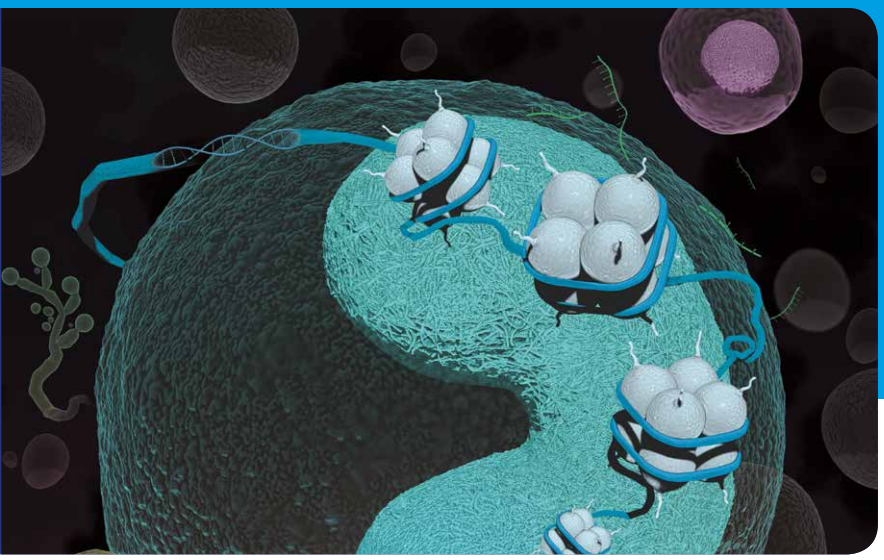


Bringing epigenetics into the world of medical practice

For years scientists have known that epigenetic changes that cells undergo during their lifetimes influence individuals' susceptibility to cancer and other diseases and their responses to treatment. Yet the use of such knowledge in clinical practice has been slow.

Now four new studies by members of the international BLUEPRINT consortium, headed by Christoph Bock of the CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences (Vienna) and Stephan Beck of the University College of London have brought epigenetics much closer to practical applications in clinics and hospitals for the benefit of patients world-wide.

Over a lifetime, the normal development of cells and their interactions with the environment comes with a long term epigenetic memory of cells. Cells "learn" to partition their DNA in regions where genes are active (and epigenetically accessible) and other regions where genes are silent (and epigenetically inaccessible). This epigenetic memory provides each cell with a specific repertoire of functions that not only influences its development but also its potential to respond to threats from diseases and its responses to treatment. Epigenetic features of cell types memorize challenges over a person's lifetime, for example changes in health status, a medical treatment, long periods of stress, and other events.

The current studies follow such memory changes and address two types of epigenetic signatures. The first is DNA methylation, a process in which chemical tags are directly attached to specific DNA sequences. These markings influence whether DNA and the molecules bound to it (called chromatin) are in an open configuration, making a gene accessible to factors that can activate it, or closed and inaccessible. The study looked at one mark of each type.

To observe these types of changes a second method, called ATAC-seq, comes into play that reveals the accessibility of genes. This method to scan chromatin structures represents the second landmark achieved in the studies, and it will be applicable to many other diseases as well. ATAC-seq directly distinguishes between different types of chromatin structures, and can reveal how regions around a gene change in states of health and disease. Cancers such as chronic lymphocytic leukemia, which are highly variable in the disease course, are often accompanied by stereotypical changes in chromatin that might be useful in diagnosing and treating individuals.

Robust and cost-effective epigenetic technologies

Changes in DNA methylation represent well-known hallmarks of particular diseases and a person's susceptibility to them. The changes can be scanned by a method called whole-

genome bisulfite sequencing (WGBS), which is the "gold standard" for detecting methyl tags at any position along the entire genome of a cell. However, finding the important points in the genome where the tags undergo changes (called differentially methylated positions, or DMPs), requires repeated and expensive applications of the method. This incurs high costs that prohibit its routine application to many individual cases. Additionally, WGBS produces complex maps with huge amounts of data, and extracting relevant information from them is a great challenge.

BLUEPRINT partners from Austria, Germany, the UK, US, Spain and the Netherlands carried out several studies to find a way to determine the quality and robustness of alternative (and cheaper) methods in comparison to the "gold standard" WGBS method, in hopes of making clinical applications much more feasible. First, Christoph Bock's group coordinated a comprehensive review of the technologies currently used for WGBS to identify those that fit several criteria: they should be fast, inexpensive, highly reliable, and widely available to scientists and clinicians. Did the methods identify specific biomarkers equally well?

"The aim was to target specific locations, analyze their methylation patterns, and compare the performance of many different technologies," Bock says. "We shipped 32 reference samples to 18 laboratories in seven countries. Each was asked to examine diagnostically relevant methylation patterns in a set of predefined regions, and to contribute the results to an integrative, consortium-wide benchmarking analysis."

This comparison of technologies between laboratories revealed that most currently used methods produce consistent results. Two methods – called amplicon bisulfite sequencing and bisulfite pyrosequencing – performed best overall. "Our results will help scientists and clinicians select the best targets and optimize their technologies to develop biomarkers and use them in clinical diagnoses," Bock says.

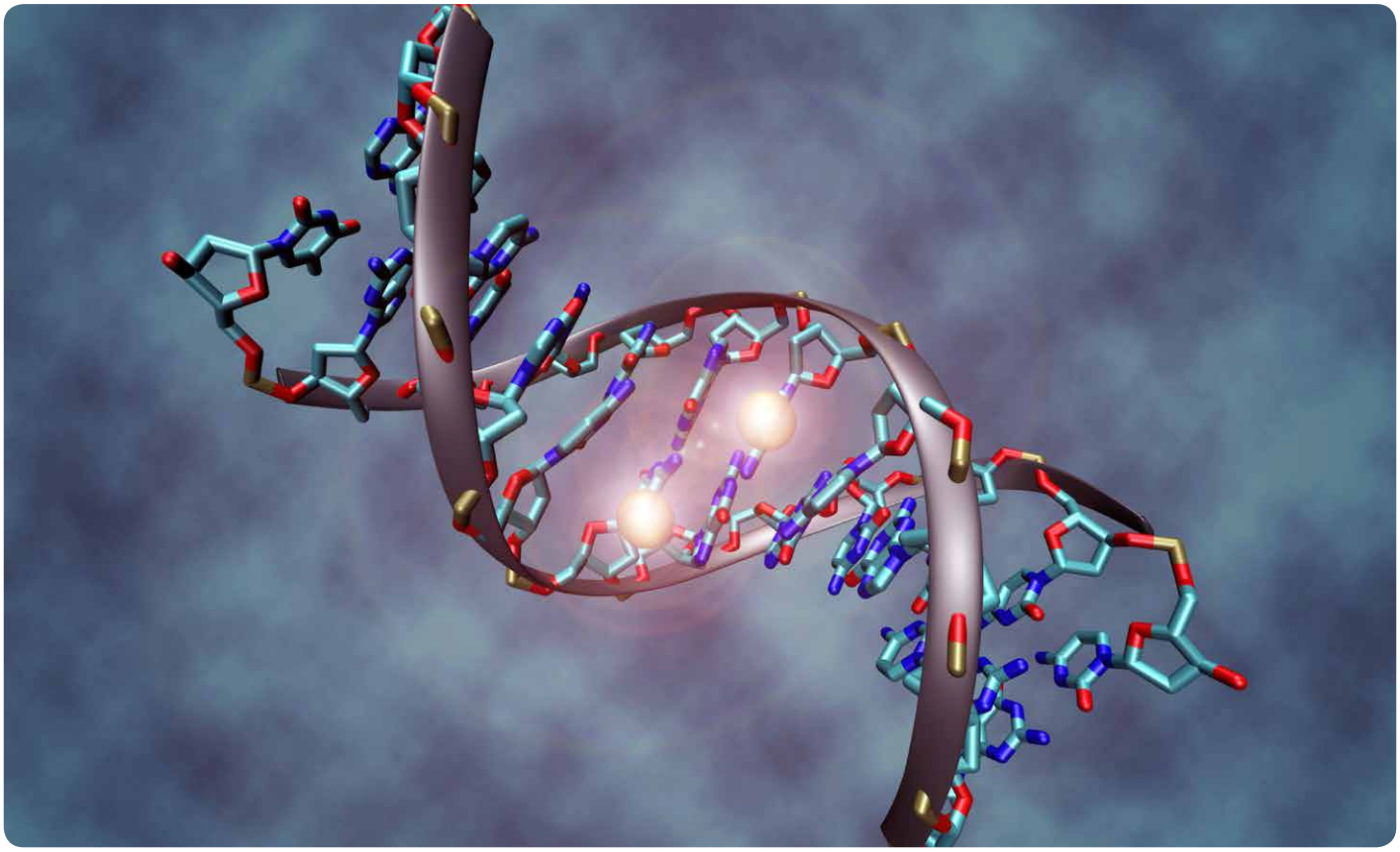
Data analysis

Extracting the relevant epigenetic data points from WGBS for a clinical application has a lot to do with data analysis – how many regions of DNA must be sequenced, in how many types of cells, how many times? "Until now, the standard recommendation for recovering sufficient diagnostic information in most cases has been to sequence the genome 30 times," says Stephan Beck. "This is not only very costly but produces a massive amount of data that must be analyzed and ultimately understood." Even with such repeated coverage, Beck says, up to 50% of information about single positions in the genome is routinely lost.

Perhaps there was a more efficient way to capture detailed information. Emmanuele Libertini and other members of Beck's London lab, collaborating with other groups in the UK, Austria, the US, the Netherlands and Spain, pored through the data to identify blocks of comethylation, or COMETs. The aim was to determine whether the changes in the methylation of particular positions in the genome (DMPs) could be reliably associated with other changes.

"If that were the case, determining what happened at one sequence could be used to predict what happened at others," Beck says. "This would reduce the number of rounds of sequencing and allow us to recover lost data."

The scientists created two new computational tools called COMETgazer and COMETvintage that could break blocks of the genome down into segments which could be used to predict methylation events at single positions in the genome. By comparing their predictions with reliable sets of data, they discovered that about 30% of information that was lost when the genome was only sequenced five times could be faithfully recovered by their COMET approaches. This represents another huge step toward lowering the costs of epigenetic tests. The scientists have made their tools freely available on-line.



Methylated DNA molecule. DNA methylation plays an important role for epigenetic gene regulation in development and cancer.
Source: © Christoph Bock, CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences

Epigenetic signatures of disease

DNA methylation and other epigenetic marks determine how chromosomes are organized and packed within the nucleus – determining which genes are accessible and which are not.

Accessible genes can be reached by factors that will transcribe them into RNA molecules and ultimately translate them into proteins, usually the most critical molecules for cell functions. Which genes are "on" and which genes are "off" in a cell are therefore a key determinant for the way a cell functions.

Changes in the organization of chromosomes are likely to be linked to chronic lymphocytic leukemia (CLL), the most common form of leukemia in the Western world, which affects mainly the elderly. A key question for scientists is whether disease-specific changes in chromatin structure can be used for CLL diagnoses, possibly even in distinguishing between different subtypes of CCL cancers. This would provide a new route for CLL-specific and patient-specific treatments of the disease.

"This common disease manifests itself in a number of sub-types that have different patient outcomes, making it a challenge to diagnose and determine appropriate treatments for individuals," Christoph Bock says. "Currently to achieve an adequate diagnosis, clinics have to analyze a range of features such as mutations in genes, serum markers, levels of proteins in specific blood cells, and so on. Epigenetic studies can be an important complement to existing methods."

Scientists had a number of hints that epigenetic changes played a key role in the development of CLL. ATAC-seq reliably exposes the looser regions of the genome. Comparing maps of these regions obtained from CLL patients with those of healthy people, or those suffering from different forms of the disease, might provide a useful diagnostic tool.

The scientists used the method to analyze samples from 55 CLL patients and compared the resulting "chromatin profiles" to data from other types of epigenetic analyses. „We developed a method to link our data to clinical observations and other types of molecular diagnoses," Bock says. "Then we examined the genes in interesting regions to correlate them with molecules that are known to be disrupted in CLL." Similar analyses were carried out on cells from patients that were undergoing treatment or had suffered a relapse of the disease.

The result, he says, is a set of data that is accessible to clinicians, researchers and the public. Bock and his colleagues believe that further analysis of the regions will permit scientists to develop new hypotheses about the causes of various types of CLL and the mechanisms that cause them to arise. All of this information has key relevance for the diagnosis and treatment of the most common type of leukemia world-wide. The general method is equally applicable to patients suffering from a range of other diseases.

References:

- *Christoph Bock, Florian Halbritter, Francisco J Carmona, Sascha Tierling, Paul Datlinger, Yassen Assenov, María Berdasco, Anke K Bergmann, Keith Booher, Florence Busato, Mihaela Campan, Christina Dahl, Christina M Dahmcke, Dinh Diep, Agustín F Fernández, Clarissa Gerhauser, Andrea Haake, Katharina Heilmann, Thomas Holcomb, Dianna Hussmann, Mitsuteru Ito, Ruth Kläver, Martin Kreutz, Marta Kulis, Virginia Lopez et al. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. Nature Biotechnology 34, 726–737 (2016) doi:10.1038/nbt.3605*
- *Emanuele Libertini, Simon C Heath, Rifat A Hamoudi, Marta Gut, Michael J Ziller, Javier Herrero, Agata Czyz, Victor Ruotti, Hendrik G Stunnenberg, Mattia Frontini, Willem H Ouwehand, Alexander Meissner, Ivo G Gut & Stephan Beck. Saturation analysis for whole-genome bisulfite sequencing data. Nature Biotechnology (2016) doi:10.1038/nbt.3524*
- *Emanuele Libertini, Simon C. Heath, Rifat A. Hamoudi, Marta Gut, Michael J. Ziller, Agata Czyz, Victor Ruotti, Hendrik G. Stunnenberg, Mattia Frontini, Willem H. Ouwehand, Alexander Meissner, Ivo G. Gut & Stephan Beck. Information recovery from low coverage whole-genome bisulfite sequencing. Nature Communications 7, Article number: 11306 doi:10.1038/ncomms11306*
- *André F. Rendeiro, Christian Schmidl, Jonathan C. Strefford, Renata Walewska, Zadie Davis, Matthias Farlik, David Oscier & Christoph Bock. Chromatin accessibility maps of chronic lymphocytic leukaemia identify subtype-specific epigenome signatures and transcription regulatory networks. Nature Communications 7, Article number: 11938 doi:10.1038/ncomms11938*