Gene Expression Changes and Micro RNA Regulation in Embryonic Stem Cells

David Stroncek, MD
Chief, Cellular Therapies
Department of Transfusion Medicine
Clinical Center, NIH
Bethesda, MD, USA

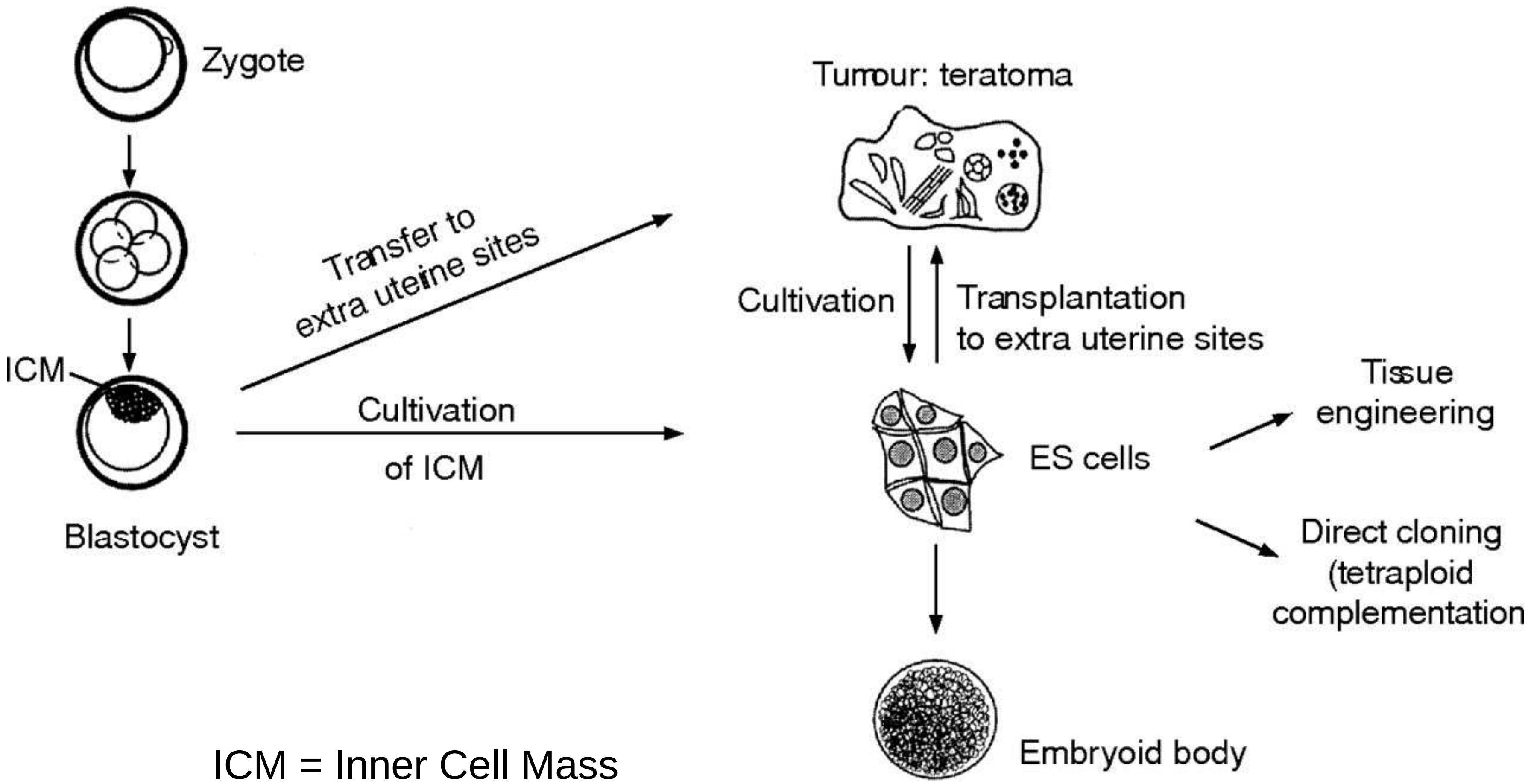
Presenter Disclosure

David Stroncek, MD

The following relationships exist related to this presentation:

No Relationships to Disclose

Embryonic Stem (ES) Cells



ICM = Inner Cell Mass

Denker, H-W J Med Ethics 2006;32:665-671

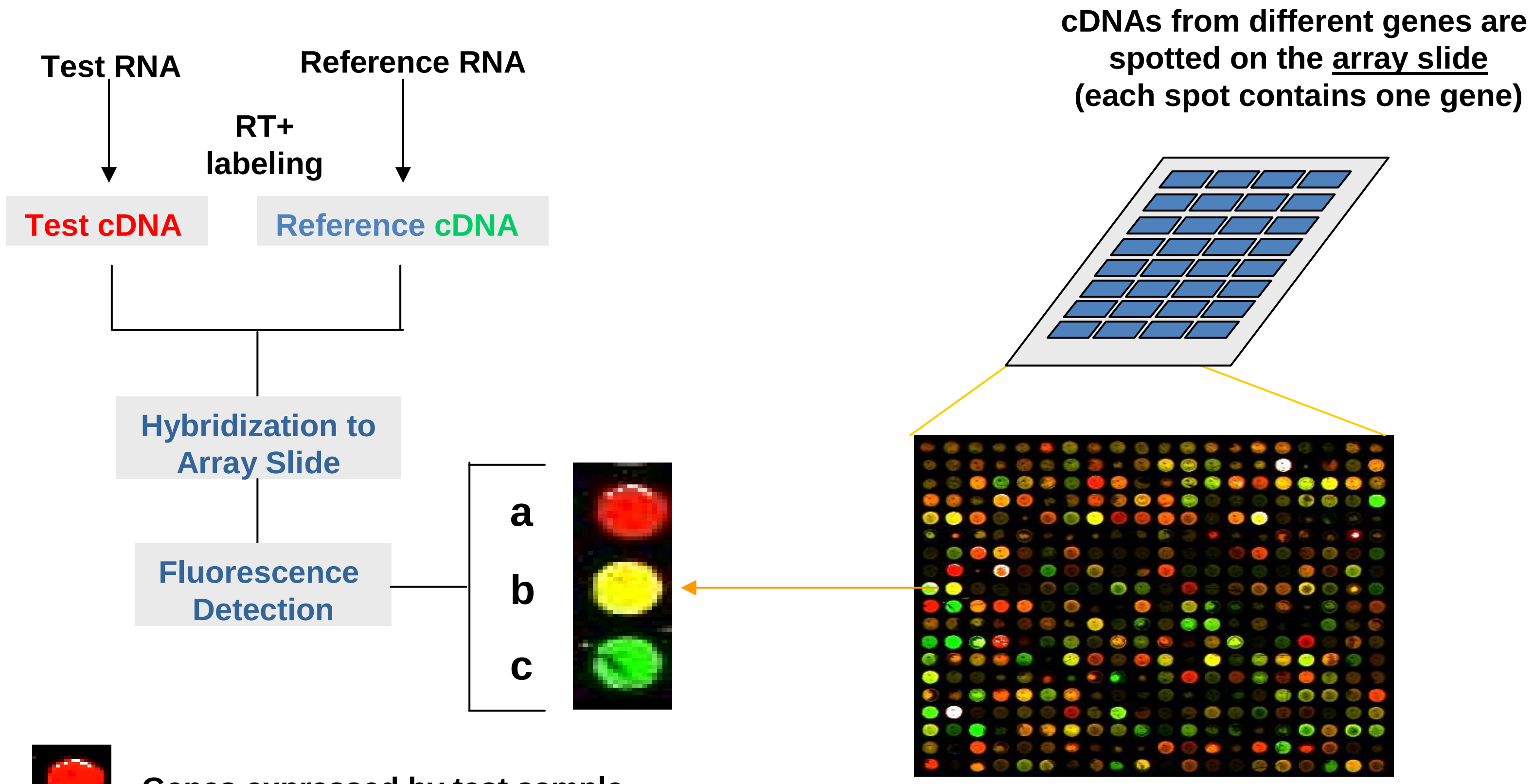
Copyright ©2006 BMJ Publishing Group Ltd.

Gene and micro RNA Expression by Human Embryonic Stem Cells

Early studies looked for the molecular basis of pluripotency

- What's expressed?
- What does it mean?

Principles of Gene Expression Analysis with Microarrays



Test RNA

Reference RNA

RT+ labeling

Test cDNA

Reference cDNA

Hybridization to Array Slide

Fluorescence Detection

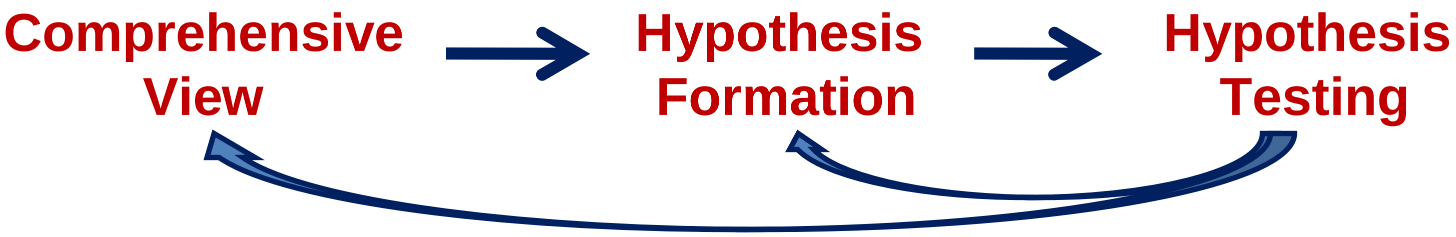
cDNAs from different genes are spotted on the <u>array slide</u> (each spot contains one gene)

a = 🔴 **Genes expressed by test sample**

b = 🟡 **Commonly expressed genes**

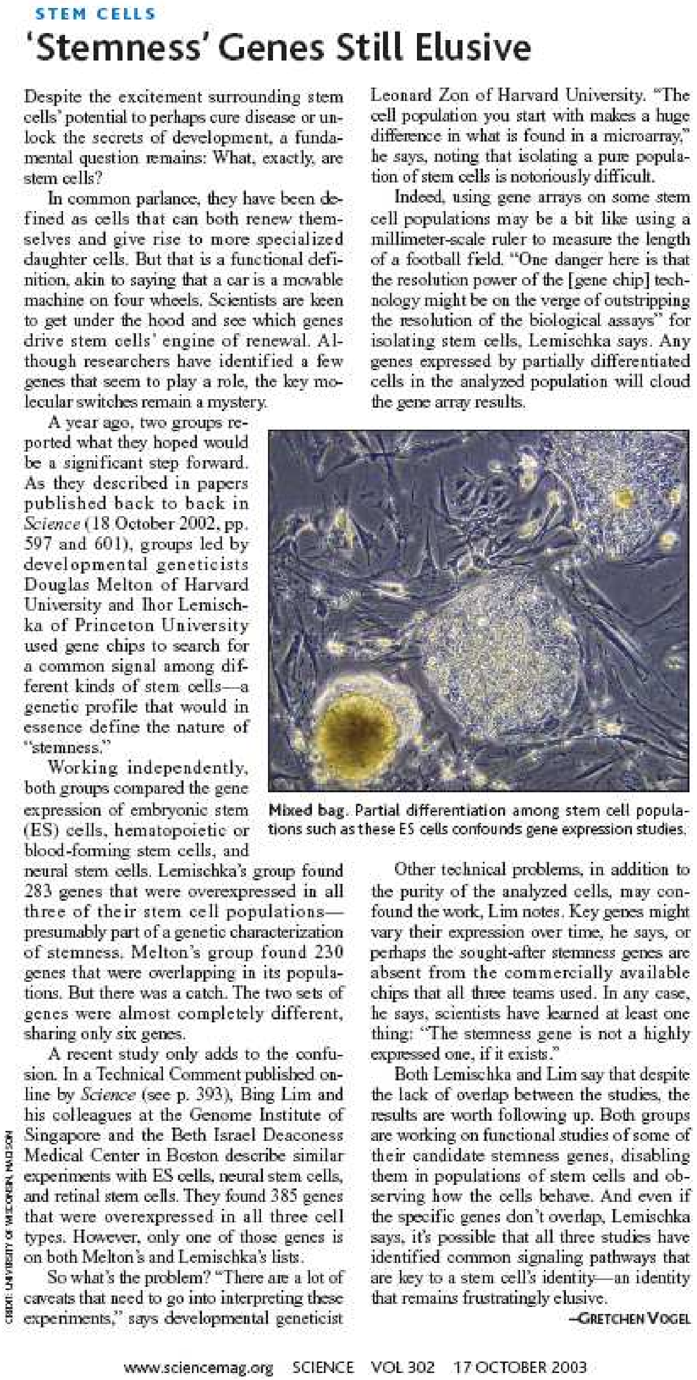
c = 🟢 **Genes expressed by reference sample**

Wang et al. ASHI Quarterly, 2001



Comprehensive View → Hypothesis Formation → Hypothesis Testing

Embryonic stem cell gene expression data has been difficult to interpret



STEM CELLS

'Stemness' Genes Still Elusive

Despite the excitement surrounding stem cells' potential to perhaps cure disease or unlock the secrets of development, a fundamental question remains: What, exactly, are stem cells?

In common parlance, they have been defined as cells that can both renew themselves and give rise to more specialized daughter cells. But that is a functional definition, akin to saying that a car is a movable machine on four wheels. Scientists are keen to get under the hood and see which genes drive stem cells' engine of renewal. Although researchers have identified a few genes that seem to play a role, the key molecular switches remain a mystery.

A year ago, two groups reported what they hoped would be a significant step forward. As they described in papers published back to back in *Science* (18 October 2002, pp. 597 and 601), groups led by developmental geneticists Douglas Melton of Harvard University and Ihor Lemischka of Princeton University used gene chips to search for a common signal among different kinds of stem cells—a genetic profile that would in essence define the nature of "stemness."

Mixed bag. Partial differentiation among stem cell populations such as these ES cells confounds gene expression studies.

CREDIT: UNIVERSITY OF WISCONSIN, MADISON

Working independently, both groups compared the gene expression of embryonic stem (ES) cells, hematopoietic or blood-forming stem cells, and neural stem cells. Lemischka's group found 283 genes that were overexpressed in all three of their stem cell populations—presumably part of a genetic characterization of stemness. Melton's group found 230 genes that were overlapping in its populations. But there was a catch. The two sets of genes were almost completely different, sharing only *six* genes.

A recent study only adds to the confusion. In a Technical Comment published online by *Science* (see p. 393), Bing Lim and his colleagues at the Genome Institute of Singapore and the Beth Israel Deaconess Medical Center in Boston describe similar experiments with ES cells, neural stem cells, and retinal stem cells. They found 385 genes that were overexpressed in all three cell types. However, only one of those genes is on both Melton's and Lemischka's lists.

So what's the problem? "There are a lot of caveats that need to go into interpreting these experiments," says developmental geneticist Leonard Zon of Harvard University. "The cell population you start with makes a huge difference in what is found in a microarray," he says, noting that isolating a pure population of stem cells is notoriously difficult.

Indeed, using gene arrays on some stem cell populations may be a bit like using a millimeter-scale ruler to measure the length of a football field. "One danger here is that the resolution power of the [gene chip] technology might be on the verge of outstripping the resolution of the biological assays" for isolating stem cells, Lemischka says. Any genes expressed by partially differentiated cells in the analyzed population will cloud the gene array results.

Other technical problems, in addition to the purity of the analyzed cells, may confound the work, Lim notes. Key genes might vary their expression over time, he says, or perhaps the sought-after stemness genes are absent from the commercially available chips that all three teams used. In any case, he says, scientists have learned at least one thing: "The stemness gene is not a highly expressed one, if it exists."

Both Lemischka and Lim say that despite the lack of overlap between the studies, the results are worth following up. Both groups are working on functional studies of some of their candidate stemness genes, disabling them in populations of stem cells and observing how the cells behave. And even if the specific genes don't overlap, Lemischka says, it's possible that all three studies have identified common signaling pathways that are key to a stem cell's identity—an identity that remains frustratingly elusive.

—GRETCHEN VOGEL

G Vogel. Science 2003;302:17

Limitations of Gene Expression Profiling

Platforms vary

Number of genes analyzed has varied: Increase with time

Results depend on types of cells analyzed

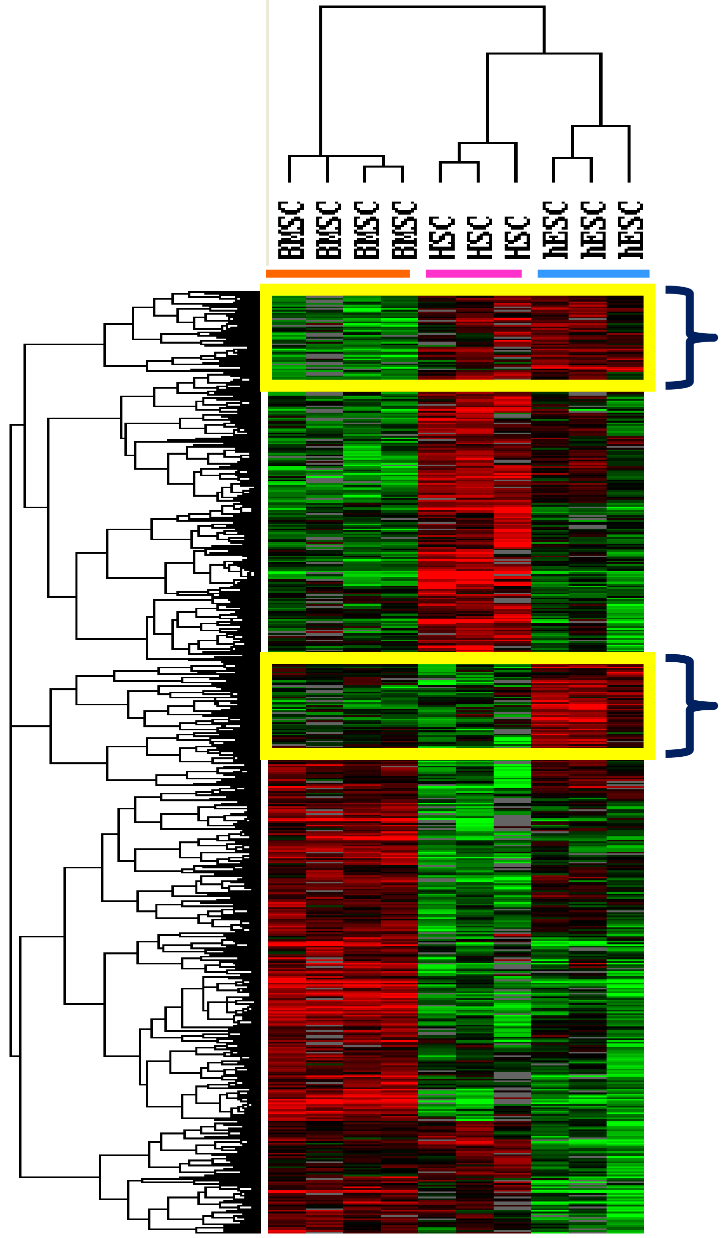
- ESCs vs adult cells
- ESCs vs hematopoietic stem cells
- Human ESCs vs mouse ESCs
- ESCs vs embryoid bodies

Cell state affects gene expression

Difficult to interpret the data

Uniquely Expressed Embryonic Stem Cell Genes:

Hierachical Clustering Analysis of Genes Differentially Expressed by Embryonic Stem Cells, Hematopoietic Stem Cells and Bone Marrow Stromal Cells



🟧 Bone Marrow Stromal Cells

🟪 Hematopoietic Stem Cells

🟦 Human Embryonic Stem Cells

Expressed by ESCs and HSCs

Unique to ESCs

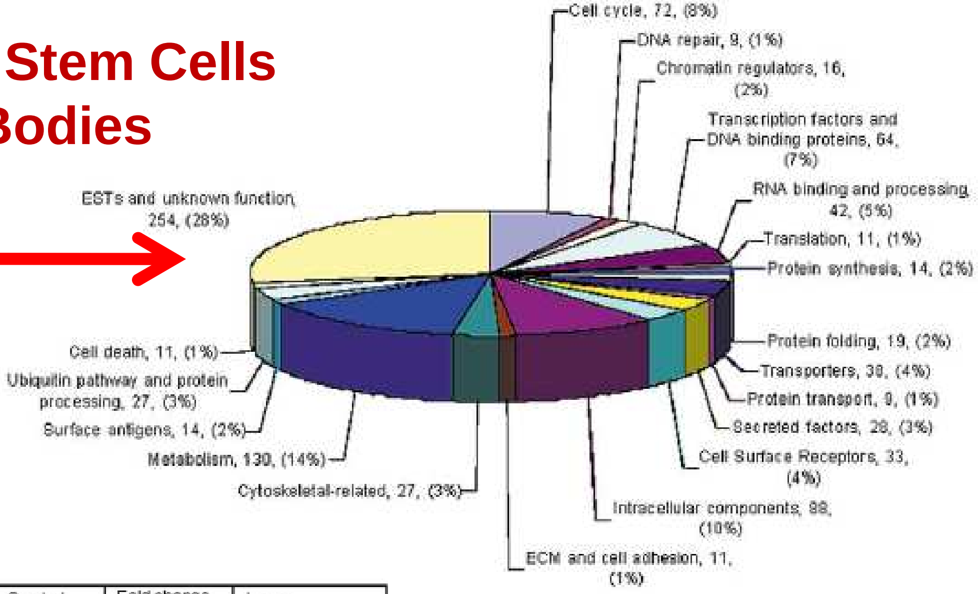
ESCs express a relatively small number of genes

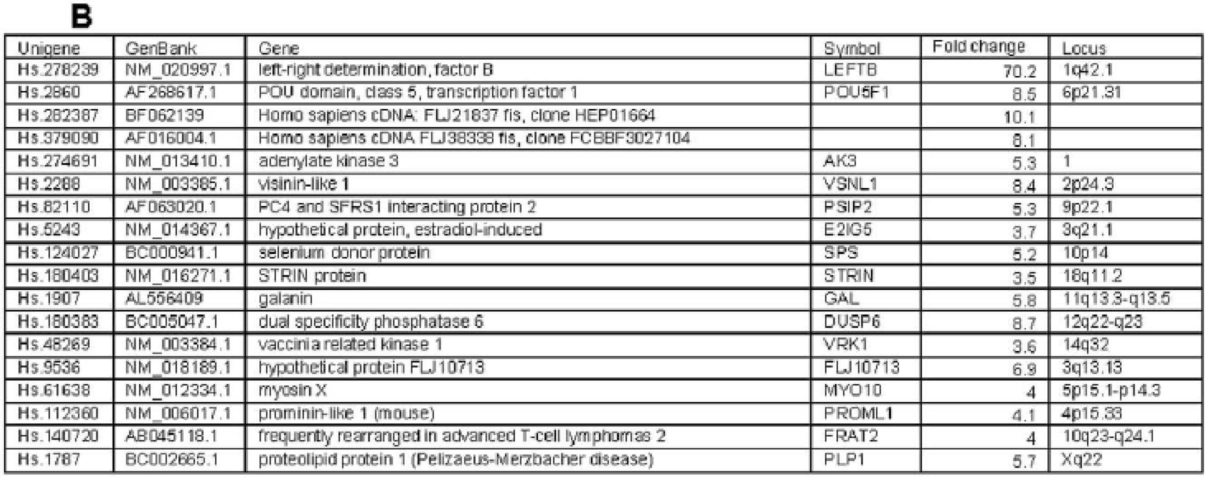
supervised cluster, 2800 genes, F-test, $p \leq 0.05$

Genes Enriched in Embryonic Stem Cells Compared to Embryoid Bodies

Functional Annotation

- Transcription Factors
- DNA repair
- Cell cycle
- RNA binding and processing





B

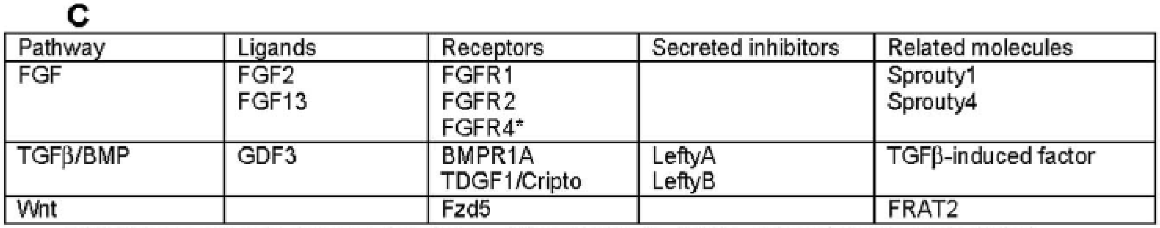
Unigene	GenBank	Gene	Symbol	Fold change	Locus
Hs.278239	NM_020997.1	left-right determination, factor B	LEFTB	70.2	1q42.1
Hs.2860	AF268617.1	POU domain, class 5, transcription factor 1	POU5F1	8.5	6p21.31
Hs.282387	BF062139	Homo sapiens cDNA: FLJ21837 fis, clone HEP01664		10.1	
Hs.379090	AF016004.1	Homo sapiens cDNA FLJ38338 fis, clone FCBBF3027104		8.1	
Hs.274691	NM_013410.1	adenylate kinase 3	AK3	5.3	1
Hs.2288	NM_003385.1	visinin-like 1	VSNL1	8.4	2p24.3
Hs.82110	AF063020.1	PC4 and SFRS1 interacting protein 2	PSIP2	5.3	9p22.1
Hs.5243	NM_014367.1	hypothetical protein, estradiol-induced	E2IG5	3.7	3q21.1
Hs.124027	BC000941.1	selenium donor protein	SPS	5.2	10p14
Hs.180403	NM_016271.1	STRIN protein	STRIN	3.5	18q11.2
Hs.1907	AL556409	galanin	GAL	5.8	11q13.3-q13.5
Hs.180383	BC005047.1	dual specificity phosphatase 6	DUSP6	8.7	12q22-q23
Hs.48269	NM_003384.1	vaccinia related kinase 1	VRK1	3.6	14q32
Hs.9536	NM_018189.1	hypothetical protein FLJ10713	FLJ10713	6.9	3q13.13
Hs.61638	NM_012334.1	myosin X	MYO10	4	5p15.1-p14.3
Hs.112360	NM_006017.1	prominin-like 1 (mouse)	PROML1	4.1	4p15.33
Hs.140720	AB045118.1	frequently rearranged in advanced T-cell lymphomas 2	FRAT2	4	10q23-q24.1
Hs.1787	BC002665.1	proteolipid protein 1 (Pelizaeus-Merzbacher disease)	PLP1	5.7	Xq22

Most Significantly enriched

- LeftyB
- Oct4

Enriched Signaling-Related Genes

- FGF
- TGFb
- Wnt



C

Pathway	Ligands	Receptors	Secreted inhibitors	Related molecules
FGF	FGF2 FGF13	FGFR1 FGFR2 FGFR4*		Sprouty1 Sprouty4
TGFβ/BMP	GDF3	BMPR1A TDGF1/Cripto	LeftyA LeftyB	TGFβ-induced factor
Wnt		Fzd5		FRAT2

*FGFR4 was strongly downregulated upon differentiation ($p=0.012$), although it was not included in the enriched genes ($p<0.01$).

Sato N et al. Developmental Bio. 2003;260:404-413

Genes Enriched in Embryonic Stem Cells Compared to Embryoid Bodies

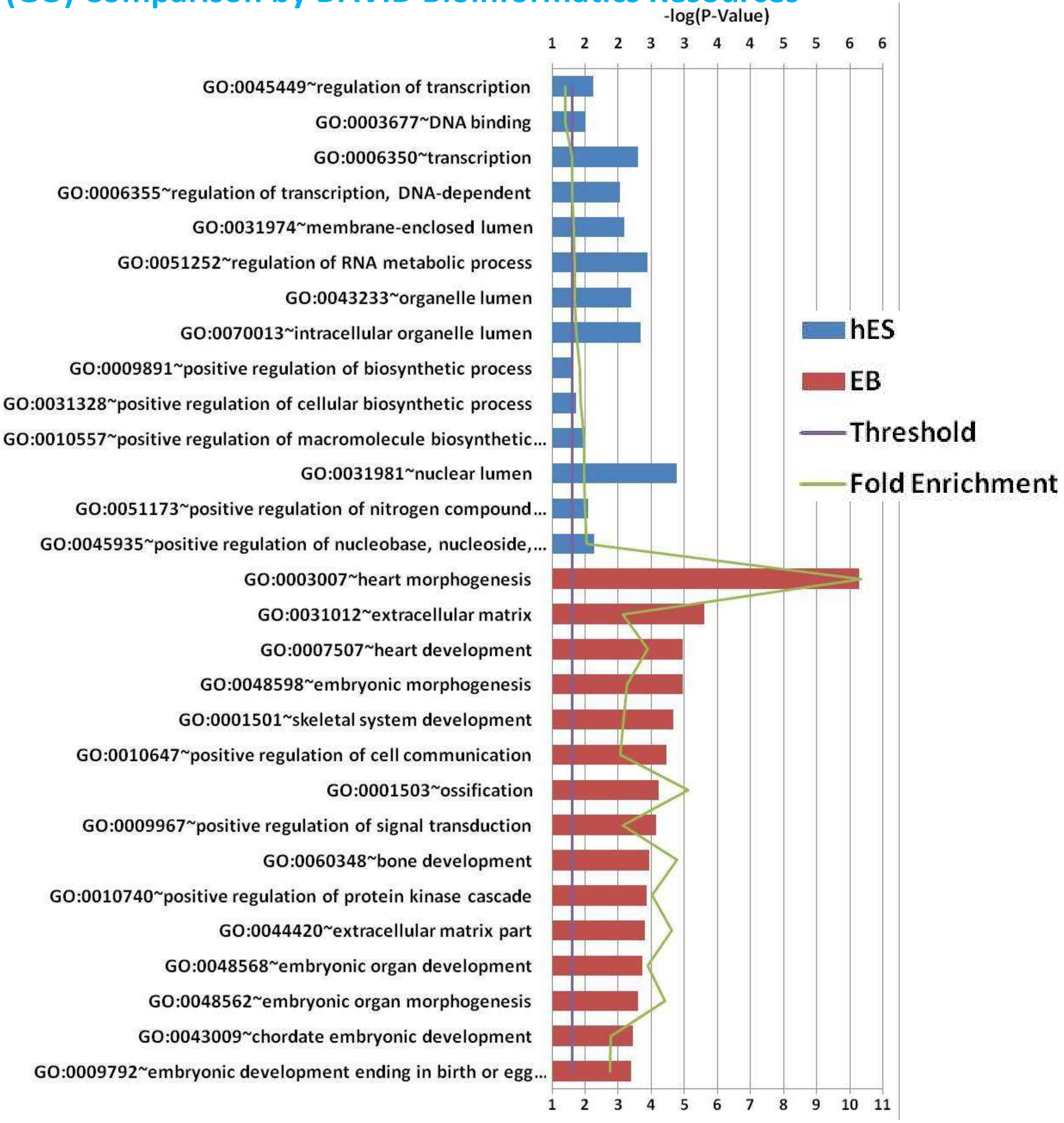
Gene Ontology (GO) Comparison by DAVID Bioinformatics Resources

Enriched in ESCs

- Regulation of transcription
- DNA binding
- RNA metabolism
- Biosynthetic processes

Enriched in EBs

- Organ Development
- Tissues development
- Embryonic development



Comparison of Embryonic Stem Cells with Bone Marrow Stromal Cells (BMSCs) (Ingenuity)

Molecular and Cellular functions

Enriched in hESCs	Enriched in BMSCs
RNA post-transcriptional modification	Cellular movement
Cellular growth and proliferation	Cellular growth and proliferation
DNA replication, recombination and repair	Cell death
RNA damage and repair	Cell morphology
Cell cycle	Cell-to-cell signaling and interaction

Networks

Enriched in hESCs	Enriched BMSCs
RNA post-translational modification, hematological system development and function, nervous system development and function	Organismal functions, dermatological diseases and conditions, inflammatory diseases
Protein synthesis, gene expression, RNA trafficking	Opthalmic disease, genetic disorder, inflammatory disease
Gene expression, post-translational modification, protein folding	Protein synthesis, organismal injury and abnormalities, cancer
Post-translational modification, protein folding, respiratory system development and function	Skeletal and muscular system development and function, cellular development and function, cellular development, gene expression
DNA replication, recombination, and repair, cellular assembly and organization, infection metabolism	Digestive system development and function, hepatic system development and function, lipid metabolism

Canonical Pathways

Enriched in hESCs	Enriched BMSCs
DNA methylation and transcriptional repression signaling	Regulation of Actin-based motility by Rho
Role of Oct4 in mammalian embryonic stem cell pluripotency	Actin cytoskeleton signaling
Ran signaling	Ephrin receptor signaling
Role of BRCA1 in DNA damage response	Clathrin-mediated endocytosis
Nitrogen metabolism	Hepatic fibrosis/hepatic stellate cell activation

Genes Expressed by hESCs

Genes

Transcription Factors

- Oct4
- Nanog
- Sox2
- Rex1
- Utf1

Others

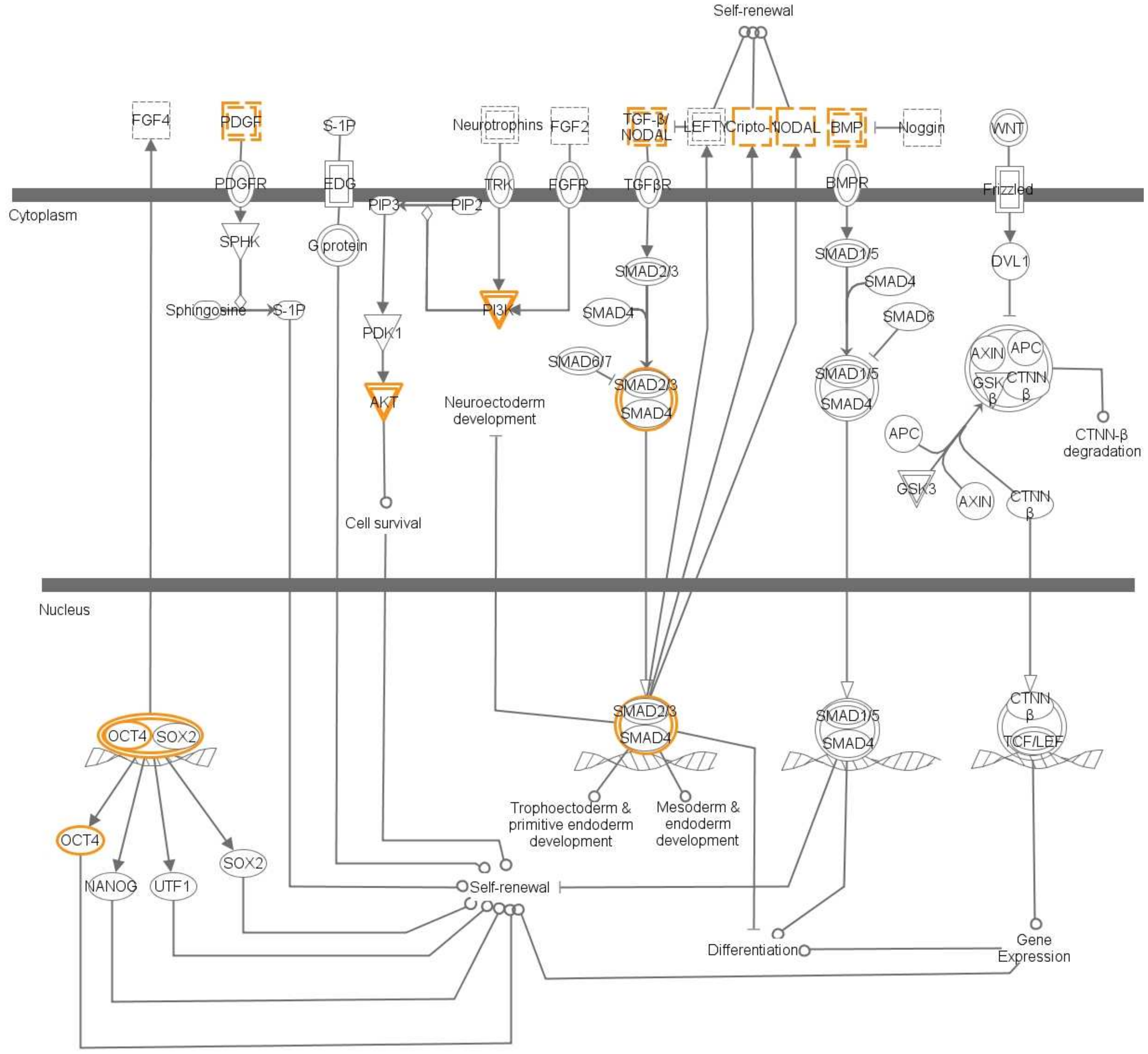
- LeftyB
- Nodal
- TDGF1 (Cripto)- epidermal growth factor-related protein
- DNMT3B - DNA methyltransferase 3B
- GDF3 - growth differentiation factor 3
- GJA1 - gap junction protein, alpha 1

Signaling Pathways

TGFβ
Wnt
BMP
FGF
PDGF
Neurotrophins

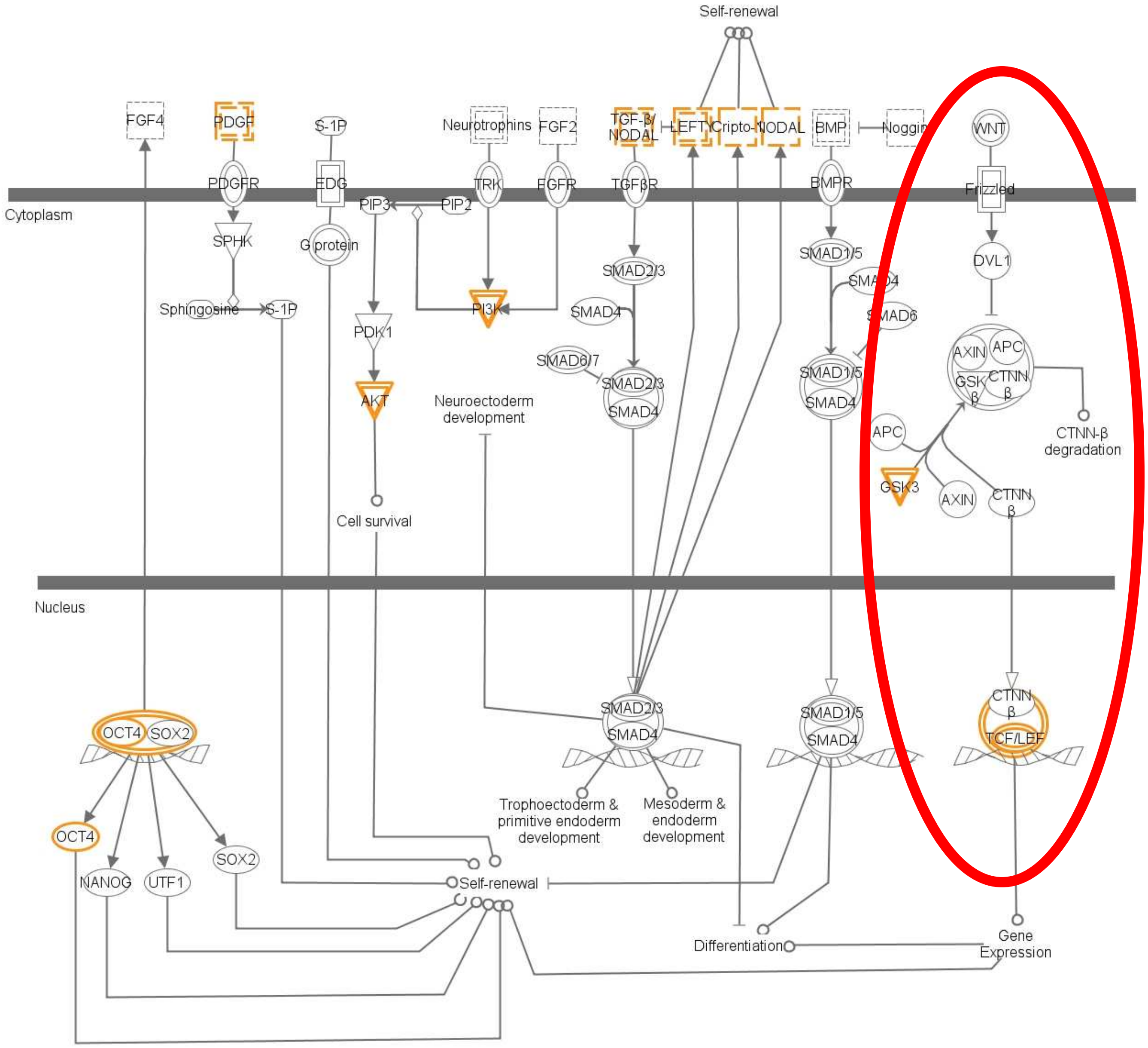
Human Embryonic Stem Cell Pluripotency

Genes Expressed by Embryonic Stem Cells Compared to Embryonic Bodies

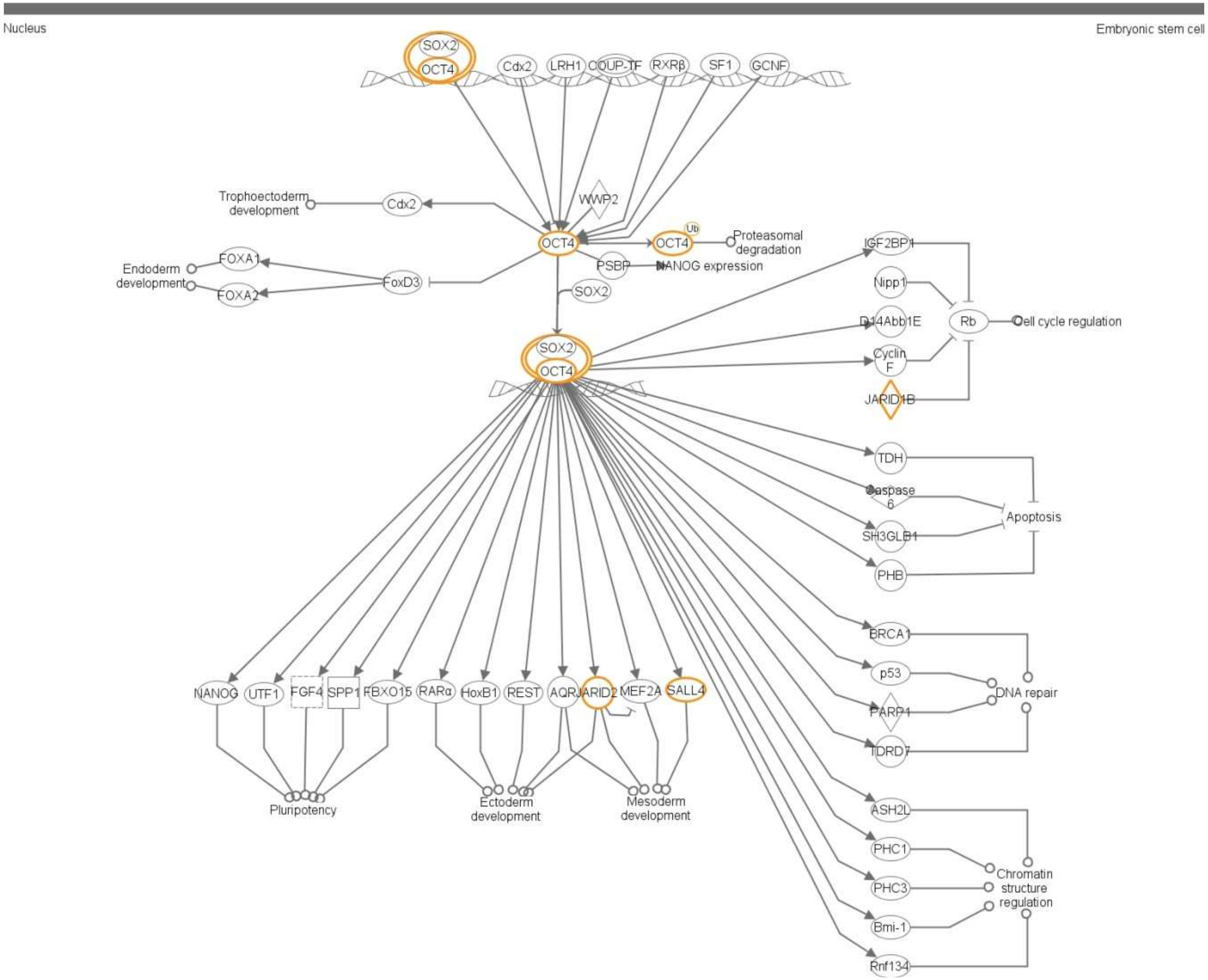


Human Embryonic Stem Cell Pluripotency

Genes Expressed by Embryonic Stem Cells Compared to Adult Cells



Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency



Summary

- Genes expressed by hESCs are involved with DNA replication, recombination and repair, RNA damage, and repair RNA post-transcriptional modification, cellular growth and proliferation, and cell cycle
- ESC express several transcription factors including Oct4, Sox2, and nanog
- Genes in several signaling pathways are expressed including wnt, TGFβ, FGF, and BMP

MicroRNA

Endogenous non-protein coding RNA

Small – 19 to 23 nucleotides

Negatively regulate the expression of protein encoding genes

- Reduce transcript levels
- Reduce translation

MicroRNA

There are >1000 in the mammalian genome

Phylogenetically conserved

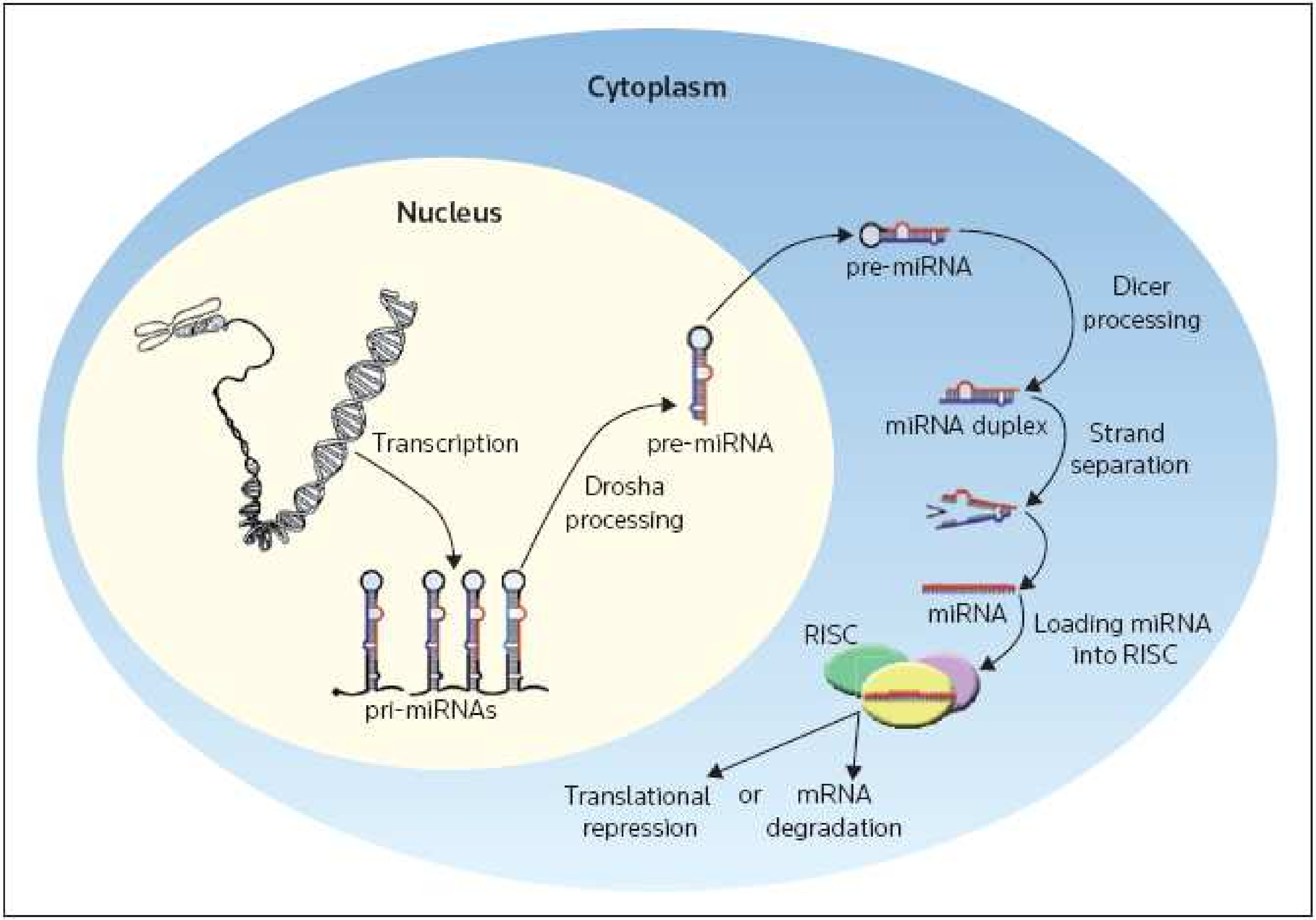
Some are encoded in clusters and transcribed polycistronicly

One miRNA may influence many genes

One gene may be influenced by more than one miRNA

Makes interpretation of results difficult

Micro RNA (miR)

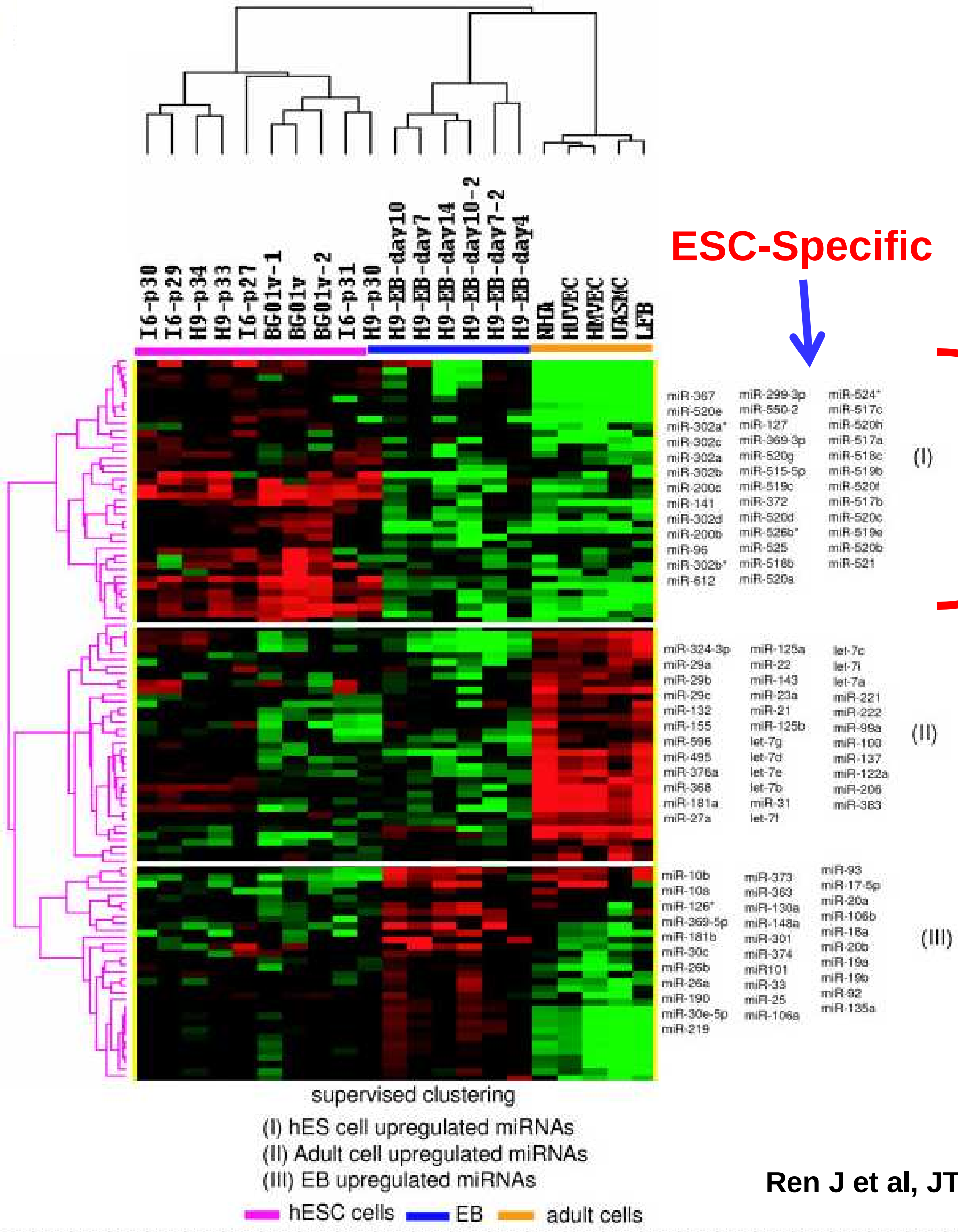


Sartipy P et al. IDrugs 2009;12:492-96

Dicer Knock out → Severe growth and differentiation defects

Embryonic Stem Cell-Specific miR:

Hierachical Clustering Analysis of 104 miR Differentially Expressed by Embryonic Stem Cells, Embryoid Bodies and Adult cells



supervised clustering

(I) hES cell upregulated miRNAs
(II) Adult cell upregulated miRNAs
(III) EB upregulated miRNAs

ESC-Specific

miR-302 Cluster - Chromosome 4

miR-302a
miR-302b
miR-302c
miR-302d
miR-367

miR-520 Cluster - Chromosome 19

miR-515-5p
miR-517a
miR-517b
miR-517c
miR-517e
miR-518b
miR-518c
miR-518d
miR-519c
miR-520a
miR-520b
miR-520c
miR-520d
miR-520e
miR-520f
miR-520g
miR-520h
miR-521
miR-524
miR-525-3p
miR-526b

miR 372 Cluster - Chromosome 19

miR-371
miR-372
miR-373

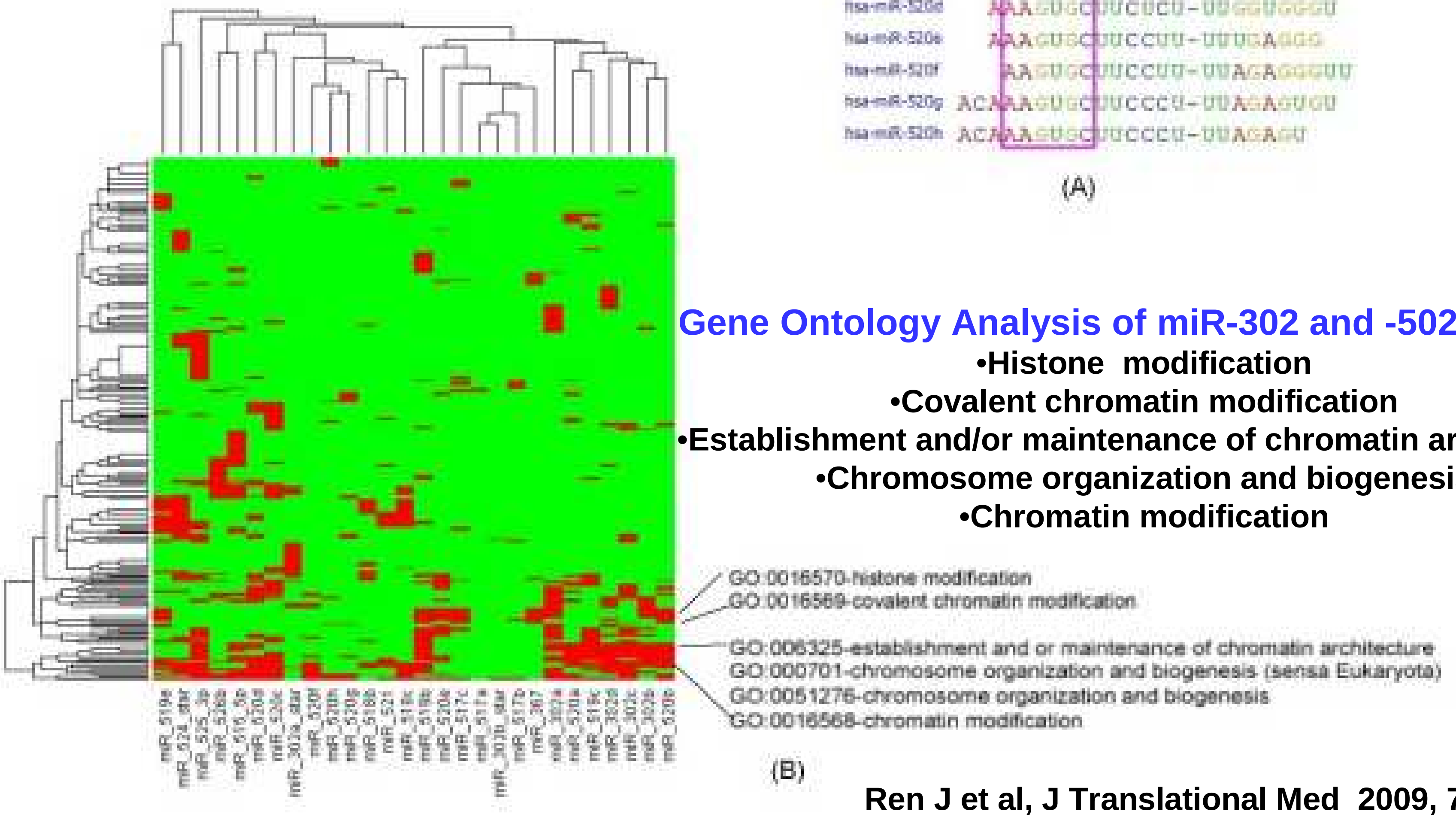
Ren J et al, JTM 2009,7:20

Several members of the miR-302 and -520 clusters share the same 7 nucleotide seed sequence

(A)

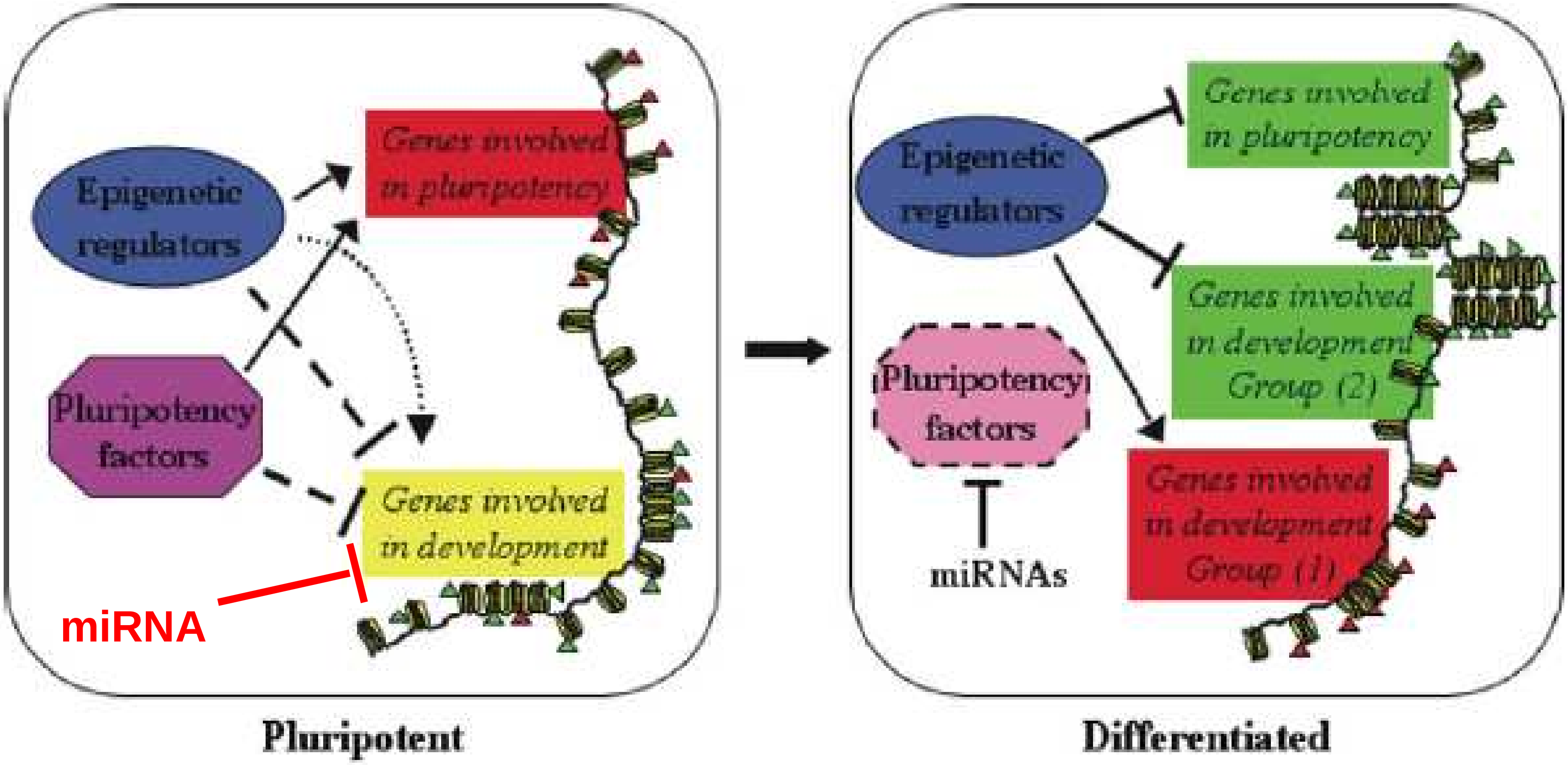
Gene Ontology Analysis of miR-302 and -502 clusters

- Histone modification
- Covalent chromatin modification
- Establishment and/or maintenance of chromatin architecture
- Chromosome organization and biogenesis
- Chromatin modification



(B)

Ren J et al, J Translational Med 2009, 7:20



L Chen and GQ Daley. Human Mol Genetics. 2008;17:R23-R27

Questions

- How do these genes and miR maintain pluripotency?
- What genes and miR are involved with differentiation?
- How can these genes and miR be manipulated for laboratory and clinical applications?

Acknowledgments

Department of Transfusion Medicine, Clinical Center, NIH

Marianna Sabatino
Jiaqiang Ren
Ping Jin
Ena Wang
Franco Marincola
Virginia David-Ocampo
Ji Feng
Sara Civini