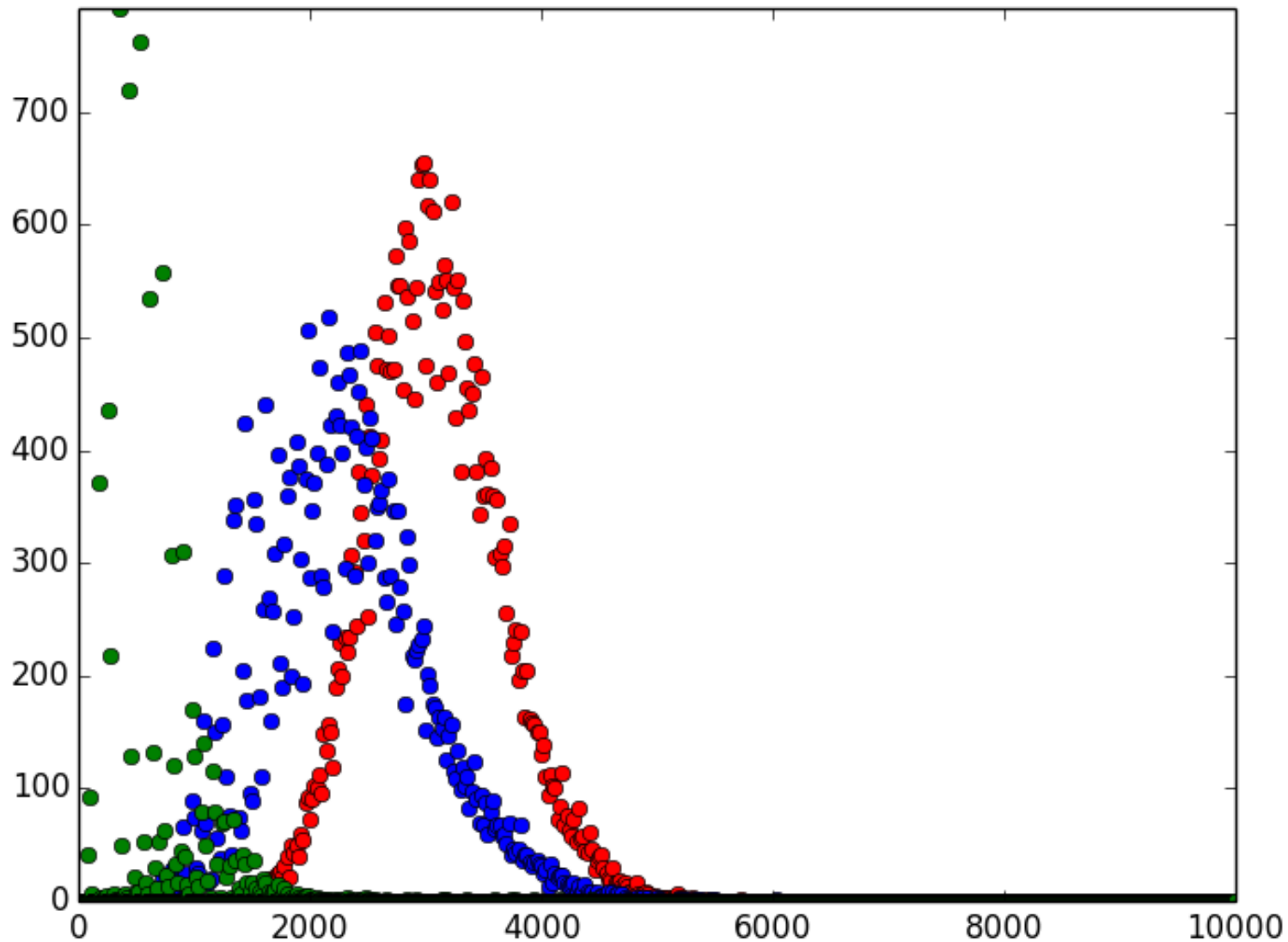


# Консервативность LTR/ERV1 повторов



По горизонтали:  
номер нуклеотида в  
последовательности  
По вертикали:  
консервативность

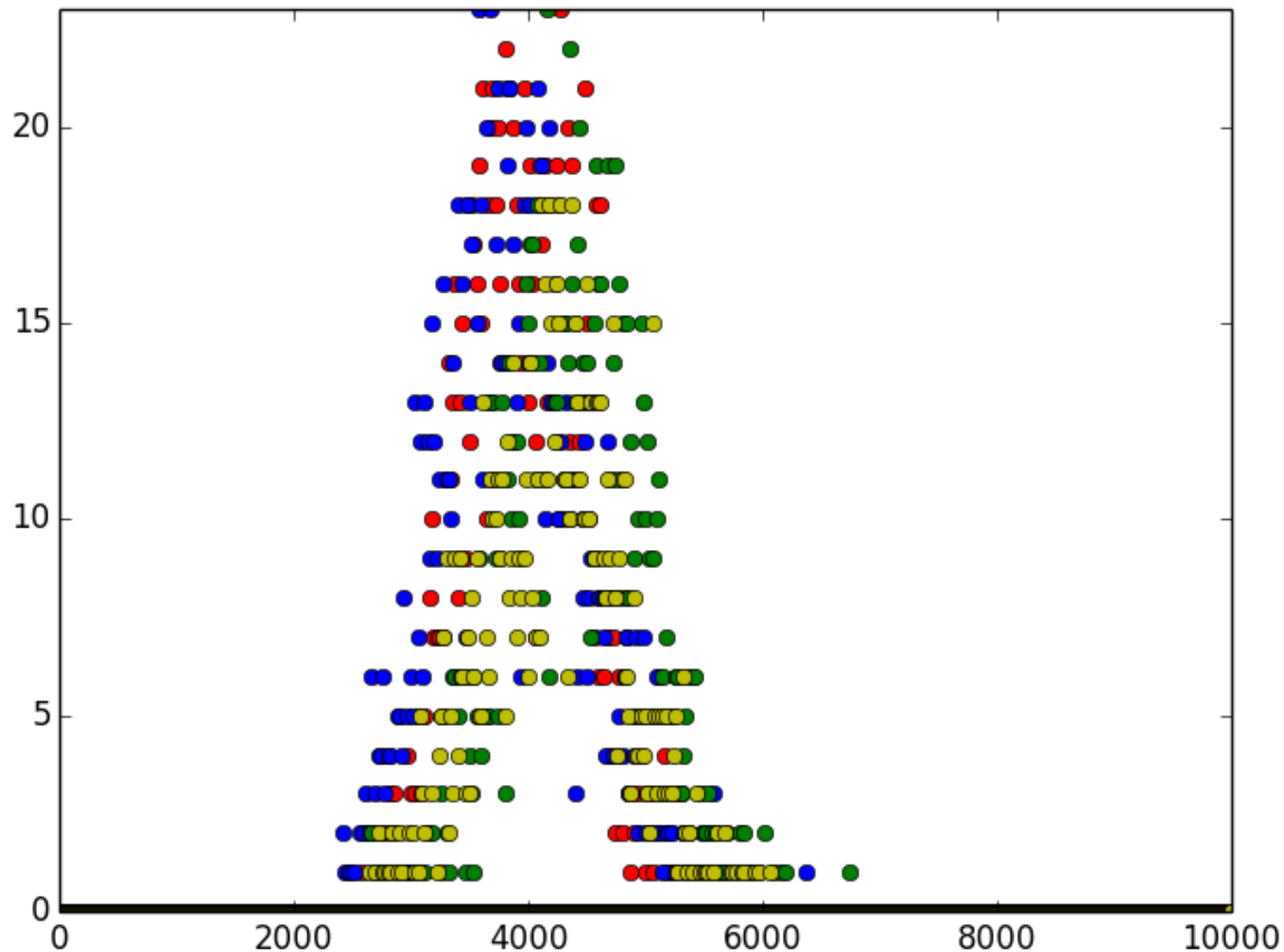
# SINE/tRNA повторы



По горизонтали:  
дивергенция от 0 до 50  
По вертикали:  
количество повторов

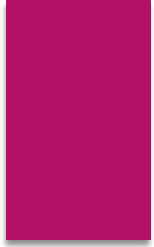
SINEC\_Fc3 - красным,  
SINEC\_Fc2 - синим,  
SINEC\_Fc - зеленым.

# LTR/ERV1 повторы



По горизонтали:  
дивергенция от 0 до 50  
По вертикали:  
количество повторов

MLT1A0 - красным,  
MLT1B - синим,  
MLT1D - зеленым,  
MLT1C - желтым.



Спасибо за внимание..... Спасибо за внимание