

GO-a-GO: an R package for Gene Ontology enrichment analysis of gene pairs

Aleksander Jankowski
University of Warsaw



useR! 2026 Conference
Warsaw, July 9, 2026

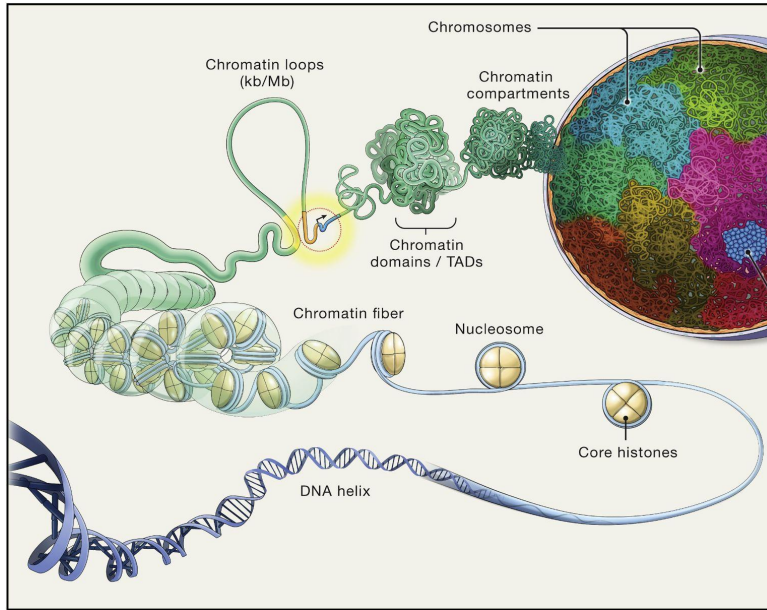


Daryna Yakymenko
Jagiellonian University
of Cracow

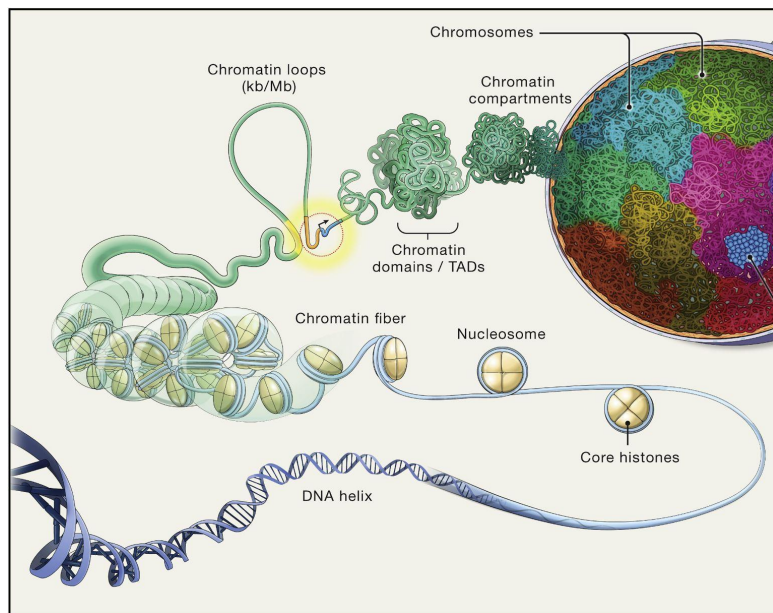


Teresa Szczepińska
Warsaw University
of Technology

Motivation: functional interpretation of chromatin loops

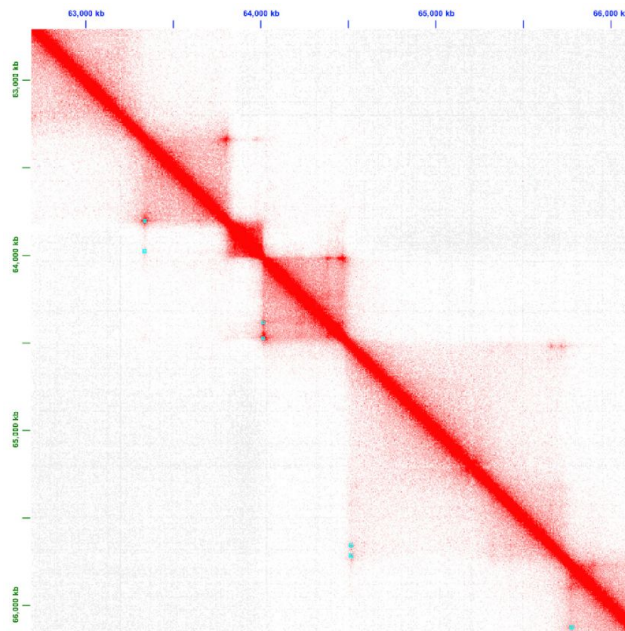


Motivation: functional interpretation of chromatin loops



Misteli, *Cell* 2020

Computational perspective: Hi-C contact maps



Human
lymphoblastoid
cell line GM12878

9,448 loops

1,581 gene-gene
loops

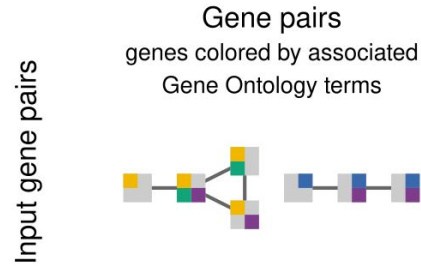
1,743 distinct
gene pairs

Loop dataset from Rao et al., *Cell* 2014

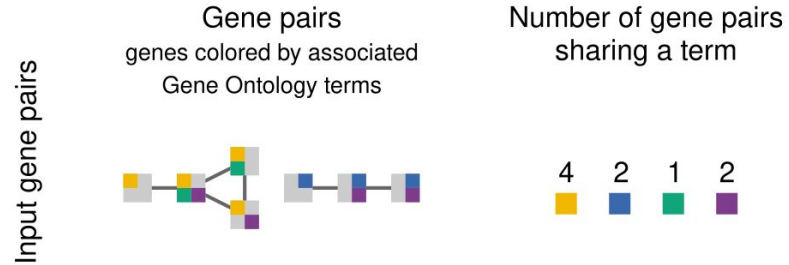
How to functionally annotate genes linked by chromatin loops?

- Identification of overrepresented Gene Ontology terms in a set of *genes* is a standard approach to obtain functional annotations.
- However, to annotate *gene pairs* defined by chromatin loops, it would make sense to use the information that the genes we are annotating are coupled.
- Our approach quantifies overrepresentation of gene pairs that are associated with a shared Gene Ontology term (i.e. both genes have the association).
- The original set of gene pairs is compared against randomized sets, in which the structure of the pairs is preserved, but the gene identities are permuted.

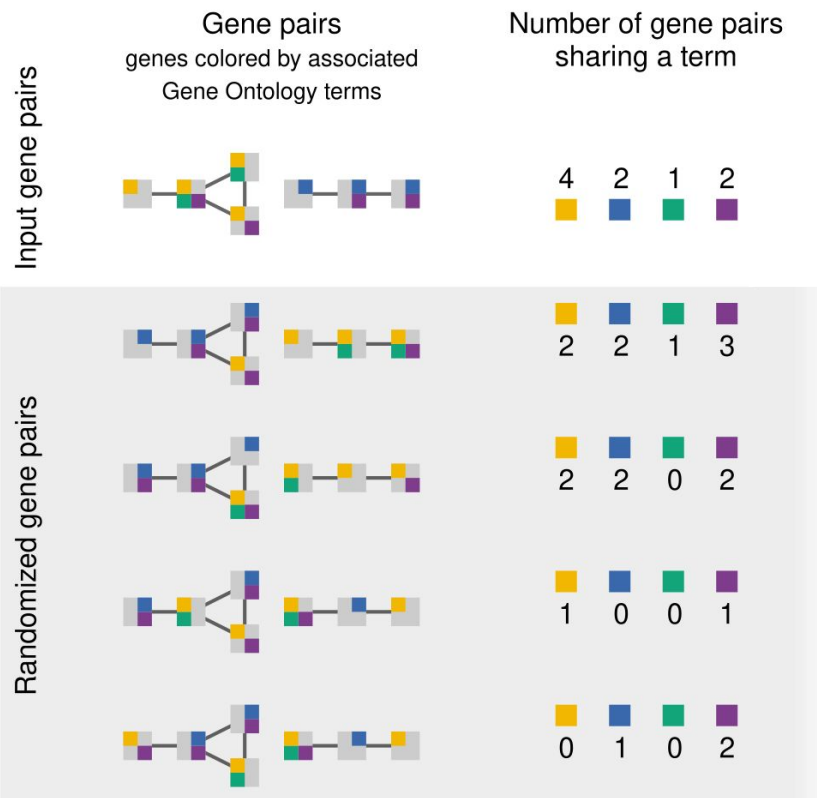
Counting gene pairs and comparing against randomized ones



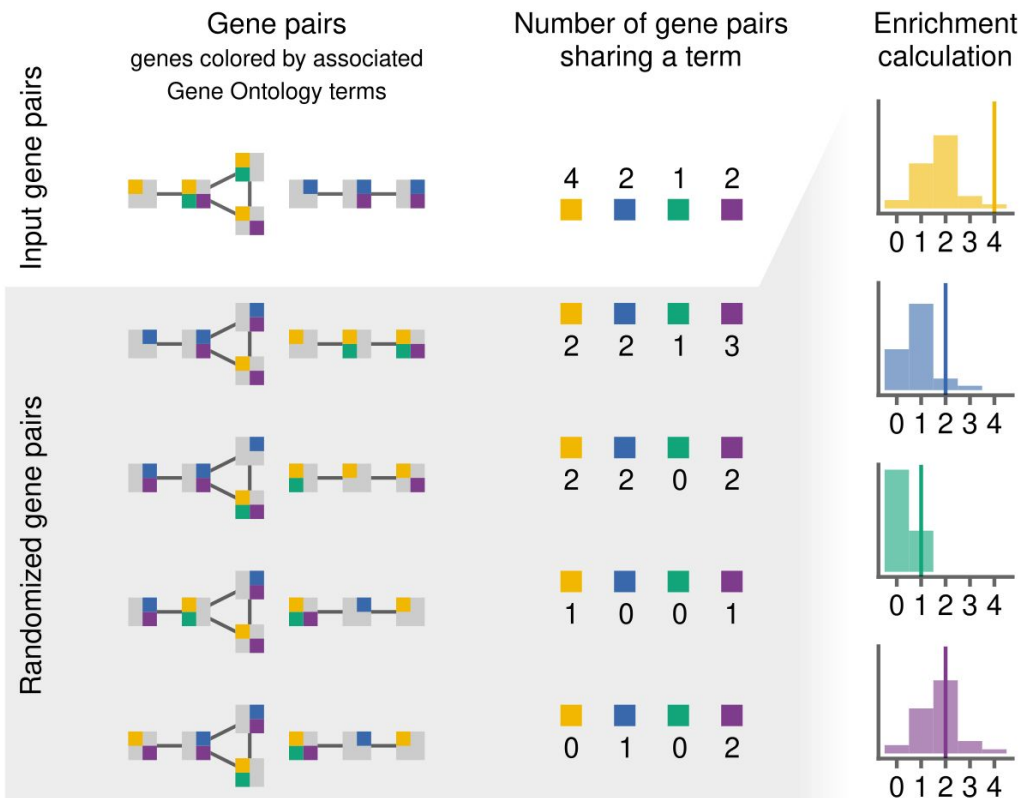
Counting gene pairs and comparing against randomized ones



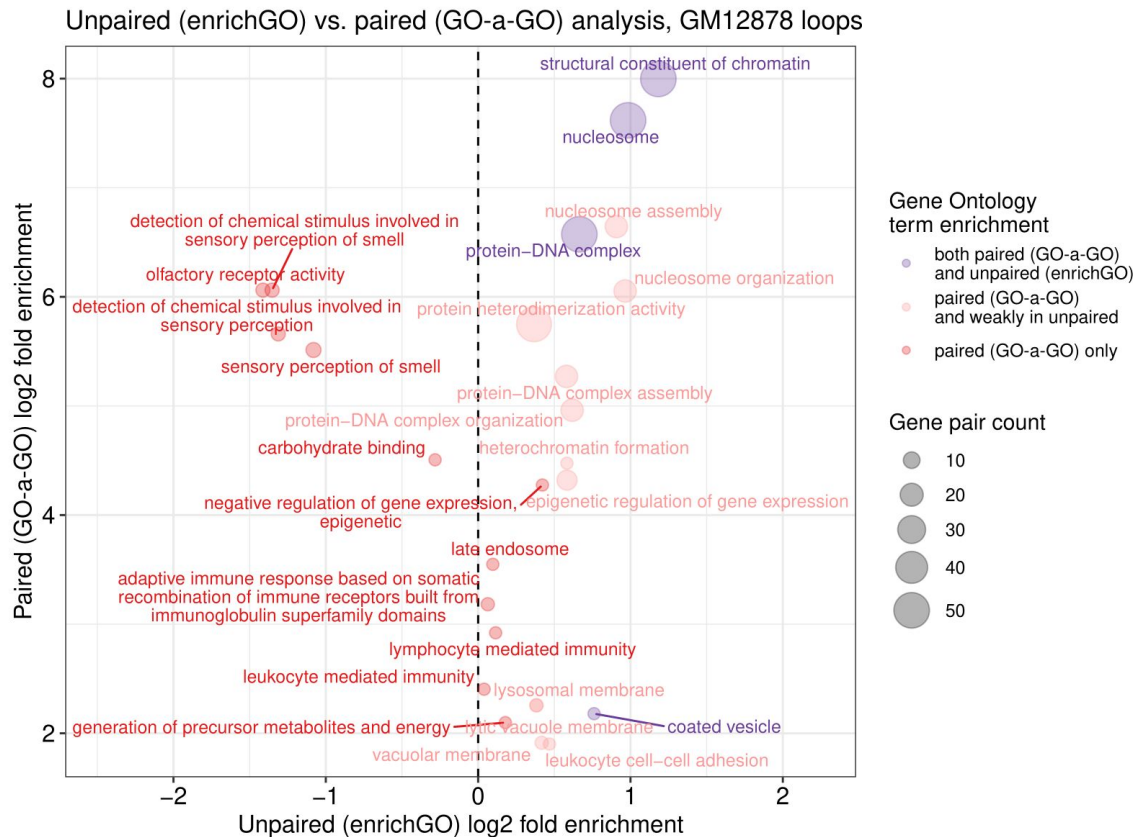
Counting gene pairs and comparing against randomized ones



Counting gene pairs and comparing against randomized ones



Paired analysis reveals more enriched terms than unpaired analysis



GO-a-GO is available as a Bioconductor package



Gene Ontology enrichment analysis of gene pairs with GO-a-GO

Aleksander Jankowski

6 May 2026

Package

GOaGO 1.0.1

Contents

- 1 Introduction
- 2 Installation
- 3 Quick start to using GO-a-GO
 - 3.1 Example dataset of gene pairs
 - 3.2 Enrichment analysis of the example dataset
 - 3.3 Plotting the enriched Gene Ontology terms
 - 3.4 Comparison to unpaired Gene Ontology enrichment
 - 3.5 Extracting gene pairs associated with terms of interest
 - 3.6 Obtaining additional statistics
- 4 Extracting gene pairs from paired genomic regions
- 5 Citing GO-a-GO
- 6 Session information

<https://bioconductor.org/packages/GOaGO>

Vignette contains code examples:

```
library(GOaGO)
```

```
library(org.Hs.eg.db)
```

```
goago <- GOaGO(genePairsGM12878,  
  keyType = "ENTREZID", OrgDb = org.Hs.eg.db, ont = "ALL"  
)  
## Warning in uniqueGenePairs(genePairs): removing 509 duplicated gene pair(s)  
## Warning in uniqueGenePairs(genePairs): removing 87 gene pair(s) containing the  
## same gene twice
```

```
DotPlot(goago)
```

Summary



- An attempt to characterize the chromatin loops has inspired a general method to identify functional terms that are overrepresented in a given set of gene pairs.
- Our method identifies terms that get preferentially paired in a set of gene pairs, addressing a complementary problem to the identification of enriched terms in a set of genes.
- GO-a-GO can be used to study the function of chromatin contacts of various flavours, e.g. chromatin loops or differential chromatin contacts, but can also be applied to any other set of gene pairs.