



# FUNCTIONAL ANALYSIS OF MULTI-MAPPED READS IN DNA-RNA INTERACTOME STUDIES

Kosimov M.N. (mihamsik00@gmail.com) Zharikova A.A. Mironov A.A.  
Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University;



## Introduction

Standard bioinformatics pipelines, customized to analyze DNA-RNA interactomes (Figure 1), involve the utilization of only uniquely mapped reads in various experiments. The data loss associated with neglecting multiply mapped reads can account for more than half of all contacts. Consequently, this leads to a diminished understanding of how repetitive elements behave in DNA-RNA contacts.

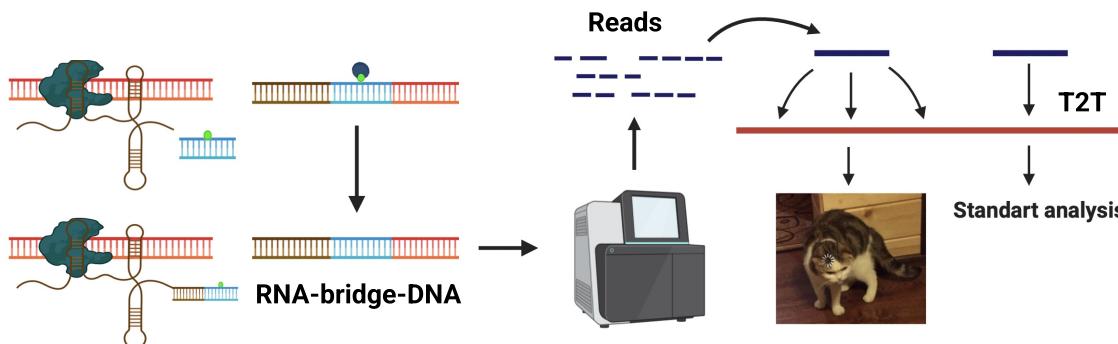


Figure 1. Protocol of DNA-RNA interactome procedure

## Aims

Development of a protocol for analyzing DNA-RNA interactome using multiply mapped reads to reveal the functional significance of genome repetitive elements.

## Methods

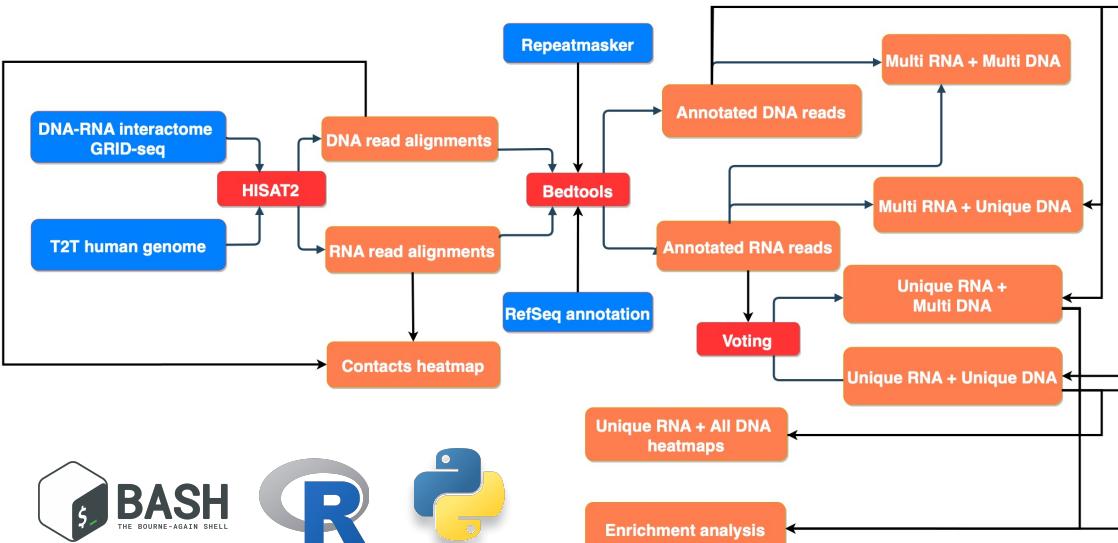


Figure 2. Bioinformatics pipeline

DNA and RNA reads were independently mapped to the human (T2T) genome, with allowance for up to 20 mappings. It has been shown that most of the contacts arise from pairs where either the DNA or RNA part was uniquely mapped, and the number of contacts decreases with the increasing number of times the reads are mapped (Figure 3.A). The distribution of contacts among chromosomes was also observed (Figure 3.B). When dividing genes according to the reads that map to them, it is evident that most genes contain both uniquely and multiply mapped reads (Figure 4.A). Moreover, if only highly contacted genes (more than 500 contacts) are selected for further analysis, genes with only uniquely mapped reads will be excluded from consideration (Figure 4.B).

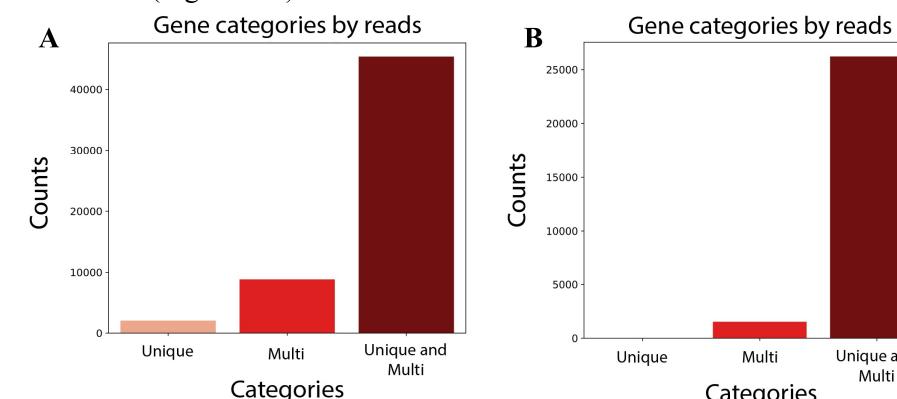


Figure 4. All genes (A) and genes having more than 500 contacts (B) that have only uniquely mapped reads, only multiply mapped reads, or both types

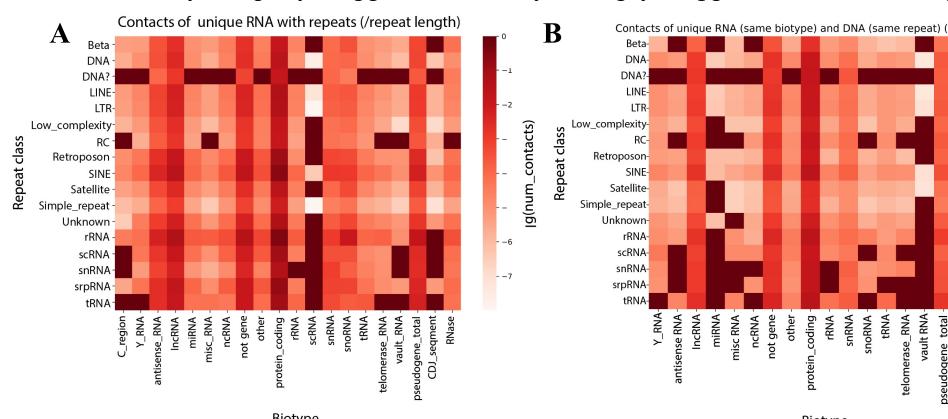


Figure 5. All (A) contacts between RNA reads classified based on biotypes and DNA reads classified based on repeat class, and those where DNA part maps to the same class of repeats

## Results

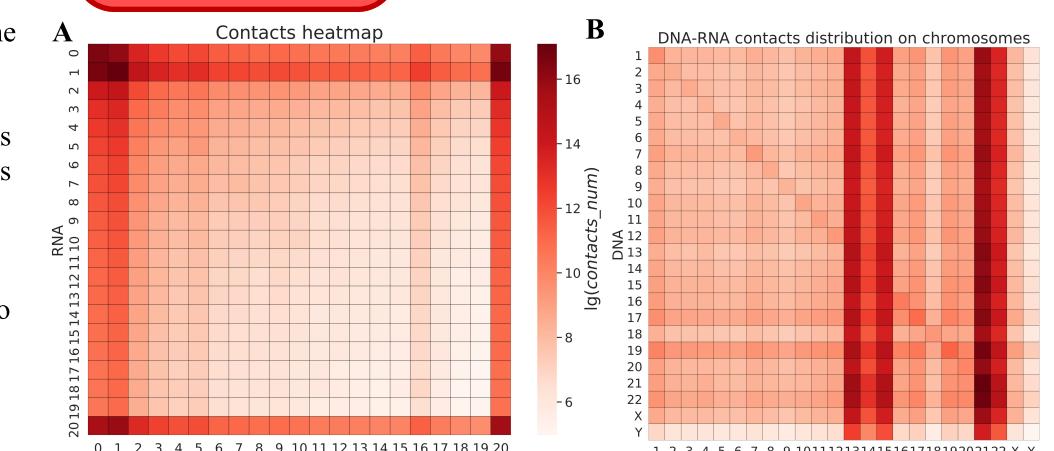


Figure 3. DNA-RNA contacts of reads mapped up to 20 times (A) and mapped to different chromosomes (B)

Subsequently, all contacts between uniquely mapped RNA reads, categorized by biotypes and subjected to a voting procedure, and DNA reads, divided by repeat classes, were analyzed (Figure 5.A). Among these contacts, those in which the multiply mapped DNA segment maps to the same class of repeats were inspected (Figure 5.B). Enrichment analysis was then performed to identify the preferences of different RNA biotypes for contacts with repeat classes (Figure 6).

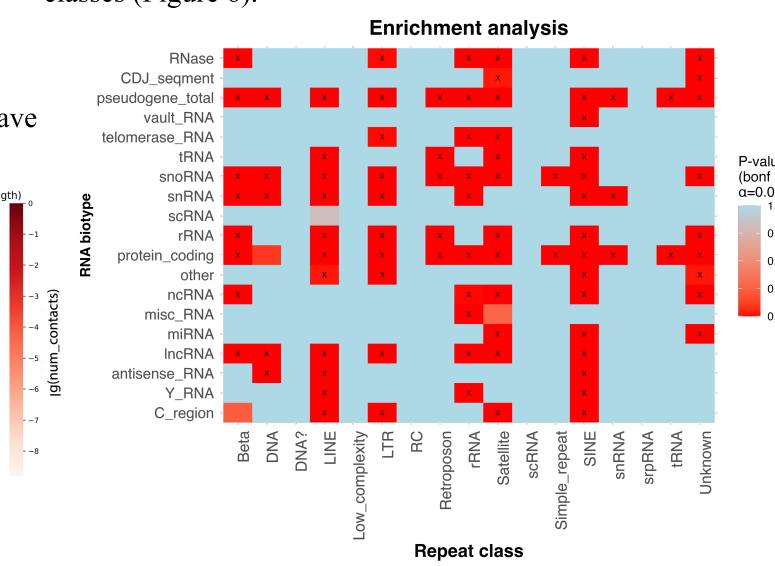


Figure 6. Enrichment analysis of uniquely mapped RNA reads categorized by biotype and DNA reads categorized by repeats classes. Fisher exact test was performed with bonferroni correction.

## Conclusions

Taking multiply mapped reads into account has contributed to a broader study of repetitive human genome elements participation in DNA-RNA interactome. Statistically significant preferences of RNA biotypes in contacts with repeat classes were revealed.