BIOLOGY REPORTS

Published: 15 March 2010

Mapping chromatin conformation

Peter J Shaw

Address: Department of Cell and Developmental Biology, John Innes Centre, Colney Lane, Colney, Norwich, NR4 7UH, UK

Email: peter.shaw@bbsrc.ac.uk

F1000 Biology Reports 2010, **2**:18 (doi:10.3410/B2-18)

Abstract

Chromatin conformation capture (3C) and related techniques have become well-established methods to examine which distal DNA sequences are spatially located close to a locus of interest. Hi-C is a new addition to the growing set of 3C-based techniques that has extended the approach to a genome-wide analysis of nuclear chromatin using high-throughput DNA sequencing to generate global interaction maps for the entire genome.

Introduction and context

The chromatin in eukaryotic nuclei is clearly visible by microscopy during mitosis and meiosis as well-defined separate chromosomes, but chromosome and chromatin organization during interphase has been much harder to determine. It was predicted by early cytologists such as Boveri and Rabl that interphase chromosomes would have a definite structural organization, and clear evidence was obtained for this by UV irradiation studies by Cremer and colleagues in the early 1980s [1], but it required the advent of *in situ* hybridization for the detection of specific DNA sequences to really begin to probe the organization of interphase chromosomes. This has given us a picture in which each chromosome occupies a separate interphase chromosome territory [2]. These territories were originally proposed to be non-overlapping but the extent to which the territories really intermingle has proven to be a somewhat controversial question [3,4]. *In situ* hybridization is a time-consuming and demanding technique, especially when reasonable three-dimensional (3D) structural preservation is required, and becomes more and more difficult as single-gene resolution is approached. This has limited its application.

Within chromosomes, there is a hierarchy of structural organization. At the lowest level, the DNA is packaged into nucleosomes by the core histones, resulting in fibres about 10 nm in diameter. The nucleosomes interact to form higher-order structures, regulated by histone H1 and other chromatin proteins, and by various post-translational histone modifications [5]. Apart from the 10-nm fibres whose existence *in vivo* is well established, virtually all other higher-order chromatin structures are controversial to some degree [6]. Furthermore, the 3D organization of the chromosomes or chromosome territories, the degree of condensation of different regions, the histone modifications and histone variants present, and even the relative positioning of different chromosome territories are highly dynamic, changing as a function of development and transcriptional regulation.

Chromatin immunoprecipitation (ChIP) has been very successful at determining the associations of different histone variants and modifications, as well as other proteins, with specific genome sequences (see [7] for a recent review). The basic ChIP technique is now routinely coupled with hybridization of the resulting DNA to whole-genome arrays (ChIP-chip) or more recently with high-throughput DNA sequencing to obtain a whole-genome view of where in the genome the proteins are bound or the specific modifications are found [8]. This, however, produces only a 1D map of what we know from structural studies is a 3D problem (or 4D if a temporal axis is included).

Chromatin conformation capture (3C) and related methods were developed as a way of investigating the

long-range interactions of a DNA sequence in the nucleus [9,10]. In 3C, chromatin is cross-linked by formaldehyde and digested by restriction enzymes to leave sequences held together by the cross-links. These are then ligated under dilute conditions that favour the ligation of only DNA fragments held together by the cross-linking. The resulting ligated DNA fragments contain sequences that were thus in close physical proximity at the time of the cross-linking. In conventional 3C techniques, the pool of ligated sequences is analysed by polymerase chain reaction using primer pairs from the target sequence and potential interacting sequences.

Major recent advances

In a recent paper by Lieberman-Aiden *et al*. [11], the 3C technique is taken one further stage to provide an unbiased map of genome-wide interactions in a technique they have named Hi-C. In this method, the cross-linked chromatin is digested as in 3C, using an enzyme that leaves a 5′ overhang, which is then filled, including a biotinylated nucleotide, and the blunt-ended fragments are ligated in dilute conditions as in 3C. The resulting ligated DNA is sheared and biotin-containing sequences are selected with streptavidin beads to yield a library of fragments containing sequences from pairs of interacting loci. These fragments are then subjected to massively parallel DNA sequencing to give a genome-wide catalogue of interactions. In the current work, the interactions were grouped together in regions of either 1 Mb or 100 kb and are thus comparatively low in resolution.

The resulting dataset clearly contains an enormous amount of information that in general is related to the 3D structure of the genome. However, there are problems in understanding such a dataset in terms of structure. First, the data are statistical in nature – the sequences are the result of interactions taking place in many cells, and in all probability no one cell will display more than a fraction of the total. Second, it is not clear how we go from a list of interactions to a 3D map of some sort that can be compared with other nuclear structural data. Despite these caveats, the Hi-C approach adds a powerful new tool for probing the intranuclear organization of chromosomes, as Lieberman-Aiden *et al*. [11] have demonstrated.

They first tested whether the data are consistent with the existence of chromosome territories by calculating the average intrachromosomal contact probability as a function of genomic distance on each chromosome. This showed that the probability of detecting an interaction between sequences on the same chromosome was always much larger than between sequences on different chromosomes, consistent with the segregation of chromosomes into territories. Furthermore, the interaction probability decreased the further apart the sequences were along the chromosome, suggesting that the 3D distance between loci increases with increasing genomic separation along the chromosome. In another test of their data, the authors showed that interaction frequencies measured by Hi-C agreed reasonably well with 3D distance measured by *in situ* labelling.

Lieberman-Aiden *et al*. [11] then concentrated on the substructure of individual chromosomes by calculating an interaction matrix of each chromosome with itself. This gave an overall 'plaid' patterning, which the authors interpreted as suggesting that the chromosomal sequence was divided into two sets of interspersed blocks (denoted A and B). A blocks interact with A blocks and B with B, whereas the interactions between A and B are less strong. The boundaries of these blocks in the plaid pattern corresponded strikingly with the boundaries between gene-rich regions, which showed higher levels of specific histone methylations and greater DNAse1 sensitivity, and gene-poor regions. Thus, the most obvious feature of the data was to divide the chromosome into regions probably broadly comparable to active chromatin and inactive chromatin. The interaction frequencies at a given genomic separation were greater in the gene-poor regions of chromatin (assumed to be more condensed), presumably because the DNA strands are physically closer together.

Finally, the authors used a modelling approach to examine the possible structure of the chromosome by calculating the intrachromosomal contact probability as a function of genomic distance (denoted as s). Between 500 kb and 7 Mb (roughly the size found for the open and closed chromatin domains), the probability followed an approximate curve of s^{-1}. The expected contact probability curves were proposed by modelling to be consistent with what the authors call a 'fractal' globule, in which a series of smaller globules like beads on a string are crumpled together in a hierarchical series of structures, but inconsistent with alternative arrangements [12].

Future directions

The Hi-C method opens up a number of possibilities. For a start, with successively larger sequencing datasets, progressively finer interactions should be detectable. For example, the existence of regular higher-order structures in interphase nuclei, such as the 30-nm fibre, has long been proposed but is still controversial *in vivo* [6]. It is debatable whether such regularities will be detectable in this type of data; at short range, random

collisions may generate too much non-specific background to determine any regularities [13], but ways to detect such interactions may be found if they are regular and specific enough. Similar considerations apply to the detection of regular packing in mitotic chromosomes, but differences through interphase may show up, as should systematic differences between cell types. Other recent methods have combined ChIP with unbiased high-throughput 3C methods. Fullwood *et al*. [14] used such a method, which they called ChIA-PET (chromosome interaction analysis by paired-end sequence tagging), to analyse chromatin sequences bound to human oestrogen receptor-alpha. Hopefully, future developments in these methods will reduce the background noise level and increase the sensitivity and specificity (see [15] for a detailed discussion of this).

Using various assumptions, several groups are now attempting to model chromosome territorial organization and chromosome dynamics [12,16-19]. As yet, there are relatively few data to test these models against, and the data that do exist are relatively low in resolution, such as overall positioning of chromosomes and shapes of chromosome territories. 3C, Hi-C and related techniques offer many more extensive and detailed data to compare with predictions from modelling studies.

Two of the most attractive aspects of the Hi-C and similar techniques are that they are, in principle, applicable to any species with a sequenced genome (e.g., see Louwers *et al*., 2009 [20] for the application of 3C methods to plants) and that they use sequencing methods that are rapidly becoming more routine and cheaper. *In situ* techniques, on the other hand, have to be tailored specifically to each species or even different cell types within a species and have remained difficult and time-consuming. Thus, we may eventually hope to have studies of many species and cell types, which will allow much more firmly based generalizations about at least some aspects of 3D structure to be made.

Abbreviations

3C, chromatin conformation capture; 3D, three-dimensional; ChIP, chromatin immunoprecipitation.

Competing interests

The author declares that he has no competing interests.

Acknowledgments

Work in the author's laboratory is supported by the Biotechnology and Biological Sciences Research Council of the UK.

References

1. Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C: **Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments.** *Hum Genet* 1982, **60:**46-56.
2. Meaburn KJ, Misteli T: **Cell biology: chromosome territories.** *Nature* 2007, **445:**379-781.
3. Branco MR, Pombo A: **Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations.** *PLoS Biol* 2006, **4:** e138.
4. Branco MR, Pombo A: **Chromosome organization: new facts, new models.** *Trends Cell Biol* 2007, **17:**127-34.
5. Campos EI, Reinberg D: **Histones: annotating chromatin.** *Annu Rev Genet* 2009, **43:**559-99.
6. Eltsov M, Maclellan KM, Maeshima K, Frangakis AS, Dubochet J: **Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes *in situ*.** *Proc Natl Acad Sci U S A* 2008, **105:**19732-7.

   f1000 Factor 6.0 *Must Read*
   Evaluated by William Earnshaw 19 Jan 2009
7. Wong E, Wei CL: **ChIP'ing the mammalian genome: technical advances and insights into functional elements.** *Genome Med* 2009, **1:**89.
8. Gilchrist DA, Fargo DC, Adelman K: **Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation.** *Methods* 2009, **48:**398-408.
9. Dekker J, Rippe K, Dekker M, Kleckner N: **Capturing chromosome conformation.** *Science* 2002, **295:**1306-11.

   f1000 Factor 6.5 *Must Read*
   Evaluated by Regine Hengge 28 Feb 2002, Andrew Belmont 14 Aug 2002, Tamar Schlick 28 Aug 2002
10. Simonis M, Kooren J, de Laat W: **An evaluation of 3C-based methods to capture DNA interactions.** *Nat Methods* 2007, **4:**895-901.
11. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome.** *Science* 2009, **326:**289-93.

    f1000 Factor 10.2 *Exceptional*
    Evaluated by Tom Tullius 28 Oct 2009, Tamar Schlick 29 Oct 2009, Lloyd Smith 08 Dec 2009, Dean Jackson 11 Dec 2009, Thomas Ried 14 Dec 2009, Kenneth Zaret 19 Jan 2010
12. Mateos-Langerak J, Bohn M, de Leeuw W, Giromus O, Manders EM, Verschure PJ, Indemans MH, Gierman HJ, Heermann DW, van Driel R, Goetze S: **Spatially confined folding of chromatin in the interphase nucleus.** *Proc Natl Acad Sci U S A* 2009, **106:**3812-7.
13. Dekker J: **The three 'C' s of chromosome conformation capture: controls, controls, controls.** *Nat Methods* 2006, **3:**17-21.
14. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, *et al.*: **An oestrogen-receptor-alpha-bound human chromatin interactome.** *Nature* 2009, **462:**58-64.

    f1000 Factor 8.0 *Exceptional*
    Evaluated by Yi Eve Sun 08 Dec 2009, Paul Webb 18 Jan 2010

15. Fullwood MJ, Ruan Y: **ChIP-based methods for the identification of long-range chromatin interactions.** *J Cell Biochem* 2009, **107:**30-9.

16. Cook PR, Marenduzzo D: **Entropic organization of interphase chromosomes.** *J Cell Biol* 2009, **186:**825-34.

17. de Nooijer S, Wellink J, Mulder B, Bisseling T: **Non-specific interactions are sufficient to explain the position of heterochromatic chromocenters and nucleoli in interphase nuclei.** *Nucleic Acids Res* 2009, **37:**3558-68.

18. Langowski J, Heermann DW: **Computational modeling of the chromatin fiber.** *Semin Cell Dev Biol* 2007, **18:**659-67.

19. Rosa A, Everaers R: **Structure and dynamics of interphase chromosomes.** *PLoS Comput Biol* 2008, **4:**e1000153.

f1000 Factor 3.0 *Recommended*
Evaluated by Peter Shaw 10 Nov 2008

20. Louwers M, Splinter E, van Driel R, de Laat W, Stam M: **Studying physical chromatin interactions in plants using Chromosome Conformation Capture (3C).** *Nat Protoc* 2009, **4:**1216-29.

f1000 Factor 3.0 *Recommended*
Evaluated by Xing Wang Deng 04 Sep 2009