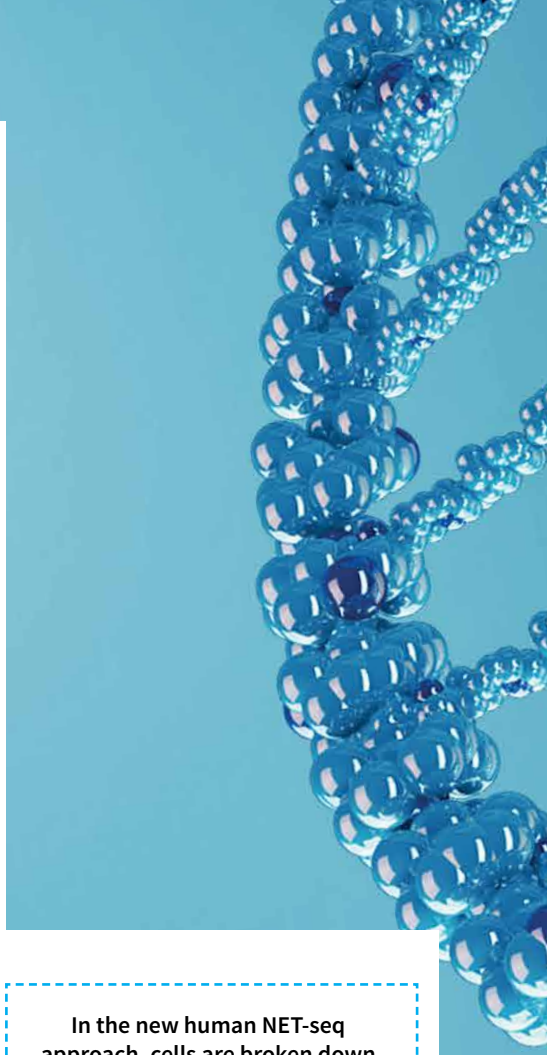
Genome Transcription Regulation During Cell Differentiation

Dr Andreas Mayer

Scientia

GENOME TRANSCRIPTION REGULATION DURING CELL DIFFERENTIATION

The mechanisms that control the transcription of DNA to produce RNA and the building blocks of life, proteins, are a fundamental cellular process in all living organisms. **Dr Andreas Mayer** at the Max Planck Institute for Molecular Genetics in Germany has spent more than a decade unravelling these complex processes. Using newly developed high-resolution genome-wide techniques, his team is discovering the vital role that RNA polymerase II transcription plays in stem cell differentiation, where a cell changes from one cell type to another usually to perform a more specialist function.



Gene Transcription and Translation

Genetic material is stored in the form of DNA in most organisms. In humans, the nucleus of each cell has 3×10^9 (9 zeros) base pairs of DNA, distributed over 23 pairs of chromosomes, and each cell has two copies of the genetic information. This is known collectively as the human genome and contains around 20,000 genes that code for proteins and at least the same number of genes that produce RNA only (non-protein-coding genes).

Converting the genetic information contained in DNA into RNA and proteins is one of the most important and highly regulated tasks performed within cells. Genes are small segments of DNA that provide the code (through combinations of four nucleotides: adenine, cytosine, guanine, and thymine) to synthesise specific RNAs and proteins using two fundamental cellular processes: transcription and translation.

Transcription is the process of producing a strand of RNA, either a messenger RNA (mRNA, protein-coding RNA) or a non-protein-coding RNA, from a DNA template. It comprises three stages: initiation, elongation, and termination. In the initiation stage, the enzyme RNA polymerase II (Pol II) attaches with the help of other proteins, so-called general transcription factors, at the promoter region on each gene. Certain general transcription factors then begin to untwist the two nucleic acid strands of the DNA double helix to allow loading of the DNA template strand into the catalytic centre of Pol II and transcription initiation to occur. During elongation, Pol II creates a new strand of precursor RNA such as a precursor mRNA (pre-mRNA), by adding the complementary nucleotides (with uracil replacing thymine) to the template strand of DNA, which then reforms back into the double helix. Sequences called polyadenylation signals on each gene signal to the Pol II that the RNA transcript is complete.

The pre-mRNA that is released from Pol II can undergo further modification through the addition of molecular 'caps' at one end and a tail to the other end, which stabilise the RNA, and through splicing, remove redundant parts of the code and re-join the remaining

In the new human NET-seq approach, cells are broken down (lysed) and transcribing human RNA polymerases are purified by chromatin fractionation. Next, the 3' ends of the nascent RNA that contain the last nucleotide that was added to the RNA chain prior to cell lysis, are concerted into a sequencing library. With modern, so called 'next-generation sequencing techniques', they can then identify the nucleic-acid sequences as well as the abundance of the ends of these RNA transcripts – and match them to the known cellular DNA sequence to establish the transcripts' and thus RNA polymerases' precise locations. This technique enables researchers to also determine all the points on a cell's genome where transcription activity was busiest, at the time the cell was lysed. Application of the human NET-seq method revealed the fine structure of RNA polymerase transcription and of transcriptional pausing along genes (Figure 2: NET-Seq Tracks).

Box 1: Human NET-Seq Approach. Credit and copyright: Dr Andreas Mayer/ Mayer Laboratory.

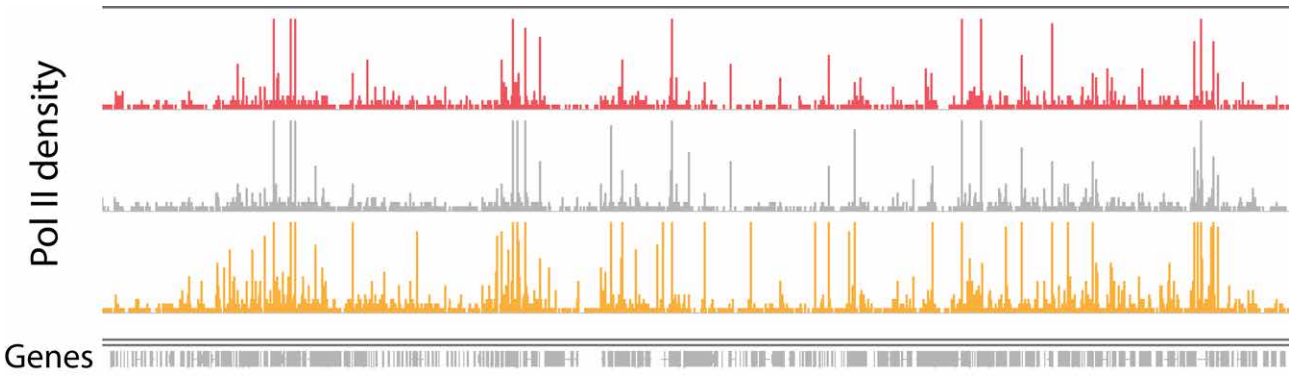


Figure 1: NET-Seq Tracks. Credit and copyright: Dr Andreas Mayer/Mayer Laboratory.

Figure 2: Pervasive Pausing. Credit and copyright: Dr Andreas Mayer/Mayer Laboratory.

pieces back together. The mRNA is then transported from the cell nucleus to the ribosome for translation, where the mRNA provides a template for protein building enzymes to build new proteins.

Cells regulate precisely the transcription of each gene and subsequent production of each required RNA and protein. However, the mechanisms controlling regulation have proven difficult to observe in living cells.

'The Captivating Complexity of Regulatory Mechanisms'

Dr Andreas Mayer, at the Max Planck Institute for Molecular Genetics (MPIMG) in Germany, describes how his interest in the 'captivating complexity of regulatory mechanisms' controlling Pol II transcription began with his PhD research, working in Professor Patrick Cramer's group at Ludwig-Maximilians University, Munich. In a summary of his career to date, he further notes, 'I have now spent more than a decade unravelling the regulatory principles governing RNA polymerase II transcription. This process is fascinating because it underlies nearly all fundamental eukaryotic cell processes.'

In Munich, Dr Mayer began using quantitative genome-wide approaches to discover new general control mechanisms of Pol II transcription in yeast. His work identified the importance of regulatory mechanisms that act downstream of transcription initiation. For decades, it was thought Pol II transcription was predominantly regulated during the initiation phase, but the available technology was not capable of providing accurate location of RNA polymerases that are engaged in transcription in living cells at a single nucleotide level resolution. Dr Mayer notes, 'Knowing the exact locations of transcribing RNA polymerases across the genome is an important step towards understanding how transcription is regulated and misregulated in diseases'.

In 2012, Dr Mayer relocated to Harvard Medical School (HMS) in Boston, Massachusetts, to pursue his goals. He explains, 'To overcome these major technical limitations and discover if post-initiation regulation played a role in mammalian cells, I joined Dr Stirling Churchman's laboratory.' The previous year, Dr Churchman and Dr Jonathan Weissman of the University of California, San Francisco, had developed a new technique for analysing transcription with high precision in yeast, called Native Elongating Transcript Sequencing (NET-Seq). NET-Seq is a high-resolution genome-wide approach, which provides a quantitative measure of newly formed RNA and that allows the genomic localisation of RNA polymerases at nucleotide resolution. However, the original NET-Seq protocol for yeast was not amenable for mammalian cells including human cells.

Dr Mayer continued, 'With my colleagues in Dr Churchman's laboratory I co-developed the human NET-Seq approach (Box 1: Human NET-Seq Approach), allowing us to define the Pol II transcriptional landscape with single-nucleotide resolution (Figure 1: NET-Seq Tracks) and, most importantly, to discover pervasive Pol II transcriptional pausing (Figure 2: Pervasive Pausing). This finding suggested that post-initiation transcription regulation seems to be much more prevalent than anticipated, leading me to wonder how we could identify the potential functions of these events.'

‘I am convinced that bridging the gaps between high-resolution functional genomics, transcription, and cell differentiation will allow important discoveries in the years to come.’

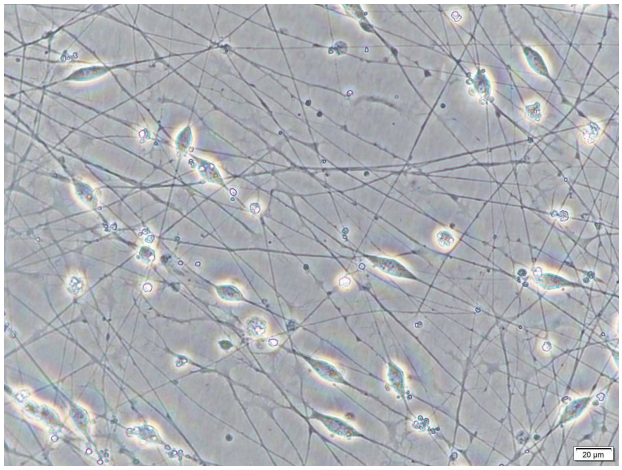


Figure 3. Neuronal Cells. Credit and copyright: Dr Olga Jasnovidova/Mayer Laboratory

The team’s critical finding that previously observed transcriptional pausing *in vivo* was much more prevalent than anticipated also expanded upon the earlier discoveries of widespread transcriptional pausing in the promoter-proximal region of genes in Drosophila and mammalian cells. RNA synthesis was observed to be a discontinuous process, during which phases of productive transcription were frequently interrupted by regular transcriptional pauses of Pol II during elongation and termination (Figure 2: Pervasive pausing). Whilst the causes and consequences of this are still poorly understood, it suggests that post-initiation transcription pausing enables many further additional opportunities for regulating gene expression after Pol II has begun transcribing the gene.

The open and creative working environment at HMS enabled a broad range of researchers from the Departments of Genetics, Genome sciences, and Medicine from HMS, and the University of Washington to collaborate on this work. Dr Mayer describes how he was able to broaden his skills and interests. ‘[The environment] enabled me to interact closely with Professor George Church’s laboratory. I was exposed to a range of cell differentiation systems and taught the practical skills required to handle and manipulate stem and differentiating neuronal cells.’

Pol II Pausing and Cell Differentiation

In 2017, Dr Mayer established an independent research group at the Max Planck Institute for Molecular Genetics (MPIMG) in Berlin, Germany. As he now explains, ‘Starting my own independent Research Group at the MPIMG…has enabled me to continue working on deciphering the regulatory mechanisms of Pol II transcription.’

The Max Planck research group’s primary goal is to reveal the principles that dynamically control and coordinate gene expression, and that drive cell differentiation using neuronal cells. Neuronal cells provide an ideal system to study, due to the extensive transcriptional changes that occur in neurogenesis, as embryonic stem cells differentiate into functional neurons (Figure 3: Neuronal Cells).

Dr Mayer enthusiastically describes the work of his new team. ‘We have successfully adapted a neuronal cell differentiation system, and I am excited that we are poised to use this model to investigate the dynamics of Pol II gene transcription and its function in cell lineage determination.’

Genome-wide approaches have already revealed that Pol II pausing is especially prevalent on genes that control cell development and differentiation, and there is a growing list of pausing regulatory proteins that tightly regulate the process. How the widespread transcriptional pausing controls gene expression and how it affects cell differentiation remains unclear, however.

The Mayer group is addressing these fundamental questions by bringing together a team with a range of interdisciplinary skills, using high-resolution genome-wide approaches, genome engineering techniques, genetics experiments, and bioinformatics tools. The team is also developing new quantitative methods to better investigate the molecular mechanisms that underlie transcription in mammalian cells.

Future Impact

The future work by Dr Mayer and his colleagues will undoubtedly contribute further insights into these unexplored and developing areas of research. The group aims to identify the general principles and dynamics of gene regulation during stem cell differentiation, in order to fully understand the complex and interrelated regulatory logic of genome transcription.

This work has broader implications too, with increasing evidence that misregulation of Pol II transcription and transcriptional pausing has a significant role in a broad range of human diseases and syndromes including cancer, autoimmunity, neurological disorders, diabetes, cardiovascular disease, and obesity. Dr Mayer’s work will address the molecular basis for many of these transcriptional defects and may ultimately contribute to new treatments to prevent or cure these conditions in the future. He concludes, ‘I am convinced that bridging the gaps between high-resolution functional genomics, transcription, and cell differentiation will allow important discoveries in the years to come.’

Meet the researcher

Dr Andreas Mayer
Max Planck Institute for Molecular Genetics
Berlin
Germany

Dr Andreas Mayer received his PhD (summa cum laude) in 2012 from Ludwig-Maximilians University (LMU), Munich, working as part of Professor Patrick Cramer's group studying genome regulation. In the same year, he was awarded the Paula and Richard von Hertwig Award from Helmholtz Zentrum München for interdisciplinary cooperation. In the following 5 years, Dr Mayer joined Dr Stirling Churchman's laboratory at Harvard Medical School in Boston, where he co-developed the human NET-seq approach as part of his specialist study of RNA polymerase II transcription. In 2012, Dr Mayer received an EMBO Long-term Fellowship and between 2013–2016, he was awarded the Human Frontier Science Program Long-term Fellowship. Since 2017, Dr Mayer has led his own independent research group at the Max Planck Institute for Molecular Genetics in Berlin. The group's work aims at bridging the gap between high-resolution functional genomics, genome transcription and stem cell differentiation.

CONTACT

E: mayer@molgen.mpg.de
W: https://www.molgen.mpg.de/mayer-lab
@mayer_lab

FUNDING

Max Planck Society
Deutsche Forschungsgemeinschaft (DFG)
Federation of European Biochemical Societies (FEBS)

FURTHER READING

A Mayer, HM Landry, LS Churchman, Pause & go: From the discovery of RNA polymerase pausing to its functional implications, Current Opinion in Cell Biology, 2017, 46, 72–80.

A Mayer, LS Churchman, A detailed protocol for subcellular RNA sequencing (subRNA-seq), Current Protocols in Molecular Biology, 2017, 120, 4.29.1-4.29.18.

A Mayer, LS Churchman, Genome-wide profiling of RNA polymerase transcription at nucleotide resolution in human cells with native elongating transcript sequencing. Nature Protocols, 2016, 11, 813–833.

A Mayer A, J di Iulio, S Maleri, U Eser, A Reynolds, J Vierstra, R Sandstrom, JA Stamatoyannopoulos, LS Churchman, Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution, Cell, 2015, 161, 541–554.