Welcome to the 6th Berlin Summer Meeting, June 13 – 15, 2013

A conference organized by the Berlin Institute for Medical Systems Biology (BIMSB) at the Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch, funded by:

BMBF, Berlin

Der Senat von Berlin

Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch

Helmholtz Gemeinschaft Deutscher Forschungszentren, Bonn/Berlin

The Berlin Institute for Medical Systems Biology

Contents

Organisation & Contact

Alexandra Tschernycheff & Michaela Langer

Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch
Robert- Rössle- Str. 10
13125 Berlin

Phone: +49 30 9406 2999 / 3720
Fax: +49 30 9406 3068 / 2206

E-Mails: tschernycheff@mdc-berlin.de, langer@mdc-berlin.de

Homepage: https://www.berlinsummermeeting.org

Welcome Address

Dear colleagues and guests,

We're looking forward to the 6th Berlin Summer Meeting where we aim to discuss general principles of gene regulation “From chromatin to RNA and back”.

Within the last decade, RNA biology has become a major focus in life sciences, with the discovery of novel families of non-coding RNA molecules that regulate cellular processes. Understanding the mechanisms of gene regulation at the level of chromatin and how RNA themselves play important roles in nuclear and cellular processes are exciting developments in this fast moving field.

The Berlin Summer Meeting traditionally invites leading scientists to present and discuss experimental and computational approaches, across a variety of biological systems, in an open atmosphere. We would like to foster intellectual exchange between various areas of biology and stimulating interactions of leading scientists and young researchers in Berlin and beyond.

The Berlin Summer Meeting is coordinated by the Berlin Institute for Medical Systems Biology (BIMSB) at the MDC, Berlin-Buch, Germany. We thank our sponsors and funding agencies for financial support, and members of the BIMSB-MDC, Max Planck Institute for Molecular Genetics, and New York University for organizing this promising scientific meeting.

With best regards,

Sevinc Ercan
Ana Pombo
Martin Vingron
Uwe Ohler
Nikolaus Rajewsky
Baris Tursun

(The Organisation Committee 2013)

Program

Thursday, 13 June 2013

08:30 - Registration + Coffee

WELCOME

09:00–09:10 **Ana Pombo**
Welcome to the 6th Berlin Summer Meeting

Session I. Chromatin processes on gene promoters and genes
(Chair: Uwe Ohler)

09:10–10:10 **Naama Barkai**
Promoter nucleosome organization and the evolution of gene expression

10:10–11:10 **Baris Tursun**
Chromatin puts on the brake: Inhibition of transcription factor-induced direct cell conversion

11:10–11:40 Coffee Break

11:40–12:40 **Sarah Teichmann**
Gene expression genomics in T-cells

12:40-13:00 **Leonie Ringrose**
Non-coding strand switching:A molecular gearbox for epigenetic regulatory elements

FOTOSESSION FOR SPEAKERS AND PARTICIPANTS

13:00 -14:30 Lunch break & poster session

Session II. Chromatin structure and organization
(Chair: Sevinc Ercan)

14:30–15:30	**Bas van Steensel** *Probing position effects in high-throughput mode*
15:30–16:30	**Amos Tanay** *The multiple scales of modular chromosome organization*
16:30–16:50	Coffee Break
16:50–17:10	**Luca Georgetti** *Structure meets function at the mouse X chromosome inactivation center*
17:10–18:10	**Ana Pombo** *From epigenetic mapping to polymer modeling: 3D folding of the HoxB locus in mouse ES cells*
19:00–21:00	Speakers Dinner

Friday, 14 June 2013

09:00 - Registration + Coffee

Session III. Transcription and co-transcriptional processes (Chair: Ana Pombo)

09:10–10:10 **Steve Buratowski**
Co-transcriptional histone modification and RNA processing

10:10–10:30 **Stefan Pinter**
Mapping of Xist RNA and Polycomb Repressive Complex 2 across the inactive X

10:30–10:50 **Stefan Muljo**
Jarid2, a Polycomb member and target of miR-155, regulates Th17 differentiation and Treg cell homeostasis

10:50–11:20 Coffee Break

11:20–12:20 **Nick Proudfoot**
Mechanisms of transcriptional termination in mammalian protein coding genes

12:20–12:40 **Michael Ziller**
The dynamic DNA methylation landscape of the human genome

12:40–14:30 Lunch break & poster session

Session IV. Non-coding RNAs and chromatin (Chair: Martin Vingron)

14:30–15:30 **Denise Barlow**
Gene silencing by macro lncRNA transcriptional interference

15:30–15:50 **Markus Kretz**
Control of epidermal tissue differentiation by long non-coding RNAs

15:50–16:20	Coffee Break
16:20–17:20	**Claes Wahlestedt** *Non-coding RNAs in Chromatin Regulation*
17:20–17:40	**Fabio Mohn** *The Rhino-Deadlock-Cutoff complex defines piRNA clusters in Drosophila*
17:40–18:00	**Alena Shkumatava** *lincRNA Cyrano hijacks miRNA/Argonaute complex for a non-conventional function indispensable to early development*
18:30–21:30	Sightseeing Tour

Saturday, 15 June 2013

9:00 - Registration + Coffee

Session V. Large-scale/chromosome-wide chromatin regulation
(Chair: Baris Tursun)

09:10–10:10	**Asifa Akhtar** *Epigenetic regulation by MSL proteins*
10:10–10:30	**Edda Schulz** *Evidence for X-chromosome gene dosage as a developmental checkpoint in females*
10:30–11:00	Coffee Break
11:00–12:00	**Sevinc Ercan** *Chromosomal recruitment and spreading of the condensin complexes in C. elegans*
12:00–13:00	**Martin Vingron** *Can we study chromation through linear models and graphical networks?*
13:00–13:10	**Uwe Ohler** *Closing remarks*
13:10	Lunch

Speaker Abstracts

(alphabetical order)

Epigenetic regulation by MSL proteins

Asifa Akhtar[1]

[1]*MPI of Immunobiology & Epigenetics, Freiburg, Germany*

Our lab is studying the chromatin and epigenetic mechanisms underlying X-chromosome specific gene regulation using Drosophila dosage compensation as a model system. Dosage compensation mechanisms regulate the expression of X-linked genes. In Drosophila this process leads to the transcriptional upregulation of the single male X chromosome and is regulated by a ribonucleoprotein male specific lethal (MSL) complex. The MSL complex localizes on hundreds of sites on the male X chromosome and is involved in hyperacetylation of histone H4 at lysine 16 by MOF, histone acetyltransferase. Interestingly, this complex is also conserved in mammals and appears to be involved in regulating diverse cellular processes. The recent progress of our work will be presented.

Promoter nucleosome organization and the evolution of gene expression

Naama Barkai[1]

[1]*Weizmann Institute of Science, Rehovot, Israel*

How readily do genes evolve in expression? What are the principle mutation types that drive expression evolution? Does chromatin play a role in the capacity to diverge in expression? I will discuss results addressing those questions, arguing for differential divergence strategies employed by different genes. Those strategies may be encoded by the organization of promoter nucleosomes, and could therefore be tuned by the cells, maintaining some genes robust while enabling rapid evolution of other genes.

Gene silencing by macro lncRNA transcriptional interference

Denise Barlow[1]

[1]*University of Vienna, Vienna, Austria*

In mammals, imprinted genes showing parental-specific expression tend to cluster with unusually long and mostly unspliced long non-coding (lnc) RNAs. In three studied imprinted clusters, the associated lncRNA has been shown to induce parental-specific silencing in cis of small groups of genes. Thus imprinted lncRNAs are an important model to understand how lncRNAs silence genes in cis and how this silencing can affect multiple clustered genes. Global transcriptome analyses in recent years have revealed that lncRNAs are not a unique feature of imprinted regions, as many thousands are transcribed throughout the mammalian genome. Imprinted lncRNAs that silence genes in cis, however, possess atypical features – such as inefficient splicing, extreme length, high repeat content, lack of transcript conservation and short half-life – that together indicate their transcription is more important than their RNA product. Due to the greater abundance of its long unspliced form, this type of lncRNA is termed a 'macro' lncRNA. In the Igf2r cluster on mouse chromosome 17, the 118kb Airn macro lncRNA overlaps and silences Igf2r promoter on the paternal chromosome, in all tissues where they are co-expressed. We have recently shown that Igf2r silencing only requires transcription of unspliced Airn across the Igf2r promoter. We are now using mouse and human 'Imprinted Region Tiling Arrays' combined with optimized RNA-seq pipelines to identify novel macro lncRNAs similar to Airn, which show additional features such as high tissue-specific expressivity, developmental or cancer-specific regulation.

Co-Transcriptional Histone Modification and RNA Processing

Stephen Buratowski[1]

[1]*Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, USA*

The C-terminal domain (CTD) of RNA polymerase II (pol II) subunit Rpb1 undergoes a dynamic series of phosphorylations during transcription. Within the repeating sequence YSPTSPS, Ser5 and Ser7 are phosphorylated by basal transcription factor TFIIH near the promoter. As the polymerase moves into elongation, levels of these phosphorylations decline while Ser2 becomes progressively more phosphorylated by one more kinases (including P-TEFb/Cdk9, Bur1, and Ctk1). Additional phosphorylations can also occur at Ser7, Thr4, and Tyr1. These phosphorylations recruit various enzymes to transcribing pol II at appropriate times. For example, capping enzymes bind Ser5P and some polyadenylation factors bind Ser2P. CTD modifications also influence whether RNA polymerase terminates via the early Nrd1/Nab3/Sen1 pathway or instead uses the polyadenylation/torpedo pathway.

The CTD phosphorylation cycle also directs the histone methyltransferases Set1 and Set2 to respectively methylate H3K4 near 5' ends and H3K36 in 3' regions of genes. My lab has been studying the roles of these methylations. H3K36me2/3 recruits the Rpd3S histone deacetylase complex (HDAC) to downstream-transcribed regions, where deacetylation inhibits transcription initiation from cryptic promoters. This pathway also inhibits transcription elongation. H3K4me2 targets a second HDAC, the Set3 complex, to 5'-middle regions of transcribed genes. We find that this pathway can also suppress internal promoters, but is also used in interesting ways to modulate gene expression. We find that genes affected by Set3 usually have overlapping non-coding transcription units and that these affect the kinetics of gene induction and repression. Finally, H3K4me3 near 5' ends is thought to be important for keeping promoter regions in a highly acetylated state with rapid nucleosome turnover. Our recent results suggest that recognition of H3K4me3 by the ING/Yng PHD finger proteins mediates a balance of acetylation/deacetylation. This balance seems to control early elongation and affects the choice between pausing/terminating and full elongation. Chromatin-mediated effects on transcription elongation are also likely to affect co-transcriptional mRNA processing events such as intron splicing.

Chromosomal recruitment and spreading of the condensin complexes in C. elegans

Sevinc Ercan[1]

[1]*New York University, Center for Genomics and Systems Biology New York, USA*

Condensins are multi-subunit protein complexes that are essential for chromosome condensation and play key roles in transcription regulation and DNA repair. Most organisms contain two types of condensins, which perform different functions and localize to different chromosomal regions. C. elegans contains a third condensin that is recruited specifically to the X chromosome and represses transcription approximately by half in XX hermaphrodites to equalize X-linked transcript levels to that of XO males. Dosage compensation condensin is recruited to the X chromosome at ~100 sites that are specified in part by a DNA sequence motif that is different for different condensin types. After recruitment, the complex spreads onto chromatin represses genes at the region of spreading. The mechanism of spreading and transcriptional repression involves the mitosis-enriched histone modification H4K20me1. Chromosomal targeting mechanisms for many important structural proteins such as condensins remain unclear. Our work is a step towards understanding how condensins bind to their chromosomal sites.

Structure meets function at the mouse X chromosome inactivation center

Luca Giorgetti[1], Guido Tiana[2], Elphège Nora[1], Tristan Piolot[1], Edith Heard[1]

[1]Institut Curie, Genetics and Developmental Biology (INSERM U934 / CNRS UMR3215), Paris, France, [2]Università di Milano, INFN and Dipartimento di Fisica, Milano, Italy

Characterizing the folding principles of mammalian chromosomes is of capital importance to understand the complexity of gene expression regulation, particularly during the major transcriptional changes occurring in development. This may help elucidating the mechanisms by which regulatory elements contact gene promoters (i.e. by looping out intervening DNA), understand what is the cell-to-cell variability of these interactions and how it does reflect transcriptional variability.
We analyzed a 4.5 Mb region of the X chromosome that includes the X-inactivation center by Chromatin Conformation Capture Carbon-Copy (5C), in order to gain insights into how chromatin structure is organized during early mouse embryonic stem cell (ESC) differentiation. We uncovered that chromatin is organized into Topologically Associating Domains (TADs), within which genomic elements preferentially interact. To fully reconstruct the statistical repertoire of chromatin conformations that give rise to these domains, we have used a combination of Monte Carlo simulations of a polymer model, high-resolution DNA FISH and quantitative RNA FISH. We show that in the TAD that contains the Tsix ncRNA (a master regulator of X-chromosome inactivation), enhancer-promoter contacts take place in a subset of cells where the whole domain is compacted, rather than resulting from stable DNA loops. In these cells, the probability of transcribing a gene is higher than in cells where the domain is in an elongated conformation. We thus show a correlation between the spatial proximity of a promoter with its cis-regulatory elements and their transcriptional activity at the single cell level.

Control of epidermal tissue differentiation by long non-coding RNAs

Markus Kretz[1, 2], Zurab Siprashvili[2], Ci Chu[2], Dan Webster[2], Arjun Raj[3], John Rinn[4], Howard Chang[2, 5], Paul Khavari[2]

[1]Present address: Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, Regensburg, Germany, [2]The Program in Epithelial Biology, Stanford University School of Medicine, Stanford, USA, [3]Department of Bioengineering, University of Pennsylvania, Philadelphia, USA, [4]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, USA, [5]Howard Hughes Medical Institute, Stanford, USA

The human genome encodes several thousand long non-protein coding transcripts >200 base pairs in length. Although this large newly discovered portion of the human transcriptome is still poorly characterized to date, a number of functional studies strongly indicate that long non-coding RNAs (lncRNAs) are highly functionally relevant and not mere transcriptional waste. To detect lncRNAs with potential function in epidermal homeostasis, we performed high-throughput full transcriptome sequencing with differentiated and progenitor human keratinocyte populations using the Illumina paired-end HiSeq platform with a read length of 101 nucleotides at >100 million mapped reads per sample. TINCR, a 3.7kb lncRNA identified in this screen, is required for induction of key differentiation genes in epidermal tissue, including genes mutated in human skin diseases characterized by disrupted epidermal barrier formation. High-throughput analyses of TINCR RNA- and protein- interactomes revealed TINCR interaction with differentiation mRNAs as well as the Staufen1 protein. TINCR-mRNA interaction occurs through a 25-nucleotide motif strongly enriched in interacting mRNAs and required for TINCR binding. Deficiency of TINCR-associated Staufen1 in epidermal tissue recapitulated the impaired differentiation seen with TINCR loss. TINCR, together with STAU1, seems to stabilize a subset of mRNAs required for epidermal differentiation. TINCR is thus a newly characterized lncRNA required for somatic tissue differentiation.

The Rhino-Deadlock-Cutoff complex defines piRNA clusters in *Drosophila*

Fabio Mohn[1], *Grzegorz Sienski*[1], *Julius Brennecke*[1]

[1]*Institute of Molecular Biotechnology, Vienna, Austria*

Transposable elements (TE) constitute large fractions of eukaryotic genomes and if unleashed cause mutations and massive DNA damage. The piRNA pathway silences TEs in the animal germline to prevent their further spreading. 23-30 nt long piRNAs serve as guides to recruit a RISC-like complex to TEs and silence them both at the transcriptional and post-transcriptional level. However, it is not known how germ cells produce the comprehensive small RNA pool for efficient defense against TEs. Previous studies have implicated Rhino and Cutoff in piRNA generation from so-called piRNA clusters, which are genomic regions enriched in TE fragments that serve as a repository to generate complementary piRNAs. Here we identify Deadlock as a third factor essential for piRNA cluster regulation and show that Rhino, Deadlock and Cutoff (RDC) form an inter-dependent complex. Yeast-two-hybrid interactions suggest that Deadlock is the bridging factor that links the putative chromatin binding activity of Rhino to the Rai-1 related protein Cutoff. In genome-wide studies we find this complex specifically localizing to piRNA clusters. Upon perturbation of any of the three factors we observe loss of localization of the RDC to chromatin and a coinciding loss of transcripts form piRNA clusters. This consequently depletes piRNAs and leads to massive upregulation of TEs, while protein coding genes are unaffected. Interestingly, we find a number of small euchromatic Rhino domains outside of major clusters. These loci give rise to piRNAs in an RDC dependent manner. Ongoing work is focusing on how the RDC is specifically recruited to its target sites and we will discuss potential novel insights at the meeting. Together, our results suggest that the RDC is required for efficient transcription of piRNA clusters. Further, we provide a novel conceptual framework of piRNA cluster definition and regulation, which might help to explain the rapid evolutionary adaptation towards newly invading TEs.

Jarid2, a Polycomb member and target of miR-155, regulates Th17 differentiation and Treg cell homeostasis

Stefan Muljo[1]

[1]*NIH, NIAID, Laboratory of Immunology, Bethesda, USA*

Great progress has been made in understanding the genome-wide roles of transcription factors in programming cellular differentiation. However, the roles of microRNAs and chromatin regulators have not been integrated into current models of genetic regulatory networks that orchestrate the required programs of gene expression. In this study, we outline an updated Th17 network that incorporates the known transcription factors, a key miRNA, histone H3K27 tri-methylation and targets thereof. Previously, miR-155 has been implicated in regulating T helper cell differentiation and regulatory T (Treg) cell fitness. By analyzing mixed bone marrow chimeras, we revealed a cell autonomous defect in Th17 cytokine expression by miR-155 deficient T helper (Th) cells in addition to a Treg defect. This phenotype is consistent with our finding that miR-155 is highly expressed in ex vivo mouse and human Th17 cultures. An analysis of the Th17 transcriptional network explains why miR-155 is highly induced in this program. Importantly, we provide evidence that Jarid2 is a consequential target of miR-155 in Th17 and Treg cells. By ChIP-seq and RNA-seq, we found that the increased expression of Jarid2 in miR-155-/- Th17 cultures led to augmented H3K27 tri-methylation catalyzed by the Polycomb Repressive Complex 2 (PRC2) and transcriptional repression of genes in the Th17 program that included *Il1r1* and *Il22*. In Jarid2-deficient Th cells, we found that PRC2 was no longer recruited to specific genomic sites, and as a result there were reduced H3K27 tri-methylation and in some cases derepression of genes. Notably, the defects in Th17 differentiation and Treg cell homeostasis in miR-155 knock-out mice can be rescued upon conditional deletion of *Jarid2*. Thus, one main function of miR-155 within the genetic regulatory networks of Th17 and Treg cells is to blunt epigenetic silencing through the targeting of Jarid2, a DNA-binding factor that is required for recruiting PRC2 to specific genomic locations.

This work was supported by the Intramural Research Program of the NIAID, NIH. The full list of authors is as follows: Thelma M. Escobar, Chrysi Kanellopoulou, David G. Kugler, Gokhul Kilaru, Cuong K. Nguyen, Ravikiran K. Bhairavabhotla, Daniel Northrup, Keji Zhao, Alan Sher, Dragana Jankovic, Jinfang Zhu, and Stefan A. Muljo.

From epigenetic mapping to polymer modeling: 3D folding of the HoxB locus in mouse ES cells

Sheila Q. Xie[2,5], Mariano Barbieri[3,5], James Fraser[4], Ines de Santiago[2], Emily Brookes[2], Josèe Dostie[4], Mario Nicodemi[3,5], Ana Pombo[1,2,5]

[1]Berlin Institute for Medical Systems Biology, Max Delbrueck Centre, Berlin, Germany, [2]MRC, Clinical Sciences Centre, Imperial College London, London, UK, [3]Department of Physics, University of Naples 'Federico II', Naples, Italy, [4]Department of Biochemistry, McGill University, Montréal, Canada, [5]contributed equally

Long-range chromatin interactions have been identified as major regulators of gene expression and chromosome organization. To investigate the mechanisms of chromatin folding and interaction, we used as model system a 1Mb genomic region containing active genes and a cluster of HoxB genes, which are repressed by Polycomb in embryonic stem (ES) cells.

We used a simple polymer model, the Strings & Binders Switch (SBS) model (Nicodemi et al. 2008, Nicodemi and Prisco, 2009, Barbieri et al. 2012), to determine the different scenarios of 3D folding of the HoxB genomic region, in the presence of active transcription factories and Polycomb bodies, separately or together. Active genes are known to interact at active transcription factories, whereas Polycomb-repressed genes associate with poised transcription factories in mouse ES cells (Ferrai et al. 2010, Brookes et al. 2012).

Using epigenetic mapping, we classified all the genes across the HoxB locus (1 Mb genomic region) according to their status in ES cells: active, Polycomb-repressed or inactive. Using high-resolution single-cell imaging by cryo-FISH, we find agreement with one of the scenarios predicted by the model, in which both active and poised transcription factories contribute to specialized 3D folding of the locus. We also identify novel examples of gene co-association at the same transcription factory, and begin to explore the effects of genomic context on the 3D interaction behavior of single genes.

Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M. (2012) Complexity of chromatin folding is captured by the Strings & Binders Switch model. **Proc. Natl. Acad. Sci. USA** 109, 16173-8
Brookes E, de Santiago I, Hebenstreit D, Morris KJ, Carroll T, Xie SQ, Stock JK, Heidemann M, Eick D, Nozaki N, Kimura H, Ragoussis J, Teichmann SA, Pombo A (2012) Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ES cells. **Cell Stem Cell**10, 157-70.
Ferrai C, Xie SQ, Luraghi R, Munari D, Ramirez F, Branco MR, Pombo A1, Crippa

MP1 (2010) Poised transcription factories prime silent uPA genes prior to activation. **PLoS Biol.** 8, e1000270.

Nicodemi M, Panning B, Prisco A (2008) A thermodynamic switch for chromosome colocalization. **Genetics** 179, 717-721.

Nicodemi M, Prisco A (2009) Thermodynamic pathways to genome spatial organization in the cell nucleus. **Biophys J**. 96, 2168-2177.

Mechanisms of transcriptional termination in mammalian protein coding genes

Nick J. Proudfoot[1]

[1]*Sir William Dunn School of Pathology, University of Oxford, Oxford, UK*

My lab focuses on the molecular mechanism of RNA polymerase II (Pol II) termination. We have identified two alternative modes of termination in mammalian protein coding genes.

1) A sub-class of genes possess G-rich transcriptional pause sites, closely following their polyA signals (PAS). These act to slow down elongating Pol II following the PAS so that cleavage at the PAS followed by Xrn2 exonucleolytic degradation elicits termination (polymerase release from the chromatin template). These G rich pause elements form RNA:DNA hybrid structures (R-loops) which require Senataxin resolution for Pol II termination (Skourti-Stathaki et al Mol Cell 2011). We now show that R-loops also promote localised antisense transcription which in turn recruits H3K9me2 heterochromatin marks via an RNAi-like mechanism to the termination region. This reinforces the termination process.

2) Alternatively the majority of Pol II genes do not terminate by a pausing mechanism but instead transcription proceeds into the gene 3' flanking region. Transcription of a co-transcriptional cleavage CoTC) element results in transcript breakage which then allows Xrn2 access to the transcript and rapid Pol II release into the nucleoplasm. Cleavage at the upstream PAS subsequently occurs on the released Pol II in the nucleoplasm. We have recently developed technology based on iCLIP to show that CoTC mediated Pol II termination occurs for many mammalian genes (Nojima et al Cell Rep. 2013) and may have significant advantage in facilitating gene expression levels.

Non-coding strand switching: a molecular gearbox for epigenetic regulatory elements

Leonie Ringrose[1]

[1]*IMBA Institute of Molecular Biotechnology, Vienna, Austria*

In flies and vertebrates, large portions of the genome are transcribed into long non coding (nc) RNAs, many of which coincide with binding sites for the Polycomb (PcG) and Trithorax (TrxG) groups of proteins. Polycomb/Trithorax response elements (PRE/TREs) can switch their function reversibly between activation and silencing, by mechanisms that are poorly understood. Here we show that a switch in strand specific nc transcription from the *Drosophila vestigial (vg)* PRE/TRE switches the function of the element. Embryonic transcription of the PRE/TRE reverse strand leads to activation, whereas larval transcription of the PRE/TRE forward strand causes *vg* gene repression. Furthermore, high ectopic transcription of specific strands leads to trans-activation or trans-silencing at the endogenous *vg* locus, demonstrating a specific role for each transcript in gene regulation. 3D FISH indicates a role for the forward strand in silencing by promoting higher order interactions. In contrast, the reverse strand binds highly specifically to the PcG protein Enhancer of Zeste, suggesting activation by removal or inhibition of PcG proteins. We name this element a "GEAR-BOX" (Gene Expression Alternating RNA) element. Finally, genome wide analysis shows that a high proportion of PcG binding sites in mouse and fly, including those of a mouse vg functional homolog, show tissue specific strand switching of ncRNAs. Thus, this work identifies a novel and potentially widespread class of PRE/TREs, that switch their function by switching between forward and reverse strand ncRNA transcription.

Evidence for X-chromosome gene dosage as a developmental checkpoint in females

Edda Schulz[1], Johannes Meisig[2], Nils Blüthgen[2], Edith Heard[1]

[1]Institut Curie, Paris, France, [2]Charité Berlin, Berlin, Germany

During early development, both X chromosomes are transiently active in female mouse embryos in the inner cell mass (ICM) of the blastocyst. Subsequently, one X is silenced to equalize X gene dosage between the sexes through the process of X chromosome inactivation (XCI). To understand how a double dose of X-linked genes contributes to sex-specific differences, we use embryonic stem cells (ESCs), which are derived from the ICM, as a model system. Transcriptome profiling of ESCs with distinct sex chromosome composition (XX, XY, XO) revealed that a large group of X-linked and autosomal genes were expressed at elevated levels in cells with two X chromosomes. We could attribute this effect to an inhibition of the Fgf/Erk signaling pathway in XX ESCs, which leads to reduced levels of the DNA methyltransferases Dnmt3a/b and thereby results in de-repression of DNA methylation sensitive genes specifically in female cells. These observations would predict that X dosage should also impact on ES cell differentiation, where Fgf/Erk signaling is known to play a key role. Comparison of the global transcriptome dynamics during early differentiation indeed revealed a delay in the down-regulation of stem cell genes in cells with two X chromosomes, in part due to reduced de novo methylation of their promoters. Finally, we have investigated, whether X inactivation, which is initiated during early differentiation will overcome the differentiation block in female ESCs. To this end, single-cell RT-qPCR revealed that cells that had undergone XCI, preferentially down-regulated stem cell genes. Moreover, ectopically induced premature initiation of XCI increased the speed of differentiation. Based on these results, we propose that the presence of two active X chromosomes blocks differentiation through inhibition of the Fgf/Erk signaling pathway until X inactivation has occurred, thereby constituting a developmental X-inactivation checkpoint.

lincRNA *cyrano* hijacks miRNA/Argonaute complex for a non-conventional function indispensable to early development

Alena Shkumatava[1], *Igor Ulitsky* [2], *David Bartel* [2]

[1]*Institut Curie, Developmental Biology and Genetics Unit, Paris, France,* [2]*Whitehead Institut/MIT, Cambridge, USA*

Thousands of long intervening noncoding RNAs (lincRNAs) have been recently identified in human and mouse. To better understand functions of these enigmatic RNAs, we identified more than 550 lincRNAs in zebrafish, an established vertebrate model for development (Ulitsky*, Shkumatava* et al., 2011). Although zebrafish lincRNAs shared many characteristics with mammalian lincRNAs, only 29 had detectable sequence similarity with putative mammalian counterparts, typically restricted to a single short region of high conservation. We showed that morpholinos targeting either conserved regions or splice junctions of lincRNAs caused similar developmental abnormalities in zebrafish embryos including brain morphogenesis defects and neurogenesis. The developmental defects could be rescued by either the mature zebrafish lincRNA or its human or mouse othologs. Our rescue experiments showed that despite limited sequence conservation, lincRNAs have conserved functions in zebrafish and mammals.

We recognized that the conserved region of one of the lincRNAs that we called *cyrano* contains a highly complementary, near perfect miR-7 site that is bound by Argonaute proteins and is highly conserved in all examined vertebrates. We show that a near perfect pairing to a miRNA is indispensable to the *cyrano* function in vivo. Although miR-7 appears to modestly regulate *cyrano* levels, it cannot explain the unusual high complementarity of the miRNA—lincRNA pairing. Loss of function and rescue assays in zebrafish embryos show that the interaction between miR-7 and *cyrano* differs from the canonical miRNA—target regulation suggesting that lincRNA—miRNA complex has an additional, novel function indispensible for normal development.

Mapping of Xist RNA and Polycomb Repressive Complex 2 across the inactive X

Pinter Stefan[1], Matthew Simon[2], Rui Fang[2], Kavitha Sarma[1], Sarah Bowman[3], Robert Kingston[3], Jeannie Lee[1]

[1]Howard Hughes Medical Institute, MGH/HMS, Stanford, USA, [2]Yale University, New Haven, USA, [3]Massachusetts General Hospital, Harvard Medical School, Boston, USA

The Xist long noncoding RNA (lncRNA) is essential for X-chromosome inactivation (XCI), the process by which mammals compensate for unequal numbers of sex chromosomes. During XCI, Xist coats the future inactive X (Xi) and recruits Polycomb repressive complex 2 (PRC2) to silence gene expression. Because Xist is a lncRNA prototype and model cis-regulatory RNA, where Xist RNA binds on Xi has been of great interest and is central to understanding the mechanistic relationship between Xist and PRC2. Here we generate allele-specific and high-resolution maps of Xist binding using CHART-seq and demonstrate broad but non-random enrichment of Xist RNA during XCI. Significantly, Xist co-localizes with and is proportional to PRC2 and H3 lysine 27 trimethylation (H3K27me3) along the Xi. Interestingly, Xist RNA targets gene-rich regions, preferentially localizes in inter-lamin domains, and is enriched at boundaries of regions that escape XCI. On the other hand, Xist is depleted from genes that escape XCI and is anti-correlated with LINE repeats and lamin-associated domains. These data invoke possible models for the spreading of XCI based on chromosomal organization and underscore the role of Xist in recruiting PRC2 during XCI.

Probing position effects in high-throughput mode.

Waseem Akhtar[1], Johann de Jong[3], Alexey Pindyurin[2], Lodewyk Wessels[3], Maarten van Lohuizen[1], Bas van Steensel[2]

[1]Division of Molecular Genetics, [2]Division of Gene Regulation, [3]Division of Molecular Biology Netherlands Cancer Institute, Amsterdam, The Netherlands

Reporter genes integrated into the genome are a powerful tool to reveal effects of regulatory elements and local chromatin context on gene expression. However, so far such reporter assays have been of low throughput. We developed a multiplexing approach for the parallel monitoring of transcriptional activity of thousands of randomly integrated reporters. More than 27,000 reporter integrations in mouse embryonic stem cells, obtained with two different promoters, show ~1,000-fold variation in expression levels. Analysis of these data reveals how local chromatin context contributes to this variation. The multiplexed reporter assay thus bridges the gap between single-locus functional/mechanistic studies and genome-wide descriptive approaches. Moreover, the assay is highly flexible in design and can be modified to query a wide range of aspects of DNA and RNA metabolism.

Chromatin puts on the brake: Inhibition of transcription factor- induced direct cell conversion

Baris Tursun[1]

[1]*MDC, Berlin, Germany*

Studying cell type conversion provides insight into mechanisms that maintain and protect specific cellular identities. Knowledge about such mechanisms has the potential to facilitate cell fate reprogramming strategies thereby supporting the development of future biomedical approaches such a tissue replacement therapies. The direct reprogramming (DR) of cells to a different cell type by ectopically expressing cell fate-inducing transcription factors (TFs) is a promising strategy to generate required tissues. DR circumvents unsafe proliferative pluripotent cell stages and *in vitro* procedures. Hitherto, DR is successful in only a few cell types and it is not well understood why most cells are refractory to DR. Recent studies provide evidence that inhibitory mechanisms play an important role in restricting cellular reprogramming. Such inhibitory mechanisms are often conveyed by factors that regulate accessibility to chromatin by, for instance, modifying histones and/or the chromatin structure. Yet, it is not well understood which chromatin-regulating factors mediate inhibition of cell fate reprogramming and importantly, whether they act the same way in different tissue types.

Our group aims to understand mechanisms that restrict DR by identifying factors involved in inhibiting direct cell fate conversion of different tissue types. We are using *C. elegans* as an *in vivo* model system and apply large-scale forward and reverse genetic screenings with high-throughput techniques. We have previously identified members of the PRC2 (Polycomb Repressive Complex 2) and histone chaperone complexes in *C. elegans* as inhibitors of converting germ cells into specific neurons or muscle cells. Our preliminary results from ongoing genetic screenings indicate that in distinct somatic tissue-types different factors and signaling pathways might be involved in order to maintain cellular identity and counteract TF-induced direct cell conversion.

Can we study chromatin through linear models and graphical networks?

Martin Vingron[1]

[1]*MPI for Molecular Genetics, Berlin, Germany*

Using linear models, one can see that histone modification patterns in promoter regions are predictive of the gene expression level of the target gene. But how are the modifications correlated with each other? How do they influence each other? In this talk we will show how to construct networks of associations between epigenetic marks. We utilize partial correlation in order to highlight direct interactions rather than consequential ones. Such networks are undirected, but by combining them with biological knowledge one can speculate about information flow as represented in the epigenetic networks.

Noncoding RNAs in Chromatin Regulation

Claes Wahlestedt[1]

[1]*University of Miami Miller School of Medicine, Miami, USA*

Much of the mammalian genome is transcribed into long and small noncoding RNAs of different categories. This lecture will primarily be concerned with long noncoding RNAs which regulate gene expression through several distinct mechanisms involving chromatin regulator protein complexes such as PRC2. Notably, inhibition/perturbation of endogenous natural antisense transcripts by AntagoNATs, in vitro or in vivo, reveals discordant regulation and results in locus specific up-regulation of conventional (protein-coding) gene expression. Implications for small molecule chromatin regulator enzyme inhibitors will also be discussed.

The dynamic DNA methylation landscape of the human genome

Michael Ziller[1]

[1]*The Harvard Clinical and Translational Science Center, Cambridge, USA*

DNA methylation is a defining feature of mammalian cellular identity and essential for normal development. Most cell types, except germ cells and pre-implantation embryos, display relatively stable DNA methylation patterns with 70-80% of all CpGs being methylated. Despite recent advances we still have a too limited understanding of when, where and how many CpGs participate in genomic regulation. Here we report the in depth analysis of 42 whole genome bisulfite sequencing (WGBS) data sets across 30 diverse human cell and tissue types. We observe dynamic regulation for only 21.8% of autosomal CpGs within a normal developmental context, a majority of which are distal to transcription start sites. These variable CpGs co-localize with gene regulatory elements, particularly enhancers and transcription factor binding sites (TFBS), which allow identification of key lineage specific regulators. Differentially methylated regions (DMRs) exhibit elevated SNP frequency and often harbor SNPs associated with cell type related diseases as determined by GWAS. Comparison of this class of dynamic CpGs to several abnormal contexts reveals that low complexity regions are preferential targets of the additional DNAme changes in cancer, Alzheimer and long-term cell culture. The results also highlight the general inefficiency of WGBS as 70-80% of the sequencing reads across these data sets provided little or no relevant information regarding CpG methylation. To further demonstrate the utility of our identified set of DMRs, we classify unknown samples and identify representative signature regions that recapitulate major DNA methylation dynamics. In summary, our results show that although in theory every CpG can change its methylation state, only a fraction does so as part of coordinated regulatory programs. As a result the identified regions will help guide novel reduced representation approaches to capture this informative fraction of CpGs more effectively as well as further pinpoint putative regulatory elements.

Poster Abstracts

(alphabetical order)

P1: The class IIa HDACs interactome: toward the identification of new functions in RNA processing

Cécile Detiffe[1], Xavier Rambout[1], Majid Cherkaoui[1], Jean-Claude Twizere[1], Franck Dequiedt[1]

[1]Laboratory of Protein Signalling and Interactions (PSI), GIGA-R (B34), University of Liege, Belgium

By modulating the acetylation level of histones, histone deacetylases (HDACs) are enzymes playing crucial functions in the control of gene expression. During the last decade, HDACs have emerged as key transcriptional regulators of major cellular processes such as cell division, apoptosis and differentiation. As such, HDACs are considered as attractive therapeutic targets for cancer and many efforts are being made in order to find therapeutic molecules modulating their activities. While HDACs inhibitors have generated a lot of hope for their antitumoral activities, these molecules have often dramatic side-effects. It is suspected that these side effects could be related to still-unknown functions of HDACs. In addition to histones, HDACs have many non-histone substrates, which may relate to a potential role for these enzymes outside of gene transcriptional regulation.
The goal of this project is to use a global, unbiased approach to identify new functions of HDACs that will help developing more efficient and specific HDAC-based antitumoral therapies.
Using a high-throughput yeast-two hybrid (Y2H) approach, we have undertaken the large scale identification of class IIa HDACs interacting partners (interactome) and established the first comprehensive interactomic map of class IIa HDACs that includes 84 class IIa HDACs protein partners. Our method uncovered 72 totally new class IIa interactors. Clusterization using the Network Analysis Tool box (NeAT) divided the class IIa HDACs interactome into 20 functional sub-networks. This led to the identification of unexpected roles for class IIa HDACs. In particular, a two-level clusterization approach revealed that class IIa HDACs are connected to several aspects of mRNA processing. Until now, we focus on 2 class IIa HDAC partners, RBPMS (RNA binding protein with multiple splicing) and RBFOX1 (RNA-binding Fox-1 homolog) that are implicated in the alternative splicing and degradation.

P2: Study of molecular mechanisms of DNA Damage induced alternative splicing: implication in cancer therapy.

Adeline Deward[1], Yves Delforge[2], Maude Gabriel[2], Jaques Piette[1], Alain Colige[2], Charles Lambert[2], Yvette Habraken[1]

[1]Laboratory of Virology and Immunology, GIGA-R, Signal Transduction Unit, University of Liège, Liège, Belgium, [2]Laboratory of conjonctive tissues biology. Department of biomedical and preclinical sciences, University of Liège, Liège, Belgium.

DNA damages have various consequences on the cell metabolism among which alternative splicing (AS). AS is an essential mechanism that is implicated in number of processes from development to genetic diseases. However, little is known about the underlying mechanisms leading to the induction of AS after a genotoxic stress and the biological role of this one. Constitutive and alternative splicing are regulated by several mechanisms upon which transcription rate, histone modifications or re-localization and post-transcriptional modifications of splicing factors. In this project, DNA lesions are inflicted by genotoxic compounds presently used in cancer therapy. We performed a deep-sequencing analysis of MCF7 cells treated with cis-platin. More than 600 genes were found to be alternatively spliced after treatment. We selected 11 genes undergoing AS for validation and confirmed the expected AS events by RT-PCR in 8 of them. Numerous studies have shown implication of kinases activated during genotoxic stress (e.g. ATM, ATR, JNK, p38 …) in splicing factors phosphorylation. We assessed the role of these kinases in the induction of AS after genotoxic stress. First, we inhibited and/or made siRNA against ATM, ATR and DNA-PK, three major kinases initiating the DNA damages responses. No difference in AS induced by DNA damage were observed. Next, we used inhibitors against JNK, p38 and MEK. Although MEK inhibitor had no effect, JNK and p38 inhibitor reduced the AS due to genotoxic stress. Role of MAP kinases in molecular mechanisms underlying AS induction by DNA damages will be further investigated.

P3: Analysis of sRNAs bound to nuclear AGO2 in HeLa human cell line

Valerio Fulci[1], Claudia Carissimi[1], Emanuela Cipolletta[1], Ilaria Laudadio[1], Silvia Gioiosa[1], Giuseppe Macino[1]

[1]Sapienza Università di Roma, Dipartimento di Biotecnologie Cellulari ed Ematologia, Rome, Italy

Argonaute proteins are the core components of RNA Induced Silencing Complex (RISC). These proteins are mainly involved in post-transcriptional gene silencing in the cytoplasm. However, several recent studies highlighted that Argonaute proteins are imported into the nucleus and play a role in transcriptional regulation of gene expression in plants and metazoa.
In order to shed light on the molecular mechanisms underlying the nuclear function of Argonaute proteins, we characterized sRNAs associated with AGO2 protein in the nucleus of human cells. By immuno-precipitation with specific antibodies, we purified from nuclear extracts of HeLa cells AGO2 protein with bound RNAs. As a control we used a mock immuno-precipitation from nuclear extracts. Our analyses highlighted that nuclear AGO2 is associated with microRNAs as well as several other short RNA molecules (20-26 nt long) which do not belong to any of the previously described families of short RNAs. Many of these novel AGO2-bound short RNAs map to regions characterized by an open chromatin epigenetic signature.
Intriguingly, we found that among nuclear AGO2-bound short RNAs there is a novel class of 22-24 nt long RNA molecules mapping to promoters and transcription start sites (TSS) of expressed genes. In agreement with recent reports highlighting bi-directional transcription at mammalian TSS, short RNAs bound by AGO2 derive from both sense and antisense strand of promoter regions.

P4: Ubiquitously active CpG-enriched genes are regulated at the transition from transcription initiation to elongation

Johannes Helmuth[1], Ho-Ryun Chung[1]

[1]Max Planck Institute for molecular Genetics, OWL Epigenomics, Berlin, Germany

Promoters are the starting site of transcription and serve as binding site for RNA polymerases. In these regions different regulatory inputs from *cis*-regulatory elements are integrated. Promoters can be classified into CpG enriched (HCP) and CpG depleted (LCP). Here, we subclassified promoters by Principal Component Analysis (PCA) of expression patterns of ubiquitously and differentially expressed genes among 16 human tissues. Then, we analyzed DNA accessibility by DNaseI-seq and also histone modification patterns, which were shown to correlate quantitatively with expression of downstream genes. Our results show that differentially expressed genes are driven by LCPs and are mainly regulated at the stage of transcription initiation. Developmental regulator genes, which are driven by HCPs, are regulated in the same way. In contrast, the majority of HCPs drive genes, which encode for essential transcripts of the cell metabolism and are regulated after transcription initiation at the level of transcription elongation.

P5: The influence of oxidative stress on transcriptional regulation

Michelle Hussong[1]

[1]*Max Planck Institute for molecular Genetics, OWL Epigenomics, Berlin, Germany*

Oxidative stress is a hallmark of many tumours and is caused by an imbalance of production and clearance of reactive oxygen species (ROS). Even many oncogenes can induce ROS, such as K-Ras and Myc, respectively. High levels of ROS can directly damage proteins, lipids and DNA resulting, among others, in chromosomal breakage and genomic instability. Low levels of ROS induce activation of signalling cascades and regulate transcription factors which result in changes in proliferation, angiogenesis and metastasis. To date, the influence of oxidative stress on transcriptional regulation is not well understood. Using next generation sequencing RNA-Seq technology we analyzed the effects of oxidative stress on transcriptional regulation to identify new oxidative stress related and potential cancer promoting pathways.

P6: Functional characterization of transcription factor mutations using ChIP-seq

Daniel M. Ibrahim[1, 2], Peter Hansen[3], Asita Stiege[2], Peter N. Robinson[1, 3], Stefan Mundlos[1, 2, 3], Jochen Hecht[1, 2]

[1]Berlin-Brandenburg Center for Regenerative Therapies, Berlin, Germany, [2]Max-Planck Institute for Molecular Genetics, Berlin, Germany , [3]Institute for Medical and Human Genetics, Charité University Hospital, Berlin, Germany

In recent years ChIP-seq technology has proven to be a powerful new method to investigate control mechanisms of transcription factors (TF) on a genome-wide scale. However, a number of technical hurdles need to be overcome in order to harness ChIP-seq methodologies to functionally characterize TF mutations involved in hereditary diseases.
We developed a novel technique that can be used to investigate a wide spectrum of TFs and mutations that have not previously been amenable to ChIP-seq experiments. In this work, we use a universal vector system coupled with ChIP-seq to investigate a missense mutation in *HOXD13* identified in a patient with a novel oligo-brachydactyly phenotype. The mutation, Q317K, alters the glutamine residue at position 50 of the homeodomain to a lysine residue, which is characteristic of bicoid-type homeodomain proteins including PITX1. We show that the Q317K mutation changes the HOXD13 binding profile towards that of PITX1, causing HOXD13^{Q317K} to bind regulatory elements bound by PITX1 as well as binding some others bound by HOXD13. Also, the induced gene expression patterns and the phenotypic effects following injection of the constructs in chicken embryos show substantial similarities between HOXD13^{Q317K} and PITX1 and provide further evidence for a global shift in the regulatory properties of the mutant HOXD13 towards that of PITX1.
We present here a robust method showing how ChIP-seq can be used to investigate the pathophysiology of mutations in TFs at a genome-wide scale. Furthermore, our findings show how a single amino-acid exchange in a TF alters the genome-wide binding and regulation of developmental programs.

P7: Popcar: Protein Occupancy Profiling of Chromatin-Associated RNAs

Marcin Kolinski[1,2], Emanuel Wyler[1], Markus Schueler[1], Guido Mastrobuoni[1], Alexandra Bergfort[1], Stefan Kempa[1], Christoph Dieterich[1], Markus Landthaler[1]

[1]Berlin Institute for Medical Systems Biology/Max-Delbrück-Center for Molecular Medicine, Berlin, Germany, [2]Computational Systems Biology Research Training Group, Humboldt University, Berlin, Germany

Differential mRNA processing plays an important regulatory role in cellular fate decisions and transformation. RNA processing is regulated by a large number of RNA binding proteins (RBPs) that bind to cis-regulatory elements in non-coding RNAs and both intronic and exonic regions of mRNA precursors. Using a photoreactive nucleoside-enhanced UV crosslinking and oligo(dT) affinity purification approach our laboratory has recently identified the mRNA-bound proteome and described the protein occupancy on mRNA transcripts [1]. This method however covers mostly mature, polyadenylated mRNAs. We therefore set out to extend our protein occupancy profiling approach to noncoding RNAs (ncRNAs) and pre-mRNAs. By using photoreactive nucleoside enhanced UV crosslinking and chromatin purification techniques in combination with next-generation sequencing we established “protein occupancy profiling of chromatin associated RNAs” (Popcar), which would allow for transcriptome-wide identification of protein sites on nascent and chromatin-associated RNAs. Popcar may become a valuable tool for studying protein-mediated RNA processing events and discovering cis-regulatory regions of pre-mRNAs and ncRNA.

Reference:

[1] Baltz AG, Munschauer M, Schwanhäusser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M. (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell. 46, 674–690.

P8: Towards Unrestricted Direct Reprogramming: A genome-wide RNAi screen to identify cell fate reprogramming-inhibiting factors in *C. elegans*

Ena Kolundzic[1], Oktay Ismail Kaplan[1], Martina Hajduskova[1], Andreas Ofenbauer [1], Alina Schenk[1], Baris Tursun[1]

[1]Berlin Institute for Medical Systems Biology (BIMSB) at Max Delbrueck Center (MDC), Berlin, Germany

Direct conversion of specific cell types into new, distinct cells has implications in basic biology as well as biomedicine. However, transcription factors (TFs) required for induction of specific cell differentiation programs are often inefficient in imposing such programs on other differentiated cells. It has recently been shown that inhibitory mechanisms could be involved in preventing these transcription factors from driving their differentiation programs in other cell types (Tursun et al. 2011, Patel et al. 2012). Here, we present a genome-wide RNAi screening approach to identify inhibitory factors of TF-induced direct conversion of different cell types in *C. elegans*. In order to study conversion into different cell types, we are using a terminal selector, CHE-1, essential for terminal differentiation of glutamatergic ASE neurons, as well as the myogenic TF HLH-1. Transgenic animals containing the ectopically expressed *che-1* or *hlh-1* gene driven by a heat shock inducible-promoter and a GFP reporter driven by the promoter of an ASE neuron or muscle fate marker are being used for the screen. Wild-type animals express GFP in one ASE neuron, or muscle cells under non heat-shocked conditions, respectively. After heat-shock, the fate-inducing TFs are ubiquitously expressed, but GFP signals of the respective fate reporters are only visible in either ASE and a few other neurons (gcy-5::gfp), or in muscle cells (unc-97::gfp). but no conversion of other cells into neuronal or muscle cells is detectable. The screen has been designed to look for factors which allow CHE-1/HLH-1 to induce expression of ASE/muscle marker in other cell types when silenced by RNAi. We present data for some of the candidates obtained from the screen.

P9: The role of Histone H3 Lysine 4 trimethylation in zebrafish embryonic genome activation

Maximilian Krause[1], Nadine Vastenhouw[1]

[1]Max Planck Institute for Cell Biology and Genetics, Dresden, Germany

Soon after revealing the coding sequence of the human genome, the scientific community recognized that knowing the DNA sequence alone does not allow to draw a human ‘blueprint’. Since transcriptional output is not only regulated by the underlying DNA sequence, it is equally important to understand molecular mechanisms of transcriptional repression and activation of genes.
DNA is compacted by histones and other proteins to a complex structure called chromatin. Post-translational modification of chromatin can affect the accessibility of DNA for DNA binding proteins and may thereby regulate transcription. Especially two modifications on Histone 3 have gained special interest: Histone H3 Lysine 27 trimethylation (H3K27me3), which was shown to be a repressive mark, as well as Histone H3 Lysine 4 trimethylation (H3K4me3), which is thought to poise genes for activation.
Interestingly, the onset of transcription during the maternal-to-zygotic transition (MZT) in zebrafish embryos is accompanied by changes in histone methylation patterns. Both H3K4me3 as well as H3K27me3 are not detected before MZT, but appear during the onset of transcription.

In my predoctoral studies, I will make use of this model system to investigate the effect of H3K4me3 on transcription initiation in zebrafish embryos. By targeted removal of H3K4me3 marks at specific loci, I want to study their transcription efficiency during MZT if H3K4me3 is absent. Additionally, I succeeded to achieve a global loss of H3K4me3 by overexpression of specific Histone demethylases as well as specific methylation-defective histone proteins. I will study the effect on transcription initiation by performing qPCR experiments and RNA sequencing at different stages of embryonic development. Possible locally restricted activation patterns will be revealed by fluorescence in-situ hybridization.
Ultimately, my studies will reveal, if the chromatin mark H3K4me3 is needed for the transcription initiation in early zebrafish embryos.

P10: Global Identification of ZFP36/TTP Specific RNA Binding and Regulation

Neelanjan Mukherjee[1], Jacobs Nicholas[1], Elizabeth Kennington[2], Jeff Nusbaum[3], Thomas Tuschl[3], Uwe Ohler[1], Perry Blackshear[2]

[1]The Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Berlin, Germany, [2]The Laboratory of Signal Transduction, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA, [3]Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, USA

Tristetraprolin (TTP/ZFP36) is one of many AU-rich RNA binding proteins (RBP) that plays a critical role in post-transcriptional regulation and specifically modulation of the immune response. TTP binds to AU-rich elements, typically in the 3' UTR of mRNAs. Upon binding it subsequently strongly promotes the degradation of its targets; a feature that can pose challenges for identifying physically interacting mRNAs, as well as for resolving secondary effects when identifying regulated mRNAs by modulating TTP levels. We combine biochemical and computational approaches to overcome these challenges. TTP-sensitive mRNAs were assayed using TTP overexpression followed by microarray. Along with directly regulated targets, there were significant secondary effects as a result of prolonged TTP overexpression. To resolve this problem, we performed PAR-CLIP to identify direct TTP-RNA interaction sites. While integrating the two types of data, we found features such as mRNA length and other RBP sites correlated with the overexpression results as strongly as TTP binding sites. Thus, we employed partial correlation analysis to account for these interdependencies and assess specific relationships. This allowed us to quantify TTP-specific effects within the overexpression experiments, while still accounting for other relationships. Furthermore, we are able to identify significant dependencies between various features, such as binding sites of TTP and HuR, another AU-rich RBP. This framework enables quantification of the specific effects of a RBP binding sites on regulation data. Additionally, it identifies and quantifies significant dependencies within mRNA features that reflect biologically relevant combinatorial action by multiple RBPs.

P11: Protein occupancy profiling globally maps putative cis-regulatory mRNA regions

Mathias Munschauer[1], Markus Schueler[1], Christoph Dieterich[1], Markus Landthaler[1]

[1]Berlin Institute for Medical Systems Biology at the Max Delbrueck Center for Molecular Medicine Berlin, Germany

A key prerequisite to understand how post-transcriptional processes are controlled and regulated by the interplay of RNA-binding proteins (RBPs) and microRNA containing ribonucleoprotein complexes (miRNPs), is the generation of comprehensive high resolution maps of protein-RNA interactions. Using photoreactive nucleotide enhanced crosslinking in combination with oligo(dT) affinity purification of mRNA, we were able to globally identify RNA contact sites of RNA interacting factors by profiling diagnostic nucleotide transitions that occur at direct protein-RNA interaction sites. Protein occupancy profiling on mRNA provides the first transcriptome-wide catalog of potential cis-regulatory elements of a human cell line. Application of our approach revealed protein-RNA contacts throughout large sequence stretches in 3'UTRs and coding sequences, with numerous putative binding sites in regions harboring disease-associated nucleotide polymorphisms. RNA sites contacted by proteins showed a significantly lower frequency of single nucleotide polymorphisms (SNPs) compared to noncrosslinked control regions, suggesting that these sites are under stronger negative selection in humans. Release of actively translating ribosomes by harringtonine or puromycin had little to no effect on the protein occupancy of coding sequences, thus the signal captured within protein coding sequences likely represents binding of non-ribosome mRNA interacting factors. In summary protein occupancy profiles narrow the genomic sequence search space for cis-regulatory elements in mRNA transcripts and we envision the identification of highly occupied mRNA sites to be valuable for the examination of rapidly emerging data on genetic variation between individuals.

P12: Analysis of fine time-course expression and DNase-seq data identifies new regulators of mammalian sex determination

Anirudh Natarajan[1], Steven Munger[2], Danielle Maatouk[1], Gregory Crawford[1], Blanche Capel[1], Uwe Ohler [3]

[1]Program in Computational Biology and Bioinformatics, Duke University, Durham, USA, [2]Center for Genome Dynamics, The Jackson Laboratory, Bar Harbor, USA, [3]Berlin Institute for Medical Systems Biology, Max Delbrueck Center, Berlin, Germany

In mammals, genetic sex is determined at fertilization by the presence or absence of a Y chromosome. In a process known as primary sex determination, this genetic difference leads to the development of a bipotential gonad into a testis (XY) or an ovary (XX). The gonadal sex difference then directs the sexual differentiation of the rest of the embryo. The unique property of the gonad, a bipotential organ capable of adopting two different fates with each fate antagonizing the other, makes it an excellent system to investigate properties of transcription networks and chromatin that are instrumental in fate initiation, canalization and maintenance across several cell fate determination processes.

However, we have limited knowledge regarding how this decision is implemented. While genetic studies have identified several genes involved in this process, mechanistic insights into how they operate still remain to be gleaned. Further, more than half the cases of human disorders of sexual development are as yet unexplained. To address these concerns, using the mouse as a model system, we conducted a fine-time course transcriptome analysis and DNase-seq to identify transcription factors (TFs), co-factors and the chromatin state that implement this fate decision.

We sampled the transcriptome at ~5 hr time points between embryonic day (E) 11.0 and E12.0, the critical fate determination window, from XX and XY mice. We then developed a Hidden Markov Model based approach to identify cohorts of dimorphically expressed genes. This approach allowed us to detect 5 sequential cascades of expression that occur in both XX and XY gonads during the fate commitment process. Interestingly, we showed that while higher expression in XY gonads is largely the result of activation in the XY gonad, higher expression in XX gonads is due to repression of these genes in XY gonads. Therefore, in addition to characterizing the known activational program in the testis, we reveal an important repressive program in the testis pathway.

We also conducted DNase-seq on E13.5 XY gonadal cells to reveal the open chromatin regions in these cells shortly after the fate decision. Using DNase-seq data from other unrelated cell types, we were able to define regions of open chromatin that are specifically open in gonadal cells. Computational analysis of these cell-type specific regions have revealed the binding sites of several known factors and enriched binding sites for novel TFs. We are currently developing techniques to discern co-binding of TFs to identify combinatorial regulation.

In order to test the predictions made by our analysis of these genome-wide datasets, we optimized a lentiviral based shRNA system to knockdown gene expression in primary gonadal cells. Specifically, we tested *Lmo4*, a co-factor that shows early up-regulation in the testis program. As predicted, knockdown of this gene resulted in down-regulation of several male pathway genes including *Sox9* and *Fgf9*. Future work will include using RNAi for functional validation of more genes involved in this process and ChIP to validate direct binding of TFs to their predicted sites.

P13: Unravelling the RNA polymerase II interactome during the transcription cycle using proteome-ChIP (pChIP)

Kedar Natarajan[1], Borislav Vangelov[2], Hiroshi Kimura[3], Mauricio Barahona[2], Ambrosius Snijders[4], Ana Pombo[1]

[1]Genome Function group, MRC-Clinical Sciences Centre, School of Medicine, Imperial College London, London, UK, [2]Department of Mathematics, Imperial College London, London, UK, [3]Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan, [4]Biomolecular Mass Spectrometry and Proteomics, MRC Clinical Sciences Centre, London, UK

Transcription is a fundamental process which regulates gene expression in all cells. RNA Polymerase II (RNAPII) transcribes genes encoding for proteins and structural RNAs and is highly modified by phosphorylation during the transcription cycle, which orchestrates the recruitment chromatin modifiers and RNA processing machinery to promote active transcription states. RPB1 is RNAPII largest subunit and contains a carboxy-terminal domain (CTD) composed of amino acid repeats with consensus sequence ($Y_1S_2P_3T_4S_5P_6S_7$). Phosphorylation of S5 (S5p) relates with initiation, and recruitment of H3K4 histone methyltransferase (HMTs) and the RNA capping enzyme, of S7 (S7p) with the transition to active elongation, and S2 (S2p) with recruitment of H3K36 HMT, splicing and polyadenylation machineries. Surprisingly, very little is known about proteins associating with specific RNAPII modifications and components of transcriptional apparatus.

We developed an unbiased approach called Proteome-Chromatin Immunoprecipitation (pChIP) to dissect the RNAPII interactome at different stages of the transcription cycle. After chromatin isolation from murine ES cells grown in SILAC (Stable Isotope Labelling by Amino acids in Cell culture) conditions, ChIP was performed with antibodies against RPB1-S5p, -S7p and -S2p. Using a multi-combinatorial design, protein enrichment with each phosphorylation state was quantified by mass spectrometry.

Using a logical binary approach and an independent novel clustering approach on the pChIP SILAC ratios of protein enrichment, we identify >400 proteins involved in transcription and co-transcriptional regulation that co-associate with chromatin-bound RNAPII with different phosphorylation marks. Remarkably, metabolic and translation proteins are found associated with active RNAPII complexes. Strikingly, we also identify >180 proteins associated only with initiating RNAPII marked by S5p, with roles in various non-transcriptional processes (DNA replication, cell cycle and Polycomb repression), which suggest additional functions for RNAPII in coordinating and integrating a multitude of nuclear processes.

In summary, we identify several new protein associations with specific phosphorylation of RNAPII, expand the interactome of chromatin-bound RNAPII, and provide novel insights into the transcription process.

P14: Deciphering the chromatin-related signaling from histone modification and chromatin modifier data

Juliane Perner[1], Julia Lasserre[1], Martin Vingron[1], Ho-Ryun Chung[1]

[1]Max Planck Institute for Molecular Genetics, Berlin, Germany

Apart from the compaction of DNA, chromatin plays an essential role in gene regulation. The chromatin landscape is characterized by various histone modifications (HMs) and chromatin-modifying proteins (CMs), e.g. histone methyltransferases, chromatin remodeler, etc. Together these components build a complex interaction network that influences gene expression. However, the precise mechanistic role of each component and their interactions are often unknown. Recent efforts to elucidate mechanistic interactions between HMs and CMs are based on the analysis of a single gene or rely on a pairwise co-occupancy analysis using Chip-seq data. However, this co-occupancy approach is likely to discover indirect interactions because HMs and CMs produce highly correlated Chip-seq profiles but in fact do only interact via a third component.
We use a recently published set of Chip-seq data of HMs and CMs in human cells to elucidate direct mechanistic interactions among them. We show that this data can be used to accurately quantify gene expression and to describe the chromatin landscape on a genome-wide scale. Further, we apply two complementary statistical techniques to decipher a high-confident interaction network representing the most direct interactions between HMs and CMs. Many of the discovered associations are known in the literature. Moreover, we reveal new associations that lead to new testable hypotheses for the role of the different HMs and CMs in regulating gene expression.

P15: Identification of Single Nucleotide Variants in short read data from paired DNA and RNA-seq experiments by ACCUSA2

Michael Piechotta[1], Christoph Dieterich[1]

[1]The Berlin Institute for Medical Systems Biology, Berlin, Germany

Direct comparisons of short read stacks are one way to identify Single Nucleotide Variants (SNVs). SNV detection is especially challenging across samples with different read depths (e.g. RNA-Seq) and high background levels (e.g. selection experiments or RNA editing studies).
We present ACCUSA2 to identify variant positions where nucleotide frequency spectra differ between two samples. To this end, ACCUSA2 integrates quality scores for base calling and read mapping into a common framework and employs a likelihood ratio test to identify variant sites. SNV calls can be filtered by the false discovery rate that is calculated by random permutations of base calls between paired read stacks. Mapping artifacts, a source for false positive variant calls especially inherent to RNA-Seq, can be handled by read stack filters that test for confounding factors such as read end bias of variant calls or exclusive associations of variants to read groups in SAM files.
We performed read simulations of short DNA and RNA read data and show that ACCUSA2 is superior to a state-of-the-art SNP caller in situations of diverging read depths and weakly diverging nucleotide frequencies. Our benchmark implements a read mixing scenario (i.e. one sample contains only reads that originate from a reference sequence and another sample contains reads from a diverged sequence and mutation-less reads) in combination with a scan over different input read coverages. We observed a superior sensitivity (DNA benchmark: avg. sensitivity of 86.71% achieved compared to 72.30% of SAMtools/BCFtools) for ACCUSA2 while being comparable in SNP calling precision (on avg. 98.76% compared to 99.99%).
We put ACCUSA2 to the test by considering publicly available data (http://www.sanger.ac.uk/resources/mouse/genomes/) from a large-scale RNA editing study of 15 mouse strains (Danecek et al. Genome Biology 2012). We assessed the accuracy of predicted RNA-DNA differences (RDDs) by comparing against a set of validated sites and contrasting with phylogenetic conservation of editing sites across multiple mouse strains.

P16: The transcriptional repressor CTBP-1 functions in the nervous system of Caenorhabditis elegans to regulate gene expression

Anna Reid[1], Duygu Yücel[1], Estelle Llamosas[1], Sashi Kant[1], Hannah Nicholas[1]

[1]School of Molecular Bioscience, The University of Sydney, Sydney, Australia

C-terminal binding proteins (CtBPs) are recruited by a variety of transcription factors to mediate gene repression. To accomplish this, CtBPs recruit chromatin-modifying enzymes such as histone deacetylases (HDACs), resulting in the down regulation of target gene expression. The Caenorhabditis elegans genome contains a single CtBP locus that encodes two variants of CtBP; CTBP-1a and CTBP-1b.
To determine where CTBP-1 functions, we tagged the C-terminus of CTBP-1 with mCherry using fosmid recombineering. We observed that CTBP-1 expression begins during embryonic development and continues throughout larval and adult stages. Also, we have found that CTBP-1 is widely expressed in the nervous system and in the hypodermis. In both tissues CTBP-1 localises to the nuclei, consistent with a role for CTBP-1 in regulating gene expression.
Chen et al., 2009 previously demonstrated that a ctbp-1 mutant had an extended life span relative to wild-type. We have determined that two additional ctbp-1 mutants have extended life spans relative to wild-type. Re-expression of CTBP-1 solely in the nervous system of a splice site mutant, called ctbp-1(eg613), restored normal lifespan, indicating that CTBP-1 may function in the nervous system to regulate lifespan.
To identify genes that are potential targets of CTBP-1-mediated repression, we performed microarray analysis on wild type and ctbp-1(eg613) mutants. 362 genes were up-regulated whilst 138 genes were down-regulated by 2-fold or greater in the ctbp-1(eg613) mutants. We found that a lipase gene lips-7(C09E8.2) was up-regulated in these microarrays and confirmed this using real-time RT PCR. Re-expression of CTBP-1 solely in the nervous system restored lips-7 expression to normal expression levels. Chen et al., 2009 had previously suggested that CTBP-1 controls life span by regulating lips-7 transcription. Our data suggest that CTBP-1 functions in the nervous system to regulate both life span and lips-7 expression.

P17: FLEXBAR - Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms

Johannes Roehr[1], Matthias Dodt[1], Rina Ahmed[1], Christoph Dieterich[1]

[1]Berlin Institute for Medical Systems Biology at the Max Delbrück Centre for Molecular Medicine, Berlin, Germany

Quantitative and systems biology approaches benefit from the unprecedented depth of next-generation sequencing. A typical experiment yields millions of short reads, which oftentimes carry particular sequence tags. These tags may be: (a) specific to the sequencing platform and library construction method (e.g., adapter sequences); (b) have been introduced by experimental design (e.g., sample barcodes); or (c) constitute some biological signal (e.g., splice leader sequences in nematodes). Our software FLEXBAR enables accurate recognition, sorting and trimming of sequence tags with maximal flexibility, based on exact overlap sequence alignment. The software facilitates fine-grained adjustment of sequence tag detection parameters and search regions. Data formats from all current sequencing platforms are supported, including color-space reads. Read pairings are maintained and separate barcode reads are processed on demand. Furthermore, varying positions of sequence tags may be considered. FLEXBAR is a multi-threaded software and combines speed with precision. Even complex read processing scenarios might be executed with a single command line call. The utility of the software is demonstrated in terms of read mapping applications, library demultiplexing and splice leader detection.

P18: Novel modes of DNA-binding identified by detailed genome-wide analysis of Glucocorticoid receptor occupancy

Stephan R. Starick[1], Michael Love[1], Alessandro Mammana[1], Morgane Thomas-Chollier[1,2], Sebastiaan Meijsing[1]

[1]Max Planck Institute for Molecular Genetics, Berlin, Germany, [2]Institut de Biologie de l'Ecole Normale Supérieure, Paris, France

The glucocorticoid receptor (GR) was one of the first transcription factors for which the DNA recognition sequence was identified now more than 3 decades ago. However, this recognition sequence only partially explains where GR binds in the genome. For example, our ChIP-seq analysis showed that for some cell-types less than half of the peaks contain the canonical motif. This raises the question: What are the sequences responsible for recruiting GR for the fraction of peaks that do not contain this motif? Other sequences have been identified that can bind GR directly or alternatively, it has been shown that GR can be tethered to the DNA by other transcriptional regulatory factors. Here we set out to identify novel sequences that can recruit GR. The analysis of ChIP-seq and ChIP-exo data resulted in the identification of several previously known and novel candidate sequences. One of these sequences is a combination of a GR half site and the recognition sequence for the transcriptional enhancer factor TEF1. We tested this sequence in luciferase reporter studies and found that it can indeed confer GR-dependent transcriptional regulation. Our data suggests that GR heterodimerizes with another factor to bind to this sequence. Future efforts are aimed at identifying this partner for example by pulldown assays using this sequence motif as a bait, combined with mass-spec. Together our studies expand the spectrum of sequences that can specify where GR binds genomically.

P19: DNA-encoded signals that prevent genomic binding of Transcription Factors

Jonas Telorac [1], Morgane Thomas-Chollier [1,2], Sergey Prykhozhij[1,3], Stefanie Schöne[1], Sebastiaan Meijsing[1]

[1]MPI for Molecular Genetics, Berlin, Germany, [2]Institut de Biologie de l'Ecole normale supérieure, Paris, France, [3]Departments of Pediatrics, Microbiology & Immunology and Pathology, Dalhousie University, Halifax, Canada

For deeper understanding of gene regulation it is necessary to gain deeper insight into the mechanisms that control transcription factor binding to DNA and thereby regulate transcription. It is known that transcription factors (TF) bind to canonical binding sites encoded in the genome whose exact binding site depends on the TF of interest. However, only a very small part of all the possible binding sites are actually bound by TFs. Often additional signals are necessary for TF binding like the presence of binding sites for cofactors. Another reason for the selective binding of TFs binding sites is the DNA accessibility, if the DNA is tightly packed into heterochromatin the canonical binding side is hidden from the transcription factors.
Since positive signals for TF binding are encoded in the genome we reasoned also negative signals could be encoded. If such signals exist, they should be depleted in regions where TFs bind.
In our lab we utilize the Glucocorticoid Receptor (GR), a hormone activated TF, as a tool to investigate gene regulation. In ChIP-Seq Data for hormone activated GR we found AT rich sequences that are depleted at loci to which GR binds. We were able to show that these sequences (NRS) have the potential to prevent GR from binding its canonical binding sites. According to preliminary results these NRS have similar effects also on other transcription factors. Additional studies in zebrafish demonstrated that the mechanism mediating the effects of NRSs is conserved across species and that NRS are active in most tissues.
How NRSs interfere with TF binding remains unclear, but we suspect that this is mediated by proteins that bind to NRSs. We were able to identify candidate proteins in Pull-Downs experiments with NRSs as bait followed by identification of associated proteins by Mass Spectrometry. In future experiments we plan to verify the binding of these candidates to NRSs in vivo chromatin immunoprecipitation. In addition we plan to prove their predicted activity with luciferase reporter assays.
Overall, our studies suggest that positive and negative sequence signals partition the genome into regions where TFs can bind, and those where it cannot.

P20: Thousands of long ncRNAs are expressed from tissue specific enhancers

Dubravka Vucicevic[1], Olivia Corradin[2], Peter Scacheri[2], Ulf Ørom[1]

[1]Max Planck Institute for Molecular Genetics, Berlin, Germany, [2]Case Western Reserve University, Cleveland, USA

Long non-coding RNAs (ncRNA) have recently been demonstrated to be expressed from a subset of enhancers and be required for the distant regulation of gene expression. Several approaches to predict enhancers have been developed based on various chromatin marks and binding of enhancer-binding proteins, and while there is rapid development of the field, still no consensus as to how to define tissue specific enhancers exist. Here, we identify 3,003 long ncRNAs that overlap tissue-specific enhancers, corresponding to 32 % of the ENCODE annotated long ncRNAs. We use a recently developed algorithm that is based on H3K4me1 marks and tissue specific expression of mRNA to predict the tissue-specific enhancers. The expression of the long ncRNAs overlapping enhancers is significantly higher when the enhancer is predicted as active in a specific cell line, suggesting a general dependency of active enhancers and expression of long ncRNA. This dependency is not identified using previous enhancer identification algorithms that do not take tissue-specific expression into consideration, suggesting that long ncRNAs are particularly involved in tissue-specific enhancer activity. In conclusion, we provide evidence for thousands of long ncRNAs that are functionally expressed from tissue-specific enhancers, suggesting a particularly important functional relationship between tissue-specific expressed long ncRNAs and enhancer activity.

P21: Posttranscriptional Regulatory Interactions during Herpes Simplex Virus Infection

Emanuel Wyler[1], Henrik Groß[2], Guido Mastrobuoni[1], Markus Schueler[1], Miha Milek[1], Christoph Dieterich[1], Stefan Kempa[1], Friederich Grässer[2], Markus Landthaler[1]

[1]Max Delbrück Center for Molecular Medicine, Berlin Institute for Medical Systems Biology, Berlin, Germany, [2]Department of Virology, Saarland University Medical School, Homburg/Saar, Germany

Herpes viruses cause a variety of human diseases with often very limited options for treatment or prevention. Thus, a greater understanding of regulatory interactions during HSV infection is crucial for the identification of novel targets for intervention. Here, we aim to analyze the dynamic interplay between RNAs and proteins expressed in human cells infected with Herpes Simplex Virus 1 (HSV-1). It has now been recognized that the interaction of proteins and small, non-coding RNAs such as microRNAs with mRNA molecules constitute an additional layer of gene regulation aside of direct transcriptional control or epigenetic modification of DNA. Only scant information is available concerning the proteins and/or non-coding RNAs associated with the mRNA expressed during the infection of human cells with HSV-1. We will apply a combination of high-throughput methods such as dynamic transcriptome analysis, SILAC-based proteomics, mRNA interactome capture, and protein occupancy profiling on mRNA to build up an atlas of RNA binding proteins (RBPs) and cis-regulatory regions involved in the post-transcriptional events during viral infection of human cells.
Apart from confirming known changes in gene expression, experiments performed so far have shown that protein occupancy profiles of dozens of mRNAs change upon HSV infection in WI-38 cells, indicating a role for protein-mRNA interactions in response to viral infection, presumably involving viral RBPs. In addition, novel transcripts and protein occupancy profiles of early viral RNAs already four hours after infection were reliably identified. These data will be particularly useful to identify candidate LNA oligonucleotides as RBP target site protectors that as such interfere with virus infection and may thus lead to novel therapeutic approaches.

P22: The dynamic DNA methylation landscape of the human genome

Michael J. Ziller[1,2,3], Hongcang Gu[1], Bradley Bernstein[1], Alexander Meissner[1,2,3]

[1]Broad Institute of MIT and Harvard, Cambridge, USA, [2]Harvard Stem Cell Institute, Cambridge, USA, [3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, USA

DNA methylation is a defining feature of mammalian cellular identity and essential for normal development. Most cell types, except germ cells and pre-implantation embryos, display relatively stable DNA methylation patterns with 70-80% of all CpGs being methylated. Despite recent advances we still have a too limited understanding of when, where and how many CpGs participate in genomic regulation. Here we report the in depth analysis of 46 whole genome bisulfite sequencing (WGBS) data sets across 34 diverse human cell and tissue types. We observe dynamic regulation for only ~20% of autosomal CpGs within a normal developmental context, a majority of which are distal to transcription start sites. These variable CpGs co-localize with gene regulatory elements and transcription factor binding sites (TFBS), which allow identification of key lineage specific regulators. Dynamic regions exhibit elevated SNP frequency and often harbor SNPs associated with cell type related diseases as determined by GWAS. Comparison of this class of dynamic CpGs to several abnormal contexts reveals that low complexity regions are preferential targets of the additional DNAme changes in cancer, Alzheimer and long-term cell culture. The results also highlight the general inefficiency of WGBS as between 70-80% of the sequencing reads across these data sets provided little or no relevant information regarding CpG methylation. To further demonstrate the utility of our identified set of dynamic regions, we classify unknown samples and identify representative signature regions that recapitulate major DNA methylation dynamics. In summary, our results show that although in theory every CpG could change its methylation state, only a fraction does so as part of coordinated regulatory programs. As a result the identified regions will help guide novel reduced representation approaches to capture this informative fraction of CpGs more effectively as well as further pinpoint putative regulatory elements.

List of Participants

Asifa Akhtar
Max Planck Institute of Immunobiology and Epigenetics
Chromatin Regulation
Stübeweg 51
79108 Freiburg
Germany
akhtar@ie-freiburg.mpg.de

Ana Babic
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
ana.babic@mdc-berlin.de

Denise Barlow
CeMM-Research Center for Molecular Medicine
Principal Investigator
CeMM GmbH, Lazarettgasse 14/AKH BT 25.3
1090 Vienna
Austria
sbaier@cemm.oeaw.ac.at

Robert Beagrie
MRC Clinical Sciences Centre
Integrative Biology
15 The Park W55NE
London
UK
rab11@ic.ac.uk

Stephen Buratowski
Harvard Medical School
Biological Chemistry and Molecular Pharmacology
BCMP/C1-210
02115 Boston
USA
steveb@hms.harvard.edu

Dorothea Busse
Max Delbrück Center for Molecular Medicine
Mathematical Modelling of Biological Processes
Invalidenstraße 143
10115 Berlin
Germany
dorothea.busse@mdc-berlin.de

Jonathan Byrne
IMB
Epigentics
Ackermanweg 4
55128 Mainz
Germany
j.byrne@imb-mainz.de

Lorenzo Calviello
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
calviello.l@gmail.com

Claudia Carissimi
Università di Roma "La Sapienza"
Biotecnologie Cellulari ed Ematologia
Viale Regina Elena,324
00161 Rome
Italy
carissimi@bce.uniroma1.it

Tao Chen
Max Delbrück Center for Molecular Medicine
Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
Tao.Chen@mdc-berlin.de

CécileDetiffe
University of Liège
Sciences
En Feronstrée, 79 bt 1
4000 Liège
Belgium
Cecile.Detiffe@doct.ulg.ac.be

Adeline Deward
University of Liège
GIGA-Virology and Immunology
Quai St-Léonard 11b bte101
4000 Liège
Belgium
adeward@ulg.ac.be

Joao Dias
MRC Clinical Sciences Centre
Integrative Biology
58 Musard Road
W6 8NW London
UK
joao.dias@imperial.ac.uk

Christoph Dieterich
MDC Berlin-Buch
BIMSB
Robert-Roessle-Straße 10
13125 Berlin
Germany
christoph.dieterich@mdc-berlin.de

Hang Du
Max Delbrück Center for Molecular Medicine
AG. Wei Chen (Genomcentrum)
Robert-Rössle-Straße 10
13125 Berlin
Germany
Hang.Du@mdc-berlin.de

KatrinEichelbaum
Max Delbrück Center for Molecular Medicine
Cell Signalling and Mass Spectrometry
Robert-Rössle-Straße 10
13125 Berlin
Germany
Katrin.Eichelbaum@mdc-berlin.de

Cipolletta Emanuela
Università di Roma "La Sapienza"
Biotecnologie Cellulari ed Ematologia
Viale Regina Elena,324
00161 Roma
Italy
cipolletta@bce.uniroma1.it

Carmelo Ferrai
Max Delbrück Center for Molecular Medicine
BIMSB - Epigenetic regulation and Nuclear architecture
Robert-Rössle-Straße 10
13125 Berlin
Germany
carmelo.ferrai@imperial.ac.uk

Cornelius Fischer
Max Planck Institute for Molecular Genetics
Otto-Warburg-Laboratory - Nutrigenomics and Gene Regulation
Selchower Straße 21
12049 Berlin
Germany
cfischer@molgen.mpg.de

Martin Franke
Max Planck Institute for Molecular Genetics
AG Mundlos
Ihnestraße 63-73
14195 Berlin
Germany
franke_m@molgen.mpg.de

Sebastian Fröhler
Max Delbrück Center for Molecular Medicine
BIMSB - Chen Lab
Robert-Rössle-Straße 10
13125 Berlin
Germany
Sebastian.Froehler@mdc-berlin.de

Alisa Fuchs
Max Planck Institute for Molecular Genetics
OWL Chung
Hohenfriedbergstraße 6
10829 Berlin
Germany
fuchs_a@molgen.mpg.de

Valerio Fulci
Università di Roma "La Sapienza"
Biotecnologie Cellulari ed Ematologia
Viale Regina Elena,324
00161 Rome
Italy
fulci@bce.uniroma1.it

Qingsong Gao
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
qingsong.gao@mdc-berlin.de

Darío Jesús García Lupiáñez
Max Planck Institute for Molecular Genetics
Mundlos
Kuglerstraße 100
10439 Berlin
Germany
lupianez@molgen.mpg.de

Luca Giorgetti
Institut Curie
Mammalian Developmental Epigenetics Group
11-13 rue Pierre et Marie Curie
75248 Paris
France
luca.giorgetti@curie.fr

Lea Haarup Gregersen
Max Delbrück Center for Molecular Medicine
Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
lea.gregersen@mdc-berlin.de

Dina Hafez
Max Delbrück Center for Molecular Medicine
BIMSB
Tieckstraße 35
10115 Berlin
Germany
dina.hafez@mdc-berlin.de

Martina Hajduskova
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
martina.hajduskova@mdc-berlin.de

Johannes Helmuth
Max Planck Institute for Molecular Genetics
OWL Epigenomics
Ihnestraße 63-72
14195 Berlin
Germany
helmuth@molgen.mpg.de

Undine Hill
Max Delbrück Center for Molecular Medicine
AG Leutz
Truseweg 2
12059 Berlin
Germany
undine.hill@mdc-berlin.de

Jingyi Hou
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
jingyi.hou@mdc-berlin.de

Michelle Hussong
Max Planck Institute for Molecular Genetics
Lehrach
Ihnestraße 63-73
14195 Berlin
Germany
hussong@molgen.mpg.de

Mahmoud Ibrahim
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
mahmoud.ibrahim@mdc-berlin.de

Sarah Kinkley
Max Planck Institute Epigenomics
Ihnestraße 63-73
14195 Berlin
Germany
kinkley@molgen.mpg.de

Antje Hirsekorn
Max Delbrück Center for Molecular Medicine
AG Ohler
Landsberger Allee 62b
10249 Berlin
Germany
antje.hirsekorn@mdc-berlin.de

Matthew Huska
Max Planck Institute for Molecular Genetics
Vingron
Ihnestraße 63-73
14195 Berlin
Germany
huska@molgen.mpg.de

Daniel Ibrahim
Max Planck Institute for Molecular Genetics
Research Group Development and Disease
Ihnestraße 63-73
14195 Berlin
Germany
ibrahim@molgen.mpg.de

Nicholas Jacobs
Max Delbrück Center for Molecular Medicine
BIMSB
Mulackstraße 6
10119 Berlin
Germany
nicholas.jacobs@mdc-berlin.de

Marcin Kolinski
Max Delbrück Center for Molecular Medicine
The Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
marcin.kolinski@mdc-berlin.de

Ena Kolundzic
Max Delbrück Center for Molecular Medicine
BIMSB - AG Tursun
Robert-Rössle-Straße 10
13125 Berlin
Germany
ena.kolundzic@mdc-berlin.de

Dorothee Kraemer
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
Dorotheekramer@web.de

Markus Kretz
University of Regensburg
Biochemistry, Genetics and Microbiology
Universitätsstrasse 31
93053 Regensburg
Germany
Markus.Kretz@vkl.uni-regensburg.de

Alexander Kukalev
Max Delbrück Center for Molecular Medicine
BIMSB - Epigenetic regulation and Nuclear architecture
Robert-Rössle-Straße 10
13125 Berlin
Germany
akukalev@gmail.com

Markus Landthaler
Max Delbrück Center for Molecular Medicine
Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
markus.landthaler@mdc-berlin.de

Wolfgang Kopp
Max Planck Institute for Molecular Genetics
Computational Biology
Ihnestraße 73
14195 Berlin
Germany
kopp@molgen.mpg.de

Maximilian Krause
Max Planck Institute for Cell Biology and Genetics
Vastenhouw Lab
Schlueterstrasse 3
01277 Dresden
Germany
mkrause@mpi-cbg.de

Sabrina Krueger
Max Delbrück Center for Molecular Medicine
BIMSB - AG Robert Zinzen
Robert-Rössle-Straße 10
13125 Berlin
Germany
sabrina.krueger@mdc-berlin.de

Scott Lacadie
Max Delbrück Center for Molecular Medicine
Computational Regulatory Genomics AG Ohler
Robert-Rössle-Straße 10
13125 Berlin
Germany
Scott.Lacadie@mdc-berlin.de

Ilaria Laudadio
Università di Roma "La Sapienza"
Biotecnologie Cellulari ed Ematologia
Viale regina Elena,324
00161 Roma
Italy
laudadio@bce.uniroma1.it

Liron-Mark Lavitas
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
liron-mark.lavitas@mdc-berlin.de

Alessandro Mammana
Max Planck Institute for Molecular Genetics
Otto-Warburg-Laboratories, Epigenomics
Ihnestraße 63-73
14195 Berlin
Germany
ale.fener@gmail.com

Sebastiaan Meijsing
Max Planck Institute for Molecular Genetics
CMB
Wiesbadener Straße 16
14197 Berlin
Germany
meijsing@molgen.mpg.de

Christin Mieth
Max Delbrück Center for Molecular Medicine
Cardiovascular and Metabolic Diseases
Robert-Rössle-Straße 10
13125 Berlin
Germany
christin.baeth@mdc-berlin.de

Miha Milek
Max Delbrück Center for Molecular Medicine
Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
miha.milek@mdc-berlin.de

Nemanja Milićević
Charité - Universitätsmedizin Berlin
International Graduate Program Medical Neurosciences
Keithstrasse 36
10787 Berlin
Germany
obnovairazvoj@yahoo.com

Fabio Mohn
Institute of Molecular Biotechnology
Brennecke lab
Dr. Bohr-Gasse 3
1030 Vienna
Austria
fabio.mohn@imba.oeaw.ac.at

Kelly Morris
Max Delbrück Center for Molecular Medicine
Epigenetic Regulation and Chromatin Architecture, BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
KellyJane.Morris@mdc-berlin.de

Jose M. Muino
Max Planck Institute for Molecular Genetics
Computational Molecular Biology
Ihnestraße 73
14195 Berlin
Germany
muino@molgen.mpg.de

Neelanjan Mukherjee
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
nelanjan.mukherjee@mdc-berlin.de

Stefan Muljo
National Institutes of Health (NIH)
Laboratory of Immunology
10 Center Drive
20892-1892 Bethesda
USA
stefan.muljo@nih.gov

Mathias Munschauer
Max Delbrück Center for Moleculare Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
mathias.munschauer@mdc-berlin.de

Anirudh Natarajan
Max Delbrück Center for Molecular Medicine
Berlin Institute for Medical Systems Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
anirudh.natarajan@gmail.com

Kedar Natarajan
MRC-Clinical Sciences Centre
Genome Function
Du Cane Road
W120NN London
UK
k.natarajan09@imperial.ac.uk

Benedikt Obermayer
Max Delbrück Center for Molecular Medicine
Systems Biology of Gene Regulatory Elements
Robert-Rössle-Straße 10
13125 Berlin
Germany
benedikt.obermayer@mdc-berlin.de

Andreas Ofenbauer
Max Delbrück Center for Molecular Medicine
BIMSB

Robert-Rössle-Straße 10
13125 Berlin
Germany
andreas.ofenbauer@mdc-berlin.de

Uwe Ohler
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
uwe.ohler@mdc-berlin.de

Ulf Ørom
Max Planck Institute for Molecular Genetics
Otto Warburg Laboratories
Ihnestraße 51
14195 Berlin
Germany
oerom@molgen.mpg.de

Hajnalka Pálinkás
Budapest University of Technology and Economics
Applied Biotechnology and Food Science
38 Kisfaludy street H-2060
Bicske
Hungary
palinkas@enzim.hu

Florian Paul
Max Delbrück Center for Molecular Medicine
Signaling and Mass Spectrometry
Robert-Rössle-Straße 10
13125 Berlin
Germany
florianpaul@mdc-berlin.de

Juliane Perner
Max Planck Institute for Molecular Genetics
Computational Molecular Biology
Ihnestraße 63-73
14195 Berlin
Germany
perner@molgen.mpg.de

Hartmut Peters
Charite - Universitaetsmedizin Berlin
Institute of Medical and Human genetics
Gartenstraße 26
10115 Berlin
Germany
hartmut.peters@yahoo.de

Michael Piechotta
Max Delbrück Center for Molecular Medicine
BIMSB - Dieterich Lab
Robert-Rössle-Straße 10
13125 Berlin
Germany
michael.piechotta@mdc-berlin.de

Stefan Pinter
Harvard Medical School, Mass. General Hospital
Genetics / Molecular Biology
50A Linnaean Street
02138 Cambridge
USA
pinter@molbio.mgh.harvard.edu

Ana Pombo
Max Delbrück Center for Molecular Medicine
Gene Regulations
Robert-Rössle-Straße 10
13125 Berlin
Germany
Ana.Pombo@mdc-berlin.de

Claudia Quedenau
Max Delbrück Center for Molecular Medicine
BIMSB - AG Chen
Robert-Rössle-Straße 10
13125 Berlin
Germany
quedenau@mdc-berlin.de

Anna Reid
The University of Sydney
School of Molecular Bioscience
5/51 Woodcourt St
2204 Sydney
Australia
arei0791@uni.sydney.edu.au

Leonie Ringrose
IMBA Institute of Molecular Biotechnology
Dr. Bohr-Gasse 3
1030 Vienna
Austria
Leonie.Ringrose@imba.oeaw.ac.at

Tiago Rito
Max Delbrück Center for Molecular Medicine
Building 89
Robert-Rössle-Straße 10
13125 Berlin
Germany
tiago.rito@mdc-berlin.de

Marta Rodriguez
Max Delbrück Center for Molecular Medicine
BIMSB - Systems Biology of Gene Regulatory Elements
Brunnenstraße 96a
13355 Berlin
Germany
marta.rodriguez@mdc-berlin.de

Johannes Roehr
Max Delbrück Center for Molecular Medicine
Bioinformatics in Quantitative Biology
Robert-Rössle-Straße 10
13125 Berlin
Germany
johannes.roehr@mdc-berlin.de

Edda Schulz
Institut Curie
Developmental Epigenetics
11-13 Rue Pierre et Marie Curie
75005 Paris
France
edda.schulz@curie.fr

Matthias Selbach
Max Delbrück Center for Molecular Medicine
Cell Signalling and Mass Spectrometry
Robert-Rössle-Straße 10
13125 Berlin
Germany
matthias.selbach@mdc-berlin.de

Gioiosa Silvia
Università di Roma "La Sapienza"
Biotecnologie Cellulari ed Ematologia
Viale Regina Elena,324 00161 Roma
00161 Rome
Italy
gioiosa@bce.uniroma1.it

Carine Stapel
Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG)
Vastenhouw lab
Louisenstrasse 60
01099 Dresden
Germany
stapel@mpi-cbg.de

Marcel Schilling
Max Planck Institute for Molecular Genetics
Vertebrate Genomics
Donizettistr. 81A
12623 Berlin
Germany
mschilli@molgen.mpg.de

Stefanie Seelk
Max Delbrück Center for Molecular Medicine
BIMSB - AG Tursun
Eberswalder Straße 16
10437 Berlin
Germany
stefanie.seelk@mdc-berlin.de

Alena Shkumatava
Institut Curie
Developmental Biology and Genetics
26, rue d'Ulm
75005 Paris
France
alena.shkumatava@curie.fr

Christoph Standfuß
Technische Hochschule Wildau
Bioinformatics Pospisil
Hohenzollerndamm 52
14199 Berlin
Germany
christoph.standfuss@th-wildau.de

Stephan R. Starick
Max Planck Institute for Molecular Genetics
Bioinformatik Ihnestraße 63-73
14195 Berlin
Germany
starick@molgen.mpg.de

Wei Sun
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
wei.sun2@mdc-berlin.de

Jonas Telorac
Max Planck Institute for Molecular Genetics
Bioiformatik
Ihnestraße 63-73
14195 Berlin
Germany
telorac@molgen.mpg.de

Elena Torlai Triglia
Max Delbrück Center for Molecular Medicine
BIMSB - AG Pombo
Robert-Rössle-Straße 10
13125 Berlin
Germany
elena.torlai-triglia@mdc-berlin.de

Bas van Steensel
Netherlands Cancer Institute
Division of Gene Regulation
Plesmanlaan 121
1066CX Amsterdam
The Netherlands
b.v.steensel@nki.nl

Alexandra Vasile
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
alexandra.vasile@mdc-berlin.de

Claes Wahlestedt
University of Miami Miller School of Medicine
Center for Therapeutic Innovation
1501 NW 10th Avenue
33136 Miami
USA
clawah@gmail.com

Thomas Wilhelm
IMB Mainz
Molecular Epigenetics
Im Münchfeld 31
55122 Mainz
Germany
t.wilhelm@imb-mainz.de

Hans Will
University Hospital Hamburg Eppendorf
Institute of Tumorbiology
Martinistrasse 52
20256 Hamburg
Germany
h.will@uke.de

Emanuel Wyler
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
emanuel.wyler@mdc-berlin.de

Xintian You
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
xintian.you@mdc-berlin.de

Henrik Zauber
Max Delbrück Center for Molecular Medicine
Signaling and Mass Spectrometry
Robert-Rössle-Straße 10
13125 Berlin
Germany
henrik.zauber@mdc-berlin.de

Bin Zhang
Max Delbrück Center for Molecular Medicine
BIMSB
Robert-Rössle-Straße 10
13125 Berlin
Germany
zb.picb@gmail.com

Michael Ziller
Harvard University
Stem Cell and Regenerative Biology
7 Divinity Avenue
02138 Cambridge
USA
michael_ziller@harvard.edu

Notes