

The Reproducibility Crisis:

A God-Sent Opportunity for STM Publishers?

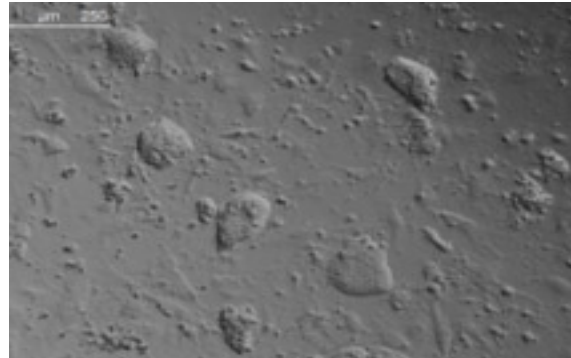
Moshe Pritsker, Ph.D.
CEO, Co-founder
JoVE

STM Annual Conference
April 25, 2017

Reproducibility: What is it?



ES cells were maintained in serum-free culture without feeder cells. ES cells were plated onto gelatin-coated plates in N2B medium and were passaged every 2–4 days. Dissociated cells were harvested in N2B medium, pelleted, resuspended in N2B and replated directly.



Reproducibility: ability to repeat (reproduce) a published experiment

The Reproducibility Crisis is Real



67 target-validation projects
in oncology,
women's health and
cardiovascular medicine

21% Replication Rate (14/67)



53 "landmark" articles in
cancer research published by
reputable labs in top journals

11% Replication Rate (6/53)

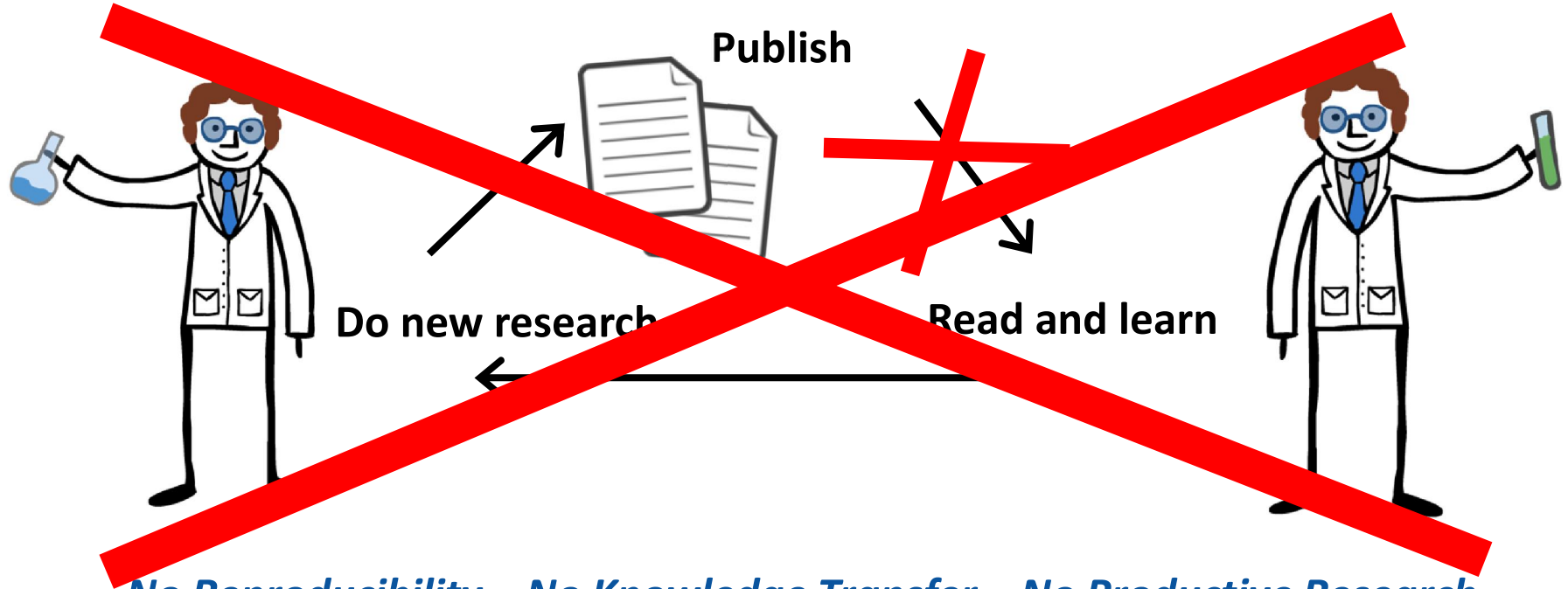


Replication of
published articles in
experimental psychology

33% Replication Rate

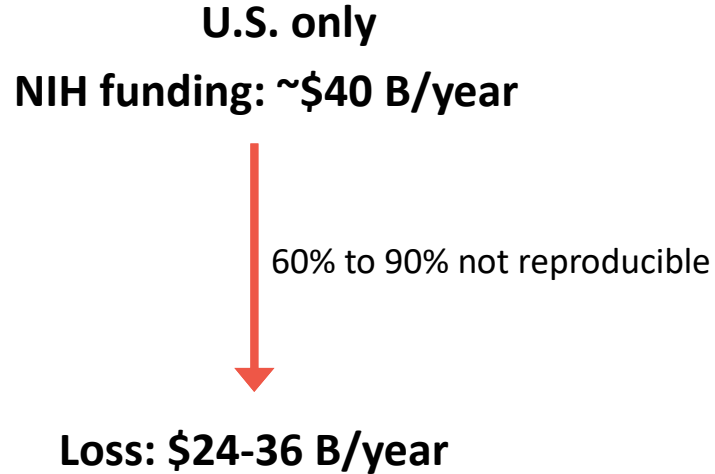
60 – 90% of published science articles are not reproducible!!!

How the Lack of Reproducibility Damages Science



No Reproducibility = No Knowledge Transfer = No Productive Research

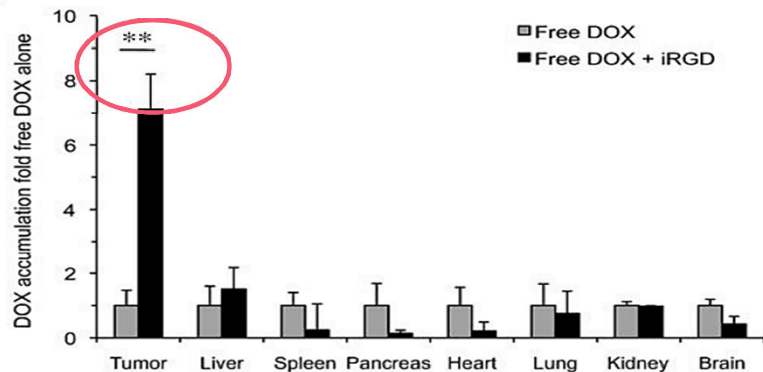
The Lack of Reproducibility is a Serious Problem



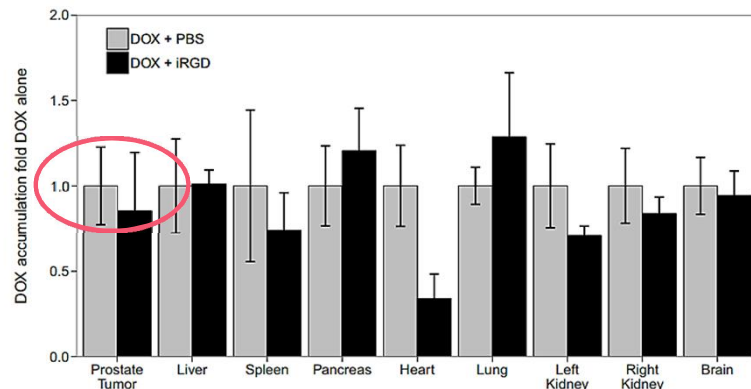
Impact:

- **For society:** delays in development of new medicines and low ROI
- **For institutions:** poor productivity
- **For scientists:** broken careers

How do Scientists Deal with Reproducibility in the Lab?



Sugahara et al. 2010 Science (Original Study)



Mantis et al. 2017 eLife (Cancer Reproducibility Project)

1. Repeat again and again until works
2. Find someone at your institution who can **show** you the experiment
3. Travel to original authors who can **show** you the experiment

Why Seeing an Experiment Works?

Text article

Position the metaphase spindle at 3 o'clock and hold it with holding pipette. Apply piezo pulses to penetrate the zona pellucida. Touch the metaphase plate with the enucleation pipette. Aspirate the spindle and withdraw the pipette.

Real life

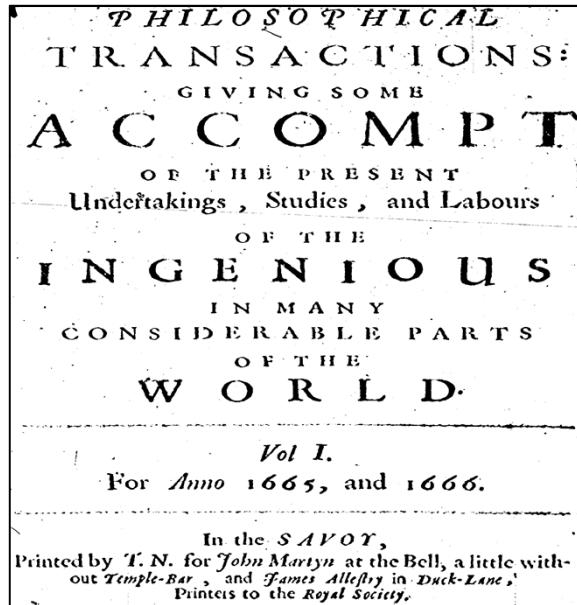


Visualization improves knowledge transfer of how-to (methods)

Traditional Science Article Format: Adequate for 2017?

1665: First scientific article

Philosophical Transactions of the Royal Society



2017: What's changed?

Genomewide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells

Moshe Pritsker*, Nicole R. Ford, Harry T. Jenq, and Ihor R. Lemischka*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Edited by Rudolf Jaenisch, Whitehead Institute for Biomedical Research, Cambridge, MA, and approved March 23, 2006 (received for review November 11, 2005)

Embryonic stem (ES) cells hold great promise for the future of medicine. To elucidate the molecular mechanisms that control ES cell self-renewal and differentiation, a comprehensive knowledge of the molecules involved in these processes is required. Here we describe an effective approach for genomewide identification of functionally active genes in ES cells. This approach combines genetic screens based on cDNA libraries with microarray detection methods to permit high-throughput functional analyses. We implement this strategy to identify genes whose overexpression can maintain phenotypic properties of undifferentiated mouse ES cells under differentiation-inducing conditions, specifically in the absence of leukemia inhibitory factor. The identified genes encode a variety of regulatory proteins whose function in ES cells was previously unknown. Moreover, our approach is capable of detecting genes whose overexpression promote differentiation or cell death. Overall, our studies establish a methodology for highly sensitive identification of genes that confer particular phenotypes on ES cells.

cDNA library | differentiation | microarray | phenotype | self-renewal

function screens identify only the "most potent" phenotype-conferring genes. For example, the only currently described gain-of-function screen in mouse ES cells identified a single gene encoding the transcription factor Nanog (5). Moreover, the clone sequencing method is incapable of negative detection, complicating the identification of gene products that promote differentiation or apoptosis (18). These limitations prevent the comprehensive identification of functional genes in mammalian systems.

In this study, we applied the microarray technology as a method of large-scale parallel analysis to conduct comprehensive gain-of-function screens in ES cells (Fig. 1A). Our approach was designed to simultaneously monitor the activity of all gene products in a cDNA library as they function to mediate a given phenotype. We implemented this approach to identify genes whose overexpression is sufficient to maintain undifferentiated mouse ES cells in the face of differentiation-inducing conditions, specifically in the absence of LIF. We also show that our approach is capable of negative detection and can identify gene products that promote differentiation or cell death.

Current Proposals to Solve the Reproducibility Crisis

- More data preservation
- Better training in statistics
- Administrative measures directed at scientists

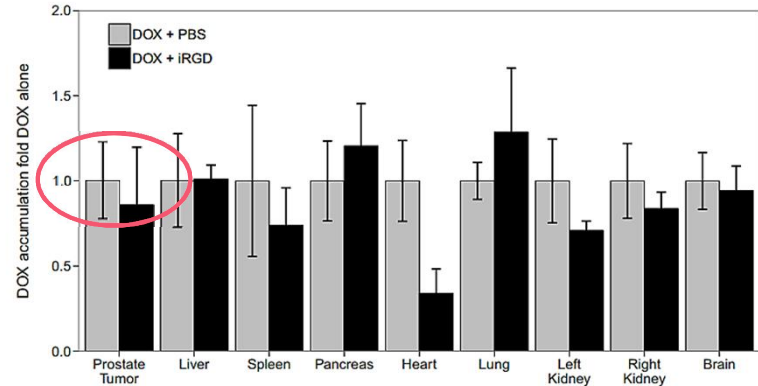
Do not improve knowledge transfer of methods

Will not solve the reproducibility crisis

Why Statistics Will Not Help Much



Sugahara et al. 2010 Science (Original Study)



Mantis et al. 2017 eLife (Cancer Reproducibility Project)

***The result is statistically significant, yet it is not reproducible.
It is all about methods.***

***The STM Industry Can Solve the
Reproducibility Crisis by Improving the
Knowledge Transfer of How-To (Methods)***

Examples of Addressing the Reproducibility Crisis by Improving Article Format

- **Visualization (Biological and Physical Sciences)**
- **Interactivity (Computer Science)**

Visualized Science Article by JoVE

8 Imaging Local Ca^{2+} Signals in Cultured Mammalian Cells

Jeffrey T. Lock¹, Kyle L. Ellefsen¹, Bret Settle¹, Ian Parker^{1,2}, Ian F. Smith¹

¹Neurobiology and Behavior, University of California, Irvine, ²Physiology and Biophysics, University of California, Irvine

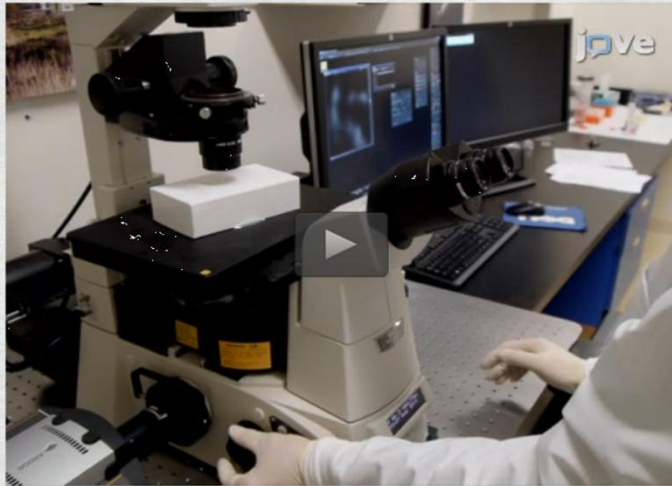
Article

Downloads

Comments

Metrics

Publish with JoVE

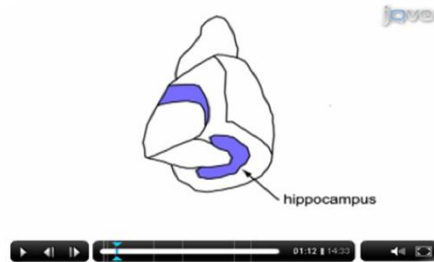


- 0:05 Title
- 2:07 Loading Cells with Membrane-permeant Cal-520, ci-IP₃ and EGTA
- 3:46 Ca²⁺ Image Acquisition
- 5:08 Automated Ca²⁺ Image Analysis
- 7:26 Results: IP₃-evoked Local Ca²⁺ Signals in SH-SY5Y Cells Loaded with Cal-520 and ci-IP₃
- 8:49 Conclusion

- Films experiments in universities around the world
- Publishes video articles
- Peer-reviewed editorial board
- Indexed in PubMed and Medline

Structure of Scientific Video Article

1. Abstract



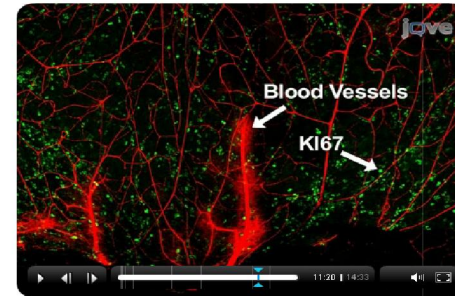
2. Introduction



3. Experimental procedure



4. Discussion of results

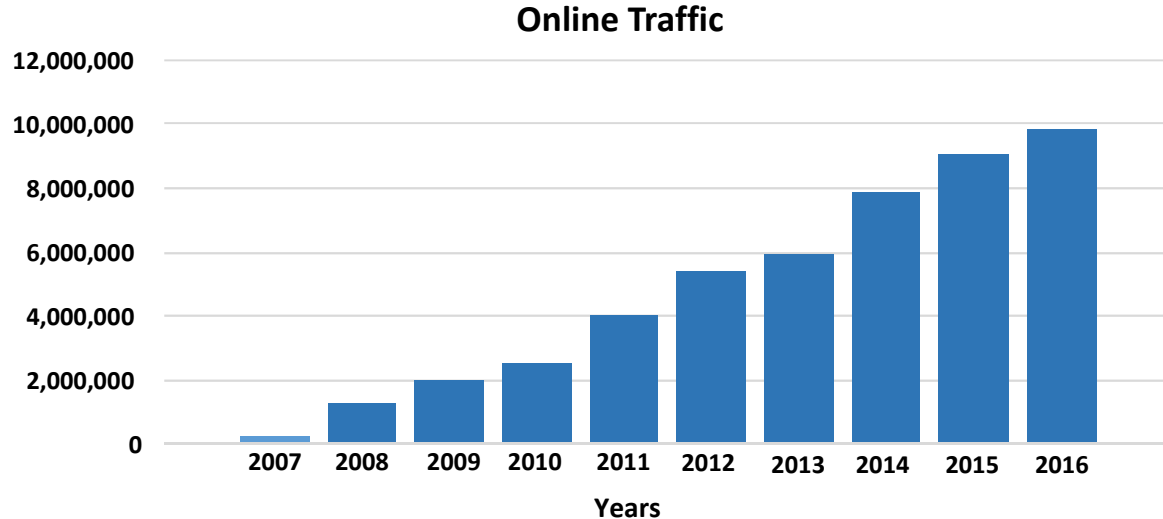


Adoption of Video Publication by Scientific Community

1,200+ peer-reviewed video articles published annually

23,000+ published authors

9,000,000+ unique users/year,
80% traffic from universities

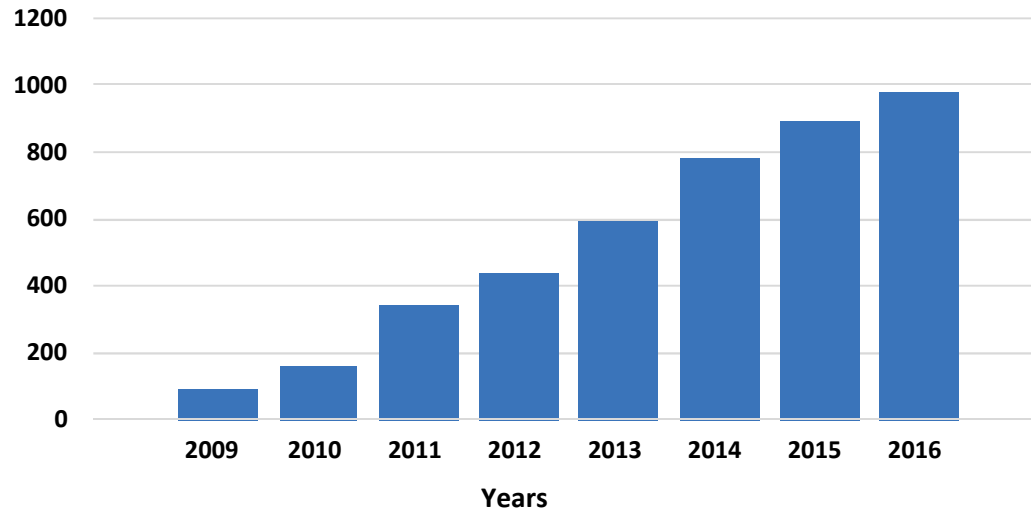


Adoption of Video Publication by Academic Institutions

1,000+ institutions subscribed

2,500 average number of views/video

Academic Institutions Subscribed to JoVE



Interactive Science Article: Executable Paper by Elsevier

The screenshot displays the Elsevier ScienceDirect website interface for an interactive article. The page is titled "Computers & Graphics" and is available online as of 1 May 2013. The article is "In Press, Corrected Proof" and is part of a "Special Section on 3D Object Retrieval".

The article title is "Data-aware 3D partitioning for generic shape retrieval", which is "Universally Available". The authors listed are Ivan Sipiran^a, Benjamin Bustos^b, and Tobias Schreck^b. The affiliations are: ^aKDW-PRISMA Research Group, Department of Computer Science, University of Chile, Chile; and ^bVisual Analytics Group, Department of Computer and Information Science, University of Konstanz, Germany. A DOI link is provided: <http://dx.doi.org/10.1016/j.sag.2013.04.002>, along with instructions on how to cite or link using DOI.

The article includes a "Highlights" section with the following points:

- We propose a simple and effective partitioning algorithm for 3D meshes.
- The use of part descriptions enhances the use of global descriptors.
- We define a distance as an optimization problem, including both linear and quadratic constraints.
- Our experiments show that the partitioning algorithm has a great influence in the final effectiveness of retrieval tasks.

The "Abstract" section states: "In this paper, we present a new approach for generic 3D shape retrieval based on a mesh partitioning scheme. Our method combines a mesh global description and mesh partition descriptions to represent a 3D".

The right-hand side of the interface features a sidebar with various interactive elements:

- Bibliographic Information**
- Citing and related articles**
- Applications and tools**
- Collage**
- More information on this application**
- Explore executable code and data belonging to this article**
- Author's introduction: Experiments Data-Aware Partitioning**
- Code 1 (Bash): Retrieval demo (Code 1)** - This section is highlighted with an orange box and contains a code editor with tabs for "Working Copy", "Retulaltes", "Original", and "Output". The code is:

```
#!/bin/sh
method="L2N"
mu="0.9"
descriptor="FANORAMA"
subList="4"

./Experiment1/retrievalExperiment.sh $object $method $mu $descriptor $subList
```
- Data 1: Query object**
- Data 2: First ranked model**
- Data 3: Second ranked model**
- Data 4: Third ranked model**
- Data 5: Fourth ranked model**
- Data 6: Fifth ranked model**
- Workspace**

Interactive Article: Reproducibility in Computer Science

- **Integration of executable code and data into science article**
- **Article-specific computer code runs while the article is read**
- **Built-in tools to visualize data such as 3D models**

elsevier.com/executablepaper

No Pain, No Gain

Improving scholarly articles with new features requires:

- Higher cost per article: new features cost money
- Higher technical expertise in-house (video, software...)
- Big changes in traditional publishing processes
- Significant upfront investment
- Author-pay business model is often not applicable

Opportunity for STM publishers!!!

Reproducibility Crisis: Opportunity for STM Publishers

Current STM industry issues

- Lower barrier to entry
- Negative perception among some users and buyers
- Government mandates challenging the traditional (subscription) business model

Improving science article format will

- Increase barrier to entry
- Improve user perception of new products
- Strengthen the traditional business model

What Is Your Next Move to Solve the Reproducibility Crisis?

Genomewide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells

Moshe Pritsker*, Nicole R. Ford, Harry T. Jenq, and Ihor R. Lemischka*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Edited by Rudolf Jaenisch, Whitehead Institute for Biomedical Research, Cambridge, MA, and approved March 23, 2006 (received for review November 11, 2005)

Embryonic stem (ES) cells hold great promise for the future of medicine. To elucidate the molecular mechanisms that control ES cell self-renewal and differentiation, a comprehensive knowledge of the molecules involved in these processes is required. Here we describe an effective approach for genomewide identification of functionally active genes in ES cells. This approach combines genetic screens based on cDNA libraries with microarray detection methods to permit high-throughput functional analyses. We implement this strategy to identify genes whose overexpression can maintain phenotypic properties of undifferentiated mouse ES cells under differentiation-inducing conditions, specifically in the absence of leukemia inhibitory factor. The identified genes encode a variety of regulatory proteins whose function in ES cells was previously unknown. Moreover, our approach is capable of detecting genes whose overexpression promote differentiation or cell death. Overall, our studies establish a methodology for highly sensitive identification of genes that confer particular phenotypes on ES cells.

cDNA library | differentiation | microarray | phenotype | self-renewal

function screens identify only the “most potent” phenotype-conferring genes. For example, the only currently described gain-of-function screen in mouse ES cells identified a single gene encoding the transcription factor Nanog (5). Moreover, the clone sequencing method is incapable of negative detection, complicating the identification of gene products that promote differentiation or apoptosis (18). These limitations prevent the comprehensive identification of functional genes in mammalian systems.

In this study, we applied the microarray technology as a method of large-scale parallel analysis to conduct comprehensive gain-of-function screens in ES cells (Fig. 1A). Our approach was designed to simultaneously monitor the activity of all gene products in a cDNA library as they function to mediate a given phenotype. We implemented this approach to identify genes whose overexpression is sufficient to maintain undifferentiated mouse ES cells in the face of differentiation-inducing conditions, specifically in the absence of LIF. We also show that our approach is capable of negative detection and can identify gene products that promote differentiation or cell death.

OR

Murine Model of Hindlimb Ischemia

Hiroshi Niiyama¹, Ngan F. Huang¹, Mark D. Rollins², John P. Cooke¹

¹Division of Cardiovascular Medicine, Stanford University, ²Department of Anesthesiology, University of California, San Francisco



0:00	Title
1:23	Introduction
2:15	Induction of Unilateral Hindlimb Ischemia
7:11	Laser Doppler Blood Perfusion
10:21	Representative Results/Outcome
11:29	Conclusion