Research Briefings

Martin Simmen

Mapping and computational modelling of nucleosome positions

EUKARYOTIC DNA IS PACKAGED INTO NUCLEOSOMES, EACH COMPRISING 147 BP OF DNA WRAPPED AROUND AN OCTAMER OF HISTONE PROTEINS PLUS A LENGTH OF DNA CONNECTING TO THE NEXT NUCLEOSOME. THE POSITIONS OF NUCLEOSOMES WITH RESPECT TO THE DNA SEQUENCE INFLUENCE GENE EXPRESSION. OUR WORK SEEKS TO MAP NUCLEOSOME POSITIONS, UNDERSTAND WHAT DICTATES POSITIONING, AND GENERATE PREDICTIVE COMPUTATIONAL MODELS OF POSITIONING.

Martin Simmen
Lecturer

School of Biomedical Sciences
Tel: 0131 651 1773

Email:M.simmen@ed.ac.uk

MAPPING THE NUCLEOSOMAL LANDSCAPE

My research, conducted in collaboration with Dr James Allan in the School of Biology, is focused on the generation and analysis of very high resolution maps of nucleosome positions over gene-sized segments of animal genomes. The data comes from in vitro experiments in which copies of the particular DNA fragment are combined with core histones; the resulting reconstituted nucleosomes are then isolated and the ≈147 bp nucleosomal DNAs characterised using next-generation sequencing. We then use computational methods to produce maps showing where on the DNA nucleosomes tend to form (Fig.1).

These maps tell us two key things. First, they demonstrate wide variation in the degree to which individual DNA segments bind the histone octamer. Second, especially near promoters the nucleosomal occupancy displays an organized periodic pattern. We are interested in this aspect, as it has previously been found in low resolution analyses of in vivo nucleosomal positions in yeast and other eukaryotes. The observation that periodic occupancy is also found near promoters in vitro suggests the in vivo pattern is – at least partly – attributable to inherent sequence-dependent differences in histone octamer binding affinity, rather than being solely due to the action of chromatin remodeling complexes or the blocking of nucleosome formation due to prior binding of transcription factors.

SEQUENCE SIGNATURES

Using these maps to identify stretches of DNA with a strong tendency to form nucleosomes, we can then ask: what are the sequence patterns that influence nucleosome positions in mammals? We

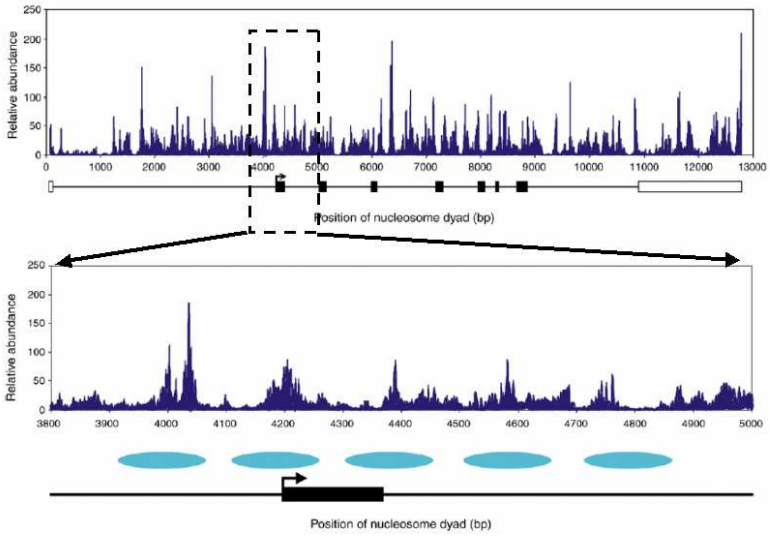

Fig.1 Nucleosomal occupancy map across sheep beta lactoglobulin gene (in vitro). Lower part zooms in on the 1.2 kbp promoter region. Solid boxes indicate exons, blue ovals the dominant nucleosomal sites found in vivo.

are finding that, as previously observed in yeast, 10 bp periodic arrangements of particular dinucleotides are important, but the key dinucleotides in the mammalian signature (GG and CC) are different from those in the yeast signature. We are interested in understanding the structural basis for this signature and in utilising it in computational algorithms designed to predict nucleosomal positioning from sequence data alone. In related work we are also exploring patterns of dinucleotide occurrence genome-wide across different animals, and assessing the role of DNA methylation in creating them (Fig.2).

SIMULATING IN VIVO MAPS

A separate strand of computational work focuses on predicting in vivo nucleosomal occupancy maps in regions

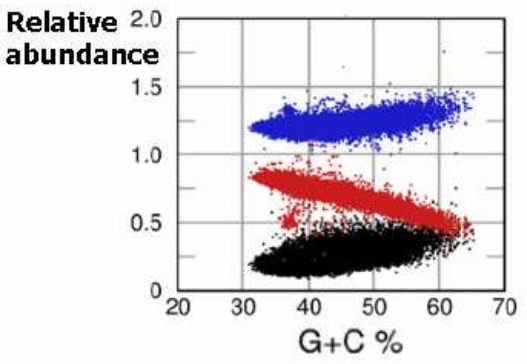

Fig.2 Normalised abundance levels of dinucleotides in 50 kbp chunks of the human genome sequence (CG-black, TG-blue, TA-red). Strong dependence on the G+C content is evident.

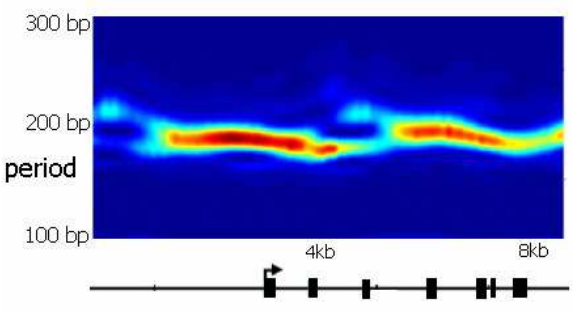

Fig.3 Spectral analysis of simulated occupancy maps using 57 nucleosomes competing for positions across sheep beta-lactoglobulin gene. Colour indicates the strength of the periodicity in the simulated occupancy maps across 1kb windows. The dominant period is seen to alter across the gene.

for which the in vitro map is known. The in vitro datasets are generated at unphysiologically low nucleosome density (one octamer per 400 bp), whereas in vivo densities are much higher. Our approach uses the in vitro data to define notional nucleosome binding energies at every 147 bp site across the region, and then computationally simulates the competition between a physiologically realistic number of histone octamers at thermal equilibrium using statistical physics techniques. The resulting (simulated) maps bear greater similarity to the in vivo data, confirming the validity of this approach.

Future work will involve constructing nucleosomal maps across 100 kb-scale human loci, and investigating the relationship between CpG methylation and nucleosome positioning

SELECTED REFERENCES

Fraser RM, Keszenman-Pereyra D, Simmen MW, Allan J (2009) High-resolution mapping of sequence-directed nucleosome positioning on genomic DNA, Journal of Molecular Biology 390: 292-305.

Simmen MW (2008) Genome-scale relationships between cytosine methylation and dinucleotide abundances in animals. Genomics 92: 33-40.

Fraser RM, Allan J, Simmen MW (2006) In silico approaches reveal the potential for DNA sequence-dependent histone octamer affinity to influence chromatin structure in vivo. Journal of Molecular Biology 364: 582-598.