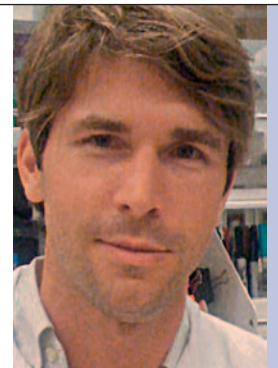


CONTENT *With their decision to fund the Roadmap Epigenomics Initiative, the National Institutes of Health (NIH), USA, has paved the way for the most comprehensive project ever embarked upon for the research of epigenetic modifications and their significance in genetic regulation and cell differentiation. Critics have expressed doubt as to whether the ends justify the means and whether the project warrants the US \$190 million that it will cost. Alexander Meissner (Department of Stem Cell and Regenerative Biology of Harvard University in Cambridge, USA) is involved in this project. In this interview, he defines the goals of and opportunities presented by the Roadmap project and comments on some of the controversy.*

A roadmap for epigenetics

**Interview with Alexander Meissner,
Harvard University and Broad Institute of MIT and
Harvard, Cambridge, USA**



Dr. Meissner, ever since Conrad Waddington coined the term 'epigenetics' in the early forties, the definition of the word has been used in many different contexts. What do we understand by the term epigenetics nowadays?

• The prefix 'epi' means 'in addition to' or 'on top of' genetics. Waddington used it to describe interactions in the genome that eventually lead to the development of a certain phenotype, such as is found in the differentiation of cells. Later, in the mid-seventies, when DNA methylation was attributed a role in genetic regulation, the term 'epigenetics' stood for mechanisms which are heritable via cell division, i.e. mitosis, and which therefore stably influence gene expression without altering the DNA sequence.

The modern definition, as used in the Roadmap Epigenomics Initiative of the National Institutes of Health, and which I would agree with, is somewhat more general. Here the word epigenetics embraces a whole

spectrum of different regulatory mechanisms that do not change the DNA sequence.

So what is the bone of contention?

• From my point of view, it is largely a semantic discussion regarding the definition of the word. It revolves around whether or not the modifications are necessarily hereditary. We at the Broad Institute believe that for many of the goals of the project it is irrelevant whether a modification is passed on from one cell to the next, as long as it has a gene-regulatory effect. Even a purely descriptive map that provides new insights of a particular cells 'normal' state, can have an important impact on our understanding of human diseases. This applies to the classical DNA methylation as well as to histone modifications and possibly even non-coding RNAs. Just recently, Science magazine published two papers linking non-coding RNAs to the regulation of histone modifications.

But only DNA methylation is broadly accepted nowadays...

• That is correct, although evidence has emerged that at least some histone marks are also mitotically heritable. In the 1970s, DNA methylation was postulated as the classical epigenetic modification involved in the regulation of gene expression. The methyl group is covalently attached to the cytosine base in the DNA. The enzymes responsible for the establishment and maintenance of DNA methylation patterns were identified in the nineties. When the methyl group is added, it usually has a repressive effect.

It is also clear that the methylations can be passed on from one cell to another, i.e. that they are heritable. One well-known class of epigenetically regulated loci are imprinted genes: here the maternal or paternal copy of a gene is silenced through DNA methylation in a parent-of-origin-specific manner.

And this criterion of heredity does not necessarily apply to histone modifications?

- It seems to apply to a few, but many other cases are still unclear. When it comes to the histone modifications, things get considerably more complicated. Here we are faced with a whole series of different modifications such as acetylation, methylation, phosphorylation and ubiquitination – each of which is attached to various amino-acid residues at the N-terminal tail of the histone proteins, for example to lysine (K) or arginine (R) residues. Only the combination of various marks allows a meaningful interpretation of the so-called histone code. For example, H3K4-trimethylation is a modification that suggests an active state and this is true if the mark is present alone. However, if combined with a second mark, H3K27-trimethylation, than the meaning changes and the gene is inactive, but ‘poised’ for later activation. This explains the complexity and the need for a comprehensive project that will provide as many marks as possible. It would be beyond the scope of this interview to try to list all the histone modifications that have been identified.

Could you perhaps give us a few more details on some prominent examples of histone modification?

- The two examples mentioned above, are among the best studied. H3K27-trimethylation, the addition of three methyl groups to the lysine 27 residue on histone H3, has a repressive effect, and the enzyme responsible for this modification is well known. The histone methyltransferase Ezh2 is part of the polycomb repressive complex 2, in short PRC2. It was originally identified in *Drosophila* and has been conserved from fly to mammals. There is some evidence that this modification could be heritable and that it therefore falls within the classical definition of epigenetics.

Another mark is H3K9-trimethylation, which is found at a different position in the genome than H3K27-trimethylation, but has a similar repressive effect and also appears to be heritable, as experiments in yeast have indicated.

Does that mean other histone modifications do not fit the classical definition?

- Well, in many cases we simply don’t know. For example, we don’t know whether or not the H3K36-trimethylation is heritable. But regardless, one thing is for sure; the H3K36-trimethylation is invariably found at points in the genome that are in the process of being transcribed, making it a good marker for active genes and RNA transcription. In other words, it could provide us with plenty of information on where transcription takes place in the genome. That is precisely what makes it so valuable to us – regardless of whether or not it is hereditary. For instance, the H3K36-trimethylation provides us with new tools in our search for the origins of micro-RNAs. It is known that these originate from long primary transcripts, but it proved very difficult to locate these based on the very short sequence of the micro-RNAs. Now we can very effectively use the H3K36-trimethylation mark together with the above-mentioned mark for active promoters, H3K4-trimethylation, and simply search for regions that are transcribed, but are non-coding.

The H3K36-trimethylation also leads to a better understanding of what genetic regulatory mechanisms are all about?

- Yes, that is the crux of the matter as far as I am concerned. We examine these epigenetic modifications because we want to learn more about the mechanisms of gene regulation. We want to understand how the same genome can generate this variety of cell types, and stably maintain their

identity without altering the actual DNA sequence. We are convinced that we are on the right track. That’s why I also find it rather futile to get bogged down in discussions about definitions. That doesn’t get us anywhere. In principle, I guess we could simply avoid this and give it a new name.

The NIH has now committed itself to the most comprehensive project ever launched for the investigation of epigenetic modifications. Why does this venture attract such a tremendous amount of attention?

- That is difficult to say. First of all, let us not forget that the NIH Roadmap is a special funding mechanism that supports a variety of initiatives, epigenetics being one of the most recent of these projects. One reason why the decision fell to support it now could be related to the fact that a number of projects and networks in the last few years were dedicated to the epigenome. A considerable amount of pioneering research has therefore already been carried out, particularly in Europe. More and more scientists recognized that it was time to launch a large-scale project to go about things more systematically. Remember, not only do we have many different epigenetic modifications, but every cell type also has its own epigenome. Such a large and complex field can be best explored if the community as a whole comes together. Another reason is that, through recent technology developments, we are now finally able to tackle such a major project in the first place. Up until a year ago, that would have been unrealistic. Thanks to new sequencing techniques, several million sequences can now be examined simultaneously and large numbers of sequencing reads is the basis for all of the techniques that investigate these modifications on a large scale. This way, a large number of different cell types and modifications can be examined at the same time on a genome-wide scale.

What does the Roadmap Epigenomics Initiative actually investigate? What is the goal of the project?

- As the name of the project suggests, the idea is to draw up a map – to gain a survey – of the modifications at hand and their contribution to genetic regulation. We aim to identify the differences between most types of cells, i.e. those modifications peculiar to a certain phenotype, a differentiation status or a cell type. This should provide a better understanding of what happens during the normal development of a cell and what changes in diseases like cancer. This will involve two complementary components. Some of the groups involved in the project will examine all known modifications in a small number of cell types, whereas the other groups will closely examine only a smaller number of modifications that are known to be particularly informative in a maximal number of different cell types. This way, the project can cover a reasonable part of the vast epigenomic space – many modifications that differ among many cell types. In the future this will need to be further expanded to also account for inter-individual variation as well as changes with age.

That all sounds as if an immense amount of data is generated. However, critics doubt whether these measures will have the desired effect. In the journal Nature, Kevin Struhl from the Harvard Medical School in Boston, Massachusetts, while agreeing that epigenetics are of importance, still expressed doubt as to whether such a large-scale project costing US \$190 million is really justified.

- First of all, I think that this discussion fails to take into account several basic benefits of such large-scale projects. Smaller ventures, with only a small number of scientists, lose the economy of scale that is required for the development of high-throughput pipelines and technologies. The point is that technological progress – and

thus scientific progress – will be promoted more rapidly. At the end of the day, this is doubtless to the advantage of the whole scientific community, in particular the individual investigator, who will gain access to new technologies and large amounts of data. As for the roadmap, I am convinced that by examining and comparing many different epigenomes from a large array of cell types, that we can gain new insights into normal development and gene regulation that will provide a better understanding of human diseases.

But here again, some scientists argue that, since DNA modifications are highly dynamic and the state of modification in a cell can change quickly, it would be more prudent to tackle the factors or transcription to elucidate gene regulation.

- Well, the truth probably lies somewhere in-between, which is why it shouldn't be a case of everyone working for himself. If we combine our epigenetic maps with the knowledge on transcription factors, then we'll start getting somewhere. In many ways, the transcription factors and epigenetics problem is rather like the question of the chicken and the egg – what came first? We know that transcription factors play an important role in genetic regulation. However, they too are controlled – either by other transcription factors or epigenetic modifications. We know, for example, that a whole series of transcription factors are repressed in embryonic stem cells. These are then later activated at various points during the differentiation, according to the cell type. Others must be deactivated for differentiation to take place to begin with. Both of these classes are known to be regulated through epigenetic modifications.

Can you give us an example of this?

- One classical example is *Oct4*, one of the typical embryonic transcription factors that must be deactivated

to allow proper differentiation. This involves the DNA methylation of the *Oct4* promoter, which ensures that the transcription factor remains stably repressed.

These days, however, precisely these embryonic stem cells, or, to be more precise, the reprogramming of somatic cells, are claimed to emphasize the importance of regulatory proteins. In 2006, Takahashi and Yamanaka from the University of Kyoto showed that four proteins are sufficient to reprogramme a somatic cell.

- That's correct and I don't intend to contradict this. However, we must also bear in mind that these factors were artificially introduced into the cell. That doesn't correspond to natural conditions in the organism. If it were that easy, we could take a certain cell A with, let's say 100 different transcription factors and a cell B with 100 others and then swap their transcription factors around. Although we would expect cell A to become cell B and *vice versa*, this would not be the case, since epigenetic modifications are cell-type-specific and will – at least temporarily – provide a barrier for such a cell-state switch. That is one explanation why reprogramming with the four Yamanaka factors takes several weeks.

How did your own results compare when you carried out similar experiments?

- At the Broad Institute, we analysed the whole process in more detail and compared the reactivation of inactive genes with the underlying epigenetic maps. Here an obvious trend could be observed; genes that had been stably repressed by H3H27-trimethylation or DNA methylation reacted only very slowly, or not at all, to the expression of the four 'Yamanaka' factors. However, genes found in 'accessible' chromatin were activated quickly and efficiently. This illustrates just how much valuable infor-

mation epigenetic maps contribute to our understanding of biological processes. At the same time, it shows how we can selectively improve the process. The efficiency of reprogramming can be considerably improved if the DNA is demethylated, for instance by treating the cells with drugs such as 5-azacytidine, or by increasing the acetylation with histone deacetylase inhibitors that make chromatin more accessible. This clearly shows that both the transcription factors and the epigenetic modifications play a decisive role.

Does this also explain why nucleus transfer frequently fails in mice and other species?

- Exactly. If the transcription factor environment were the only factor involved, then one would expect the nucleus transfer from a somatic cell into an enucleated oocytes to work perfectly – 100% as long as there is no physical damage. But that is not the case. We know that many cloned embryos cannot develop because reprogramming is incomplete and the methylation pattern does not correspond to that of earlier embryos or stem cells. For example, in many clones, *Oct4* remains DNA methylated and inactive. As I had mentioned earlier the purpose of the DNA methylation at the *Oct4* promoter is to ensure that it stays silent. It resists the influence of the transcription factors, at least for a while.

How else can we expect to profit from the knowledge gained by the initiative? What are the major long-term goals?

- The goal is to gain a fundamental understanding of the underlying

mechanisms of gene regulation and how it changes during the differentiation process. Once we know that we can tackle the more complex question of what causes certain illnesses. In some kinds of cancer, for example, erroneous epigenetic regulations are responsible for the methylation of tumour suppressor genes. In neurological illnesses with a genetic component, such as Parkinson's disease, an alteration in gene expression might possibly be linked to epigenetic alterations. Knowing all these disease-associated changes can help in many ways. Changes could be used as powerful markers for disease detection and prognosis or for the development of more targeted therapies.

Are there already any tangible indications that this might work?

- One example that I personally find particularly interesting is that of the myelodysplastic syndrome, a disease of the haematopoietic system in which the formation of the red and partially also of the white blood cells is disrupted. This leads to anaemia and to lack of iron. The susceptibility for this disease increases with age and epigenetic mechanisms may well be involved. The best evidence for that is that the standard treatment includes the use of a DNA methyltransferase inhibitor – i.e. an agent that acts on the level of epigenetic modification. We do not yet have any further evidence that epigenetic mechanisms are the cause of this disease, but it does seem very reasonable. But if we now compare the epigenetic pattern of healthy and of diseased cells and examine what has changed on this level, we might gain some more specific insights on the mechanisms of the illness and search for better inhibitors that intervene more selec-

tively in the process. One of the major goals of the Roadmap project is to have reference maps for any normal cell type, so that individual investigators can simply compare their specific epigenetic pattern, at their gene of interest, in their disease sample to the 'normal' map.

Is such applied research envisaged by the NIH-Initiative?

- Yes indeed, the first phase in which we participate will generate the normal reference maps. The next round of research projects granted will be devoted to specific medical issues. The applications have already been received and these projects are scheduled to begin in about a year's time. I am certain that we will recognize before long just how much we are going to gain from the information provided by the Roadmap Initiative. The way I look at things is this; most large-scale initiatives and projects in the last years have seen major objections in the initial phase. The Human Genome Project and the HapMap Project both had their fair share of negative publicity before they got off the ground. After completion, scientists have clearly recognized these ventures as tremendous breakthroughs. Criticism is also useful, because it makes us pause to reflect critically on our own work and to check that we are still on the right track. But one thing is for sure; at the end of the day, all of these projects, including the epigenome project, invariably benefit the science community at large.

Dr. Meissner, thank you very much for talking to us.