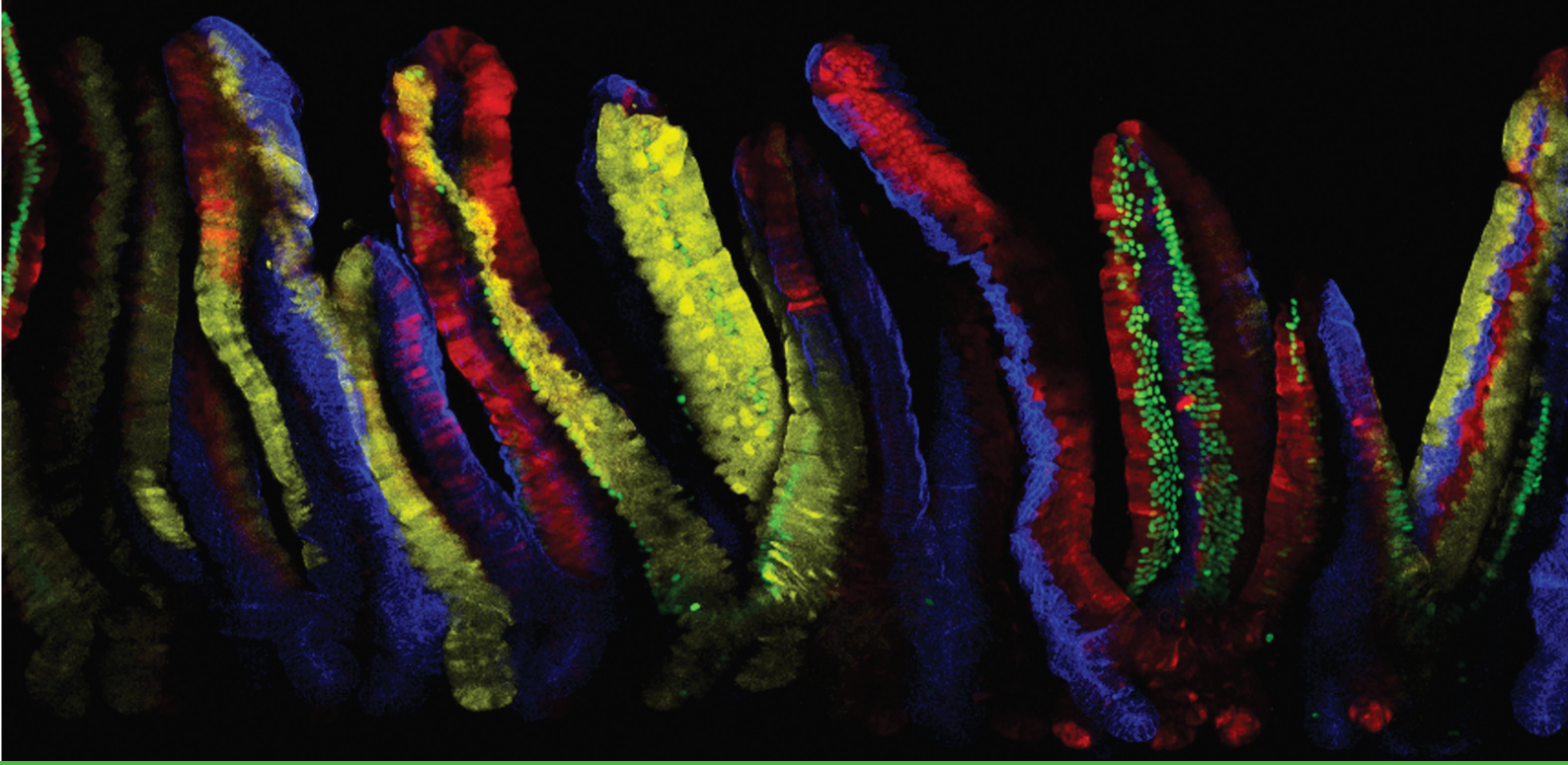
Abstract book

Quantitative Methods in Gene Regulation

22–23 September 2011
Institute of Physics, London, UK

http://gene.iopconfs.org



IOP Institute of Physics

Programme

Thursday 22 September

9:00 – 10:00	Registration (Opening Remarks)
10:00 -10:40	**Unfolding gene epistasis: vulva patterning in C.elegans** Eric Siggia, Rockfeller
10:40 -11:00	**The dynamics of supply and demand in mRNA translation** Chris Brackley, Aberdeen
11:00 – 11.30	Break
11:30 – 12:10	**The role of microRNA – mediated feedforward loops in the human regulatory network** Michele Caselle, Turin
12:10 – 12:30	**Buffering of the intracellular ribosome pool by slow codons** Philip Greulich, Edinburgh
12:30- 13:30	Lunch
13:30 – 13:50	**Modelling the effect of transcription factories-mediated interchromosomal contacts in the organisation of nuclear architecture** Andrzej Stasiak, Lausanne
13:50 – 14:10	**Organising principles of long-range chromosomal DNA contacts in a eukaryotic genome** Kai Kruse, MRC-LMB
14:10 – 14:30	**Optical stretching and digital holography quantify physical phenotypes of differentiating cells** Andrew Ekpenyong, Cambridge
14:30 – 15:10	**A model for all genomes: the role of RNA polymerases fixed in specialized factories** Peter Cook, Oxford
15:10-15:45	Break
15:45 – 16:45	Live-poster presentations (3 minutes per presentation)
16:45 – 17:05	**The helix-turn periodicity and length dependence of DNA mechanical properties** Agnes Noy, Oxford
17:05 – 17:25	**Moving beyond Watson-Crick coarse-grained models of DNA** Kevin Dorfman, Minnesota
17:25 – 18:05	**Single-molecule DNA sequencing by force: forward and backward motion of replicative polymerases and their coupling with helicases** Vincent Croquette, ENS Paris

Friday 23 September

9:40 – 10:00	**Mitochondrial variability as a source of extrinsic cellular noise** Iain Johnston, Oxford
10:00 – 10:40	**A general decomposition of variation in gene expression** Peter Swain, Edinburgh
10:40 – 11:00	**A Polycomb-based switch underlying quantitative epigenetic memory** Andrew Angel, John Innes Centre
11:00 – 11:30	Break
11:30 – 12:10	**Uncovering the interplay between transcription factor and nucleosome binding codes** Sarah Teichmann, MRC-LMB
12:10 – 12:30	**RNA sequencing reveals two major classes of gene expression levels in metazoan cells** Daniel Hebenstreit, MRC-LMB
12:30 – 13:30	Lunch
13:30 – 13:50	**Quantitative analysis of the fission yeast transcriptome and proteome** Samuel Marguerat, UCL
13:50 – 14:10	**A computational modelling approach to identifying mechanistic networks in RNA-Seq data** Vijay Chickarmane, CalTech
14:10 – 14:30	**Genome partitioning and large-scale transcriptional regulation** Marco Cosentino Lagomarsino, CNRS
14:30 – 15:10	**Unraveling the rules of transcriptional regulation using large-scale libraries of designed promoter sequences** Eran Segal, Weizmann
15:10 – 16:45	Poster session with coffee
16:45 – 17:05	**Gene expression noise as a mechanism for bacterial persistence** Andrea Rocco, Surrey
17:05 – 17:45	**Mechanisms and evolution of Mammalian Transcriptional Control** Duncan Odom, Cancer Research UK

Invited Oral Presentations

Unfolding gene epistasis: vulva patterning in C.elegans

Eric Siggia
Rockfeller

Several decades of research have revealed developmental signaling networks with dozens of intertwined genetic components that are very difficult to quantify in an embryological context. Geometric reasoning allows the enumeration of a hierarchy of models for differentiation in space and time. Vulval development in the worm C.elegans is a classic example of pathway integration, specifically EGF/Ras with Notch/Delta. One geometric model is favored and its minimal parameters can be mostly fit to genetic data for the penetrance of the three allowed terminal states of the vulva precursor cells (VPC) in various genetic backgrounds and after anchor cell (AC) ablation. The parameters are much closer to observable phenotypes than in genetic models and it is transparent which parameters can only be fit in combinations. The geometric description suggests variables in terms of which the two pathways additively combine without interaction, yet the penetrance phenotype implies epistasis, e.g., hypomorphic alleles of the two pathways that individually have no phenotype show a large effect when crossed. Our model predicts the fates of an isolated VPC, from the data on AC ablation and more generally the phenotype of the matrix of crosses resulting from all the single allele/condition data we fit to.

The role of microRNA-mediated feedforward loops in the human regulatory network

Michele Caselle
Turin University

MicroRNAs are endogenous non-coding RNAs which negatively regulate the expression of protein-coding genes in plants and animals. They are known to play an important role in several biological processes and, together with transcription factors, form a complex and highly interconnected regulatory network. Looking at the structure of this network, it is possible to recognize a few overrepresented motifs which are expected to perform important elementary regulatory functions. Among them, a special role is played by the microRNA-mediated feedforward loop in which a master transcription factor regulates a microRNA and, together with it, a set of target genes.
We show analytically and through simulations that the incoherent version of this motif can couple the fine-tuning of a target protein level with an efficient noise control, thus conferring precision and stability to the overall gene expression program, especially in the presence of fluctuations in upstream regulators. Among the other results, a nontrivial prediction of our model is that the optimal attenuation of fluctuations coincides with a modest repression of the target expression. This feature is coherent with the expected fine-tuning function and in agreement with experimental observations of the actual effect of a wide class of microRNAs on the protein output of their targets.
We also discuss the impact on noise-buffering efficiency of the cross-talk between microRNA targets that can naturally arise if the microRNA-mediated circuit is not considered as isolated, but embedded in a larger network of regulations. Finally as an example of our results we discuss in detail the miRNA mediated FFLs involving Myc as master Transcription factor.

A model for all genomes: the role of RNA polymerases fixed in specialized factories

Peter R Cook
University of Oxford

A parsimonious model for all genomes involving one major architectural motif will be presented: DNA/chromatin loops are tethered to transcription factories through active RNA polymerases and/or transcription factors. The polymerases are immobile and produce their transcripts by reeling in the DNA; this contrasts with the conventional view where polymerases track like locomotives down the template. We tested experimentally whether active polymerases are immobile using 3C (chromosome conformation capture) and human genes switched on rapidly and synchronously by tumor necrosis factor . This potent cytokine signals through NF B to stimulate and repress many genes. Within ten minutes after stimulation, responsive genes come together to be transcribed in specialized "NF B" factories; contacts are invariably between just those sequences being transcribed at that particular moment. Some factories further specialize in transcribing responsive genes encoding micro-RNAs that target down-regulated mRNAs. Super-resolution microscopy confirms that nascent transcripts (detected by RNA fluorescence in situ hybridization) co-localize at relevant times. As transforming growth factor also induces its responsive genes to co-associate, we suggest cytokines typically signal through analogous specialized factories. These results are consistent with polymerases being immobilized when active.

Single-molecule DNA sequencing by force: forward and backward motion of replicative polymerases and their coupling with helicases

Vincent Croquette, Maria Manosas, Fangyuan Ding, Spierring Michelle, Steven Benkovic
ENS Paris

In the replisome the polymerase collaborates with the helicase to drive the leading strand synthesis. In vitro, the helicases alone appear to unwind the replication fork with a rate far slower than the replicative one. Polymerases alone are even worse when working on a fork substrate that is in strand displacement. On the other hand, coupling the two enzymes leads to a fast and processive synthesis. The understanding of this coupling is far from perfect and was the motivation of our study.
We investigate replicative polymerases and their helicase coupling in single molecule assays using a DNA hairpin in an unzipping configuration. In this simple fork model we assist these enzymes by an external force which provides a control parameter. Modulating the force between 0 and 15pN is a mean to assist a molecular motor opening the fork and thus to replace a partner in the collaborative work. This strategy has already been used to evidence that T7 and T4 helicases display an unwinding rate increasing exponentially with the force. We have conducted this assay with different replicative polymerases, we show that strand displacement polymerization is possible with substantial assisting force level. When the force is reduced we show that the exonuclease activity is dominant.
This finding leads to the development of a Cyclic Polymerase Assay (CPA) where a polymerase is periodically switched from polymerization to strand degradation by modulating the assisting force. Such an assay is very convenient to study polymerase activity. Moreover, we present here a single molecule version of the Sanger sequencing method. We shall also discuss other sequencing assay using a DNA hairpin.
Finally, the same assay can also be carried out to study the coupling between helicase and polymerase. We find that the coupled system is very efficient and advances at maximum rate, in stark contrast to the case of the isolated polymerase. We explain this result by a collaborative model where both the helicase and the polymerase are described by the Betterton-Julicher helicase model (Betterton 2003). The helicase is described by a fast passive helicase while the polymerase corresponds to a slow weakly active helicase.

A general decomposition of variation in gene expression

Peter Swain
University of Edinburgh

To understand how cells control and exploit biochemical fluctuations, we must identify the sources of stochasticity, quantify their strength, and distinguish information flow from confounding `noise'. I will present a general theory to decompose fluctuations of gene networks into their component parts and discuss its application to recent measurements of osmosensing in budding yeast.

Uncovering the interplay between transcription factor and nucleosome binding codes

Sarah Teichmann
MRC-LMB

Transcription factors (TFs) and histone octamers are two abundant types of nuclear proteins that bind to DNA. Do they bind to overlapping or different sites on genomic DNA? Here, we demonstrate that two thirds of all known TFs in budding yeast show a sequence preference to regions on naked DNA that are also preferred by nucleosomal histones. These TFs tend to be associated with gene activation, and presumably have to compete with histones to bind to genomic DNA. Conversely, TFs with sequence preferences significantly different to those of histones tend to be repressors, suggesting that they cooperate rather than compete with histones. Based on intrinsic DNA sequence preferences derived in vitro, a majority of computationally predicted TF binding sites are expected to be occluded by nucleosomes. We demonstrate that a significant amount of nucleosomes are repositioned to less favourable sequences upon TF binding in vivo, which results in a greater proportion of accessible TF binding sites. In addition, our findings suggest that the combinatorial binding of histone-competing TFs to proximal sites is an important mechanism that facilitates global nucleosome repositioning in vivo, possibly in a cooperative manner. Taken together, our study provides genomic level insights into the interplay between DNA sequence, nucleosome occupancy, and TF binding in the regulation of gene expression.

Unraveling the rules of transcriptional regulation using large-scale libraries of designed promoter sequences

Eran Segal
Weizmann

Proper control of mRNA levels is critical in nearly all biological processes. Since much of this control is encoded within non-coding regulatory regions, deciphering the details of this mapping between DNA sequence and mRNA expression levels is key for understanding transcriptional control. Such an understanding could allow us to predict gene expression from DNA sequence, with far reaching implications. Most notably, genetic studies in a broad range of human diseases found a substantial contribution of genetic variation in non-coding regions to phenotypic diversity, and many expression changes have in turn been linked to disease states. However, without a ‘regulatory code’, comparable to the ‘genetic code’, we cannot tell which sequence changes cause the observed expression changes, and by what mechanism. To address this challenge, we designed and synthesized over 6000 different regulatory sequences in which we systematically varied several regulatory building blocks, including location, number, affinity and organization of transcription factor binding sites and of nucleosome positioning sequences, and measured the expression of each of these promoters with high accuracy. Our analysis of this library provides many insights into transcriptional control and paves the way towards the ultimate goal of predicting expression changes among human individuals from the genotype information that is rapidly being collected for them.

Mechanisms and evolution of Mammalian Transcriptional Control

Duncan Odom
Cancer Research UK

The relationships among tissue-specific transcription, genetic sequence, and evolution are complex and surprisingly plastic, and there is remarkable divergence in tissue-specific transcriptional regulation among closely related mammals. I will be presenting our recent data on the evolution of transcription factor binding, insulator binding elements, and polymerases, as well as new approaches we have developed to discover hidden conservation within these dynamic and rapidly changing processes.

Contributed Oral Presentations

Buffering of the intracellular ribosome pool by slow codons

Philip Greulich[1], Rosalind J. Allen[1], Luca Ciandrini[2], Mamen C. Romano[2]
[1]University of Edinburgh, [2]University of Aberdeen

In cells, mRNAs compete for a finite number of ribosomes. This suggests that the protein synthesis levels of genes might be mutually negatively correlated: If the transcription rate of one gene increases significantly, the new mRNAs are expected to bind ribosomes, thereby decreasing the concentration of available free ribosomes for translation. This should lead to an effective down-regulation of other genes, preventing the cell from independently responding to different external stimuli.
We propose a mechanism that effectively buffers the free ribosome pool, making it independent of fluctuations in mRNA-number and total amount of ribosomes. The effect can prevent negative correlations in the expression of different genes. Slow codons act as bottlenecks to protein synthesis, which can lead to ribosome queues on the mRNA molecules. We show that these queues can buffer the free ribosome pool since ribosomes are "stored" in the queues and are automatically and rapidly released when the concentration of free ribosomes drops to a given level, without the need for explicit regulation of the ribosome pool. The concentration at which the free ribosome pool is buffered is determined by the ribosome stepping rate at the slow codon bottleneck.
Our results may have significant implications for cells´ ability to respond independently to multiple demands on the ribosome pool.

The dynamics of supply and demand in mRNA translation

Chris Brackley
University of Aberdeen

We examine the balance of supply and demand of resources in a driven diffusion model of mRNA translation. The asymmetric simple exclusion processes (ASEP) is a paradigmatic model from non-equilibrium physics, and has been used to describe many systems, including protein production in cells. This process involved the translation of mRNA molecules (which are transcribed from DNA) by molecular machines called ribosomes, which move along the mRNA adding different amino acids to a growing chain that will become a protein.
I will discuss a recent extension to the ASEP model of translation [1] that includes the use and replenishment of tRNA molecules – the carriers that provide the amino acids to the ribosome, and match specific amino acids to the coding sequence of the mRNA. I will show simulation results and analytic predictions for groups of mRNA sequences taken from the yeast genome [2]. Changes in the supply of and demand for different tRNA species can lead to different production rates of different proteins, and I will

show that bottlenecks to translation induced on some mRNA can lead to the freeing of resources required for the transcription of others.
[1] C A Brackley et al (2010) Phys Rev Lett. 105 078102; Phys Rev E 82 051920
[2] C A Brackley et al (2011) "The Dynamics of Supply and Demand in mRNA Translation" Under Review.

Modelling the effect of transcription factories-mediated interchromosomal contacts in the organization of nuclear architecture

Andrzej Stasiak, Julien Dorier
University of Lausanne

Using numerical simulations, we investigate the underlying physical effects responsible for the overall organization of chromosomal territories in interphase nuclei. In particular, we address the following three questions: 1. Why are chromosomal territories with relatively high transcriptional activity on average, closer to the centre of cell's nucleus than those with the lower activity? 2. Why are actively transcribed genes usually located at the periphery of their chromosomal territories? 3. Why are pair-wise contacts between active and inactive genes less frequent than those involving only active or only inactive genes? Our simulations indicate that transcription factories-mediated contacts between active genes belonging to different chromosomal territories are instrumental for all these features of nuclear organization to emerge spontaneously due to entropic effects arising when chromatin fibres are highly crowded.

Organizing principles of long-range chromosomal DNA contacts in a eukaryotic genome

Kai Kruse, Sven Sewitz, Madan Babu
MRC Laboratory of Molecular Biology

In most eukaryotic genomes, DNA is packed into a higher-order structure of layered and hierarchically organized chromatin. Given the enormous complexity of molecular events that need to be coordinated in space and time during transcription, one can assume that chromatin organization is tightly constrained within the nucleus. Recent advances in the mapping of long-range chromatin interactions enable us to construct a genome-scale network of links between genomic DNA segments in Saccharomyces cerevisiae. Through integration of genome-wide datasets we can now uncover the relationships between nuclear architecture and various genomic properties.
At the DNA level, we show that meiotic recombination hotspots are enriched on highly linked DNA segments, whereas the transcription-associated Z-DNA and G-Quadruplex structural motifs appear particularly frequently on apparently unconnected segments. An analysis of the transcriptional and chromatin regulatory networks reveals that many inducible genes (e.g. stress and nutrient response) do not engage in any long-range DNA-contacts and are highly sensitive to deletion of chromatin regulators, suggesting that they may be poised for transcriptional activation. Finally, we show that the topological properties of the genomic interaction network are similar to other biological networks (i.e. scale-free and small-world) and that the clustering behavior of DNA segments differs strongly between inter- and intra-chromosomal contacts. These properties permit the possibility of discovering specific "organizing centers" for maintaining the genomic architecture.

Optical stretching and digital holography quantify physical phenotypes of differentiating cells

Andrew Ekpenyong[1], Kevin Chalut[1], Graeme Whyte[1], Franziska Lautenschläger[1], Jochen Guck[1], Ada Olins[2], Donald Olins[2]
[1]University of Cambridge, [2]Bowdoin College, USA

Biological cells in a state of stemness continue in their state of stemness unless compelled by internal and external changes in their niche to respond otherwise. The response, which may be differentiation and or death, is driven by the way each cell deploys its genome. We focused on the process of differentiation, which involves points of transition where chromatin structures and transcription factors combine at particular genes. It involves unchanging DNA sequence (except in very rare cases) at one end and changing physical characteristics at the other end. In between, it means that the phenotypic attributes of a cell depend on how proteins resulting from gene expression proceed to interact with one another. We have used two new techniques, an optical cell stretcher and a digital holographic microscope, to monitor the changes in mechanical and structural properties of HL-60/S4 cells (malignant human hematopoietic precursor cells) as they differentiate into monocytic, granulocytic and macrophage forms. We found significant differences in the compliance of HL-60/S4 cells and the differentiated forms, which are consistent with known cytoskeletal and karyoskeletal changes. We also found lineage specific modulation of the refractive index distribution of differentiating cells.
Our combination of novel marker-free optical micromanipulation and imaging techniques to quantify phenotypic changes in differentiating cells promises to be a useful method for direct correlation of gene expression with the proteins that result therefrom. This may help to establish a better understanding of the mechanism of hematopoietic stem cell differentiation.

Moving beyond Watson-Crick coarse-grained models of DNA

Kevin Dorfman, Margaret Linak, Richard Tourdot
University of Minnesota

The local, sequence dependent structures of DNA play important roles in gene regulation. The length and time-scales characterizing the dynamics of such structures are beyond the capabilities of all-atom molecular dynamics simulations and thus require coarse-grained models. Most existing models only incorporate standard Watson-Crick hydrogen bonds, which limits their applicability to the canonical B-form double helix or minor perturbations therefrom. To further extend the utility of coarse-grained models of DNA, we developed an experimentally parameterized model that includes both Watson-Crick and non-Watson-Crick hydrogen bonding and simulate the DNA dynamics using Brownian dynamics. The model captures the microscopic features of double-stranded DNA, without the need for explicit constraints, as well as the experimental melting curves for a number of short single-stranded DNA hairpins. These properties allow us to smoothly move between single-stranded, double-stranded and partially denatured structures. We demonstrate the utility of the model by simulating more complex tertiary structures such as the folding of the thrombin apatmer, which includes G-quartets, and strand invasion during triplex formation. Our results highlight the importance of non-canonical interactions in coarse-grained DNA models.

The helix-turn periodicity and length dependence of DNA mechanical properties

Agnes Noy, Ramin Golestanian
University of Oxford

Although single-molecule force-extension methods described DNA with a persistence length of 50 nm and a stretch modulus of 1100 pN (1), novel experimental techniques have suggested DNA might be more

bendable at smaller scales (2). More controversial was the last small-angle x-ray scattering (SAXS) measure of stretch deformability using fragments shorter than 40 base-pairs which provided a value one order of magnitude lower than previous estimations achieved through a cooperative behavior (3,4). Precisely, this is the biologically relevant scale for DNA packaging, transcription and gene regulation where frequently DNA needs to be severly distorted. Here, we study how physical properties of DNA change as we travel from base-pair level to near 4 DNA-turns, showing for the first time a DNA-turn periodicity of the persistence length due basically to the static persistence length. In addition, we find an apparent cooperative behaviour for the stretch parameter caused by the superposition of various linear modes, as well as a match of the torsion elastic constant to the long range experimental values from half DNA-turn.

C Bustamante, SB Smith, J Liphardt and D Smith. Curr Opin Struct Biol, 10, 279–285 (2000)
PA Wiggins, TVD Heijden, F Moreno-Herrero, A Spakowitz, R Phillips, J Widom, C Dekker and PC Nelson. Nature Nanotech, 1, 137–141 (2006).
R Mathew-Fenn, R Das and P Harbury. Science, 322, 446 (2008).
NB Becker and R Everaers. Science, 325, 538 (2009).

Mitochondrial variability as a source of extrinsic cellular noise

Iain Johnston
Oxford University

We present a study investigating the role of mitochondrial variability in generating noise in eukaryotic cells. Noise in cellular physiology has been found to play an important role in many fundamental cellular processes, including transcription, translation, stem cell differentiation and response to medication. However, the mechanisms through which random influences affect these processes have yet to be clearly elucidated. Here we present a mechanism by which variability in mitochondrial volume and functionality is linked to variability in transcription rate and hence has a profound effect on downstream cellular processes. We expose a knitting of cell cycle variability, transcription rate variation and mitochondrial variation. Our model mechanism is supported by an appreciable volume of recent experimental evidence, and we present the results of several new experiments with which our model is also consistent. We find that noise due to mitochondrial variability can sometimes dominate over other extrinsic noise sources (such as cell cycle asynchronicity) and can significantly affect large-scale observable properties such as cell cycle length. We also explore two recent regulatory network-based models for stem cell differentiation, and find that extrinsic noise in transcription rate causes significant variability in the behaviour of these model systems. These results suggest that mitochondrial and transcriptional variability may be an important mechanism influencing a large variety of cellular processes and properties.

A Polycomb-based switch underlying quantitative epigenetic memory

Andrew Angel, Jie Song, Caroline Dean, Martin Howard
John Innes Centre

Although much is known about Polycomb-induced regulation in the field of epigenetics, key mechanistic questions remain concerning how silencing is initiated and maintained. Vernalization, the perception and memory of winter in plants to aid in the correct timing of flowering, is a classic epigenetic process which, in Arabidopsis, involves polycomb repressive complex 2 (PRC2) based silencing of the floral repressor FLC. Vernalization is a slow process, with memory of cold being built up quantitatively over weeks. These features make vernalization an ideal experimental system to investigate epigenetic silencing. Here, using mathematical modelling, chromatin immunoprecipitation (ChIP) and an FLC:GUS reporter assay, it is shown that the quantitative nature of vernalization is generated by H3K27me3-mediated silencing in the warm in a subpopulation of cells whose number increases with length of cold. Simulations predict that an

experimentally-measured spatially-limited region of nucleated H3K27me3 marks can switch individual cells into a silenced state and the resulting bistable expression pattern is verified using the FLC:GUS reporter system. The proposed switching mechanism is likely to be of wide relevance in epigenetic reprogramming.

RNA sequencing reveals two major classes of gene expression levels in metazoan cells

Daniel Hebenstreit[1], Muxin Gu[1], Varodom Charoensawan[1], Miaoqing Fang[2], Alexander van Oudenaarden[2], Sarah Teichmann[2]
[1]MRC Laboratory of Molecular Biology, [2]Massachusetts Institute of Technology

The expression level of a gene is often used as a proxy for determining whether the protein or RNA product is functional in a cell or tissue. Therefore, it is of fundamental importance to understand the global distribution of gene expression levels, and to be able to interpret it mechanistically and functionally. Here we use RNA sequencing (RNA-seq) of mouse Th2 cells, coupled with a range of other techniques, to show that all genes can be separated, based on their expression abundance, into two distinct groups: one group comprised of lowly expressed and putatively non-functional mRNAs, and the other of highly expressed mRNAs with active chromatin marks at their promoters. These observations are confirmed in many other microarray and RNA-seq data sets of metazoan cell types.

Quantitative analysis of the fission yeast transcriptome and proteome

Samuel Marguerat[1], Sandra Codlin[1], Juerg Baehler[1], Alex Schmidt[2], Ruedi Aebersold[3]
[1]University College London, [2]Basel University, [3]ETH, Zuerich

Over a decade of genome-wide studies have made important contributions to our understanding of the dynamics of eukaryotic transcriptomes and proteomes. However, absolute concentrations of individual transcripts and proteins and their relationships remain unexplored on a global scale, and are likely to provide unique insight into cell regulation. We have acquired average absolute transcript counts per cell using RNA-seq data calibrated with absolute concentrations of selected mRNAs, along with corresponding protein measurements applying a proteomics approach using selected reaction monitoring and measurement of precursor ion intensities. Rapidly proliferating or quiescent fission yeast cells were used as model conditions to investigate protein and transcript concentrations in two physiological programs optimised towards cell growth or maintenance, respectively.
Our analysis revealed that during rapid proliferation a majority of coding transcripts are present at between 1 and 300 copies per cell, with a median of 2.4 copies. A small portion of transcripts were measured at less than one copy per cell; they were enriched for genes known to be tightly repressed during growth, suggesting a natural threshold for gene expression. During quiescence, when cells stop growing and become smaller, the fission yeast transcriptome shrinks in overall number but less in diversity. In both conditions, the long non-coding RNAs show expression levels mostly below one copy per cell, although dozens of these transcripts reach higher concentrations. Protein concentrations show a substantial correlation with transcript numbers during growth and quiescence, but the absolute quantities and dynamic range of proteins differ. Protein concentrations of over a million copies per cell can be detected, indicating a much larger dynamic range than for transcripts. In addition, the protein to transcript ratio was highly variable. These data suggest that even though mRNA and protein quantities are correlated overall, important regulation also occurs at the post-transcriptional level.

A computational modeling approach to identifying mechanistic networks in RNA-Seq data

Vijay Chickarmane, Sean Gordon, Elliot Meyerowitz
California Institute of Technology

Cytokinins are plant hormones with diverse roles in growth and development. Cytokinin is perceived by the cytokinin receptors AHK2, AHK3 and AHK4 which activate a multistep phosphorelay pathway leading to regulation of gene activity. Previous analysis of cytokinin receptor mutants indicates that these receptors are partially redundant but also have some specific functions. Unraveling the relative contributions of individual receptors to the regulation of specific genes is a complex challenge. In this study we perform a systematic analysis of AHK2, AHK3 and AHK4 specificity in regulation of gene transcription within the Arabidopsis shoot apex. We use RNA-Seq to generate quantitative, genome-wide gene expression data for individual receptor mutants and their response to exogenous cytokinin. We develop new computational tools to probe the underlying mechanistic specificity of individual receptors and their combined roles in gene regulation using RNA-Seq data. The most parsimonious mathematical models of gene regulation for a set of specific cytokinin regulated genes based upon the data suggest that these circuits implement feedback between the downstream effect of the receptors and the transduced signals. In particular our data and models suggest antagonism between the AHK2 and AHK3 receptors. Our approach has wide applications in the interrogation of RNA-Seq data to gain insight into specificity of pathways regulating gene expression.

Genome partitioning and large-scale transcriptional regulation

Marco Cosentino Lagomarsino,
Génophysique / Genomic Physics Group

From a coarse, large-scale viewpoint, a bacterial genome can be seen as a hierarchical repository of genes and other functional elements. Both its content and its "architecture" (how the content is organized) are shaped by a number of constraints and optimization principles of physico-chemical, biological or evolutionary nature. For example, regulation of transcription has the scope of activating the correct genes in response to a set of stimuli and environments. It can be represented as a network whose action has a complex feedback with the physical organization of the genome as a giant nucleoprotein complex. Moreover, its architecture is conditioned by the classes of metabolic processes the cell participates into, and by the evolutionary growth of the genome by horizontal transfer and gene duplication. I will describe a series of quantitative approaches guided by data and statistical physics modeling, aimed at characterizing large-scale transcriptional regulation through these constraints at the level of a single or multiple bacterial genomes.

Gene expression noise as a mechanism for bacterial persistence

Andrea Rocco, Andrzej Kierzek, Johnjoe McFadden
University of Surrey

One of the most fundamental problems in microbiology is to understand how a small fraction of a clonal population of bacteria can resist killing by antibiotics, the phenomenon known as persistence, or drug tolerance. Its characteristic signature is the biphasic killing curve, whereby the population exposed to the antibiotic is initially killed very rapidly, but shows a slow tail at long times.
The phenomenon is usually explained by assuming the coexistence of different bacterial subpopulations, each representing a different drug tolerance phenotype. However, the mechanism for the appearance of such different phenotypes remains as yet unknown.

In this context I will discuss how phenotypic heterogeneity can arise due to gene expression intrinsic noise, related to low copy numbers of protein molecules. I will propose a model defined at the single cell level as a two-stage model of gene expression, with protein inhibited cellular growth. At the population level, the protein distribution can be obtained by stochastic simulations (Gillespie), and is compatible with a Gamma distribution. This result extends the previous result obtained without growth control by Friedman et al. PRL 97, 168302 (2006).
In the limit of low expression, stochastic fluctuations and protein mixing become slow, and cells are "frozen" in growth phenotypes. This gives rise to a bimodal growth rates distribution, describing a composite population of cells in which most cells exhibit the maximal growth rate, and a minority grows more slowly. Once coupled to the well-established growth rate dependency of antibiotic killing, this model reproduces well the biphasic killing observed in persistent populations.
This simple model has major implications for efforts to develop control strategies for persistent infections.

Poster Presentations

Bistability in nutrient-sensing networks

Andrea Weisse, Nicolas Oury, Peter Swain
University of Edinburgh

Bistability is key to a wide range of cellular processes such as differentiation and neuronal memory. The galactose genetic network of Saccharomyces cerevisiae is an example of a bistable nutrient-sensing network. Upon exposure to the sugar galactose, the enzymes are expressed either at a high or at a basal, almost inactive rate. The role of bistability in microbial nutrient-sensing networks is unclear and may indicate the presence of social behaviour.
Conventionally, a bistable response is considered advantageous because it allows a population of related cells to hedge their bets with some cells being in one of the bistable states and other cells in the other. This phenotypic variation increases the fitness of the population at the possible cost of reduced fitness for some individuals. Such an argument is, however, an argument for a stochastic response, which need not be bistable.
We use stochastic models of two competing cell populations, one with a bistable and one with a graded nutrient response, to study which factors determine a winning strategy.

Analytical study of an exclusive genetic switch

Juan Venegas-Ortiz, Martin Evans
University of Edinburgh

The nonequilibrium stationary state of an exclusive genetic switch is considered. The model comprises two competing species and a single binding site which, when bound to by a protein of one species, causes the other species to be repressed. The model may be thought of as a minimal model of the power struggle between two competing parties. Exact solutions are given for the limits of vanishing binding/unbinding rates and infinite binding/unbinding rates. A mean field theory is introduced which is exact in the limit of vanishing binding/unbinding rates. The mean field theory and numerical simulations reveal that generically bistability occurs and the system is in a symmetry broken state. An exact perturbative solution which in principle allows the nonequilibrium stationary state to be computed is also developed and computed to first and second order.

Prediction of regulatory regions: from DNA sequence analysis to recognition of structural codes for specific protein-DNA binding.

Alona Sosinsky, Kathie Yeowell
Birkbeck College

Phylogenetic footprinting methods, which attempt to identify functional regulatory regions by comparing orthologous genomic sequences of evolutionary related species, have shown promising results. However, a number of limitations have so far restricted applicability of alignment-based methods. To overcome such limitations, we have developed a method called DREiVe (Discovery of Regulatory Elements in Vertebrates), which integrates motif-discovery approach with local permutation-clustering algorithm. Together they can identify large regulatory elements in vertebrate genes as evolutionarily conserved, order-independent clusters of short conserved DNA sequence motifs. We've successfully tested DREiVe performance on a set of experimentally identified regulatory elements from human genome.
DREiVe predictions can be combined with the search for clusters of binding sites for a specified set of transcription factors. Such methods usually consider site density and relative site affinity. However, the correct positioning of binding sites relative to each other is known to be important for specificity and function of many promoters and enhancers and provides another factor that can be used in prediction. To determine whether the preferential arrangements of binding sites can be revealed from computational analysis of the whole genome, we analyzed the distance preferences between pairs of binding sites from TRANSFAC collection. Our findings propose that preferential arrangement of binding sites in genome sequences can be a common feature for many factors and can serve as a test for their cooperativity.
We further attempt to refine prediction of binding sites for transcription factors with employing structural information on protein/DNA binding. We have shown that small sequence-dependent deformations of DNA structure and changing electrostatic properties of DNA can encode for specific recognition by individual members of families of transcription factors. Combination of DNA structure prediction with understanding of structural features required for specific protein/DNA recognition can improve our ability to recognize specific binding sites in genomic sequences.

Optimal origin placement in DNA replication

Jens Karschau, Alessandro P. S. de Moura
University of Aberdeen

DNA replication is an essential process and its timing must be robust so that cells can divide properly. Fluctuations in the formation of replication starting points, called origins, and the subsequent activation of proteins lead to the risk of variation in timing and loss of genomic information. I will discuss a mathematical investigation of the stochastic properties of DNA replication. The model elucidates the spatial organisation of origins on the DNA required to ensure brief and faithful duplication. Previous work has used the experimentally determined origin locations as given parameters, without seeking an understanding of the spatial arrangement.
Contrary to what one might naively expect - that placing origins as far apart as possible will achieve minimum replication time - I will show that it can actually be more efficient to place origins together in clusters for certain parameter regimes. Whether it is optimal to place origins in clusters or in isolation strongly depends on both the probability that certain sites will bind with origin forming proteins and the length of the activation time window. I will show that the same phase transition is observed in simulations which include stochasticity in origin activation and in simulations which do not, i.e. the result is independent of the precise details of the model. The work explains the basis of the spatial distribution of origins for yeast and frog embryos. Overall, the results show that the organisation of origins is an optimisation problem in nature which relies on evolution and not on randomness.

Kinetic proofreading in chromatin remodeling

Ralf Blossey
CNRS

In contrast to prokaryotes, gene activation and repression in eukaryotes requires an activated step: the remodeling of chromation structure. Dedicated motor complexes have evolved for this purpose. It is shown that chromatin remodeling involves kinetic proofreading mechanisms, both for activation and repression of genes. An essential element in these processes is the recognition of histone tails and their enzymatic modifications. Certain signals from the postulated histone code are therefore relevant for structural changes of chromatin, and hence of gene expression patterns.

Ribosome recruitment and translational regulation. A physicist's point of view.

Luca Ciandrini[1], Ian Stansfield[1], Maria Carmen Romano[1], Philip Greulich[2], Rosalind J. Allen[2]
[1]University of Aberdeen, [2]University of Edinburgh

One of the most important steps of protein synthesis is the translation of the messenger RNA (mRNA). During this process, macromolecules known as ribosomes translate the genetic information encoded by the triplets of nucleotides (codons) of the mRNA, and assemble proteins amino acid by amino acid. Several ribosomes may simultaneously translate the same messenger. It is possible to interpret this process as a driven lattice gas where particles on a uni-dimensional track represent ribosomes moving along the mRNA. We use a general approach to study mRNA translation by means of statistical-mechanical models based on the Totally Asymmetric Simple Exclusion Process (TASEP).
By applying an extended model [PRE 81, 051904 (2010)] to the more than 6,000 genes of baker's yeast we are able to study the characteristics of the whole transcriptome of S.cerevisiae, estimate the 'traffic effects' of ribosomes in real (biological) sequences, and predict the ribosome entry rate (initiation) for each mRNA.
We also investigate the competition of several lattices sharing a common reservoir of particles and we explain how a mixed population of lattices (i.e. lattices with different properties) behave differently by changing the amount of resources. The competition for the resources in the cell might be crucial in some biological conditions, e.g., to react to stress conditions.

Dynamics of bacterial chromosomes – A physical approach

Krystyna Bromek[1], Vivek Thacker[1], Eileen Nugent[1], Jurij Kotar[1], Ulrich Keyser[1], Pietro Cicuta[1], Bianca Sclavi[2], Marco Cosentino-Lagomarsino[3]
[1]Cavendish Laboratory, Cambridge, [2]LBPA, ENS Cachan, [3]Genomic Physics Group, Paris

In recent years it has become clear that the bacterial nucleoid (DNA with associated proteins) is structured into domains at various length scales. The precise structure and compaction are not fully understood but are known to be dependent on the cell cycle, growth conditions and external stimuli. We investigate the organisation in E.coli chromosomes by studying the physical properties of nucleoids from different strains of bacteria. Using optical tweezers, we have demonstrated that the nucleoids can be compressed and change their cross-section in response to external viscous forces. By studying variations in nucleoid deformations in bacteria grown in different conditions we are able to measure the structural effects of genome packaging. A study of the variation in deformations across different strains also enables us to deduce the effect of different nucleoid binding proteins in maintaining structural integrity.
Another approach we have taken is to use the resistive pulse technique with microcapillaries (3-5 μm diameter) as a sensing tool. Nucleoids from fast and slow growing strains of bacteria have different ionic

current signatures indicating a clear difference in compaction. The confined geometry of the microcapillaries makes them good sizing tools and we are working on integrating this technique with Surface enhanced Raman Spectroscopy to provide label-free detection of protein- DNA complexes.

Recognition of histone tail modifications: results from a multi-scale study

Ralf Blossey[1], Raghvendra Singh[1], Rinaldo Zucca[1], Guillaume Brysbaert[1], Marc Lensink[1], Julie Pelletier[2]
[1]CNRS, [2]USTL

The N-terminal tails of histone proteins play a crucial role in transcriptional regulation in eukaryotes; the large variety of post-translational modifications that the different tail residues can undergo has even motivated the proposition of a "histone code". However, in most cases the patterns that link histone modifications to defined regulatory events have so far only been established through statistical correlation and a detailed mechanistic understanding of the specific role of these modifications is still lacking. Recently, studies on chromatin remodeling have shown the importance of histone tail states in the initiation and activation of the displacement of nucleosomes along DNA. Spurred by these advances, we present here results from a multi-scale study of histone tail recognition, undertaken with the goal to classify modification patterns by linking functional behavior to quantifiable and physical properties of the motifs. Our study combines i) a docking study of modified histone tails to their readers, ii) a statistical analysis of data from recognition assays and iii) a functional genomics approach to reconstruct the interaction networks of histone tails.

How do cells perform arithmetic division?

Antonio Scialdone, Martin Howard, Sam Mugford, Alison Smith
John Innes Centre

Cells are able to make vital decisions by sophisticated processing of molecular inputs. In some cellular processes the implementation of arithmetic operations on two molecular inputs can be required, e.g., to eventually produce a molecule at a concentration that is proportional to the sum or the product of the concentrations of two different molecular inputs. An important question to address concerns what mechanisms can lead to the implementation of the desired arithmetic operation.
Recent work [1] has shown that arithmetic division plays a fundamental role in the normal growth of the model plant Arabidopsis Thaliana. In order to support its growth during the night, the plant relies on starch reserves that have been accumulated throughout the day: the starch is then linearly degraded during the night, at a rate that ensures the exhaustion of starch reserves at the time of expected dawn, when photosynthesis begins again. Remarkably, the plant adjusts the degradation rate in response to an unexpected early or late night: the degradation rate d, is set on the basis of the amount of starch stored, S, and the time to next dawn, t, such that $d=S/t$. How is the plant able to perform such arithmetic division?
We present different general mechanisms that the cell can exploit to implement arithmetic division between two molecular inputs, both at transcriptional and post-translational levels. Experiments to assess the specific mechanism that is working in vivo are discussed, with a focus on the case of Arabidopsis.
[1] Graf et al., PNAS 107:9458 (2010)

Gene autoregulation via intronic microRNAs and its functions

Carla Bosia[1], Mariama El Baroudi[1], Michele Caselle[1], Matteo Osella[2], Davide Corà[3]
[1]Dipartimento di Fisica Teorica and INFN University of Torino, [2]Génomique des Microorganismes, Université Pierre et Marie Curie, [3]Systems Biology Lab, IRCC, University of Torino

While many microRNAs are transcribed from their own genes, nearly half map within introns of ''host" protein-coding genes. Eventually these intronic microRNAs may regulate the expression of their host genes, forming a microRNA-mediated self loop. A recent analysis of this phenomenon suggests that autoregulation mediated by intronic microRNAs may be under positive selective pressure but the functions of this regulatory circuit are currently unknown. We used a stochastic and a deterministic analysis of the circuit validated via simulations. Our model suggests that microRNA-mediated self regulation, despite its simple topology, can perform different regulative tasks. Firstly it can alter the response time of gene expression to upstream signals, secondly it can implement fold-change detection and finally it can confer robustness to noise. Moreover, we compare the features of this circuit with the analogous transcriptional self-regulation, which is an ubiquitous network motif in different species, to highlight in what situations a post-transcriptional self-regulation can be advantageous for the cell.

Exploring the links between nucleoid physical state, growth rate and gene expression

Matteo Osella, Vittore F Scolari, Mina Zarei, Marco Cosentino Lagomarsino
Génomique des Microorganismes

Bacterial cells rely on a strong link between transcription program, chromosome physical organization and growth conditions. Growth rate is coupled to cellular macromolecular composition, and thus to the main global parameters of gene expression (such as ribosome and RNA polymerase abundance). While locally it relies on specific transcription-factor binding, transcription depends globally on the state of the mesoscopic protein/genome complex called nucleoid and the main proteins shaping the nucleoid have growth-dependent expression. Finally, nucleoid organization and growth are also linked by the process of replication.
Our work, in direct collaboration with experiments, focuses on the characterization of these links at different levels using data analysis and models. This poster presents the intermediate results of different subprojects.
Integrating several data sources (microarrays, ChIP-on-chip) we identified clusters of genes sensitive to nucleoid-perturbation that correlate with large structured regions along the chromosome that typically demix in the cell (macrodomains). In particular, we found indications that the switch between the motile and biofilm lifestyle might depend on nucleoid-mediated regulation in two large regions containing the flagella regulons.
More generally, we considered how gene functions and transcriptional interactions are influenced by chromosomal coordinates. We found that essential genes preferentially cluster in specific chromosome domains and, analyzing the transcriptional regulatory network, we found significant correlation between link topology and chromosome organization.
Finally, we are investigating how global physiological parameters such as growth rate and nucleoid organization can contribute to the cell-to-cell variability of gene expression. Starting from quantitative estimates of the cellular macromolecular composition, we find a non-trivial scaling of expression fluctuations with growth rate.

Unraveling of the hidden DNA structural/physical code provides novel insights on promoter location

Elisa Duran Moliner[1], Montserrat Soler[1], Modesto Orozco[1], Sarah Djebali[2], Roderic Guigó[2],David Torrents[3]
[1]Institute for Research in Biomedicine, [2]Center for Genomic Regulation (CRG), [3]Barcelona Supercomputing Center (BSC)

In eukaryotic organisms, DNA compaction into chromatin structure has clear functional implications since modulates the accessibility of genomic regions to regulatory proteins. The way in which transcription initiation is regulated implies great complexity and remains unclear.
Controversy is ensured when defending whether DNA sequence itself is enough to determine where transcription initiation occurs. In our group we have observed that key regulatory regions like transcription start sites (TSS) are signaled by unusual physical properties of DNA, which led us to develop an algorithm to predict TSS locations exclusively based on well-defined physical properties of DNA deformability derived from atomic simulations [1].
A comprehensive genome-wide screen of this algorithm provided us thousands of putative TSSs that had not been previously described. Since most of the promoters are located upstream of the TSS, we consequently selected regions spanning from -1000/+200bp relative to each predicted TSS for cell-based gene reporter activity assays using a high-throughput strategy. Our results confirm that 70% of these regions show promoter activity. Furhtermore, CAGE analysis was subsequently performed to map those TSSs and confirm that these regions were indeed transcribed.
Clustering of the resulting promoter regions reveals that promoters of house-keeping genes are particularly suitable to be mapped according to its physical properties and very interestingly, 80% of them could not be detected neither by regular sequence analysis nor by sequence conservation studies.
[1] Goni, J. R., A. Perez, et al. (2007). "Determining promoter location based on DNA structure first-principles calculations." Genome Biol 8(12): R263.

Toxin-antitoxin battle in bacteria

Ilaria Cataudella
Niels Bohr Institute

Toxin-Antitoxin (TA) pairs can be found in many single organisms as a mechanism to respond to starvation induced stress. One of the most well studied examples is RelE-RelB pair in Escherichia Coli. RelE-RelB pair is expressed by the same operon, but antitoxin proteins have very short half-life and compensate it by having much higher translation rate than toxins. In ordinary situations the total amount of the antitoxin is around 10 fold higher than the toxin. The antitoxin RelB can form tight complexes with the toxin, inhibiting this way the toxic activity. The toxin RelE is an mRNA interferase that cleaves mRNA at the empty ribosomal A site. When the cell is in a healthy state all the toxin is sequestered in complexes with the antitoxin and no toxic activity is performed. Toxin-Antitoxin complexes have also been shown to autoregulate the TA operon by means of a mechanism known as conditional cooperativity. When the cell faces amino-acid starvation the lower overall translation rate results in the concentration of the toxin to get higher than the concentration of the antitoxin. To test this current experimental scenario, we built a stochastic model involving a feedback mechanism working both at a transcriptional and translational level and simulate the stress response against starvation. We will address the role of conditional cooperativity in the cell's ability to recover after the starvation phase and show importance of active degradation of the antitoxin not only in its free form but also when sequestered in complexes with the toxin.

Micro-dynamics of chromosomal loci

Avelino Javer[1], Marco Grisi[1], Eileen Nugent[1], Krystyna Bromek[1], Pietro Cicuta[1], Chiara Saggioro[2], Bianca Sclavi[2], Marco Cosentino Lagomarsino[3]
[1]Cavendish Laboratory, [2]LBPA, ENS Cachan, [3]Genomic Physics Group

Previously believed to behave as an amorphous polymer, the bacterial chromosome has recently been shown to have structure at various lengthscales. Boccard et al. identified 4 different macrodomains in E.Coli, and it has been estimated that it possibly contains more than 400 independent supercoiled domains. These results have renewed interest in studies of bacterial chromosome organization and its effect on the physiological state of the bacterium. With tag proteins, specific DNA sequences can be pinpointed, allowing the study of the dynamics of particular chromosome regions in live cells in real time. This has revealed new details of chromosome segregation, and has lead to the hypothesis of physical models of chromosome structure (2010 publications by Wiggins et al., Theriot et al.). In this work we analyze 15 different bacterial strains (the Boccard collection), where in each strain a particular chromosome position is labeled. The loci positions are continuously tracked by fluorescence imaging. The mean loci displacement with time is well described by a power law, and we discuss the motility and the exponent for each individual strain and under different growth conditions.

Institute of Physics
76 Portland Place, London W1B 1NT, UK
Telephone: +44 (0)20 7470 4800
Fax: +44 (0)20 7637 4266
www.iop.org/conferences

Registered Charity Number: 293851