

## CHROMATIN GENOMICS

In every eukaryotic cell, hundreds of chromatin proteins work together to control the expression of thousands of genes. Each chromatin protein interacts with many other proteins and regulates specific parts of the genome. This network of interactions is enormously complex. To gain insight into this network and the roles in gene regulation, we take a broad integrative genomics approach, using both fruit fly and mammalian cells as model systems. We conduct our studies in the living cell, in the context of the entire genome. Thereby, we aim to identify general principles that govern gene regulation by chromatin. We develop and apply new whole-genome approaches to study the structure and composition of chromatin and the mechanisms of gene regulation.

One of our major tools is our DamID technology, which enables us to generate detailed *in vivo* genomic binding maps of a large variety of chromatin proteins and transcription factors. These binding maps provide a wealth of new insights into the roles of each protein in determining chromatin structure and gene regulation, and form the basis for systematic functional analysis of gene regulation by chromatin components. We have established a DamID “pipeline” that efficiently generates full-genome binding maps (each consisting of ~385,000 data points) for a wide range of proteins in a *Drosophila* cell line. Furthermore, a unique application of DamID is the mapping of contacts of the genome with the nuclear lamina. This enables us to study the folding of chromosomes inside nuclei of *Drosophila* and mammalian cells in detail.

**Principal chromatin types in *Drosophila* cells** The diversity of chromatin composition and the distribution along chromosomes are still poorly characterized. We have generated genome-wide high-resolution binding maps of 53 broadly selected chromatin components in *Drosophila* cells. By integrative computational analysis we demonstrated that the genome is segmented into five principal chromatin types that are defined by unique yet overlapping combinations of proteins and form domains that can extend over > 100 kb. We identified a repressive chromatin type that covers about half of the genome and lacks classic heterochromatin markers. Furthermore, transcriptionally active euchromatin consists of two types that differ in molecular organization and H3K36 methylation and regulate distinct classes of genes. Finally, we provided evidence that the different chromatin types help to target DNA-binding factors to specific genomic regions. These results provide a global view of chromatin diversity and domain organization in a metazoan cell.



Group leader Bas van Steensel

**Bas van Steensel PhD** Group leader  
**Mario Amendola PhD** Post-doc  
**Guillaume Filion PhD** Post-doc  
**Jop Kind PhD** Post-doc  
**Katy Kolodziej PhD** Post-doc  
**Alexey Pindyurin PhD** Post-doc  
**Joke van Bommel MSc** PhD student  
**Dominika Bijoś MSc** PhD student  
**Ulrich Braunschweig MSc** PhD student  
**Wouter Meuleman MSc** PhD student  
**Daniel Peric-Hupkes MSc** PhD student  
**Ludo Pagie PhD** Bioinformatician  
**Wendy Talhout BAS** Technical staff  
**Arantxa Rosado BSc** Technical staff

Figure 4: Binding profiles of 53 different chromatin components in *Drosophila* cells. A region of 1Mb from chromosome 2L is shown. Each horizontal track represents the binding pattern of a single protein. Positions of genes are indicated at the bottom.

## Publications

Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B. *Systematic protein location mapping reveals five principal chromatin types in Drosophila cells*. *Cell*. 2011;143:212-224

van Steensel B, Braunschweig U, Filion GJ, Chen M, van Bommel JG, Ideker T. *Bayesian network analysis of targeting interactions in chromatin*. *Genome Res*. 2010;20:190-200

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, Reinders M, Wessels L, van Steensel B. *Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation*. *Mol Cell*. 2010;38:603-613

van Bommel JG, Pagie L, Braunschweig U, Brugman W, Meuleman W, Kerkhoven RM, van Steensel B. *The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the Drosophila genome*. *PLoS ONE* 2010

Pauli A, van Bommel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K. *A direct role for cohesin in gene regulation and ecdysone response in Drosophila salivary glands*. *Curr Biol*. 2010;20:1787-1798

van Steensel B, Dekker J. *Genomics tools for unraveling chromosome architecture*. *Nat Biotechnol*. 2010;28:1089-1095

Kind J, van Steensel B. *Genome-nuclear lamina interactions and gene regulation*. *Curr Opin Cell Biol*. 2010;22:320-325

Filion GJ, van Steensel B. *Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells*. *Nat Genet*. 2010;42:4

**Protein targeting interactions in chromatin** Chromatin proteins often direct the genomic binding pattern of other chromatin proteins, for example by recruitment or competition mechanisms. The network of such targeting interactions in chromatin is complex and still poorly understood. Together with Dr. Trey Ideker (University of California, San Diego) we implemented Bayesian Network Inference, a probabilistic computational method, to predict the targeting interactions among a broad set of chromatin components in *Drosophila* cells, based on the DamID profiles of these proteins. Experimental and computational validation confirm the overall reliability of these predictions. For example, we found that the homologous proteins HP1 and HP1c each target the heterochromatin protein HP3 to distinct sets of genes in a competitive manner. We also discovered a central role for the remodeling factor Brahma in the targeting of several DNA-binding factors, including GAGA factor, JRA and SU(VAR)3-7. Our network model provides a global view of the targeting interplay among dozens of chromatin components and provides a framework to build an increasingly refined view of chromatin.

**Chromatin Protein Discovery Project** In 2008 we have started the Chromatin Protein Discovery Project, which aims to generate genome-wide binding maps for a large set of candidate novel chromatin proteins in *Drosophila*. The candidate proteins are selected by computational predictions that take into account protein domain structure, interactions with known chromatin proteins, and likelihood of nuclear localization. For each of these we are generating full-genome, high-resolution DamID maps, which are expected to reveal many new molecular interactions and functions of the hitherto uncharacterized proteins. We have now generated informative binding maps of 40 novel proteins. Extensive analysis of these binding maps is ongoing to predict the functions and molecular interactions for each protein. This project will broaden our view of chromatin by identifying and annotating dozens of novel components.

**Genome – nuclear lamina interactions during differentiation** The three-dimensional organization of chromosomes within the nucleus and its dynamics during differentiation are largely unknown. To visualize this process in molecular detail, we generated high-resolution maps of genome-nuclear lamina interactions during subsequent differentiation of mouse embryonic stem cells via lineage-committed neural precursor cells into terminally differentiated astrocytes. This reveals that a basal chromosome architecture present in embryonic stem cells is cumulatively altered at hundreds of sites during lineage commitment and subsequent terminal differentiation. This remodeling involves both individual transcription units and multigene regions and affects many genes that determine cellular identity. Often, genes that move away from the lamina are concomitantly activated; many others, however, remain inactive yet become unlocked for activation in a next differentiation step. These results suggest that lamina-genome interactions are widely involved in the control of gene expression programs during lineage commitment and terminal differentiation.

**Nuclear lamina interactions in Drosophila** By high-resolution DamID we found that the *Drosophila* genome is also organized in discrete lamina-associated domains (LADs), which are about five times smaller than mammalian LADs but contain on average a similar number of genes. We conducted a systematic analysis of the positions of LADs relative to the binding of several insulator proteins, which are thought to act as boundaries of chromatin domains. We found that only SU(HW) binds preferentially at LAD borders and at specific positions inside LADs, while GAF, CTCF, BEAF-32 and DWG are mostly absent from these regions. By knockdown and overexpression studies we demonstrated that SU(HW) weakens genome – NL interactions through a local antagonistic effect, but we did not obtain evidence that it is essential for border formation. These results provide insights into the evolution of LAD organization and identify SU(HW) as a fine-tuner of genome – NL interactions.