The Three-Dimensional Architecture of a Bacterial Genome

Umbarger MA[1], Toro E[2], Wright MA[1], Porreca GJ[1], **Baù D**[3], Hong S-H2,[4], Fero HJ[2], Marti-Renom MA[3], McAdams HH[2], Shapiro L[2], Dekker J[5], and Church G[1]

1 Department of Genetics, Harvard Medical School, Boston, MA 02115 USA.
2 Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305 USA.
3 Structural Genomics Laboratory, Centro de Investigación Príncipe Felipe, 46012 Valencia, Spain.
4 Department of Physics, Stanford University School of Humanities and Sciences, Stanford, CA 94305 USA.
5 Department of Biochemistry and Molecular Pharmacology, UMASS Medical School, Worcester MA 01605 USA.

PRINCIPE FELIPE
CENTRO DE INVESTIGACION

UMASS

Summary

We determined the three-dimensional (3D) architecture of the *Caulobacter crescentus* genome by combining genome-wide chromatin interaction detection, live-cell imaging, and computational modeling. Using the Chromosome Conformation Capture Carbon Copy (5C, http://my5c.umassmed.edu)[1] technology (**Fig. 1**) in combination with the Integrative Modeling Platform (IMP, http://www.integrativemodeling.org)[2], we derived ~13 Kb resolution 3D models of the *Caulobacter* genome. These models illustrated that the genome is ellipsoidal with periodically arranged arms (**Fig. 2**). The *parS* sites, a pair of short contiguous sequence elements involved in chromosome segregation, were positioned at one pole of this structure, where they nucleated a compact chromatin conformation (**Figs. 3** and **4**). Repositioning these elements resulted in rotations of the chromosome which changed the sub-cellular positions of nearly all genes. However, such chromosomal rotations did not lead to large-scale changes in gene expression, indicating that genome folding does not strongly affect gene regulation. Our approach provides an experimental paradigm for deriving insight into the cis-determinants of 3D genome architecture. This work is in press at *Molecular Cell*.

Methods

Restriction fragments derived from our 5C experiment (**Fig. 1**) were modeled as points connected by springs. The distance derived from the contact frequency between pair of fragments (**Fig. 2**) was used to define the equilibrium distance between the points representing the connecting fragments. The 3D coordinates of all points were randomly initialized and optimization was performed to derive a structure that minimally violates these equilibrium distances. The initialization and optimization procedure was repeated thousands of times to generate an ensemble of structures, which were then superimposed and grouped based on their structural similarity. This yielded clusters of models in which the 3D coordinates of the restriction fragments were very similar. The clusters were then represented as 3D density maps (**Fig. 3**).

Results

Our contact-based 3D models illustrate that the *parS* sites define the global structure of the *Caulobacter* genome (**Fig. 3**). Our data indicate that these sites reside at the pole of the wild-type swarmer chromosome/cell and that moving a 10 Kb region containing them elsewhere in the genome yields large-scale rotations of the chromosome that reposition these elements at the cell/structural pole. We therefore propose that our findings are a consequence of the *parS* sites being the first genomic elements to segregate to the opposite pole where they become anchored and thereby establish the global orientation of the genome. Our findings also indicate that the global orientation of the chromosome may be defined by the order of segregation and that the rotated global arrangements of the genome observed in inversion strains are the result of a perturbation of the order of segregation of loci that results from the movement of the *parS* sites.

Conclusions

The work presented here illustrates how a comprehensive study of genome 3D architectures can provide insight into the roles of sequence elements and fundamental DNA-based processes in defining this structure. The experimental paradigm we introduce is general and could be used in conjunction with genetic perturbations to elucidate the roles of nucleoid associated proteins, cis-elements, and DNA-templated processes such as transcription and replication in shaping the folding of the genome. With additional advances, including decreases in DNA sequencing costs, such a paradigm could also be applied to larger eukaryotic genomes to further elucidate the complex relationships between genome sequence, structure, and function.

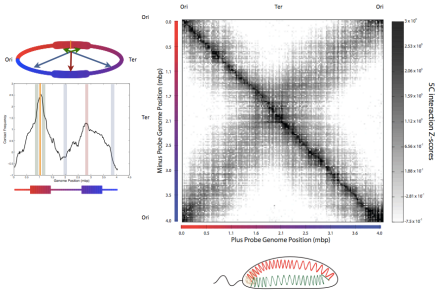

Figure 1. 5C interaction matrix
ELLIPSOID *for* Caulobacter cresentus

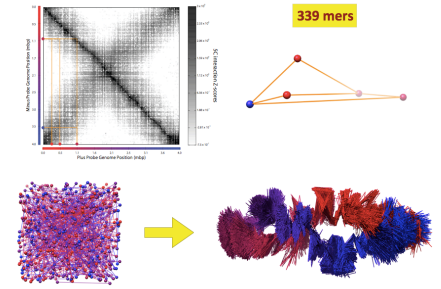

Figure 2. IMP Model Building
HELICAL *for* Caulobacter cresentus

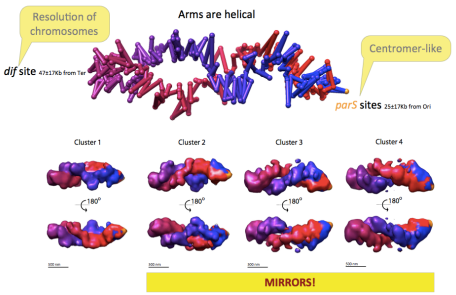

Figure 3. Genomic architecture
dif and parS sites correctly placed

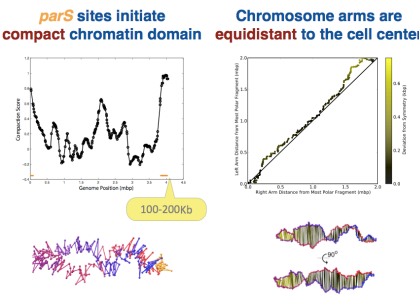

Figure 4. Model analysis
Arms compaction and equidistance to long axis

References

1. J. Dostie et al., Genome Res 16, 1299 (2006).
2. Baù and Marti-Renom. Chromosome Research 19. 25 (2011)