

Novel Combination of Chromatin Modifications Governs Embryonic Stem Cell Maintenance

Harvard Medical School's Bradley Bernstein talks with UT Southwestern's Ryan Weil about how bivalent methylation may control the timing of cell division and differentiation in stem cells

By Megha Satyanarayana

Harvard researchers, led by Bradley Bernstein and Eric Lander, have discovered a novel feature of chromatin in embryonic stem (ES) cells that keeps these cells poised to become different tissues, once they get the molecular go-ahead signal. This finding is critical to understanding what makes stem cells

differentiate into the myriad tissues of the body.

The research team used custom tiling microarrays to examine 56 genomic regions of mouse embryonic stem cells. They found that the protein spools, or histones, around which DNA is wound into chromatin contained two methylated

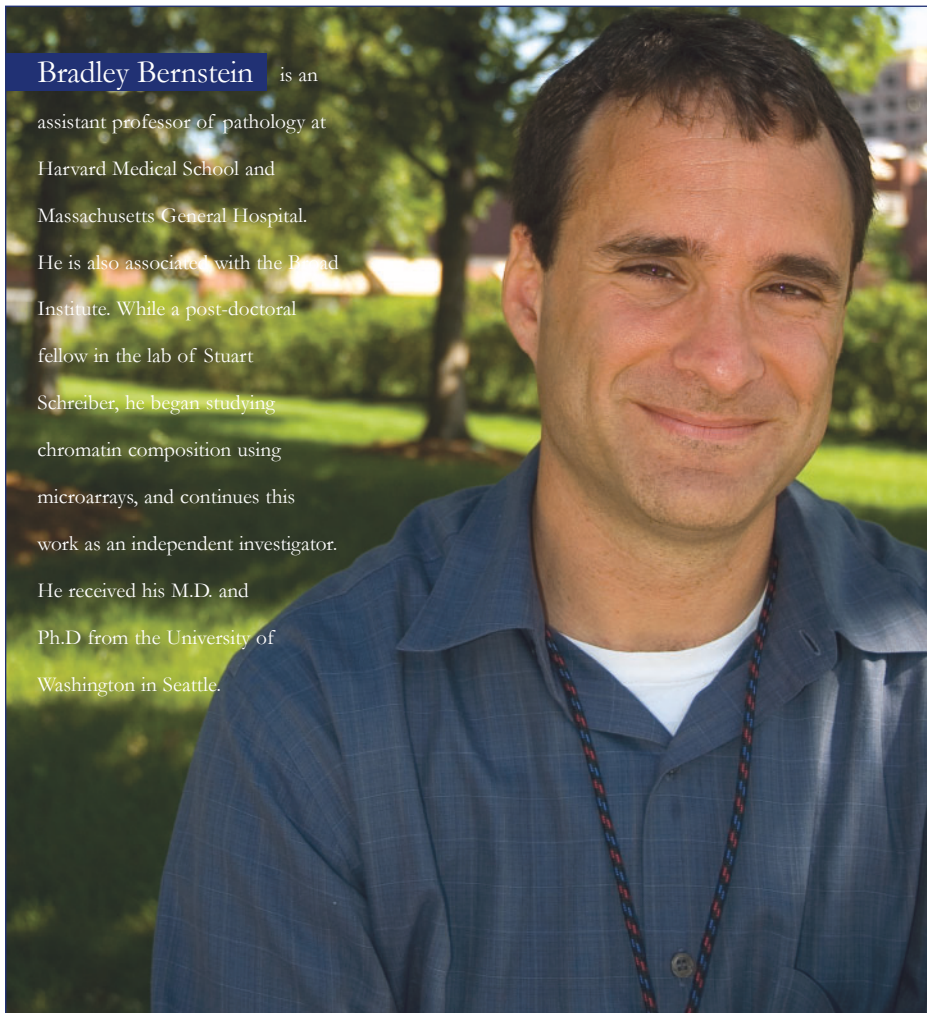
lysines, known as K4 and K27. The group's work was published in the April 21, 2006 edition of *Cell*.

"The really surprising result was that when we looked in embryonic stem cells, where developmental genes have yet to be committed to defining the lineage, there was a different epigenetic structure, which we called "bivalent," said Bernstein, assistant professor of pathology at Harvard Medical School, and lead author of the study. "It had characteristics both of active and silent chromatin."

Methylation at K27 usually serves to keep the DNA bound tightly to the histone, keeping it unavailable for transcription. K4 methylation loosens the hold of the histone on the DNA, allowing genes to be transcribed. The researchers believe that the combination of methylated lysines in ES cells may be the key to keeping developmental genes turned off in embryonic stem cells while ensuring that these cells maintain their pluripotency and remain at-the-ready.

The team also found that the methylation patterns in ES cells correlated with specific DNA sequences. K4 methylation occurs in CpG islands and K27 methylation occurs in regions free of transposable elements.

Bernstein recently spoke with Ryan Weil, a recent graduate of UT Southwestern Medical Center in Dallas, about using tiling arrays to study chromatin in ES cells. Weil's graduate work focused on developing a microarray-based system that allows researchers to directly measure the



Bradley Bernstein is an

assistant professor of pathology at Harvard Medical School and Massachusetts General Hospital.

He is also associated with the Broad Institute. While a post-doctoral fellow in the lab of Stuart Schreiber, he began studying chromatin composition using microarrays, and continues this work as an independent investigator.

He received his M.D. and Ph.D from the University of Washington in Seattle.

chromatin condensation state of DNA.

The two discussed:

- How DNA modifications may act in concert to control cell differentiation
- Development of a custom tiling array to look at chromatin structure
- Future studies and the call for an “epigenome project”

Tiling arrays and histone modifications

Weil: Would you give a brief introduction to chromatin, its modifications and how you started studying these processes?

Bernstein: Sure. Eukaryotic genomes are organized into a higher ordered structure called chromatin, which consists of DNA associated with histone proteins. Histones are subject to chemical modifications, such as methylation, that regulate the accessibility and consequently, the expression of the genes they are associated with. One of the reasons we are so interested in histone modifications is that evidence in model systems suggests that these modifications are maintained through cell divisions and may now represent an additional layer of heritable information. This is the basis of epigenetics. We think that histone modifications may play an important role in maintaining gene expression patterns and in defining and maintaining cell identity.

In our earlier studies using tiling arrays, we found that the Hox gene cluster contained broad domains of modified histones — very large regions of up to 100 kilobases in size — that we thought contributed to the epigenetic maintenance of Hox gene expression or repression in a particular cell type. So, we went back to

The really surprising result was that when we looked in embryonic stem cells, where these developmental genes have yet to commit to a specific lineage, there was a different epigenetic structure, which we called “bivalent”. It had characteristics both of active and silent chromatin. The regions had large domains of K27 methy-

“One of the reasons we are so interested in histone modifications is that evidence in model systems suggests that these modifications are maintained through cell divisions and may now represent an additional layer of heritable information.”

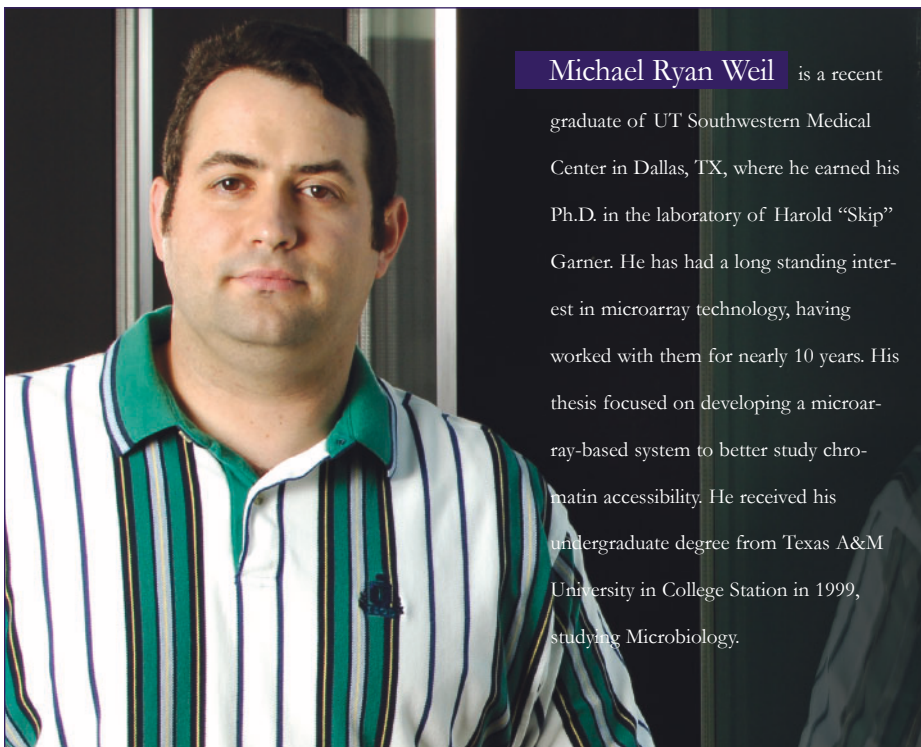
tiling arrays and looked at a number of developmental loci in differentiated cells, to see whether such broad domains of chromatin were a more general organizational feature of mammalian genomes. Indeed, we found that in differentiated cells, there are many broad domains of chromatin that exhibit characteristics of either active chromatin, in that they show K4 methylation or condensed chromatin with K27 methylation.

lation overlapping smaller sites of K4 methylation and tended to coincide with the transcription starts of key developmental genes or, as one might call them, the master regulator genes. So, this and supporting data really suggest that in embryonic stem cells, master regulator genes are regulated by a unique epigenetic structure that has characteristics of both active and repressed chromatin and this seems to silence the master regulators in the embryonic stem cells while maintaining their potential to activate during development.

Weil: In order to study this model system in a fashion that is both high resolution and high throughput, you needed to custom design tiling arrays for your regions of interest. What challenges did you face when designing probes and the array itself?

Bernstein: We were one of the first groups to make custom tiling arrays and of course there were a number of glitches along the way that had to be dealt with. Fortunately, we caught them by the bioinformatics stage. By the time we made a mask and had the arrays synthesized, the sequences were accurate and the resulting product was exactly what we were shooting for.

We used both perfect match and mismatch probes. We are doing tests to



Michael Ryan Weil is a recent graduate of UT Southwestern Medical Center in Dallas, TX, where he earned his Ph.D. in the laboratory of Harold “Skip” Garner. He has had a long standing interest in microarray technology, having worked with them for nearly 10 years. His thesis focused on developing a microarray-based system to better study chromatin accessibility. He received his undergraduate degree from Texas A&M University in College Station in 1999, studying Microbiology.

figure out whether we need both of them or whether we can get by with perfect match alone. We want to determine at what level we can rely on these data. We know that when we set a high threshold and use the statistical analyses developed by both our group and Tom Gingeras' group at Affymetrix, we can identify modified regions using the arrays, a very high proportion of which validate by rtPCR. We are currently working to understand the degree to which the data are quantitative, and whether weakly modified sites can be identified through improved analysis techniques.

To interpret the data, we used a method where an approximately 500-base pair window is defined. The signal within that window is quantified and compared to the same set on a control array and, using a probabilistic model, one can identify the degree to which that genomic region was enriched in the immunoprecipitation step. That is essentially how we identified enriched sites using the array method for modified regions of the genome.

Modeling mechanism and association with genomic features

Weil: After reading the paper, I started thinking of a metaphor for describing a potential mechanism for bivalent chromatin controlling early transcription. In drag-racing, the cars wait for the green light, the driver's foot is on the gas, but the transmission not yet engaged. That way when the driver gets a green light, the engine's revved and ready to go, and all that needs to be done to get off the line quickly is put the car in gear, so the wheels start turning. Is that how you see bivalent chromatin domains controlling transcription, with the genes ready to go, and just waiting for the signal to start transcription?

Bernstein: I think there are a number of analogies one can use, and yours is perfectly reasonable. Existing data really indicates that bivalent domains function to repress critical master regulator genes. Yet they clearly allow their subsequent activation during development. My sense is that these developmental loci are where

we find the bivalent domains because they control some of the most critical genes in the organism. We think that the combination of K27 and K4 methylation helps to precisely regulate the expression of these genes, to keep them silent in the embryonic stem cell and to be sure the proper epigenetic controls are put in place upon differentiation.

Weil: Have the triggers for switching back to monovalent chromatin been identified?

Bernstein: We've done a number of experiments in differentiated cells. What we see is that essentially all the bivalent domains resolve to either K4 methylation alone or K27 methylation alone. This state correlates on a global level with the transcriptional status of the genes in the region, whereby active genes are associated with the K4 methylation and silent genes become associated with K27 methylation. But we really don't know much about the molecular mechanisms behind this yet. You can imagine possible mechanisms might include histone replacement, histone demethylation or more passive processes, but we don't yet have data to address this.

Once differentiation is triggered, the process isn't homogenous, so you end up with a heterogeneous population of cells. Because you need a synchronous population to do the experiment, it's quite difficult to look at the dynamics with which these domains resolve or to identify the triggers and regulatory factors that are involved.

Weil: What is the link between bivalent domains and CpG islands in transcriptional control?

Bernstein: Good question. CpG islands are these CG-rich clusters defined by genomic sequence. In general, we understand from evolutionary principles that CpG islands should actually be DNA methylation-free. To be clear here, the DNA methylation is distinct from the histone methylation which we examined. We found that histone methylation at K4 actually occurs mostly at CpG islands in the ES cells. Interestingly, the proteins that methylate histones at K4 have affinity for un-methylated CpG dinucleotides.



Bradley Bernstein

So, it proposes a nice model in which CpG islands that are un-methylated at the DNA level recruit proteins that promote methylation of histones at K4.

Weil: Interesting. So how do highly conserved noncoding elements (HCNEs) fit in here?

Bernstein: What we have learned in this paper is that the epigenetic modification patterns in the embryonic stem cells correlate exquisitely with specific DNA sequences in genomic features. As I mentioned, K4 methylation correlates very nicely with CpG islands. We also see that there is a specific genomic feature that correlates very nicely with K27 methylation in that there are these large regions without transposable elements that are highly correlated with K27 methylation. We really don't know the mechanism yet, though it may be that the transposons are acting as boundary elements, stopping the spread of K27.

These same regions also have a very high number of these highly conserved noncoding elements (HCNEs). But we really don't yet know what the HCNEs are doing. The HCNEs tend to be enriched across regions of hundreds of kilobases, if not a megabase — much larger than the bivalent domains. So HCNEs really can't simply be explained as being necessary for bivalent domains to occur. They must be doing something else. We think they may have some role in higher order chromatin or chromosomal



Michael Ryan Weil

organization, but it remains a mystery exactly what these highly conserved non-coding elements are doing.

Weil: So, given the gene expression similarities in stem cell maintenance and oncogenesis of differentiated cells, would you expect to find renewed bivalency in tumor cells?

Bernstein: This is an incredibly important question to ask. We want to do those experiments, but there are some real technical hurdles that we face. One of the clear limitations of the ChIP-on-chip method is that we need to start with a very large population of homogeneous cells — on the order of 10 million cells to get enough material for the tiling arrays. To ask whether the cancer stem cell has epigenetic characteristics that are reflective of either embryonic stem cells or of

adult stem cells, we will need a few technological breakthroughs.

Future studies and the epigenome

Weil: So, with what you have learned about embryonic stem cells and bivalent domains, what are you going to do next?

Bernstein: We are very interested in the initial epigenetic state, how modifications in the ES cells are established, and in particular, the role of the DNA sequence in the establishment of the epigenome. We are looking at ES cell mutants that lack particular chromatin enzymes and seeing whether these have defective bivalent domains and what are the consequences of these mutations on the regulation of the associated genes.

We are also interested in seeing how pervasive the bivalent domain may be in adult stem cells and cancer stem cells. Our initial studies suggest that bivalent domains are a unique signature to embryonic stem cells, but at a more limited level, some bivalent domains may persist in adult stem cells or other multipotent tissues.

Weil: So, Peter Jones, past president of the AACR, is calling for a large-scale effort to sequence the epigenome in the same way that the regular genome was sequenced. There is a lot of interest right now in high throughput methods for studying the epigenome. Where do you see your method fitting into this work?

Bernstein: I think this is an incredibly important point that Dr. Jones is making. Epigenetic deregulation is broadly implicated in human disease. Especially in cancer where many chromatin regulators are known oncogenes or tumor suppressors, and inhibitors of histone modifying enzymes are promising therapeutic agents. So, for so many reasons, it would be informative to obtain a comprehensive inventory of the epigenetic state, both of normal and diseased tissues, including and beyond cancer. I think the technology is there or nearly there to do the experiments, and in particular, tiling arrays are a promising method for looking at histone modifications on the genome-wide level.

AFFYMETRIX MICROARRAY BULLETIN

Editorial Staff

Wes Conard, *Editor-in-Chief*
wes_conard@affymetrix.com
Tommy Broudy, *Managing Editor*
thomas_broudy@affymetrix.com
Rachel Shreter, *Editor*
rachel_shreter@affymetrix.com
Kamalia Dam, *Associate Editor*
Stacey Ryder, *Associate Editor*
Stephen Zins, *Assistant Editor*
Stephen Antonopoulos, *Assistant Editor*
Marva Maida, *Graphic Designer*

FOR MORE INFORMATION

Contacts

■ Bradley Bernstein, M.D., Ph.D.
Assistant Professor of Pathology
Harvard Medical School
149 13th Street, Rm 7148
Charlestown, MA 02129
bbernstein@partners.org

■ Michael Ryan Weil, Ph.D.
McDermott Center for Human Growth and Development
UT Southwestern Medical Center in Dallas
6000 Harry Hines Blvd.
Dallas, TX 75390
ryan.weil@gmail.com

Companies

■ Affymetrix, Inc. - <http://affymetrix.com>

Organizations

■ Harvard Medical School -
<http://hms.harvard.edu/hms/home.asp>

- UT Southwestern Medical Center in Dallas - <http://www.utsouthwestern.edu>
- Mc Dermott Center for Human Growth and Development - <http://www.utsouthwestern.edu/utsw/home/research/mcdermott/index.html>
- AACR (American Association for Cancer Research) - <http://www.aacr.org/>

People

- Peter Jones, Ph.D., D.Sci.
Past President, American Association of Cancer Research http://www.usc.edu/programs/pibbs/site/faculty/jones_p.htm
- Rudolph Jaenisch <http://web.mit.edu/biology/www/facultyareas/facresearch/jaenisch.shtml>

Further Reading

- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. A bivalent

- chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006 Apr 21;125(2):315-26.
- Huebert DJ, Bernstein BE. Genomic views of chromatin. *Curr Opin Genet Dev*. 2005 Oct;15(5):476-81. Review.
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005 Jan 28;120(2):169-81.
- Weil MR, Widlak P, Minna JD, Garner HR. Global survey of chromatin accessibility using DNA microarrays. *Genome Res*. 2004 Jul;14(7):1374-81.