

# Pattern Discovery in 3D Genome Data

An Independent Study with Dr. Welch  
In Collaboration with The Pombo Lab (1)

## Overview:

Using both computational and biological methods, explore the 3D structures [1] and organization of embryonic stem cells of both mice and people, as well as F123 cells of humans. Genome Architecture Mapping (GAM) [2] and Statistical Inference of Co-segregation (SLICE) [3] will be used and extended to measure and identify various features of the three-dimensional structures. This information will then be analyzed to gain new insights into these types of cells and their genomic structures.

## Objective:

Learn, implement, and invent computational methods for the modelling of the three-dimensional structure of genomes; apply these methods to gain new insights into the genomic organization of embryonic stem cells.

## Learning Goals:

- (1) computational modeling (including graph analytics and feature selection);
- (2) statistical significance testing;
- (3) data cleaning, preprocessing, and exploration;
- (4) visualization; and
- (5) interpretation.

## Schedule Outline:

Phase I			Phase II			Midterm
Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
Usage	Design	Code	Usage	Design	Code	Present (Draft 1)

Phase III			Phase IV		
Week 8	Week 9	Week 10	Week 11	Week 12	Week 13
Usage	Design	Code	Usage	Design	Code

Final		
Week 14	Week 15	Week 16
Present (Draft 2)	Finalizing	Present (Final Draft)

**Usage:** Using bioinformatics pipelines to analyze genomic data, or to define requirements for a pipeline.  
**Design:** Designing new code pieces or tools to enhance a pipeline for further analysis.  
**Code:** Implementing new designs into the pipeline.

## Grading:

**Phases I-IV:** 60% (15% per Phase: Usage 5% Design 5% Code 5%)

**Midterm:** 15% (Present progress to lab)

**Final:** 25% (Present progress to lab)

## Grading Scale:

**A** 94-100 **A-** 90-93

**B+** 87-89 **B** 84-86 **B-** 80-83

**C+** 77-79 **C** 74-76 **C-** 70-73

**D+** 67-69 **D** 64-66 **D-** 60-63

## References:

[1] The 4D nucleome project - J. Dekker et al.

<https://www.mdc-berlin.de/research/publications/4d-nucleome-project>

[2] Complex multi-enhancer contacts captured by Genome Architecture Mapping (GAM) - A. Pombo et al.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5366070/>

[3] On the nature of chromatin 3D organization - M. Nicodemi et al.

<https://www.mdc-berlin.de/research/publications/nature-chromatin-3d-organization>

## GRADING RUBRIC

PROJECT PHASE: \_\_\_\_\_

DATE: \_\_\_\_\_

	<i>points</i>	<i>Comments</i>
<b>COMPLETENESS</b>	(45)	
<b>CORRECTNESS</b>	(45)	
<b>QUALITY</b>	(10)	