AFFYMETRIX
MICROARRAY
BULLETIN

SPECIAL STEM CELL ISSUE
EXPRESSION PROFILING ■ COPY NUMBER ■ SEQUENCE ANALYSIS
VOLUME 1 ■ ISSUE 3 ■ $10.00

www.microarraybulletin.com
September 2005

IN THIS ISSUE

Copy Number/Sequence Analysis
Aravinda Chakravarti and Anirban Maitra from Johns Hopkins and Mahendra Rao at the National Institute on Aging discuss the use of high-resolution genomic methodologies for quality control of cell lines

Expression
University of Vermont's Jeff Spees and Tulane's Carl Gregory discuss the challenges of classifying and purifying mesenchymal stem cells for cell-based therapies

Princeton's Ihor Lemischka and Virginia Commonwealth's Raj Rao discuss new data on how stem cells maintain the ability to differentiate into various body tissues.

COPY NUMBER/SEQUENCE ANALYSIS

Human Embryonic Stem Cell Study Uncovers Genetic Mutations in Federally-approved Lines

Aravinda Chakravarti and Anirban Maitra from Johns Hopkins and Mahendra Rao at the National Institute on Aging discuss the use of high-resolution genomic methodologies for quality control of cell lines

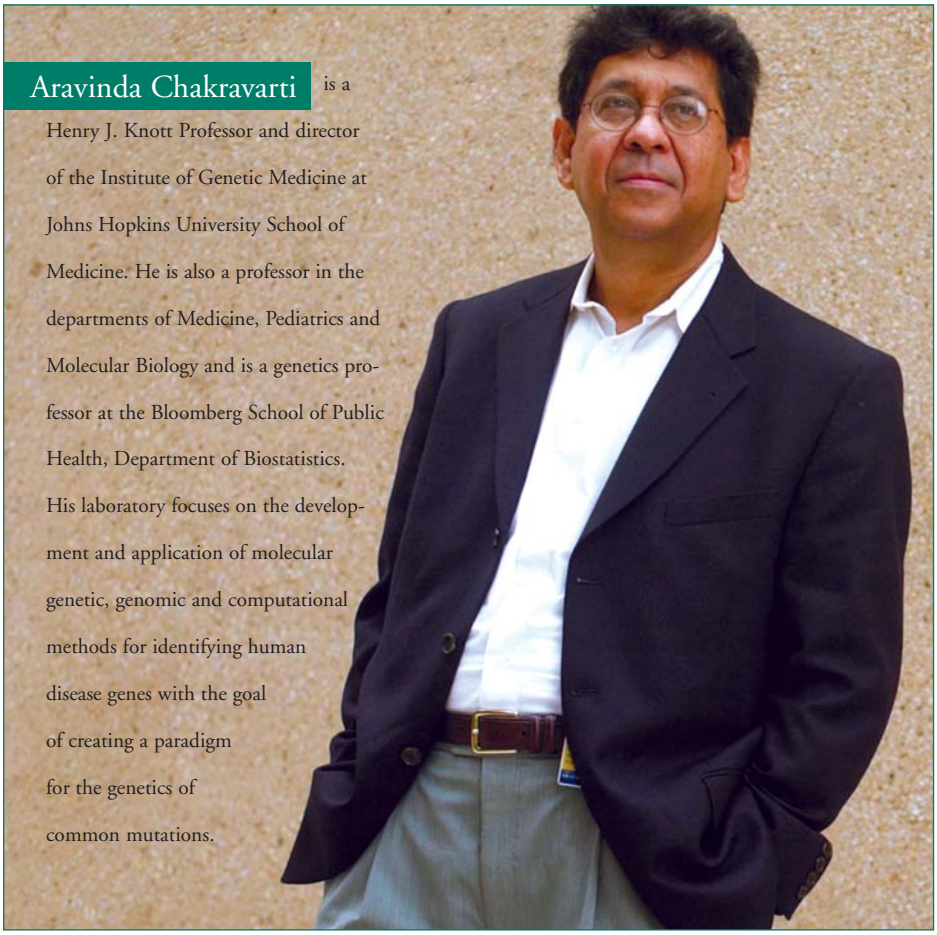
Scientists led by Dr. Aravinda Chakravarti, director of the McKusick-Nathans Institute of Genetic Medicine at Johns Hopkins Medical Center and Mahendra Rao, head of the Stem Cell Group in the Laboratory of Neuroscience at the National Institute on Aging, report that they have uncovered mutations in federally approved embryonic stem (ES) cell lines that have been cultured in the laboratory over extended periods of time.

Identification of genetic abnormalities in these cell lines cast doubt on their safety and efficacy for therapeutic use and has many implications for biological studies on them.

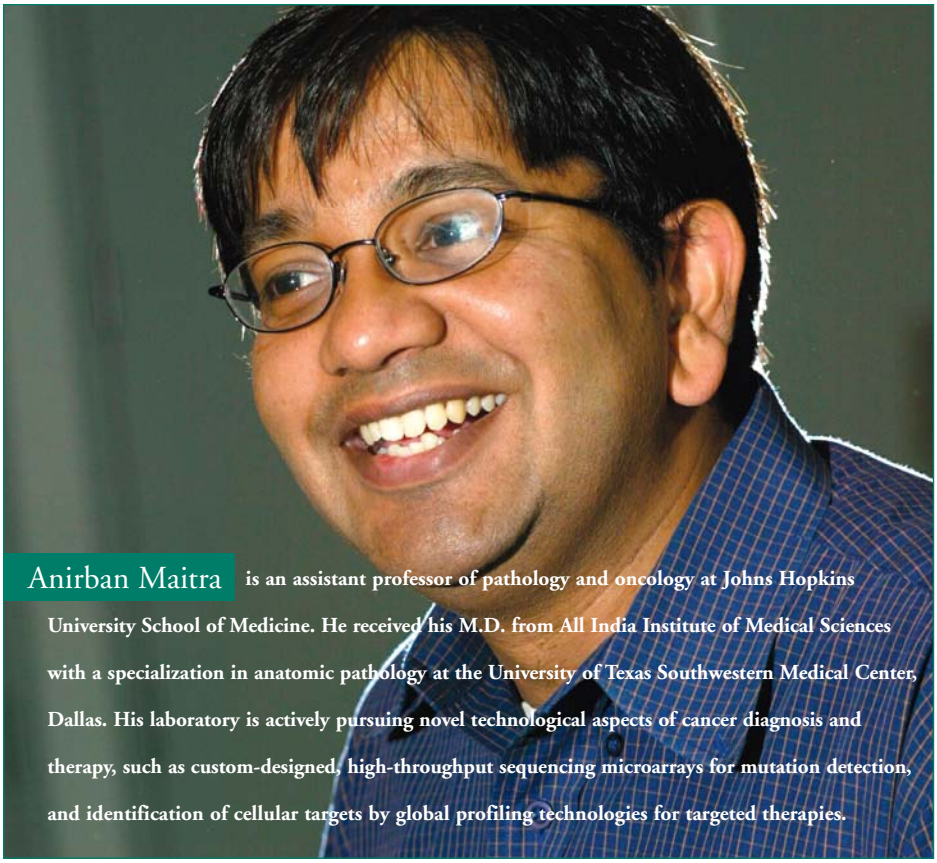
"These mutations may be widespread in all kinds of cultured cells, but genetic changes are particularly important in embryonic stem cells," said Chakravarti. "Everybody is hoping to use these cells in some therapeutic protocol one day, so we need to know exactly what the cells are, what kinds of genetic changes they have, and whether those changes make any biological difference or not."

The use of embryonic stem cells (ESCs) has been a controversial issue since 2001, when the Bush administration restricted publicly funded embryonic stem cell research to what are now 22 existing cell lines, a move that many scientists believe stifles important advances in stem cell research.

The work, which was done largely using the Affymetrix 100K SNP and Mitochondrial Resequencing Arrays, was published in the September 4th online version of the journal *Nature Genetics*. Conventional methods for analyzing the genetic integrity of cell lines relied on low-resolution methodologies such as karyotyping.



Aravinda Chakravarti is a Henry J. Knott Professor and director of the Institute of Genetic Medicine at Johns Hopkins University School of Medicine. He is also a professor in the departments of Medicine, Pediatrics and Molecular Biology and is a genetics professor at the Bloomberg School of Public Health, Department of Biostatistics. His laboratory focuses on the development and application of molecular genetic, genomic and computational methods for identifying human disease genes with the goal of creating a paradigm for the genetics of common mutations.



Anirban Maitra is an assistant professor of pathology and oncology at Johns Hopkins University School of Medicine. He received his M.D. from All India Institute of Medical Sciences with a specialization in anatomic pathology at the University of Texas Southwestern Medical Center, Dallas. His laboratory is actively pursuing novel technological aspects of cancer diagnosis and therapy, such as custom-designed, high-throughput sequencing microarrays for mutation detection, and identification of cellular targets by global profiling technologies for targeted therapies.

"I think at least one of the things that we have shown is that karyotyping is not adequate," said Anirban Maitra, assistant professor of pathology and oncology, at Johns Hopkins. "Some cell lines that we looked at in this paper were karyotyped and found to be 'normal,' yet we found localized regions of amplification and deletion, as well as epigenetic changes in these cells."

Chakravarti, who sits on the scientific advisory board of Affymetrix, Maitra, and Rao recently spoke to *AMB* Editor Rachel Shreter. The research team discussed:

- The types of changes they found in ES cell lines and why those changes occur
- Why they chose a microarray-based approach to look at the quality of today's federally-approved stem cell lines
- Future research that will provide more information about the types of changes that accumulate in these and other cell lines and how to prevent them from becoming prevalent
- The importance of wisely using the data wisely from new studies to influence federal research policies

Genetic Alterations in ES Cell Lines

Shreter: You have been looking at late passage human embryonic stem cell lines obtained from federally-approved providers and found a number of mutations in them. Are these cell lines more susceptible to mutation than others?

Chakravarti: I think we're using fairly powerful technologies on what are basically old problems and seeing new things. I'm not convinced these cells lines are any more unstable than other cell lines. It is quite possible, and even likely, that since we've never done these kinds of experiments—looking at gene copy number and genome-wide methylation or sequencing specific genes—on other cell lines, that all cell lines suffer these kinds of changes. A cell line normally considered to be a pure, unadulterated cell line, might not be that at all. In fact, it might be a population of very closely related cells with many genetic changes.

We would hope the large majority of those changes are benign and don't affect the function of those cells. They're probably mostly changes in non-coding DNA with just a small fraction in coding sequences for proteins and in their regulatory regions. Even then, there is probably just a small minority of those mutations that make a difference to the phenotype of these cells. Then, it depends on how many cells in those lines have a particular kind of mutation and what the nature of the mutation really is.

Shreter: What do you think causes these genetic changes?

Chakravarti: Well, the basic answer is, we don't know. Mutations occur during cell division in every kind of cell. So I think that's the cause. I don't think embryonic stem cells or any other cell lines are particularly vulnerable to mutation, since we know that DNA replication is not perfect and that mutations will occur.

We've always assumed that in a given cell line, the cells that have these mutations are all different from one another, so no particular change overrules any other kind of change. I think our studies show that there are at least some genetic changes that have occurred in substantial number, or have allowed the mutant cells to overtake some of the non-mutant cells. So some speculation has been made in our recent *Nature Genetics* publication about the quality of these cell lines.

Rao: In nature, there's a relative inefficiency in the replication process. Fidelity is never 100 percent. And if you have cells dividing for a long period—whether it's in culture or in the body—a small number of them will accumulate mutations that may be in a critical region. The accumulation of mutations can be directly related to the number of times the cells have divided.

Maitra: Different culturing methods also lead to different rates of mutation. There are two established culture methods for ESCs—enzymatic separation and mechanical dissociation. There is some published data that suggests the enzymes used may have an adverse impact on the chromosomes.

In fact, in our limited panel, the results suggest that mechanically dissociated ES cells had a lower incidence of copy num-

ber alterations than those enzymatically dissociated.

It is important to note, however, that this dichotomy was erased when we took epigenetic changes, like promoter methylation, into consideration. Cells separated by either technique demonstrated evidence of methylation. The ultimate effect of the dissociation technique on stem cell stability is something that will need to be examined in greater detail in a larger panel of ES cells.

Minimizing Mutations

Shreter: Is there any evidence that suggests that these changes can be slowed down, stopped or even reversed as the cells are passaged?

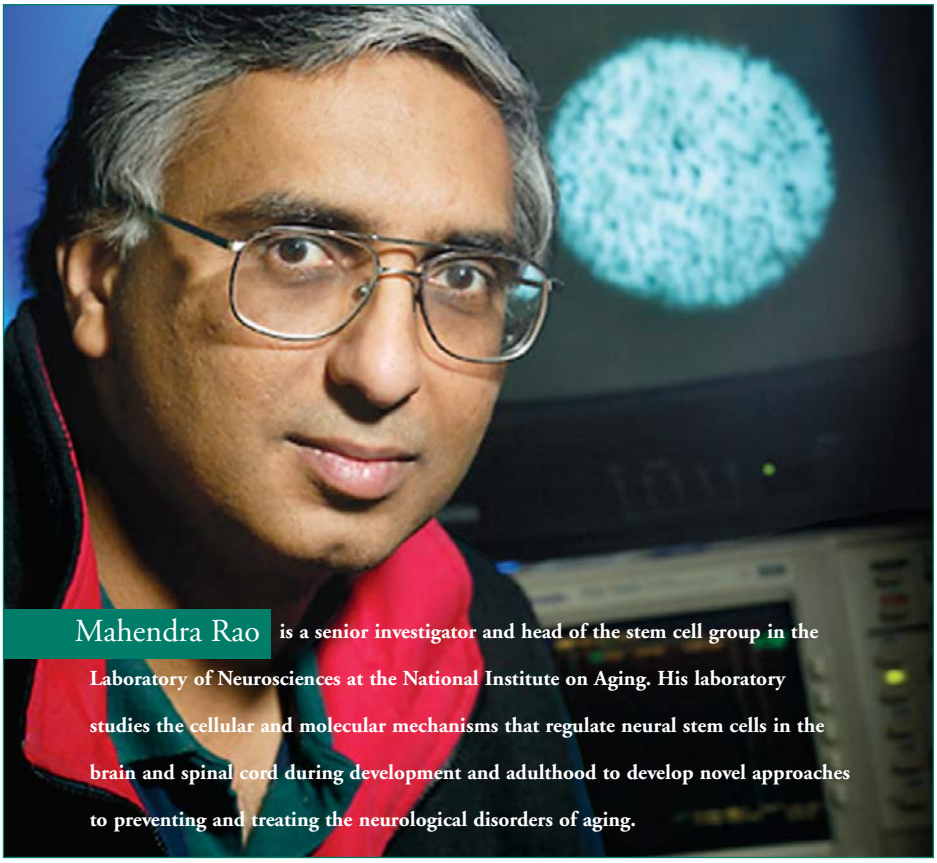
Rao: Well, if you figure out a way to identify the cells that don't have changes, you can select or reselect or subclone out the normal population. People clearly need to look at the cells and identify changes as early as they occur, because then the chance of selecting cells without those changes is much better. So careful monitoring is one way of making sure that you can propagate a line, even in the face of stochastic changes.

Chakravarti: Often, when we do find multiple changes, they are distributed across various cells. I think the only reversion that could happen is a back mutation of exactly the same kind. That's very unlikely with the kind of major changes we are talking about.

On the other hand, some cells may have had major chromosomal changes that we don't see. Those cells have died off because the mutation was lethal to that cell. In our study, I think the mutations we saw were all randomly occurring. Some were benign mutations that have simply been carried along with no functional consequence whatsoever. But, some mutations could have been lost or given a somewhat small growth advantage to the cell. Since we always pick cells that grow well for the next generation in cell culture, it's possible that we inadvertently applied some kind of selection to those with good growth properties.

Using an Array-based Approach

Shreter: Why did you choose to use high-density microarrays for this study?



Mahendra Rao **is a senior investigator and head of the stem cell group in the Laboratory of Neurosciences at the National Institute on Aging. His laboratory studies the cellular and molecular mechanisms that regulate neural stem cells in the brain and spinal cord during development and adulthood to develop novel approaches to preventing and treating the neurological disorders of aging.**

Chakravarti: We chose an approach that was available to us and could sample the genome at very high density to look for particular kinds of mutations. Microarray technology is now at a stage where we can screen for a certain class of changes. Even then, I'm fairly sure we did not detect all of the relevant copy number changes. The technology is still evolving. The technologies will improve and there's no doubt in my mind that we will find more changes, not only in these cell lines, but in other cell lines as well.

Maitra: The current gold standard used for many of the embryonic stem cells is conventional karyotyping. I think at least one of the things that we have shown is that karyotyping is not adequate. Some of cell lines that we looked at in this paper were karyotyped and found to be 'normal,' yet we found localized regions of amplification and deletion, as well as epigenetic changes in these cells.

Rao: The logic we used when we discussed our approach came from Dr. Chakravarti's and Dr. Maitra's experience with what happens in cancer cells. When cells change, the changes can be in the genome, epigenetic changes, changes in mitochondria, or changes in the microRNA that regulates gene expression. We needed to figure out which method would allow us to sample each one of these qualities of a cell. And we had to choose from existing technology for which there was expertise and there was a reasonable cost-benefit to us. Given the microarray expertise in Dr. Chakravarti and Dr. Maitra's laboratories, we thought that using oligonucleotide microarrays was the best option right now.

This test certainly isn't exhaustive and we will need to do things at a higher resolution. We have improved the resolution for looking at these cells by using newer technology, but the resolution needs to get better still.

Investigating microRNAs

Shreter: Dr. Rao, you mentioned microRNAs. Have you started looking at those and the role they may play in stem cell regulation?

Rao: Several people have started looking at that. It turns out that microRNAs seem to be specifically enriched, and specific subsets seem to be uniquely expressed in embryonic stem cells and other stem cell populations. So investigators have begun using a variety of strategies including arraying microRNAs, sequencing them by enriching for small-sized RNAs, or actually going in and looking at microRNA bind-

ing sites in highly expressed genes in a stem cell population. These strategies have yielded converging results. It appears there are several different classes of microRNAs that regulate several different aspects of function, including regulation of gene expression, regulation of the promoter, and regulation of RNA degradation.

Maitra: There are several competing technologies out there that we can use to study microRNAs. Most of them are array-based. What differs is the detection technique being used—whether it's radionucleotide-based detection or it's a biotin label. The 'good' thing about microRNAs is that there are only about 200 known, although a larger number is suspected. That makes them easier to study than, say, mRNA transcripts.

Future Directions

Shreter: Where do you see your work going in the future?

Maitra: Well, we talked about the available technology and why it has given us an unprecedented view of the stem cell genome. That is clearly not the entire story and we'd like to take a more magnified look at what's going on and elucidate on a global scale the sort of aberrations that these cultured cells have.

We'd also like to understand the functional implications of these mutations. Mahendra talked about looking at microRNAs and additional changes at the epigenetic level. We'll be looking at those and we'd also like to play a contributory role in determining the benchmarks required to pronounce these cells stable or not, and useable or not.

Chakravarti: There's still so much more nuts-and-bolts work to be done. There's no doubt that we need to look at this issue at a much higher resolution. I think the technology for reading the Affymetrix array—the software, the bioinformatics of all of that—will need to keep improving. As arrays obtain higher density, our ability to detect subtle changes will depend on much, much better methodology.

As you add more oligonucleotides and your feature size goes down, you'll need to address how the arrays can be read on different kinds of scanners, how the data can be filtered and how to interpret the data. It's an entirely technological issue having nothing to do with biology, but it's informed by biology. The majority of changes that we will need to look at, and have to detect will be heterozygous changes, small deletions or small copy number changes. We're still not very good at detecting those because of a host of technical issues that need to be addressed.

Influencing Federal Policy

Shreter: How would you ultimately like to see your work affect federal policy?

Chakravarti: It is important that we use this work, and many other kinds of work, to show that our existing cell lines are insufficient for us to do any biology, let alone contemplate therapeutic use. Although I think this field has tremendous amount of promise, it's years away from any kind of therapeutic application.

There's absolutely no doubt that this country is very divided over stem cell issues and we need to work towards a more informed consensus. At the same time, as scientists, we need to be much more honest about the current limitations of stem cells and what we still need to learn for clinical applications. I'm certainly not saying the federal government should stick to its policy, but I think we scientists also have to be a little bit more realistic about the time frame for applications.

Rao: When Dr. Zerhouni went before the U.S. Senate, they asked him whether we need more cell lines. And he said, more cell lines could not hurt. I think that's very true, and I think the data here suggest that it would be useful to have a lot more cell lines.

FOR MORE INFORMATION

Contacts

■Aravinda Chakravarti, Ph.D.
Johns Hopkins University School of Medicine
Institute of Genetic Medicine
733 N. Broadway
Broadway Research Building, Suite 579
Baltimore, MD 21205
aravinda@jhmi.edu

■Anirban Maitra, MBBS
Affiliate, McKusick-Nathans Institute of Genetic Medicine
The Sol Goldman Pancreatic Cancer Research Center
Johns Hopkins University School of Medicine
Ross Bldg 632
720 Rutland Ave
Baltimore, MD 21205
amaitra1@jhmi.edu

■Mahendra Rao, M.D., Ph.D.
National Institute on Aging
National Institute of Health
Laboratory of Neurosciences
5600 Nathan Shock Drive
Baltimore, MD 21224
raomah@grc.nia.nih.gov

Companies

■Affymetrix Inc. – http://www.affymetrix.com/

Organizations

■Johns Hopkins University School of Medicine – http://www.hopkinsmedicine.org

■National Institute on Aging – http://www.nia.nih.gov

Further Reading

■Maitra A, Arking DE, Shivapurkar N, Ikeda M, Stastny V, Kassauei K, Sui G, Cutler DJ, Liu Y, Brimble SN, Noaksson K, Hyllner J, Schulz TC, Zeng X, Freed WJ, Colman A, Sartipy P, Matsui SI, Carpenter M, Gazdar AF, Rao M, Chakravarti A. Genomic alterations in cultured human embryonic stem cells *Nature Genetics* 2005; (Sept 4th online edition)

■Maitra A, Cohen Y, Gillespie SE, Mambo E, Fukushima N, Hoque MO, Shah N, Goggins M, Califano J, Sidransky D, Chakravarti A. The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. *Genome Res*. 2004; 14(5):812-9.

■Cutler DJ, Zwick ME, Carrasquillo MM, Yohn CT, Tobin KP, Kashuk C, Mathews DJ, Shah NA, Eichler EE, Warrington JA, Chakravarti A. High-throughput variation detection and genotyping using microarrays. *Genome Res* 2001; 11(11):1913-25.

■Noaksson K, Zoric N, Zeng X, Rao MS, Hyllner J, Semb H, Kubista M, Sartipy P. Monitoring differentiation of human embryonic stem cells using real-time PCR. *Stem Cells* 2005; Aug. 4 [Epub ahead of print].

People

■Elias A. Zerhouni, MD
Director, National Institutes of Health
http://www.nih.gov/about/director/

Scientists Use Gene Expression Arrays to Profile Adult Stem Cells for Regenerative Medicine

University of Vermont's Jeff Spees and Tulane's Carl Gregory discuss the challenges of classifying and purifying mesenchymal stem cells for cell-based therapies

Scientists at the University of Vermont and Tulane have discovered new gene expression and protein profiles that enable them to classify adult stem cell differentiation in more detail than ever before. Understanding these molecular pathways is a key step towards developing new drugs that control adult stem cell differentiation and new regenerative therapies for diseases such as multiple myeloma.

Dr. Carl Gregory, an assistant professor in the department of medicine at Tulane University, studies how mesenchymal stem cells (MSC)—largely accepted as the source of bone repair—are affected in multiple myeloma and how healthy cells could be used to treat disease.

"On top of using stem cells as a *bona fide* cytotherapeutic, we are also using the MSCs as a model for *in vitro* assays," said Gregory. "We use these assays to find drugs that may be able to reverse multiple myeloma degeneration."

Dr. Jeff Spees, assistant professor of medicine at the University of Vermont, directs the university's adult stem cell core facility, where he studies the potential for using stem cells derived from bone marrow in cell-based therapies. Working with Dr. Darwin Prockop at Tulane's Center for Gene Therapy, he was among the first to show that adult stem cells can fuse with differentiated cells and take on their characteristics.

Spees and Gregory recently sat down with the *AMB* to discuss their stem cell research. Topics covered include:

- Discovering new stem cell differentiation pathways and inhibitors of those pathways
- Identifying and characterizing new classes of mesenchymal stem cells for regenerative medicine
- Improving the accuracy of stem cell classification through integrative genomic and proteomic studies

Discovering Stem Cell Differentiation Pathways

Spees: You've been actively examining the role of the Wnt pathway and its activity in the proliferation and differentiation of adult stem cells, namely MSCs from bone marrow. What has expression analysis told you about this pathway?

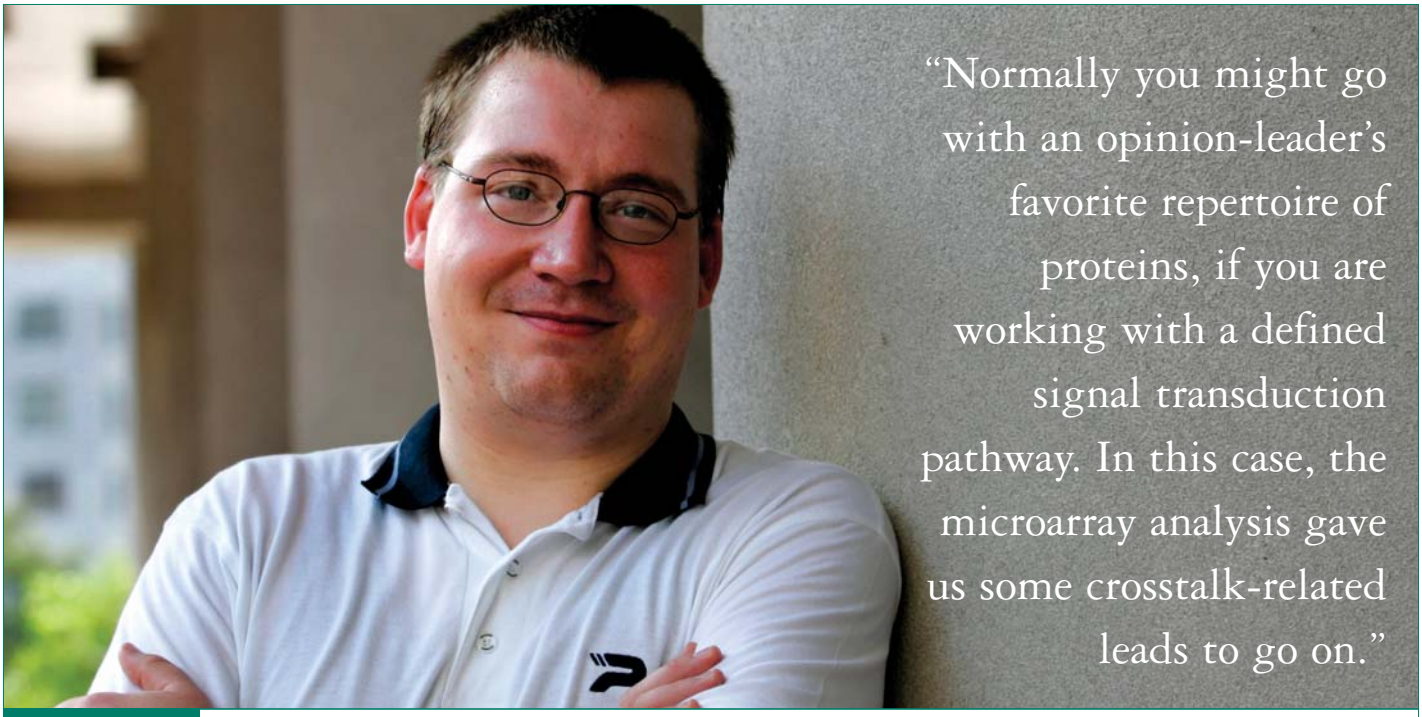
Gregory: Initially, when we were working on the Wnt pathway and Wnt inhibitors, we didn't know exactly what was going on. The Wnt pathway is extremely complex with many inhibitors, positive regulators and everything in between. There are many different isotypes and they behave in different ways. Using the Affymetrix arrays confirmed our focused experiments and gave us some ideas on what proteins to look at further. This data allowed us to look at glycogen synthetase kinase (GSK) inhibitors and other inhibitors that may play roles in modulating growth and differentiation of MSCs, which is along the lines of John Shaughnessy's work on Dikkopf-1 (Dkk1) at Arkansas. It gave us the bigger picture and brought up things we wouldn't have normally thought of. Normally you might go with an opinion-leader's favorite repertoire of proteins if you are working with a defined signal transduction pathway. In this case, the microarray analysis gave us some crosstalk-related leads to go on.

Spees: So why is Dkk1 secretion important in the context of multiple myeloma?

Gregory: In the cells we work with, it's fairly well accepted that there is a mutual exclusion between proliferation and differentiation. We found that inhibition of the Wnt pathway, by virtue of Dkk1 expression, actually prevented inappropriate differentiation to bone and allowed cells to proliferate; it's something like a checkpoint. Cells are proliferating, but because the Dkk levels are high, they don't differentiate in a Wnt-mediated manner. For myeloma cells, instead of autoregulating Dkk expression, the myeloma cells are expressing Dkk and inhibit the mesenchy-

Jeff Spees is an assistant professor of Medicine and the Director of the Adult Stem Cell Core in the the Cardiovascular Research Institute at the University of Vermont. His research interests include elucidating the differentiation capacity, migration patterns, cytokine/growth factor secretion, and fusion of adult stem cells with tissue endogenous cells to determine their ability to repair tissues.





Carl Gregory is an assistant professor at the Tulane University Center for Gene Therapy under the directorship of Darwin Prockop. Gregory's recent work has highlighted the importance of Wnt signaling in MSC-mediated bone repair in health and disease. He received his doctorate in 1999 from the Wellcome Trust Centre for Cell-Matrix Research (University of Manchester, United Kingdom) where he worked on the biochemistry of type X collagen assembly in heritable disorders of the skeleton.

mal stem cells from producing osteoblasts that repair bone. And what happens is that the myeloma cells produce a growth-permissive environment within the bone in the form of osteolytic lesions, which are basically big holes in the bone marrow. While the body would normally respond to that by repair, the myeloma cells secrete Dkk1 and block that process.

Spees: Do you plan on using expression analysis to look at the mesenchymal stem cell responses to glycogen synthetase kinase inhibitors like BIO?

Gregory: For the most part, the Wnt pathway is very well characterized and everyone agrees with how it functions, but there are problems. First of all, these drugs have certain other side effects, some on the physiology of the cell, and the efficacy is altered by the crosstalk with other signal transduction pathways. We would like to use microarrays to confirm, on a global level, that the drug is doing what it is intended to, and to know about other pathways affected by the inhibitors. For instance, the PPAR gamma axis that controls adipogenesis crosstalks with the Wnt pathway, which controls osteogenesis. They are mutually exclusive as well. We can look at how we are modulating the signal transduction pathways and how the ratio of the magnitude of the signals would affect the cell. So we can look at how the combinations of drugs—PPAR gamma modulators and GSK3 modulators—affect the Wnt pathway. We use microarrays to give us a global representation of what's happening, where we are going wrong, or where we are making progress.

Identifying and Characterizing New Classes of Mesenchymal Stem Cells

Gregory: I'm intrigued by what you call the alternative populations of mesenchymal-like cells that you can recover from the bone marrow. The list of cells gets longer every month or so, but CD133 positive cells and p75 cells certainly appear to have more multipotentiality than the transiently amplifying MSCs that we play with normally. Are they better than classically derived MSCs? Are they more multipotent? What's in store for them in the future?

Spees: In terms of the transitory amplifying cells, I don't know if they are necessarily more multipotent; we are still defining the different cells they can differentiate into. It appears that MSCs are really a mixed population of non-hematopoietic cells that are adherent when you isolate them from the bone marrow—which include cells with the propensity to make endothelium, some cells more suited for smooth muscle differentiation, and others better for osteoblast differentiation. It may be possible to use the surface epitopes to pull out particular sub-populations for different therapeutic applications. That said, it may be that the different populations are useful for direct cell differentiation, but also that the different cell populations may have different repertoires of secreted proteins, like neurotrophins or growth factors, for example. And it's becoming apparent now that a large part of reparative ability of the cells is the secreted repertoire or profile of proteins. So if different populations secrete different profiles, we might use them for different applications, like central nervous system repair or vascular repair.

Gregory: When you take a bone marrow aspirate and recover the cells, is it feasible to do array analysis and look at the repertoire of different epitopes being expressed and maybe choose antibodies available to separate those cells? Is this how you've been going about things so far?

Spees: Exactly. For example, using Affymetrix arrays, we were able to identify that both the CD133-derived transient

"...the only way you could link a given surface epitope with a given repertoire of trophic factors is probably by large-scale transcriptome analysis."

amplifying cells and MSCs express the gene for the sialomucin protein CD164, which is related to CD34, a classical hematopoietic antigen. In the future, it will probably be useful to pull non-hematopoietic progenitor cells out of bone marrow. Basically, we are using the arrays to look at time courses of magnetically separated CD133-derived cells, p75 low-affinity nerve growth factor receptor-derived cells, and MSCs, and comparing these cells to see what they are expressing on the surface and also what growth factors they might be secreting. We'll confirm the array data by real-time PCR, and then start to look of secreted proteins in the cells and the conditioned medium with Westerns and ELISAs.

Gregory: I'm probably going to come back to dovetailing proteomics with transcriptome analysis, but the only way you could link a given surface epitope with a given repertoire of trophic factors is probably by large-scale transcriptome analysis.

Spees: I agree. Going at one protein at a time is probably not the way to do it.

Gregory: So the ultimate goal of this venture is to separate the cells out into different populations, predict their repertoire of secreted agents, and tailor-make a therapeutic for a given disease?

Spees: Yes, that's the right idea.

Integrating Genomics and Proteomics for Cell Classification

Gregory: The problem right now with transcriptome analysis is validation at the level of RNA because the hybridization technology is great, but it's not perfect. It also needs to be linked to proteomics to confirm absolutely, that what you're seeing is what you're getting. So how do you think the Affymetrix microarray data, real-time PCR, and maybe proteomics can be linked to open the future? I mean, do you think that's possible?

Spees: In general, when we run microarrays from Affymetrix, and then we run, for example, fluidics cards with panels of transcripts, the things that are up on the Affymetrix chip are generally up by real-time PCR also, just in a more quantitative fashion. And it's only really low abundance transcripts that you'll miss by a large-scale hybridization. But even transcripts that are expressed at relatively low levels often come up on the Affymetrix chip. The question would be, how do we link that to large-scale proteomic techniques to take it to the next level.

Gregory: Our Dkk1 data was derived from the old-school methodologies of radiolabeling, 2D separation, purification, and then tryptic mapping. In the future, if monoclonal antibodies become available, it might be possible to engineer a chip that corresponds, as best it can, with microfluidics cards and Affymetrix chips, and that would result in the bottleneck being removed.

Spees: To look simultaneously at 15 or 16 Wnt proteins?

Gregory: Even with the technology as it stands, you can do dot blotting with hundreds of proteins using the limited number of monoclonal antibodies available. So you combine the Affymetrix microarrays with real-time confirmation and then follow with proteomics, you'll have a much better idea of what's going on. Again, you are getting away from one opinion-leader's favorite repertoire of proteins.

If you can take a mixed population of cells, find a repertoire of different epitopes on the surface, separate them out accordingly, and then check the characteristics of each cell type, then you have a therapeutic with different flavors for different applications.

FOR MORE INFORMATION

Contacts:

■Carl Gregory, Ph.D.
Tulane Center for Gene Therapy
Tulane University Health Sciences Center
J. Bennett Johnston Building
1324 Tulane Avenue, SL-99
New Orleans, LA 70112-2699
cgregory@tulane.edu

■Jeffrey L. Spees, Ph.D.
Department of Medicine
Cardiovascular Research Institute
University of Vermont
208 South Park Drive, Suite 2
Colchester, VT 05446
Jeffrey.Spees@uvm.edu

Companies:

■Affymetrix Inc. – http://www.affymetrix.com

Organizations:

■University of Vermont - http://www.med.uvm.edu/

■Tulane University - http://www.tulane.edu/

Further Reading:

■Spees JL, Olson SD, Ylostalo J, Lynch PJ, Smith J, Perry A, Peister A, Wang MY and Prockop DJ. Differentiation, cell fusion and nuclear fusion during ex vivo repair of epithelium by human adult stem cells from bone marrow stroma (hMSCs). *Proc Natl Acad Sci USA.* 2003, 100 (5):2397-2402.

■Spees, JL*, Gregory CA*, Singh H, Tucker HA, Peister A, Lynch PJ, Smith J and Prockop DJ. Internalized antigens must be removed to prepare hypo-immunogenic mesenchymal stem cells for cell and gene therapy. *Molecular Therapy. 2004,* 9 (5): 747-756. (*) equal contributors

■Gregory GA, Singh H, A. S. Perry AS and Prockop DJ. The Wnt signaling inhibitor Dickkopf-1 is required for reentry into the cell cycle of human adult stem cells from bone marrow. *The Journal of Biological Chemistry.* 2003, 278: 28067-28078.

■Gregory CA, Gunn WG, Reyes E, Smolarz AJ, Munoz J, Spees JL and Prockop DJ. How Wnt signaling affects bone repair by mesenchymal stem cells from the bone marrow. *Ann NY Acad Sci. 2005,*1049: 1-10.

■Gregory CA, Prockop DJ, and Spees JL. Non-hematopoietic bone marrow stem cells: Molecular control of expansion and differentiation. *Experimental Cell Res. 2005,* 306: 330-335.

EXPRESSION

Microarrays Used to Understand How Stem Cells Maintain Pluripotency and Regulate Cell Fate Decisions

Princeton's Ihor Lemischka and Virginia Commonwealth's Raj Rao discuss new data describing how stem cells maintain the ability to differentiate into various body tissues

Scientists led by Princeton University's Ihor Lemischka have used microarrays to discover previously unknown proteins that maintain pluripotency and regulate cell fate decisions in embryonic stem (ES) cells. Control of the self-renewing capability and the differentiation pathways of stem cells is necessary for developing regenerative therapies.

Lemischka recently presented preliminary work at the annual meeting of the International Society for Stem Cell Research detailing his studies using genomic strategies to analyze the regulation of stem cell fate decisions.

"Our recent results are just the tip of the iceberg," said Lemischka, professor of molecular biology at Princeton. "There are many, many, many more molecules playing interesting roles in stem cells."

Raj Rao, assistant professor of Chemical and Life Science Engineering at Virginia Commonwealth, studies human and primate embryonic stem cells in order to develop better regenerative therapeutic treatments and biomaterials for culturing and differentiating stem cells.

AMB Editor Rachel Shreter recently spoke to Lemischka and Rao about:

- The use of microarrays to understand stem cell function
- Known factors involved in maintaining pluripotency in stem cells
- Data analysis techniques in the stem cell community
- Future directions for studying maintenance of pluripotency

Microarrays in Embryonic Stem Cell Research

Lemischka: For us, the value of arrays is that we should be able to take away one gene at a time in a controlled way. We could do that with inhibitory RNA techniques and then watch what happens to the biological fate of the cells over time—to the transcriptome, the proteome and to chromatin structure. We can perturb the system in as minimal a way as possible and then see what happens on multiple levels simultaneously. Hopefully, that gives us some real insight.

The comparative approach is really a powerful one. Unfortunately it doesn't yet show up that much in the stem cell field, except for the work our lab, Raj's and what several others are doing. Some people in the stem cell field do not yet understand what the cancer people and the yeast people have understood for a long time—that using microarrays takes careful statistics and it's not as simple as looking at a gene list. Different ways of treating the data can give you very different answers and people need to be aware of that. I think that this situation is now improving.

Complicating matters in the stem cell field is that the power of the microarray is in the ability to detect two-fold or even lower differences in gene expression. Unless the biology of your system is such that you can detect two-fold changes in cell fate or biological outcome, the power of microarrays is lost.

Rao: Stringent statistical approaches are important. And things can be especially confusing when we talk about fold-change values and fold-change analysis. We've done most of our analysis using a mixed-linear model, so we get a fold-change value based on the specific treatment, which is differentiation in our case.

With most of the existing statistical techniques, you get the global fold-change value, and that doesn't present the true picture. So it's extremely important to use models of statistical methods that give you the fold-change value specific to your treatment of interest.

Shreter: Do you think further advancement of microRNAs and alternative transcripts will help elucidate the mechanisms responsible for maintaining pluripotency or differentiation?

Lemischka: Oh, definitely. We just completed a computational/experimental project looking at genome-wide alternative splicing patterns in [embryonic stem] ES cells and in hematopoietic stem cells. It would be great if microarray companies would develop arrays to look at alternative splicing patterns on a genome-wide level. I know there has been some progress along those lines.

Rao: I think it's important to ascertain the functional relevance of many of these candidate genes as they pertain to pluripotency. I have some preliminary microarray data that compares normal and abnormal cells. We see some skewed chromosomal distribution of genes on particular chromosomes. So I'm interested in looking at functional relevance of some of those genes. Hopefully, I'll be able to follow that up with transcriptional profiling using microarrays.

Lemischka: We are doing similar things in our lab. The libraries available for inhibitory RNA experiments are really extensive. They approach genome-wide coverage in both mouse and human. MicroRNAs and what they normally do in cell differentiation are really interesting. I will be looking at that in the near future. It would certainly be useful to have microarrays that cover microRNAs.

Current Understanding of Pluripotency

Shreter: What is currently known about the mechanisms regulating pluripotency?

Lemischka: All the embryonic stem cell work done in my lab uses mouse embryonic cell lines. In mouse ES lines, we know leukemia inhibitory factor (LIF) and bone morphogenic protein (BMP) signaling are required for pluripotency.

There are two homeodomain proteins, Nanog and Oct4, that are also required to maintain pluripotency in these cells; but LIF and BMP appear to be the only two

external cytokine or growth factor signaling pathways required to maintain cells undifferentiated in vitro.

There is a little bit of information that suggests that Wingless or Wnt signaling might be important to pluripotency as well. Those data are intriguing and probably

> "We can perturb the system in as minimal a way as possible and then see what happens on multiple levels simultaneously."

correct, but I think the involvement of Wnt signaling remains to be seen.

In terms of other genes that might be involved, SOX2 is a transcriptional regulator that works in concert with Oct4 at least part of the time.

So we know quite a bit, but we haven't been able to put it all together yet.

Rao: I think we are still at the tip of the iceberg. But there is a difference between self-renewal mechanisms in mouse and human embryonic stem cells. In human embryonic stem cells there is still no established role for LIF, but there is definitely a role for FGF (Fibroblast Growth Factor) and BMP (Bone Morphogenetic Protein) signaling.

It has been shown that the addition of basic FGF in conjunction with inhibition of BMP4 signaling maintains these stem cells in an undifferentiated state. As Dr. Lemischka mentioned, there are also roles for the transcription factors, Oct4 and Sox2, in maintaining the undifferentiated state. In fact, in the global pluripotent gene list generated by different groups, many of these transcription factors, and many of the key molecules that are components of both the FGF signaling and BMP4 signaling pathways, show up. So there is a role for all these molecules in maintaining the undifferentiated state.

Flavors of Pluripotency

Shreter: Dr. Rao, you recently did a comparative study of two established human embryonic cell lines. What was the rationale behind your study and what did you find?

Rao: The rationale behind the research was pretty straightforward. We have different embryonic stem cell lines that are currently being subjected to diverse passaging conditions. We started by comparing cell lines grown in two different labs. We obtained a microarray data set from Dr. Ali Brivanlou's lab at Rockefeller, which corresponds to the H1 cell line. The other data set, which corresponds to



Dr. Ihor Lemischka is a professor in the Department of Molecular Biology at Princeton University. His laboratory is focused on using high-throughput, functional genomic approaches to investigate the regulation of cell fate in a variety of stem cell types, including hematopoietic, neural and embryonic stem cells. His work includes using microarray technology combined with high-throughput inhibitory RNA techniques to understand cell fate decisions and to distinguish stem cells from their more mature progeny.



Raj Rao is an assistant professor in the Department of Chemical and Life Science Engineering at Virginia Commonwealth University. The work in his laboratory focuses on transcriptional profiling of human embryonic stem cells with the goal of understanding the pluripotent state and characterizing early differentiation events. Earlier this year, his laboratory published their findings on the genomic integrity of presidential approved stem cell lines.

the BG01 cell line, originated in our lab while we worked with Dr. Steve Stice's group at University of Georgia.

We decided to use a stringent uniform statistical approach with a mixed-linear model and we found that there are both similarities and differences between cell lines. And although we found a total of 133 genes common between the cell lines, we found many genes unique to each cell line. It's worth noting that where there were genes unique to each cell line, there were a lot of similarities in functional ontologies.

Based on this information, you can say that there are flavors of pluripotency with a lot of similarities to each other. From that perspective, it's important to combine all these gene lists and look at their functional relevance as it pertains to pluripotency.

> "...it's extremely important to use models of statistical methods that give you the fold-change value specific to your treatment of interest."

Lemischka: I like the idea of different flavors of pluripotency. We have also been doing a considerable number of studies collecting both human and mouse stem cell gene expression data sets from different laboratories. What Raj said is absolutely correct. There are differences, but it appears that—using some algorithms or some computational tools—you can come up with broad similarities as well.

I think, especially if you take the analysis beyond the individual gene or sequence level and start looking at conservation of similar pathways, you begin to see fairly deep similarities. Those data go hand in hand with what people are finding, say, when they compare what kinds of signals work in regulating hair follicle stem cells or intestinal stem cells, and even bone marrow stem cells.

Shreter: Do you think the differences in cell lines are related to the underlying genetic differences between the embryos they were derived from or are they due to mutations that accumulate with passaging?

Lemischka: Probably a little bit of both. For example, we grow clonal populations of embryonic stem cells from the mouse, but it's difficult to say if that truly represents an absolutely homogeneous cell population, even under the best of circumstances.

It's essential to have a standard set of control markers so you know for sure what the cells are, what they can do, and whether they changed in culture before using them in an array experiment or any other kind of high-throughput experiment. There are a lot of variables in a microarray experiment design and more variables in the biological design. These variables must be controlled as carefully as possible.

Rao: The differences we observe right now could also be because of the diverse cell culture and passaging conditions that have been adopted for human embryonic stem cells. It's extremely important for the stem cell community to be aware of that and to develop methods that are more uniform across laboratories.

Differences in Analysis

Shreter: Is it important for everyone to be on the same page in terms of how they analyze their data?

Lemischka: I think it might actually be bad if everybody were on the same page right now. I don't think there is one optimal method of statistical analysis. In fact, for different experiments and different systems, the optimal statistical analytical techniques are likely to vary.

First of all, we need to make the data sets available. That is happening in the stem cell community, but there are still some people who don't realize how important it is for different people to use different statistical tools to analyze the same sets of data obtained from different laboratories. That's done all the time in the yeast and the cancer communities.

I'm not sure the kinds of analysis that the yeast community and the cancer community do with array data sets will be possible in the stem cell field, however. No single stem cell lab is ever going to generate array data or genome-wide array data across sev-

eral hundred or thousand points. That happens for cancer and yeast because there are so many different labs doing work in those areas.

Rao: I think it's extremely important for researchers in the stem cell community to share the array data. We were lucky to get the H1 array data from Brivanlou's lab. It allowed us to compare their statistical techniques to ours. We found that the gene list produced differs depending on the statistical techniques used.

Although there are a lot of similarities, they found 918 genes, while our statistical technique gave us 1,300. Out of those, close to 700 were common. So, you may get a different gene list with different statistical techniques, but overall there will be a lot of similarities. That's important to note. And again, as Dr. Lemischka said, the statistical techniques you use depend on the experimental design.

Future Directions

Shreter: What are the next steps to take before you can apply the information you've obtained from these studies?

Lemischka: We don't know how the regulatory network works in space and time inside these cells, or how they talk to, and interact with signals coming from the outside. I think it will require a systems biology approach to stem cells to figure these things out.

From a basic scientific point of view, I think the mammalian stem cell community could take some lessons from what's being done in fruit flies, *C. elegans*, and even in yeast, or in the humble bacteriophage lambda. In the bacteriophage, lambda makes the cell fate choice. It either lyses the bacterium or becomes a passenger in the bacterial chromosome. That's a biological fate decision that's fairly well understood right now, but it is only understood because people have started to model it and use systems biology approaches, in addition to traditional experimental avenues.

Rao: Since my background has been at the interface of biology and engineering, my interests differ slightly. We are working on developing the enabling technologies for stem cell research. I'd like to use some of the results from our microarray studies and integrate them with existing engineering design principles. For example, what we'd like to do is use biomaterials to develop so-called feeder-free systems for growing stem cells. I am hoping we can integrate the lessons we learn into biomaterials to help develop these different approaches for growing stem cells and maintaining them in an undifferentiated state.

> "There are a lot of variables in a microarray experiment design and more variables in the biological design. These variables must be controlled as carefully as possible."

FOR MORE INFORMATION

Contact Information

■Ihor Lemischka, Ph.D.
Professor of Molecular Biology
Lewis Thomas Lab-210
Princeton University
Washington Road
Princeton, NJ 08544
ilemischka@molbio.princeton.edu

■Raj R. Rao, Ph.D.
Assistant Professor
Department of Chemical and Life Science Engineering
Virginia Commonwealth University
601 West Main Street, Room 414
P.O. Box 843028
Richmond, VA 23284
rrrao@vcu.edu

Companies

■Affymetrix Inc. – http://www.affymetrix.com

Organizations

■Virginia Commonwealth University – http://www.vcu.edu

■Princeton University – http://www.princeton.edu

■Rockefeller University – http://www.rockefeller.edu

■University of Georgia – http://www.uga.edu

Further Reading

■Lemischka IR. Stem Cell Biology: A View toward the Future. *Ann N Y Acad Sci* 2005; 1044:132-8.

■Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR. A stem cell molecular signature. *Science* 2002; 298(5593):601-4.

■Mitalipova MM, Rao RR, Hoyer D, Jones K, Johnson J, Dalton S, Meisner L, Stice SL. Preserving the genetic integrity of human embryonic stem cell lines. *Nature Biotechnology.* 2005; 23(1): 19-20.

■Rao RR and Stice SL. Gene expression profiling of embryonic stem cells leads to greater understanding of pluripotency and early developmental events. *Biol Reprod.* 2004; 71(6):1772-1778.

■Rao, RR, Calhoun, Qin, X, Rekaya, R, Clark, J, and SL Stice. Comparative transcriptional profiling of two human embryonic stem cell lines. *Biotechnology and Bioengineering.* 2004; 88(3):273-286.

People

■Ali Brivanlou, Ph.D.
Rockefeller University
http://www.rockefeller.edu/graduate/faculty/research/abstract.php?id=127

■Steven L. Stice, Ph.D.
The University of Georgia
http://www.biomed.uga.edu/mem_stice_steve.htm

WWW.MICROARRAYBULLETIN.COM

LOG ON FOR A FREE SUBSCRIPTION

AFFYMETRIX

MICROARRAY BULLETIN

Affymetrix Microarray Bulletin (AMB) is a quarterly newsletter published by Affymetrix Inc. that highlights the work of thought leaders in molecular biology and microarray research from around the world. AMB is a scientist-to-scientist publication, written for researchers, statisticians, and software and information technology engineers. Our goal is to create a forum where the research community can discuss its work and share methods, tools, insights and information.

The online version of this newsletter can be accessed at http://www.microarraybulletin.com

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*

Affymetrix Inc.
3420 Central Expressway
Santa Clara, CA 95051
408.731.5000

AFFYMETRIX®