Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Higher order assembly: folding the chromosome

Sven A Sewitz[1], Zahra Fahmi[1] and Karen Lipkow[1,2]

The linear molecules of DNA that constitute a eukaryotic genome have to be carefully organised within the nucleus to be able to correctly direct gene expression. Microscopy and chromosome capture methods have revealed a hierarchical organisation into territories, domains and subdomains that ensure the accessibility of expressed genes and eventually chromatin loops that serve to bring gene enhancers into proximity of their target promoters. A rapidly growing number of genome-wide datasets and their analyses have given detailed information into the conformation of the entire genome, allowing evolutionary insights, observations of genome rearrangements during development and the identification of new gene-to-disease associations. The field is now progressing into using computational models of genome dynamics to investigate the mechanisms that shape genome structure, placing increasing importance on the role of chromatin associated proteins for this process.

Addresses
[1] Nuclear Dynamics Programme, Babraham Institute, Cambridge CB22 3AT, UK
[2] Cambridge Systems Biology Centre, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

Corresponding author: Lipkow, Karen (karen.lipkow@babraham.ac.uk, KL280@cam.ac.uk)

Current Opinion in Structural Biology 2017, **42**:162–168

This review comes from a themed issue on **Folding and binding**

Edited by **Jane Clarke** and **Rohit V Pappu**

For a complete overview see the Issue and the Editorial

Available online 8th March 2017

http://dx.doi.org/10.1016/j.sbi.2017.02.004

Introduction

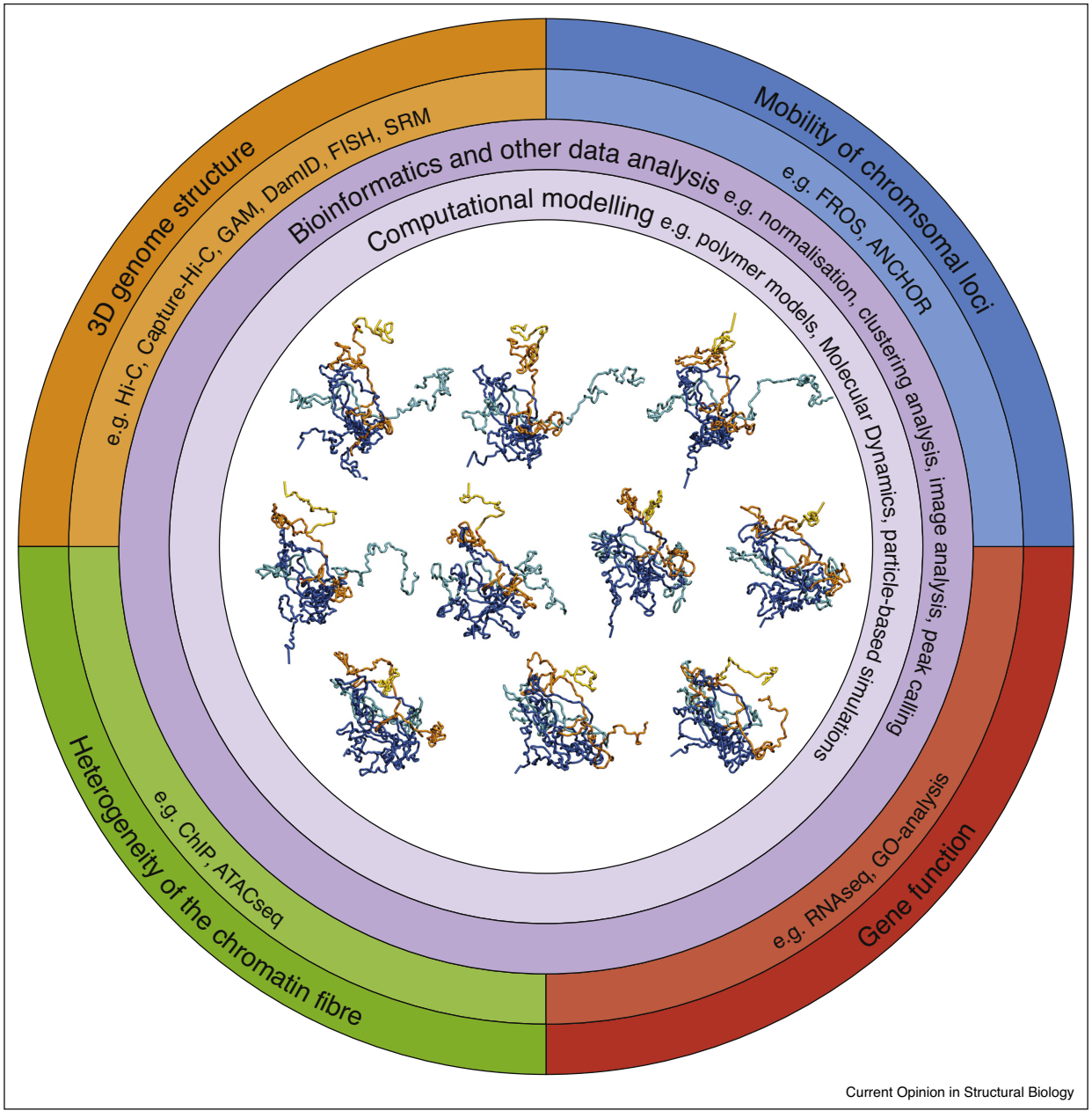
The linear DNA molecule of eukaryotic genomes has to be compacted into the nucleus by a factor of over several hundred thousand (>300 000 fold for humans), even during interphase. This is an astonishing level of compaction. At the same time, the genome remains highly organised on several scales. It is now clear that this organisation is both non-random and vital. Even subtle mistakes in arrangement can lead to severe developmental defects [1•] or disease [2,3•]. This organisation, which consists of several hierarchies, has been studied intensively by microscopy, and increasingly by conformation capture methods and computational methods (Figure 1) (see Ref. [4] for a technical review of the experimental approaches and data analysis methods). The dynamics of the genome are also being investigated at ever increasing resolution [5••]. While significant amounts of data have been collected describing the details of genome organisation, the mechanisms that shape it are only starting to emerge.

Domain structure of genomes

In the cells of higher eukaryotes, the largest structural units are chromosome territories [6]: Every interphase chromosome has its preferred location, and preferred neighbouring chromosomes, within the nucleus. Localised experimental proof came first from Fluorescence *In-situ* Hybridisation (FISH) work on human lymphocyte nuclei, which showed consistent locations of the short arms of chromosome 18 and 19 [7]. Each chromosome territory is separated by an interchromatin compartment (IC), a small uninterrupted space between chromosomes. While territories can be visualised, their boundaries are not rigid but changeable [8]. Contacts between chromosomes, or *trans* interactions, occur frequently and in certain cases coincide very well with frequencies of chromosomal translocations, linking physical contacts to observations of cytogenetic abnormalities [9]. The existence of chromosome territories on a genome-wide scale was shown comprehensively in 2009 using Hi-C [10]. This study furthermore demonstrated the spatial segregation of open, active regions from closed, inactive genomic regions: Alternating blocks of ~5 Mb come together in preferentially interacting compartments: the active A compartment and the inactive B compartment. The peripheries of territories are mostly in the A compartment [11], remain structured, and are defined by higher numbers of *trans*-chromosomal contacts [12]. At a smaller scale, topologically associated domains (TADs) are stretches of ~500 kb, which have very strong internal interactions and are insulated from other TADs [13].

The overall structure of the mammalian interphase genome is believed to be a fractal globule [14]: A chain of small globules (TADs), which organises into larger globules (regions of A/B compartments) and into territories. First derived from theoretical polymer physics [15], evidence for the fractal organisation now comes from analysis of Hi-C data [10], modelling [16,17]), and microscopy [6]. It ensures that the genome can be replicated without entanglement [17], and that each region can easily unfold and become accessible [16]. The differences in accessibility have been observed

Figure 1



The dynamics of the genome and methods of analysis.
Four main aspects of studying genome organisation are shown in the outer circle, together with select experimental techniques. To arrive at an understanding about genome conformation and dynamics, bioinformatics and other data analysis methods and increasingly computational modelling are necessary (two innermost circles). These two areas are in contact with all experimental methods. The centre shows a time series of the movement of two simulated budding yeast chromosomes within a whole-genome simulation (as in Ref. [38]). Each chromosome is tethered via the centromere (red) to the spindle pole body (not shown). The chromosome arms are colored separately, light blue (Chr IV L), dark blue (Chr IV R), yellow (Chr VIII L), and orange (Chr VIII R). The images are produced from 10 consecutive time points, with 100 simulation time steps in between any two figures and demonstrate the dynamics of chromosome conformations over a short period of time.

using superresolution microscopy (SRM), and this has also shown that parts of the genome which corresponding to different epigenetic domains are characterised by their own scaling exponent describing the packaging density of chromatin [5••].

Over the course of evolutionary time, the overall domain structure in mammals is highly conserved, while the structure within domains is more dynamic and can be modulated as a result of local sequence evolution [18•]. Also, gene paralogs are often found within the same TAD