



# Всё как у людей

## Amoeba's chromatin domains are similar to human's

Сергей Исаев, Екатерина Храмеева

### Abstract

In multicellular eukaryotes, chromatin is organized in compact structures called topologically associating domains (TADs). The frequency of contacts within TADs is much higher than between TADs, as seen in the contact maps. We have analyzed features genes situated in TADs or between TADs in slime mold *Dictyostelium discoideum*, and demonstrated that features such as localization of housekeeping and stage-specific genes in this organism are similar to those of animals (*Drosophila*, chicken, human).

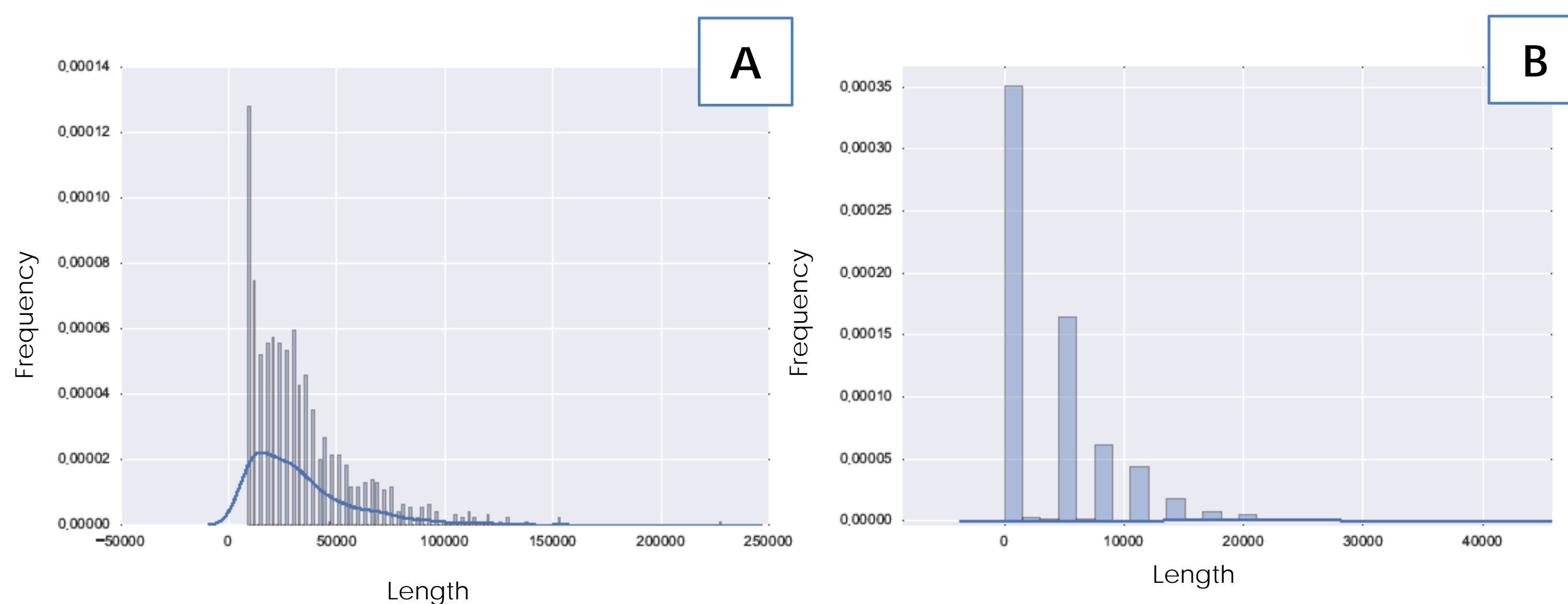
### Введение

Хроматин существует в виде топологически ассоциированных доменов (ТАДов), внутри которых вероятность контакта ДНК значительно выше, чем вне (что видно на картах контактов). Особенности распределения генов на регионы генома, которые свёрнуты или не свёрнуты в ТАДы, могут сказать о функциональной роли такой упаковки ДНК. Анализ этих особенностей уже неоднократно проводился для человека, курицы, дрозофилы, однако неясно, актуальны ли изученные закономерности для других эукариотических организмов, особенно филогенетически далёких.

### Задачи

Основная задача проекта – проанализировать распределение генов амёбы *Dictyostelium discoideum* в контексте пространственной организации генома и определить, имеет ли место корреляция между положением гена и особенностями его экспрессии.

### Распределение длин ТАДов и межТАДов



Гистограмма распределения длин ТАДов (А) и межТАДов (В). Средняя длина ТАДа (33784) больше, чем средняя длина межТАДа (5848).

Histograms of the length distribution of TADs (A) and interTADs (B). The average TAD length (33784) is larger than the average interTAD length (5848).

### Количество белок-кодирующих и белок-некодирующих генов в ТАДах, межТАДах

	В ТАДах	В межТАДах
Белок-кодирующие гены	10660	2459
Белок-некодирующие гены	404	79

Значение критерия  $\chi^2$ : 1.75, p-value > 0.05. Зависимость не выявлена.

There is no significant differences in the number of protein-coding and non-coding genes in TADs and interTADs (p > 0.05, Chi-squared test).

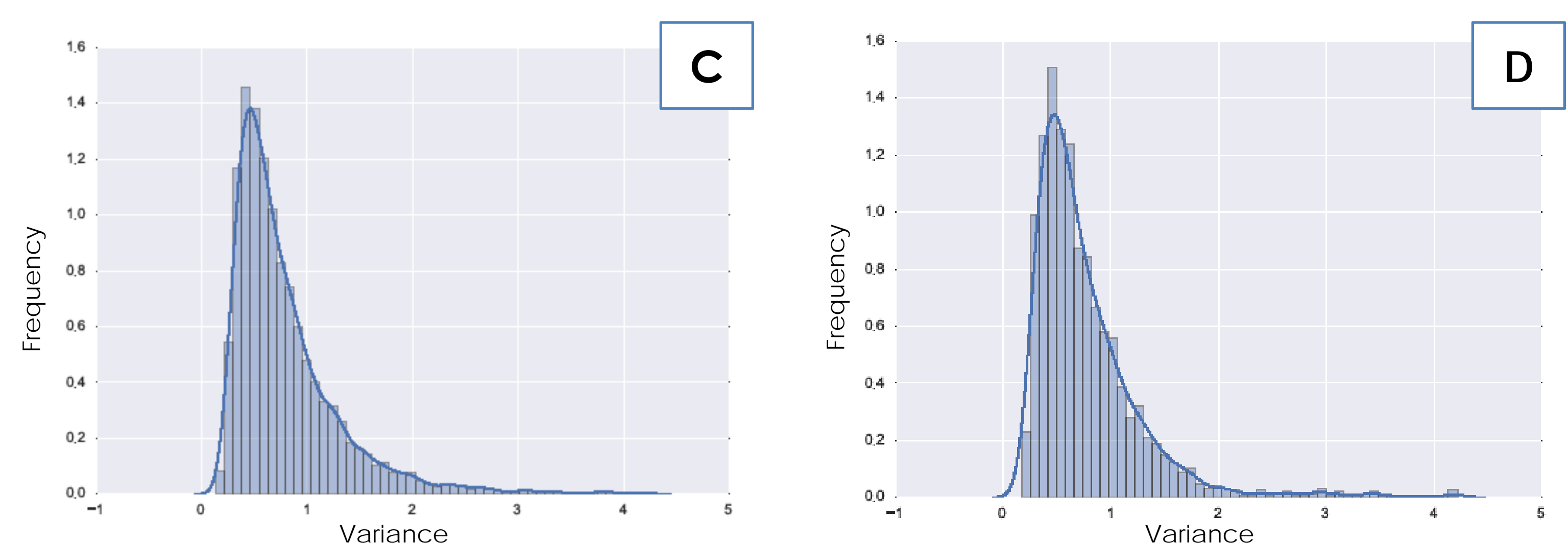
### Особенности расположения канонических генов «домашнего хозяйства» в ТАДах, межТАДах

	В ТАДах	В межТАДах
Гены «домашнего хозяйства»	2882	780
Все гены	11196	1845

Значение критерия  $\chi^2$ : 110.415, p-value < 0.01. Канонические гены «домашнего хозяйства» более склонны находиться в межТАДах, чем в ТАДах.

Housekeeping genes tend to be localized in TADs compared to interTADS (p < 0.01, Chi-squared test).

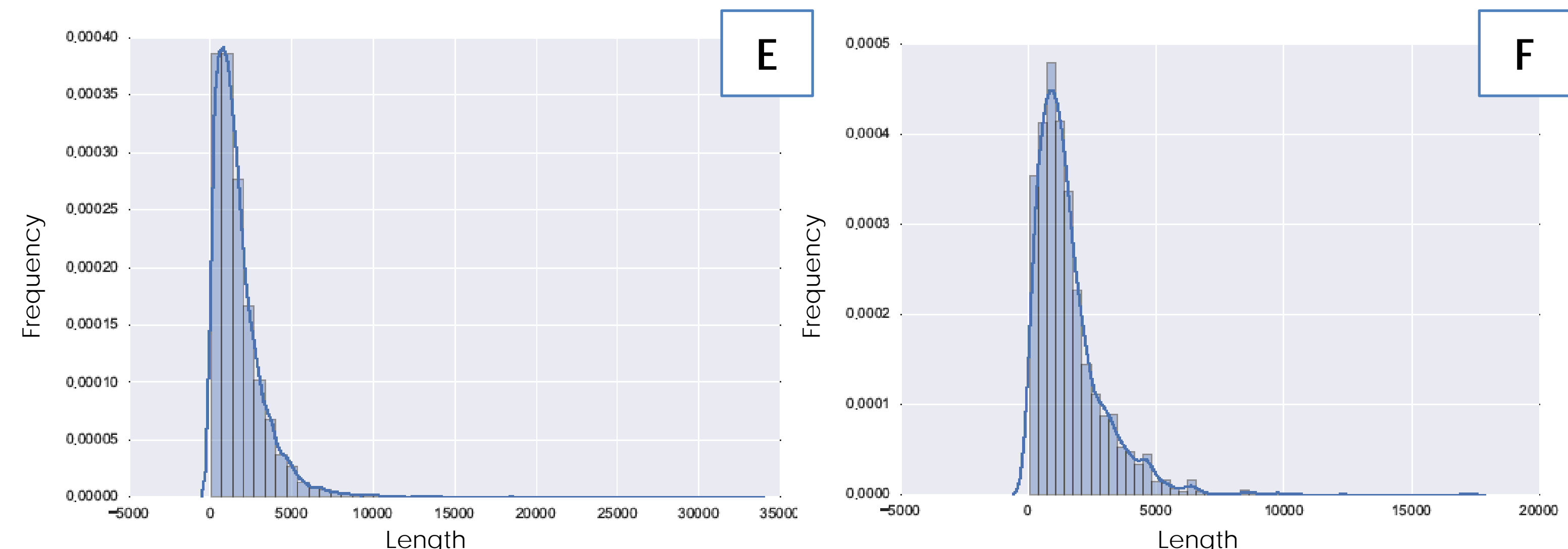
### Особенности расположения генов с различной дисперсией по экспрессии



Среднее значение дисперсии в ТАДах (С) - 0.8, в межТАДах (D) - 0.77. Разница была проверена U-критерием Манна — Уитни, p-value < 0.05, статистическая значимость, по-видимому, достигается из-за большого количества наблюдений, поэтому зависимость не выявлена.

Genes with high variability of expression at different states demonstrate a significant, but weak tendency to be localized in TADs (p < 0.01, Mann-Whitney U-test). However, the average variance of expression of TAD (C) genes (0.8) is only slightly larger than that of interTAD (D) genes (0.7), so the biological significance of this observation is questionable.

### Распределение длин генов в ТАДах, межТАДах



Средняя длина гена в ТАДах (E) 1739.23, в межТАДах (F) - 1615.61. Разница была проверена U-критерием Манна — Уитни, p-value < 0.05, статистическая значимость, по-видимому, достигается из-за большого количества наблюдений, поэтому зависимость не выявлена.

TAD (E) genes tend to be slightly longer than interTAD (F) genes (p < 0.01, Mann-Whitney U-test). However, the average length of TAD genes (1739) is only slightly larger than that of interTAD genes (1615), so the biological significance of this observation is questionable.