BIST Barcelona Institute of Science and Technology

BIST IGNITE PROJECT
FINAL REPORT

GENSTORM^2

Striking back to determine how genes fold and work in space and time

Project coordinators:
Maria Victoria Neguembor (CRG – BIST)
Pablo D. Dans (IRB Barcelona – BIST)

Collaborative project:
Pia Cosma lab (CRG – BIST)
Modesto Orozco lab (IRB Barcelona – BIST)

CRG Centre for Genomic Regulation

IRB BARCELONA INSTITUTE FOR RESEARCH IN BIOMEDICINE

The human genome is a two meters long molecule capable of folding in 3D to fit into the microscopic nuclear volume. Every cell in our body carries the same genetic information but genes can fold in different ways, giving rise to cell-specific patterns of gene expression. Importantly, gene-folding defects can also alter gene function and lead to disease. Understanding how this spectacular folding is achieved is one of the most exciting and challenging research areas in biology. Yet, **the three-dimensional anatomy of genes is still largely unknown**. Recently, chromosome conformation capture methods (3C, 4C, 5C, HiC) allowed the detection of genomic regions in physical proximity and the generation of genome-wide maps of DNA contacts with kilobase resolution. Likewise, complementary techniques such as ChIP-seq and MNase-seq revealed the importance of histones (which wrap the DNA) and architectural proteins for the 3D conformation of the genome. Still**, the three-dimensional conformation at the single gene scale is mostly unexplored due to intrinsic limitations in resolution of the abovementioned techniques, to the lack of super resolution (SR) microscopy studies and to the limited integration of orthogonal technologies**. The recent development of SR microscopy breaks the 200 nm limit of resolution of optical microscopy down to 20 nm; thus opening exciting possibilities in the field of genome architecture. This is indeed the size scale at which molecular events regulating gene expression occur.

The GENSTORM^2 project aimed to solve the 3D conformation of individual genes at near atomic resolution to investigate the relation between gene folding and gene function. We have approached this goal from a **multidisciplinary perspective** by integrating the highest number of technologies available and developing *ad-hoc* methodologies. We combined **STochastic Optical Reconstruction Microscopy (STORM) with the innovative oligoSTORM technique for DNA labeling, HiC and MNase-seq data along with cutting edge mesoscopic and molecular simulations.**

We successfully obtained the **first STORM images of expressed and repressed genes with 20 nm resolution in human cells** (Figure 1). We provided exciting insights regarding nuclear positioning, shape and compaction of genes (Figure 1). Moreover, we developed **immuno-oligoSTORM** (**ioS**), a labeling strategy for simultaneous DNA and protein detection to evaluate the role of DNA-binding proteins and other regulatory proteins right in their site of action: the genes.

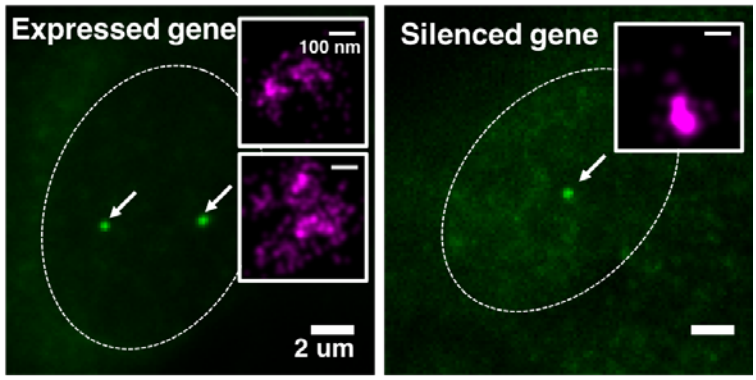

Figure 1. *oligoSTORM images of expressed and repressed genes (indicated with arrows). In green conventional images, in magenta Super-resolution image.*

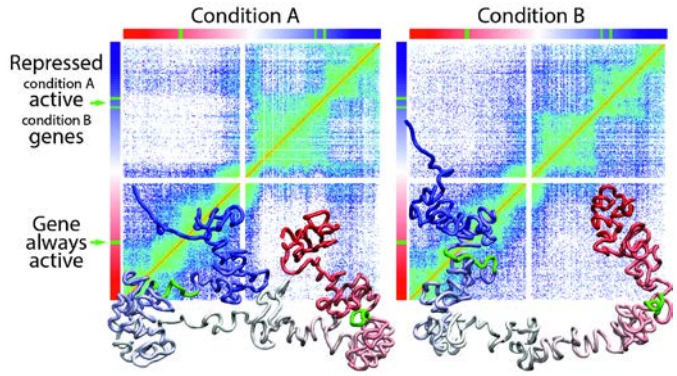

Figure 2. *HiC images of expressed and repressed genes (indicated with arrows). The 3D structures determined are over-imposed to the HiC contact matrices.*

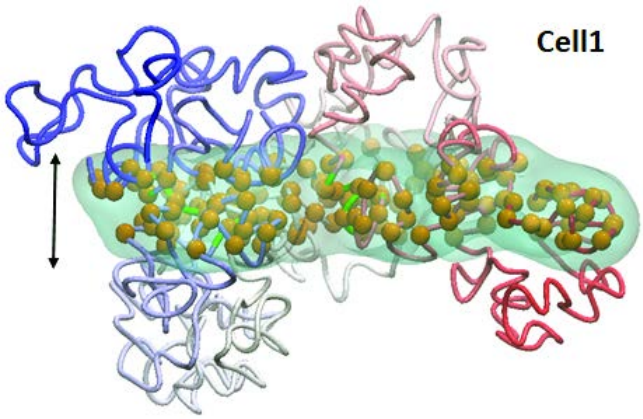

Figure 3. *Fitting between the HiC models and the immuno-oligoSTORM localizations (orange spheres) in the confocal plane (transparent blue region).*

In parallel, we produced capture HiC maps of chromosomes containing the genes of interest in different conditions, and developed a strategy to use the HiC data to obtain the 3D structure, conformation and dynamics of chromosome segments. The **3D models were obtained by transforming the HiC contacts into spatial distances in a coarse grain model** (Figure 2), which in turn is coupled to an algorithm to perform simulations. We then integrated this knowledge with the structural information provided by immuno-oligoSTORM imaging. The results are the **first representations of chromosome paths connecting genes through optical localizations combining microscopy, HiC and modelling** (Figure 3).

We then moved to higher resolution by looking at the chromatin level, which strongly depends on the positions of its nucleosomes. We conducted in-house MNase-seq experiments, **obtaining the nucleosome positioning along the sequence for the genes of interest.** This information was then integrated into a **second model of DNA with base-pair resolution**, which was coupled to a Monte Carlo algorithm **allowing us to sample chromatin conformations in 3D**. Finally, we choose from the pool of sampled structures those that better fitted with oligoSTORM localizations generating the **first 3D representations of genes at near atomic resolution that fulfil all cutting-edge technologies available to study gene folding** and chromosome conformations (Figure 4).

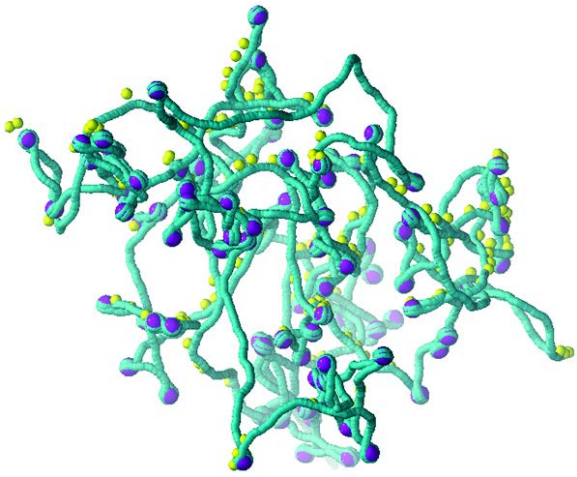

Figure 4. *3D chromatin structure (nucleosomes in violet) of a repressed gene obtained from the coarse grain model based on MNase-seq and the fitting to oligoSTORM localizations (yellow spheres).*

In summary, we have established the basis for **the most comprehensive method to study gene conformation**, integrating several sources of experimental knowledge (immuno-oligoSTORM, HiC, MNase-seq) with advances modelling technologies.