NEW FINDINGS CHALLENGE THE OLD PARADIGMS OF GENE TRANSCRIPTION

The genome is often described as an instruction manual containing all the information needed but requiring an accurate reading to create an organism. The use of information encoded in genes is regulated by a complicated system of DNA sequences and DNA binding factors. This system chooses which genes will be turned on, meaning transcribed into mRNA, at any given time and place. The regulation of gene transcription is one of the most fundamental processes to the molecular mechanics and evolution of life. While we understand quite a bit about transcription, MCB scientists are uncovering some surprises that make us question accepted paradigms.

The following five articles describe MCB research into aspects of gene transcription and regulation. Professor Robert Tjian’s work has demonstrated that what was thought of as the universally conserved make-up of the transcriptional machinery can actually vary depending on cell type. Professor Mike Levine has found that, for many developmental genes, releasing a paused polymerase is a more critical regulatory step than recruiting the polymerase. Professor Qiang Zhou is deciphering the biochemical processes of paused polymerase release. Assistant Professor Rachel Brem has developed a technique that detects evidence of polygenic evolution of cis-regulatory factors, a process that has been hypothesized but rarely tested. Associate Professor Michael Eisen has found that the rules of enhancer architecture are much more flexible than generally believed.

Through this work, the emerging picture of transcriptional mechanisms and regulation is looking much different than the common textbook vision. As is often the case with living systems, transcription is more complicated and more interesting than we imagined.
Some scientific paradigms prevail simply because they aren’t challenged. The roles of the core factors of the RNA polymerase II (Pol II) general transcription machinery were thought to be the same for every cell type, since the mechanism is so faithfully conserved between evolutionarily disparate eukaryotes. While this is a reasonable assumption, recent work in MCB Professor Robert Tjian’s laboratory has shown that this is emphatically not the case.

Historically, experiments to unravel the mechanisms of gene transcription were mainly performed in three cell types: yeast, flies, and humans. The conclusion from these experiments was that the role of the core transcription factors was to drive transcription, while separate factors were responsible for regulation. But, out of technical convenience, these experiments used rapidly dividing cell lines, not differentiated cell lines that are no longer dividing and replicating, such as skeletal, muscle, fat, or nerve, even though differentiated cells make up the overwhelming majority of a human body.

“The assumption for most of us for the last 20 or 30 years has always been that it didn’t really matter whether you looked at yeast or you looked at a humans, the nucleus of all those cells are going to have the same core machinery,” says Tjian about the RNA polymerase II transcription machinery. “Given this premise, it didn’t seem particularly rational or logical to ask whether different cell-types in your body would have dramatically different transcriptional machinery.”

When the Tjian lab did look at Pol II transcription in differentiated cell lines, they were surprised to discover dramatic differences in the composition of one of the most evolutionarily conserved components of the core machinery—the TFIID complex, which includes the TATA box binding protein (TBP) plus other transcription factors. Tjian’s graduate student Maria Deato looked for TFIID complex subunits in human skeletal muscle cells and came up empty-handed. While this general transcription complex is so well conserved between species and found in practically all cell-types that had been studied, she found it is largely dispensed with in these differentiated muscle cells [Genes & Development 21:2137-2149].

“This is the first time we had ever seen a cell type that retains the ability to actively transcribe genes but where TFIID was being severely down regulated, actually proteolytically degraded and dramatically reduced in concentration,” says Tjian. “That really began the whole excitement. Maybe it’s not just that you add a new subunit or two to change the gene specificity of the core machinery, but you actually replace the entire 15 subunit complex of the old TFIID with something new, so you really switch the machinery. It’s a new paradigm.”

Using distinct sets of transcription machinery in specialized cells may be an efficient way to change the regulation of many genes at once. The different core promoter transcription factors recognize different promoters perhaps coordinately up- and/or down-regulating many genes at once.

Tjian wondered whether other differentiated cell lines also have specialized TFIID-like complexes, or if the skeletal muscle cells are a special case. Subsequent experiments looking at a variety of cell types, including liver, fat, and nerve cells, have shown at least three examples of other cell types that lose the canonical TFIID complex and instead use different core promoter recognition complexes. In some of the cell types examined (ie. motor neurons) the canonical TFIID continues to be used but one or more cell-specific subunits may be added to the complex.

“We haven’t figured out enough about what’s going on between a motor neuron and a fat cell to say why in certain lineages you see a clear dramatic loss of TFIID and in other lineages you don’t,” says Tjian.

Future experiments will explore this question of why different cell lines completely change the TFIID complex while others simply add a factor or two. Tjian’s lab is also exploring the roles of the transcription factors and why so many different factors are needed.

“We are at the very, very early stages of peeling back this story,” says Tjian.
MCB Professor Mike Levine admits that his recent discoveries regarding the regulation of developmental genes in *Drosophila* fruit flies are a bit controversial. Levine has found that in some genes, polymerase II binds the promoter sequence, transcription starts, then the polymerase pauses after transcribing 30-50 nucleotides. It is as if the polymerase is primed and waiting for another signal. Levine’s lab determined that the pausing behavior is specific and is dependent on the sequence of the promoter, although the mechanistic details are not yet fully understood.

In the currently accepted model, called the recruitment model, activator proteins bind to an enhancer DNA sequence, sometimes located far away from the promoter and the transcription start site, to recruit the polymerase. Once the polymerase is bound to the promoter, it transcribes the gene straight through.

“I’m not challenging the conventional view that a gene is switched on by recruiting RNA polymerase to the promoter,” says Levine. “There’s no question that’s an important process. But it’s not always true that once polymerase is brought to the promoter that gene is automatically on, that the polymerase will now run through the gene and transcribe it.”

Serendipity played a role in the initial discovery of paused polymerases at developmental gene promoters. Levine’s collaboration with Rick Young’s group at MIT’s Whitehead Institute prompted Julia Zeitlinger, a postdoc in Young’s lab, almost on a whim to test a variety of polymerase antibodies in whole genome chip assays. She found that polymerase was binding to a number of developmental genes that did not have activator protein present [Nat. Genet. 39:1512-6]. This was completely unexpected according to the current model, which holds that the polymerase should require activators in order to bind the promoters. In the three years since this discovery, other types of experiments have confirmed that the polymerases stall at a specific place on many developmental genes in the absence of activator proteins.

Work done by Levine’s graduate student Alistair Boettiger suggests the advantage to using a paused polymerase as a regulatory control. He found that genes with paused polymerases are turned on more swiftly and uniformly than their conventional counterparts. While it took the genes lacking the paused polymerases around 20-30 minutes to activate, the genes with the paused polymerases were activated in about 2-3 minutes [Science 325:471-473].

Although surprising, this polymerase pausing model does have a precedent. Heat shock protein genes use a paused polymerase mechanism so that they are primed for an immediate SOS response to cell stress. Similarly, being held at the ready would also be advantageous for genes important for developmental processes, which need to be turned on at a very specific time and place. If mistakes occur in this timing, a drastic error like a head cell maturing in the tail region may result.

Levine estimates that a large percentage (possibly 60-80%) of developmentally critical genes use a paused polymerase mechanism versus around 7-8% in the entire *Drosophila* genome. “It looks like it’s a fairly general mechanism of developmental gene control, at least in the *Drosophila* embryo,” says Levine. “Genes are repressed, but poised for rapid induction. And that’s what the paused polymerase does. The gene is off, but it’s ready to go. It’s like the engine is on, but it’s in neutral.”

To test the model that paused polymerases are a regulatory mechanism important for embryonic development, Levine’s group is embarking on experiments in which promoter sequences that cause promoter pausing will be replaced by promoters that do not cause pausing in a handful of genes critical to developmental stages. The researchers will be looking to see if replacing these promoters causes problems during embryonic development.
MCB Professor Qiang Zhou points out the importance of polymerase pausing as a general mechanism of transcriptional control. “We call it suspended transcription,” says Zhou. “It makes a lot of sense when you need a fast response. It doesn’t make sense to take extra time to remodel chromatin, modify histones and set-up your pre-initiation complex—that takes too much time. When you have everything ready to go and just wait for the final signal, releasing paused polymerases will give you really fast responses.”

Zhou’s work focuses on the positive transcription elongation factor P-TEFb, which is a key regulatory factor in releasing paused polymerases and allowing them to continue transcribing the genes. P-TEFb, composed of CDK9 and its partner cyclin T1, works by phosphorylating and antagonizing two negative transcription elongation factors and modifying RNA polymerase itself. Recent data from Zhou and others show that it can interact and cooperate with another elongation factor, ELL2, to synergistically stimulate polymerase elongation.

P-TEFb is governed by a complicated regulatory system. Zhou has found that P-TEFb in our cells is divided between three separate complexes, one of which is responsible for rescuing suspended transcription. Diseases like HIV infection and cancer can affect the distribution of P-TEFb among the various complexes. An article about Zhou’s work regarding the relationship between P-TEFb and HIV is in the last (Fall 2009) issue of Transcript. “It happens that the HIV gene expression and replication are particularly sensitive to elongation control,” says Zhou. “That’s one of the earliest model systems for studying transcription elongation.”

Since that article, Zhou’s group has discovered that the Tat protein of HIV is able to enhance the P-TEFb-ELL2 interaction to increase the expression of viral genes [Mol. Cell 38:428-38]. “The current data show that up to 80% of human genes are actually controlled at the elongation level,” says Zhou. “About 10-15 years ago people thought that HIV must be just one crazy exception and that pretty much everything else is regulated at the initiation stage. But now we think that HIV is not an exception; it is merely one good example illustrating how the expression of many mammalian genes is controlled”.

Evolution is driven by the creation and accumulation of mutations in DNA. Differences in genes cause differences between species and between individuals within a species. Phenotypic changes can be caused by mutations within the coding part of the gene, which end up as changes in a gene product, or by mutations in the regulatory sequences of the gene, which affects when and to what extent a gene product is transcribed. Evolutionary changes that affect transcription of a given gene can either act in cis, meaning that the mutation is in the promoter or other sequences close to the physical position of the gene, or in trans, where the mutation is located in a transcriptional regulator that acts on the gene.
A current theory in the field posits that cis-acting changes should be more prevalent than trans-acting changes: a mutation that is likely to only affect one gene has less potential for detrimental effects, and therefore more likely to be maintained, than a change in a transcription factor that might affect many different targets. MCB Assistant Professor Rachel Brem is tracking cis-acting mutations that underlie phenotypic changes between individuals or among species. She has developed a novel technique to find and study cases of polygenic evolution, where a species has, over time, acquired cis-regulatory mutations in multiple genes encoding related proteins that work synergistically. Brem’s work has transformed the possibility of polygenic evolution from theoretical to factual.

Brem used mixed-species diploid cells created from haploid genomes of two related yeasts. The *Saccharomyces cerevisiae* and *Saccharomyces bayanus* heterozygous cells can live and divide but are not able to undergo meiosis. In the diploid, a given gene is thus present in two copies, one derived from the *S. cerevisiae* parent and the other from *S. bayanus*. Brem uses a technique called RNA-seq to sequence short regions of millions of mRNA transcripts from these cells. Many of these sequences can be identified as having been transcribed from one parent’s allele or the other due to genetic differences. Counting the transcripts for a given gene shows whether one species’ allele is associated with up- or down-regulation compared to the other species. Because the two haploid genomes are in the same cellular environment and are acted on by the same mixture of trans-acting factors, differences in their transcription levels can only be due to cis-acting sequence differences.

From this data, collections of genes that code for proteins that function together were analyzed for evidence of polygenic evolution. Out of the roughly 250 gene groups studied, six or seven exhibited this behavior. One of these systems is an RNA processing complex called the exosome, in which eight of the constituent proteins were up-regulated in *S. bayanus* and one was down-regulated, compared to *S. cerevisiae* [PNAS 107: 5058-5063].

“When you see enough of these cis-regulatory changes that govern expression of different genes whose products function together, it starts to not look like an accident anymore,” says Brem. “Variants arise by accident all of the time, but if they were maintained by accident in a given species, sometimes they would drive expression up and sometimes they would drive expression down. And that’s not what this data looks like. This data looks like coherent changes that are all going in the same direction.”

Brem says this is evidence of polygenic evolution. The many mutations in the exosome genes arose separately over many generations—whether in *S. cerevisiae* or *S. bayanus* is not yet known—and were selectively kept.

“So what interests us about these results is that it suggests that sometimes evolution really does act by making sustained changes in a bunch of different genes instead of just in one gene,” says Brem. “No one knows how prevalent this so-called polygenic evolution might be.”

Since Brem pioneered this technique to look for polygenic evolution, other laboratories have begun to use it in their research to study variation in individuals within a species as well as between species. Brem’s lab is using the technique to look at other yeast species, and she is also collaborating with MCB Assistant Professor Craig Miller to look at divergence between separated populations of fish.
The increasing availability of genomic sequences has changed our understanding of how gene transcription is regulated. In the absence of comparative sequence analyses, simple and rigid rules were assumed to govern the placement of transcription factor binding sites within enhancer sequences. New evidence suggests that this is not the case. MCB Associate Professor Michael Eisen is trailblazing by taking advantage of the most modern techniques to decipher new rules for transcriptional regulation.

The Drosophila homeobox even-skipped (eve) gene is a well-studied model system for transcriptional regulation. Eve’s enhancer sequences (discovered and characterized by MCB’s Michael Levine) are responsible for turning the gene on in certain locations of the fly embryo in very early embryo development as a response to a variety of signals. Eisen likens the mechanism by which this gene can tell at what segment in the embryo it is located to that used by a GPS system. As with distance readings from a few satellites, the molecular concentrations of a few important protein factors can be combined to give location information.

"Every gene within each nucleus has the job of figuring out: am I in a place where I should be on or am I in a place where I should be off?" says Eisen. "As that happens, you start to get genes coming on in patterns, and the ability to make more and more complex patterns increases. And that’s basically how development proceeds. What interests us is trying to understand how those patterns are created."

Eisen’s work focuses on the stripe 2 enhancer sequence of eve. Despite the near perfect conservation of eve expression in other flies, comparison of eve stripe 2 sequences from different fly species revealed that many of the functional transcription factor binding sites they contain are not conserved. For example, specific sequences that can not be altered or moved in D. melanogaster are often not present in the homologous enhancer of another species—in the most extreme cases almost none of them are conserved—yet swapping the whole enhancer sequence into melanogaster creates no ill effects.

"The way people used to think about it is that there is some sort of cis regulatory grammar that says you have to have A next to B and B is 50 base pairs away from C, because they interact with each other in a complicated grammar," says Eisen. "We no longer think that there is much of a grammatical constraint. There’s an amazing amount of flexibility in how you can create an enhancer. It remains a deeply mysterious problem."

To compound the mystery, the transcriptional factors’ binding sites are such short sequences that they could potentially bind promiscuously throughout the genome. Eisen points out that the sequence only tells part of the story. The factors’ binding kinetics, binding order, and local environment can also affect where and when the factors bind the enhancers.

Eisen plans to continue detangling the myriad of variables involved in eve regulation by comparing the developmental processes of around 20 different fly species from three fly genera. The diversity represented by these organisms will give an evolutionary angle to understanding how the regulatory sequences function at a molecular level. Because of modern sequencing techniques, he has the luxury of choosing certain species for their ease of growing and using in a laboratory setting and is not limited to model organisms others’ have used. This is a dramatic new way of designing experiments, only made possible by the lowered costs of genome sequencing.

"We can really do something that’s never been possible before," says Eisen. "We are in the era of experimental comparative genomics, where every organism—every individual—is in play as an experimental resource. The practical limitations that used to restrict species you worked on and individuals you worked on and experiments you did are now completely gone. It’s hard to overstate the ways in which having access to cheap sequencing is transforming the way that we do this comparative biology."
2010 AWARDS

OUTSTANDING GRADUATE STUDENT INSTRUCTORS

The following GSIs for MCB courses were among those honored by the Graduate Division in a May 5 event at the International House for outstanding teaching.

- Lisa Dennison [Kristin Scott lab]
- Mark Herzik [Michael Marletta lab]
- Hania Koever [Shaowen Bao lab]
- Nikki Kong [Robert Tjian lab]
- Zhe Liu [Robert Tjian lab]
- Katherine Miller [Susan Marqusee lab]
- Kate Monroe [Russell Vance lab]
- Suruchi Nandu [Astar Winoto lab]
- Adrienne Pigula [Georjana Barnes lab]
- Brock Roberts [Nicole King lab]
- Gail Sondermeyer [Public Health]
- David Steakley [Jasper Rine lab]
- Julie Ullman [Lu Chen lab]

DEPARTMENTAL AWARDS

Departmental Citation
- Dipankan Bhattacharya [Richard Harland lab]
- Outstanding Scholar
- Isaac Max Oderberg [Gary Karpen lab]

DIVISION OF BIOCHEMISTRY & MOLECULAR BIOLOGY

Grace Fimognari Memorial Prize
- Joo Yeon (Jenny) Ryu [Joseph Napoli lab]
- Kazuo Gerald Yanaba & Ting Jung Memorial Prize
- Shaunak Adkar [Sharon Amacher lab]
  Jesse Rabinowitz Memorial Prize
  for outstanding junior in BMB
- Gloria Wu [Carlos Bustamante lab]

DIVISION OF GENETICS & DEVELOPMENT

Edward Blount Award
- Dipankan Bhattacharya [Richard Harland lab]
- Spencer W. Brown Award
- Dipankan Bhattacharya [Richard Harland lab]
- Michael Goldrich [Sharon Amacher lab]
- Lisa Dennison [Kristin Scott lab]
- Mark Herzik [Michael Marletta lab]
- Hania Koever [Shaowen Bao lab]
- Nikki Kong [Robert Tjian lab]
- Zhe Liu [Robert Tjian lab]
- Katherine Miller [Susan Marqusee lab]
- Kate Monroe [Russell Vance lab]

DIVISION OF IMMUNOLOGY

Outstanding Undergraduate
- Leslie Ying Chiang [Astar Winoto lab]

Distinction in Academic Achievement
- Jeremy Wang [Nilab Shastri lab]

Excellence in Research
- Brooke M. Su [Jeremy Thorner lab]

DIVISIONS OF CELL & DEVELOPMENT BIOLOGY

Paola S. Timiras Memorial Prize
- Simina Ticau [Matthew Welch lab]

I. L. Chaikoff Memorial Awards
- Tony Chen [Gary Firestone lab]
- Cody Ender [John Forte lab]
- May Szeto [Joachim Li lab]
- Simina Ticau [Matthew Welch lab]
- Bao Anh Patrick Tran [Krishna Niyogi lab]
- Elena Zehr [Jeremy Thorner lab]

DIVISIONS OF NEUROSCIENCE

Jeffery A. Winer Memorial Prize
- Connie Wang [Lance Kriegsfeld lab]

I. L. Chaikoff Memorial Award
- Michael Chiang [Antonello Bonci lab]
- Wayland Chu [Ehud Isacoff lab]
- Jason Chung [Daniel Feldman lab]
- Heejae Kang [Dan Feldman lab]
- Connie Wang [Lance Kriegsfeld lab]
- Weichan Chris Xu [Shaowen Bao lab]

CLASS NOTES

Kevin Hybiske [PhD 2003] was appointed to Assistant Researcher in Microbial Pathogenesis at UC Berkeley, in the Division of Infectious Diseases & Vaccinology. [khybiske@berkeley.edu]

Jacob Sahakian [BA 2009] is working in the lab of Dr. Marcus Horwitz at UCLA David Geffen School of Medicine. We are developing vector-based vaccines for Mycobacterium tuberculosis and Francisella tularensis.

Julio Soto [PhD 1994] was promoted early to the rank of Professor at San Jose State University. He is the PI of grants from HHMI Science Education, NIH SCORE, and NSF-REU. [Julio.Soto@sjsu.edu]

CLASS NOTES WANTS TO HEAR FROM YOU

Do you have a bachelor’s, master’s or Ph.D. in Molecular and Cell Biology from Berkeley? Let your classmates know what you are up to by sending in a Class Note for publication in the next issue.

To send your Class Note, go to mcb.berkeley.edu/alumni/survey.html or Send e-mail to tscript@berkeley.edu
Greg Barton received the American Association of Immunologists (AAI) BD Biosciences Investigator Award, which is presented to an early-career investigator who has made outstanding contributions to the field of immunology.

David Bilder received the 2010 Harland Winfield Mossman Award in Developmental Biology from the American Association of Anatomists.

Steven E. Brenner received the 2010 International Society for Computational Biology (ISCB) Overton Prize, given to an early or mid career scientist for significant contributions in field of computational biology.

W. Zacheus Cande was elected as a Fellow to the American Association for the Advancement of Science (AAAS).

Thomas Cline received the Edward Novitski Prize from the Genetics Society of America, which honors creativity and innovation in genetic research leading to scientific achievement, for his work on Drosophila sex determination.

Kathleen Collins and Richard Harland were awarded 2010/11 Miller Professorships, which releases them from teaching and administrative duties and allows them to pursue their research full-time.

David Drubin was elected as a Fellow of the American Academy of Arts & Sciences and became Editor in Chief of the journal Molecular Biology of the Cell.

Richard Harland was elected to the American Academy of Arts and Sciences and became President of the Society for Developmental Biology.

Lin He was named a 2009 MacArthur Fellow. These $500K “genius grants” are given to people from all fields who have demonstrated “exceptional creativity in their work and the prospect for still more in the future.” He was honored for her advancements in the field of microRNAs.

Doug Koshland was elected to National Academy of Sciences.

Mike Levine was awarded an Einstein Professorship from the Chinese Academy of Sciences.

Susan Marqusee was named Director of the California Institute for Quantitative Biosciences (QB3).

Barbara J. Meyer received the 2010 Genetics Society of America Medal for outstanding contributions to the field of genetics in the last 15 years. She is also honored as the 2010-11 Distinguished Speaker for the Li Ka Shing Foundation Women in Science Program in China, where she will travel and give public lectures on her research and her career.

Craig Miller received a 2010 March of Dimes’ Basil O’Connor Starter Scholar Research Award, which provides a two-year grant to support research by young, promising scientists.