



# Влияет ли размер на число контактов?

## Does size matter?

Артём Федоренко, Маргарита Самборская

### Abstract

Chromatin structure is organised into dense globular structures called Topologically Associating Domains (TADs). The frequency of chromatin contacts within TADs is much higher than between TADs, it varies for TADs of different sizes. We studied how TAD "density" (contact frequency within TADs) and probability to interaction with other genome regions depend on TAD size. These characteristics were compared for *H. sapiens*, *M. musculus*, *D. melanogaster*, and *D. discoideum*.

### Введение

Структура хроматина организована в плотные глобулярные структуры, называемые топологически ассоциированными доменами (ТАДы). Частота контактов хроматина в ТАДах намного выше, чем в участках между ТАДами. ТАДы разных размеров варьируют по плотности упаковки и способности к взаимодействию с другими участками генома. В нашем проекте мы изучали зависимости характеристик ТАДов от их размеров для филогенетически далёких организмов.

### Ход работы

### Нахождение границ ТАДов

Границы ТАДов находились с помощью алгоритма «Armatus» при варьировании параметра  $\gamma$  для нахождения оптимальной разметки.

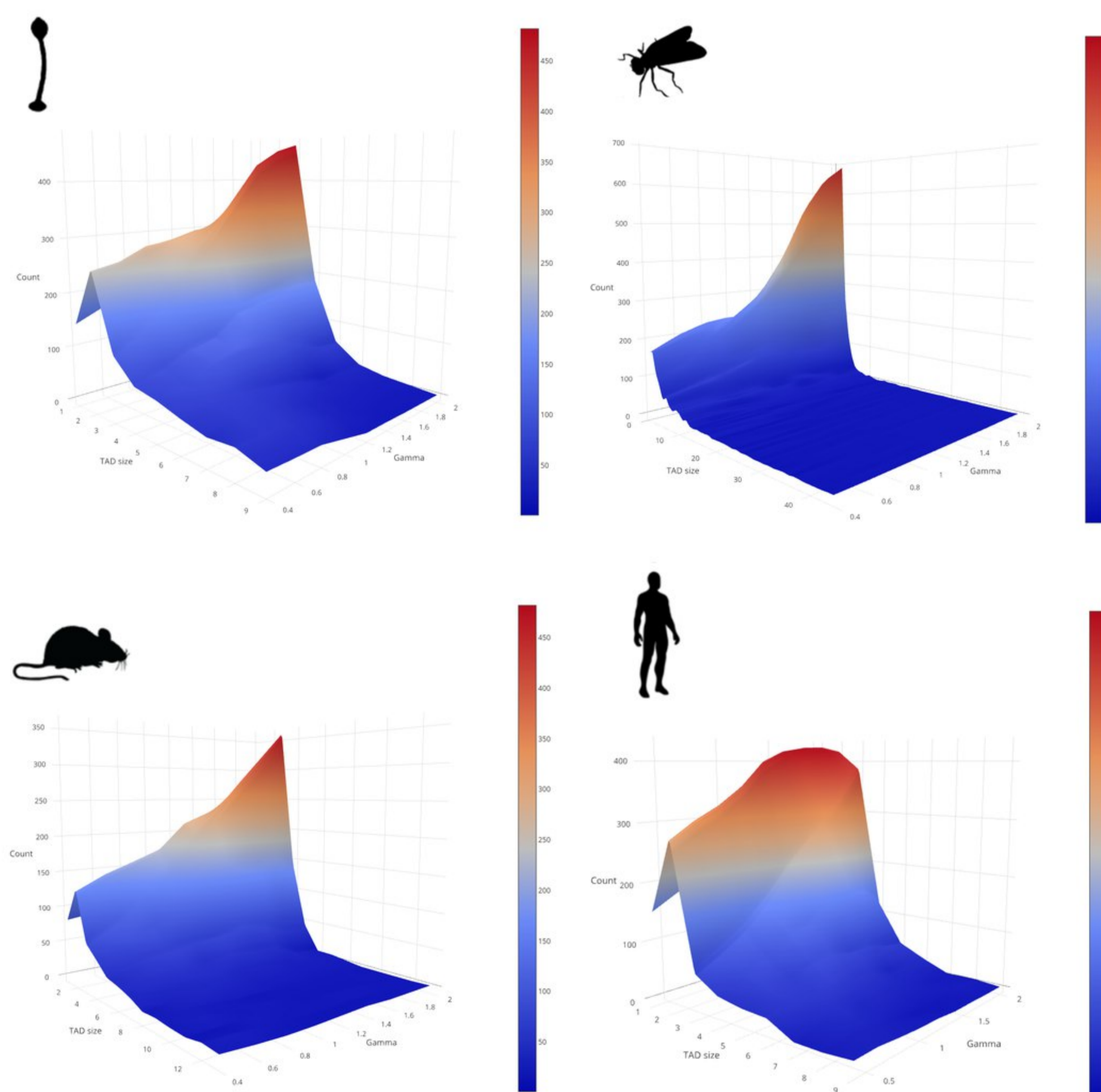


Рис.1 3D-распределение размеров ТАДов в зависимости от параметра  $\gamma$ , построенное по итеративно-корректированным данным.

Pic. 1 3D-distribution of TAD sizes depending on the  $\gamma$  parameter, based on iteratively corrected data.

### Контакты внутри ТАДа

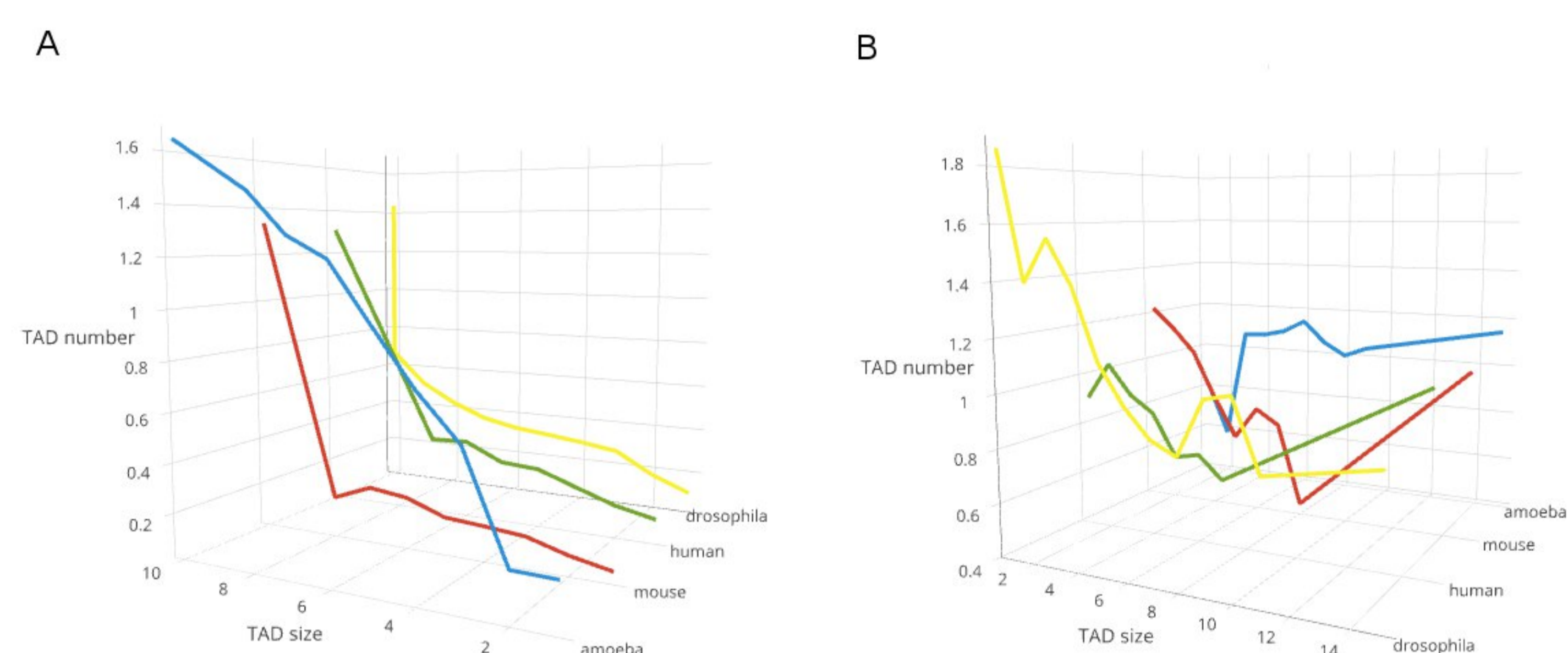


Рис. 2 Зависимость числа контактов внутри ТАДа от его размера для четырех организмов. Проводилась нормализация на среднее числа контактов для каждой диагонали (A), либо на среднее число контактов случайно выбранных подматриц, имеющих размер данного ТАДа (B).

Pic. 2 Dependence of the number of contacts within TADs on their size for four organisms. Contact frequencies were normalized by the average number of contacts for each diagonal (A), or by the average number of contacts for randomly chosen regions of the size as the TAD (B).

### Взаимодействие ТАДа с его окружением

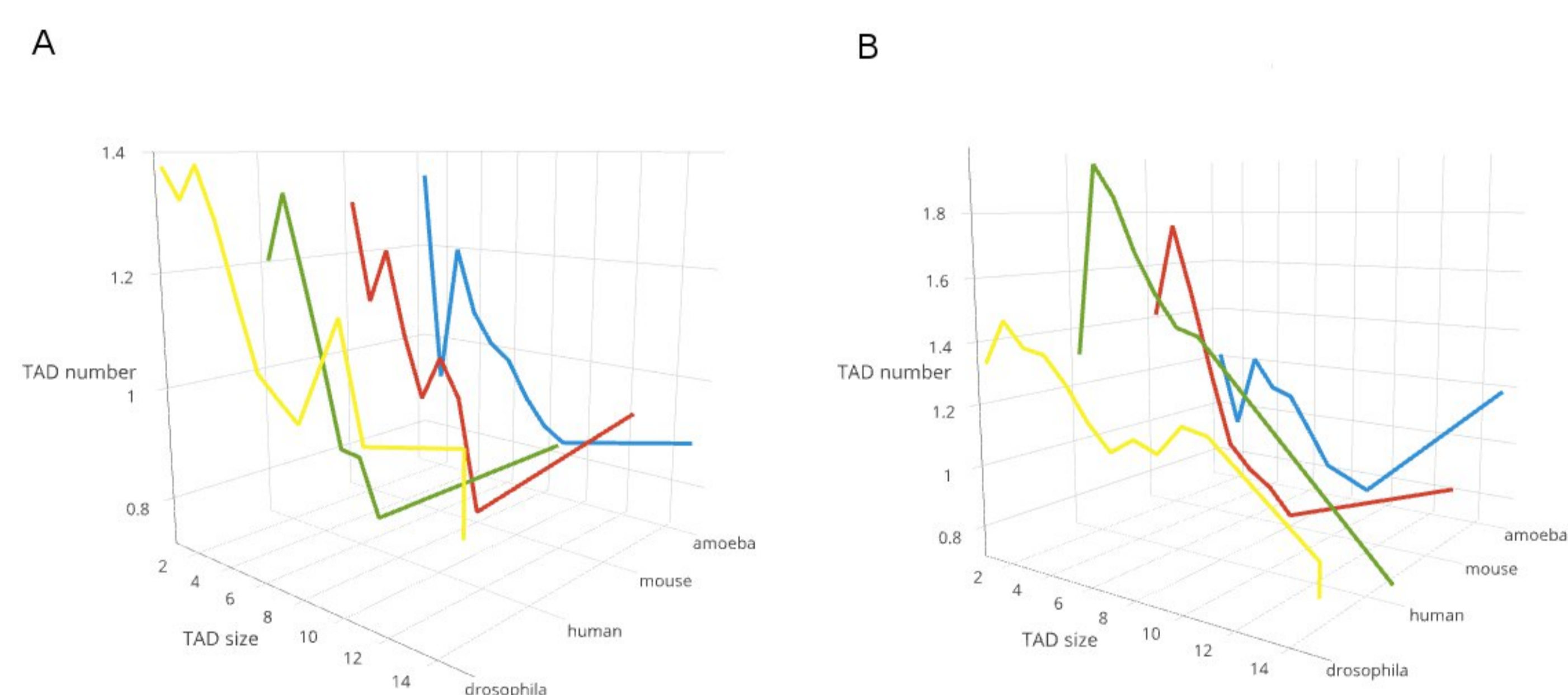


Рис. 3 Зависимость среднего числа контактов с двумя соседними ТАДами (A) и среднего числа контактов со всем геномом (B) от размера ТАДа для четырех организмов.

Pic. 3 Dependence of the average number of contacts with two adjacent TADs (A) and the average number of contacts with the entire genome (B) on the TAD size for four organisms.

### Выводы

1. Монотонность функции числа ТАДов от их размера может быть использована в качестве критерия для выбора  $\gamma$  (рис.1)
2. Границы ТАДов следует выделять по итеративно – корректированным данным
3. При нормализации на среднее значение по диагонали наблюдается быстрый рост числа контактов внутри ТАДа с увеличением его размера; это может объясняться аномальным завышением значений для периферии больших ТАДов (рис. 2A)
4. При нормализации на среднее по случайным подматрицам наблюдается уменьшение количества контактов с ростом размера ТАДов (рис. 2B)
5. Среднее число взаимодействий ТАДов с соседями падает с ростом размера ТАДа; таким образом, самыми мобильными являются небольшие ТАДы, в то время как большие ТАДы более инертны
6. Среднее число контактов ТАДа со всем геномом уменьшается с ростом размера ТАДов, то есть снижение активности является общегеномным свойством
7. Выявленные зависимости консистентны для всех организмов, однако *D. Discoideum* показывает отличные от других организмов зависимости числа контактов внутри ТАДа (среднее по подматрицам) и числа контактов с полным геномом, что может объясняться уникальными особенностями строения его хроматина.