

Microarrays and Stem Cells

Microarray Background Information

Stem cells are the building blocks that allow the body to produce new cells and repair tissues. Scientists are actively investigating the potential of both embryonic and adult stem cells as research tools. Stem cells can help researchers better understand health problems such as Alzheimer's, Parkinson's, and heart diseases, and diabetes and cancer. Increased understanding will bring the development of better treatments for these and other debilitating conditions.

Through the isolation of embryonic stem (ES) cells in culture, researchers are finding ways to identify and manipulate them. A key characteristic of embryonic stem cells is that they are able to self-renew, which means they can divide numerous times and still maintain their undifferentiated state. Even after being cultured for long periods in the undifferentiated state, they maintain the ability to differentiate into cells of all three embryonic germ layers – the ectoderm, mesoderm, and endoderm. These germ layers are the source of all the specialized cells of the body. (See Figure 1 Embryonic germ layers)



Figure 1: Embryonic germ layers

Microarrays in Stem Cell Research

Microarrays, also known as gene chips, play a key role in stem cell research. To date, microarray technology has allowed researchers to distinguish the major features that define a cell as being a stem cell. Microarrays have also been used to follow the differentiation of ES cells toward specific cell types. One product of this analysis is a catalog of the genes expressed at each step in the differentiation of embryonic endoderm cells to form mature β cells and other types of cells located in the pancreatic islets. How to manipulate and guide the differentiation of ES cells into islet β cells is the next hurdle on the way to finding a more effective treatment for type I diabetes.

What is a microarray?

A microarray is made of many different DNA sequences stuck to a flat surface. Each spot on a microarray contains many copies of the same known single-stranded DNA sequence. Different spots contain different DNA sequences. Microarrays come in different types. The more conventional version is printed on glass slides using a microarrayer robot. View the video clip “Microarrayer in Action,” www.biointeractive.org/genomics/microarray_action.html. Commercial microarrays, like the Affymetrix Gene Chips, are made using a different method. View the animation “Gene Chip Manufacturing” to learn

more about commercial Gene Chip production, www.biointeractive.org/genomics/gene_chips.html. Both of these clips can also be found on the HHMI DVD *Scanning Life's Matrix: Genes, Proteins, and Small Molecules*.



Figure 2: Microarray

Microarrays allow thousands of specific nucleotide sequences to be detected at one time on a glass or plastic slide that is about 1.5 centimeters square. (See Figure 2: Microarray)

Each microarray is made up of many small segments of single stranded DNA arranged in a grid pattern. When a sample of single stranded DNA or RNA is applied to the array, any sequences in the sample that find a match will bind to a specific spot on the array. (See Figure 3: Spots on microarray) Complementary nucleotide sequences hybridize or bond to each other. For example, a strand of DNA with the sequence -T-C-C-A-G- will hybridize to the cDNA

sequence -A-G-G-T-C- to form double-stranded DNA. A computer determines the amount of sample bound to each spot on the microarray.

Gene Expression Arrays

An expression array allows scientists to study which genes are turned on or off at any given point in time. In this type of array, mRNA is extracted from a sample. mRNA is produced when a gene is turned on or “expressed.” Complementary DNA (cDNA) is synthesized from the mRNA and tagged with a fluorescent label. This is done because cDNA is easier to work with than mRNA. cDNA is more stable.

In a typical microarray experiment, cDNA from a sample is labeled with a fluorescent dye. The fluorescent sample is then applied to a microarray that contains DNA fragments corresponding to thousands of genes. A DNA sequence that is present in the sample will bind with its complementary sequence if it is present on the array. When this occurs, the result is a fluorescent signal at that specific spot on the array. The signals are picked up using a “reader” that consists of lasers, a special microscope, and a camera. These instruments work together to create a digital image of the array. Special computer programs then calculate the fluorescence ratio for each spot. The calculated ratio reflects the relative expression of a given gene. No signal means that the gene is not expressed.

Using Microarrays to Study Embryonic Stem Cell Differentiation

Understanding the molecular mechanisms and sequential steps involved in the differentiation of embryonic stem cells on their way to become endocrine pancreas cells is an important step in the development of cell therapies for the treatment of diabetes. Potentially, gene expression analysis using microarray technology will provide researchers with the ability to direct the differentiation process.

Watch the video clip “Human ES Cells Differentiating Into Heart Cells” on the HHMI DVD *Potent Biology: Stem Cells, Cloning, and Regeneration*. The same clip is located on the *BioInteractive* site at www.biointeractive.org/stemcells/video.html with the title “Cultured Human Embryonic Stem Cells.” It provides information about how scientists are able to maintain embryonic stem cells in culture for long periods of time. When the factors that allow the stem cells to self renew are removed, they spontaneously begin to differentiate. How to regulate the differentiation of these cells into a variety of specialized cells types, such as pancreatic beta cells, is a problem that is actively being investigated.

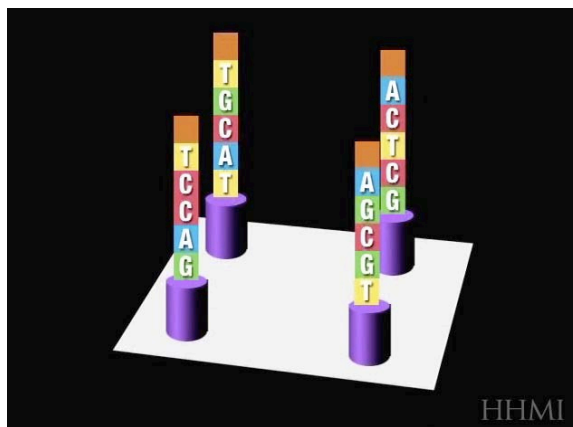


Figure 3: Spots on microarray

Definitions: The definitions below may be useful when working on the next part of this activity and when answering questions.

Cell signaling is part of the communication system that governs cellular activities and coordinates cell actions. The ability of cells to detect and respond to their environment is the basis of development, tissue repair, and homeostasis. In humans, early embryo cells exchange signals with cells of the uterus.

Genes are traditionally defined as the basic units of heredity in an organism. A more up-to-date definition of the term describes a gene as being a specific genomic sequence that corresponds to a unit of inheritance.

Genomics is the science of discovering genes, identifying where they are located in an organism, and finding out what they do.

Growth factors are substances that stimulate the growth, proliferation, and differentiation of responsive cells. They are a group of polypeptides that function as hormone-like regulatory signals.

Transcription factors are proteins that bind to specific DNA sequences. They control the transcription of genetic information from DNA to mRNA.

Questions

1. What information does the mRNA that is extracted from a sample provide researchers?

2. Explain how researchers are able to get embryonic stem cells grown in culture to spontaneously differentiate.

3. State why fluorescently-labeled nucleotides are used in the synthesis of cDNA.

4. Explain why it is important to learn more about the molecular mechanisms involved in the differentiation of embryonic stem cells to other cell types.

5. What information can microarrays provide scientists about embryonic stem cells?

What are some of the genes involved in the differentiation of ES cells to form pancreatic beta cells?

The goal of this activity is to use microarray technology to determine which genes are turned on and off at various points in the differentiation of pluripotent stem cells on their way to becoming pancreatic β cells. With this information, researchers might be able to stimulate ES cells with specific growth factors at appropriate times in the differentiation process and produce pancreatic β cells.

You will be using a simulated mini-microarray. Each of the 42 spots on the array can be used to represent a different gene. You will be working with a subset consisting of 16 selected genes. They are listed with a brief description of what they do and a short segment of coding sequence on the separate *Gene Reference Sheet*. One of the genes is classified as a *housekeeping gene*. What makes housekeeping genes unique is that they are constantly being expressed in almost every cell. They code for proteins that are needed to maintain the cell.

At the point in the investigation where you begin your work, mRNA has already been extracted from the cell cultures being studied and used as a template to synthesize single-stranded complementary DNA (cDNA). Fluorescently-labeled nucleotides were used when making the cDNA. One label color was used—red.

Recall that only the genes that were being expressed at the time the nucleic acids were extracted produced mRNA. This mRNA was used as a template to synthesize cDNA. The more a particular gene was being expressed, the greater the amount of mRNA present. This would result in a greater amount of cDNA being synthesized.

You will wash the fluorescently labeled cDNA sample (known as a probe) over the surface of the microarray. If a gene is being expressed, the cDNA corresponding to that gene will be present. It will bind to the microarray at the spot where the gene is located and fluoresce. A gene that is being highly expressed will appear brighter on the microarray since there will be more cDNA binding to that particular spot. If the gene is not being expressed, no cDNA will be present and the spot on the microarray will not fluoresce. (See Figure 4: 16 spot mini-microarray)

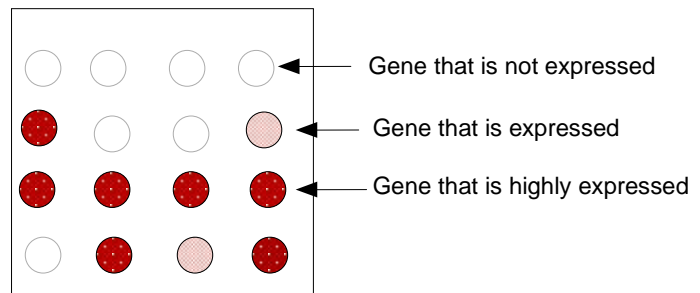


Figure 4: 16 spot mini-microarray

Materials

4 Cell Packets (1 - 4) short nucleotide sequences *Microarray Sheet for Cell Types 1 – 4*

Red and pink markers *Microarray Sheet for Cell Types 5 – 8*

Microarray Template Sheet *Gene Reference Sheet*

Microarray Summary Sheet

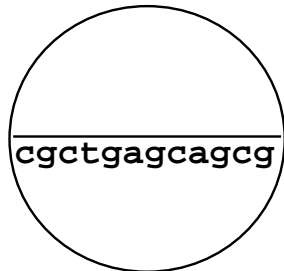
Procedure

- 1 Open the packet labeled *Cell Type 1* and remove the fluorescently labeled cDNA.
- 2 (a) Match each of the cDNA sequences to its complementary DNA sequence on the *Microarray Template Sheet*. Recall that “A” on the cDNA will bind with “T” on a microarray spot and “C” with “G.” Each cDNA that is present will bind to the spot that has a totally complementary DNA sequence. This is hybridization.
(b) Do this for all of the cDNAs in the *Cell Type 1* packet. If more than one cDNA binds to a spot, stack them one on top of the other.
- 3 For Cell Type 1 on the *Cell Types 1 – 4 Microarray Sheet*, color the spots with more than one cDNA red.
Next, color the spots with only one cDNA pink.
Leave the spots blank that have no cDNA.
- 4 Remove the cDNA sequences for Cell Type 1 from the *Microarray Template Sheet* and return them to their bag.
- 5 Repeat steps 1 - 4 for Cell Types 2 - 4.
- 6 Examine the Cells 5 – 8 *Microarray Sheet*. The hybridization results have been shaded in for you.
- 7 Using the information provided by the arrays, complete the listing of genes that are expressing for cell types 2 - 8 on the *Microarray Summary Sheet*.

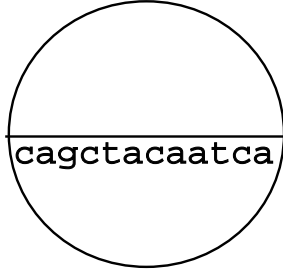
Microarray Template Sheet

Note: Do NOT color this sheet. Color the Cell Types 1 - 4 Sheet

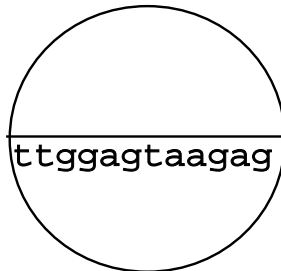
Brn4



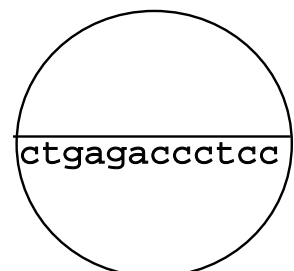
Ins



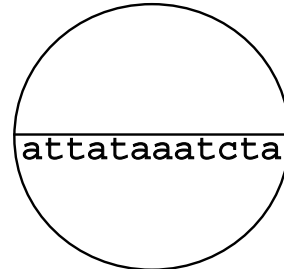
Isl1



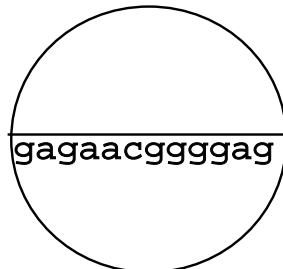
LeftyA



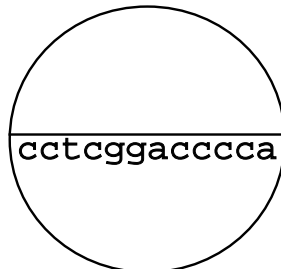
Nanog



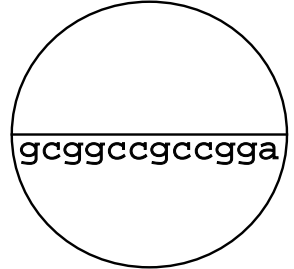
NeuroD1/Beta2



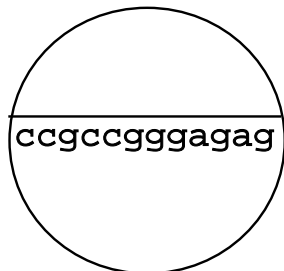
Ngn3



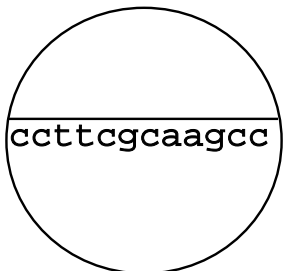
Nkx2.2



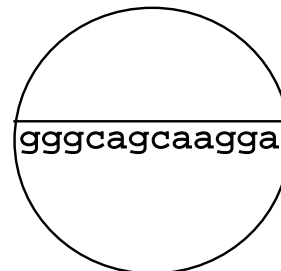
Nkx6-1



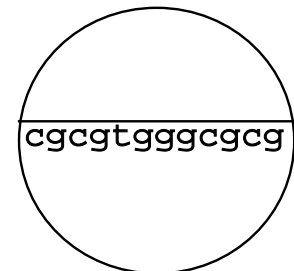
Oct-4



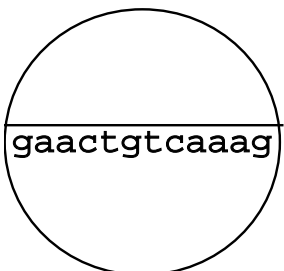
Pax4



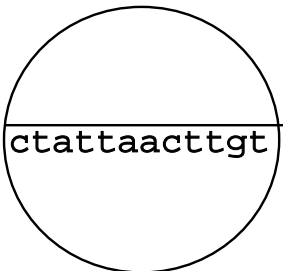
Pax6



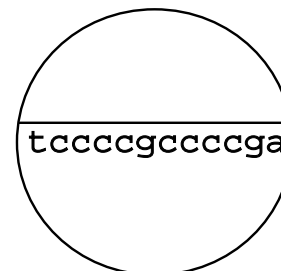
Pdx-1



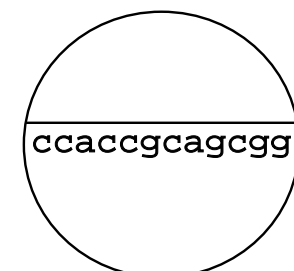
SOX2



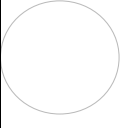
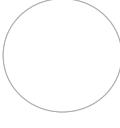
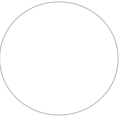
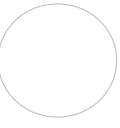
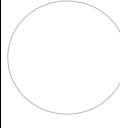
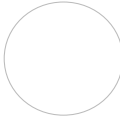
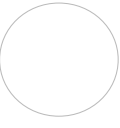
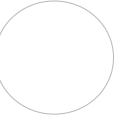
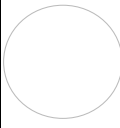
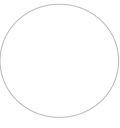
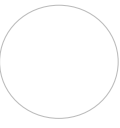
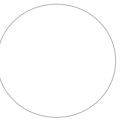

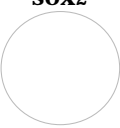
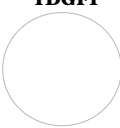
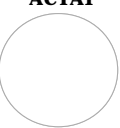
TDGF1

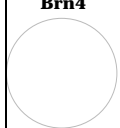
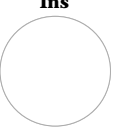


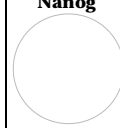
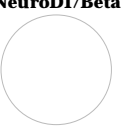


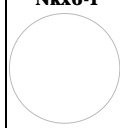
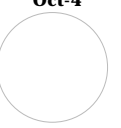
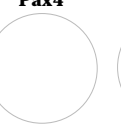
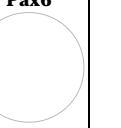
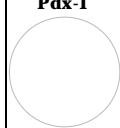
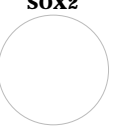
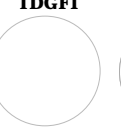
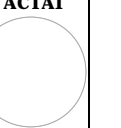


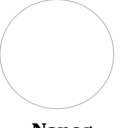
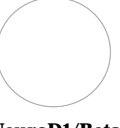
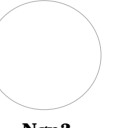

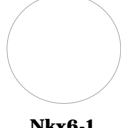


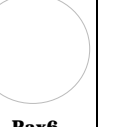
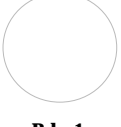
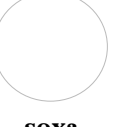
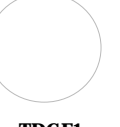

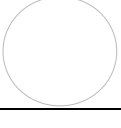
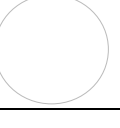

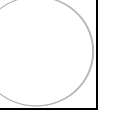
ACTA1

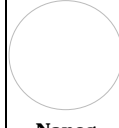



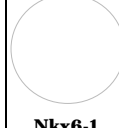

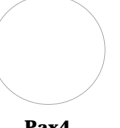

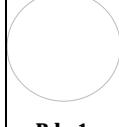
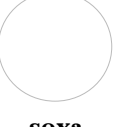
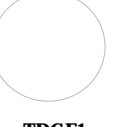
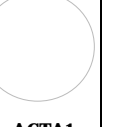
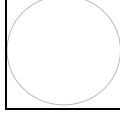
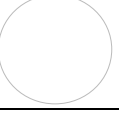
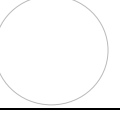
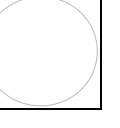


Cell Types 1 - 4 Microarray Sheet

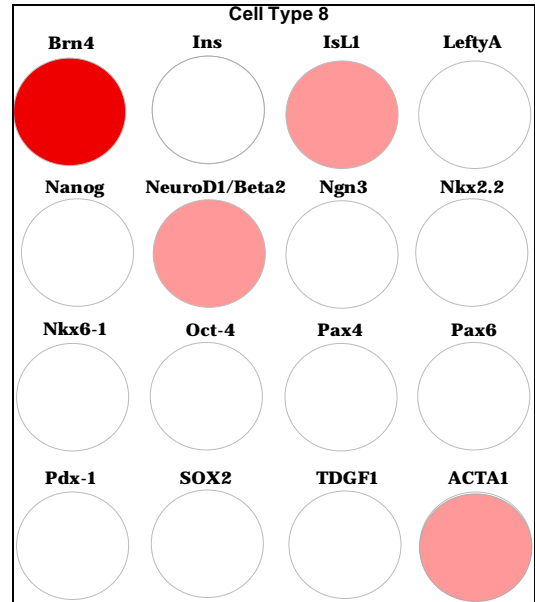
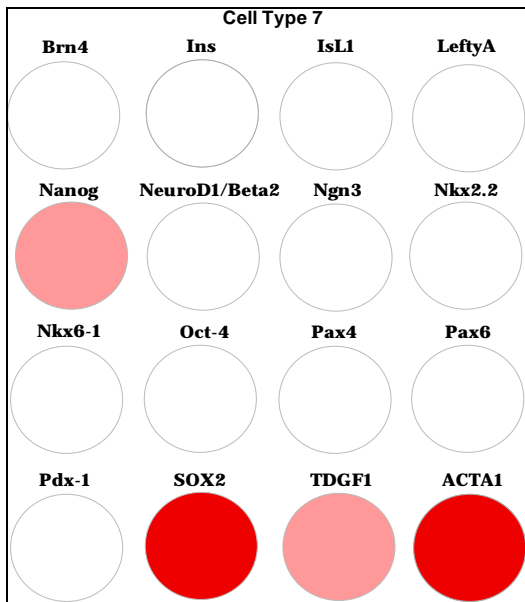
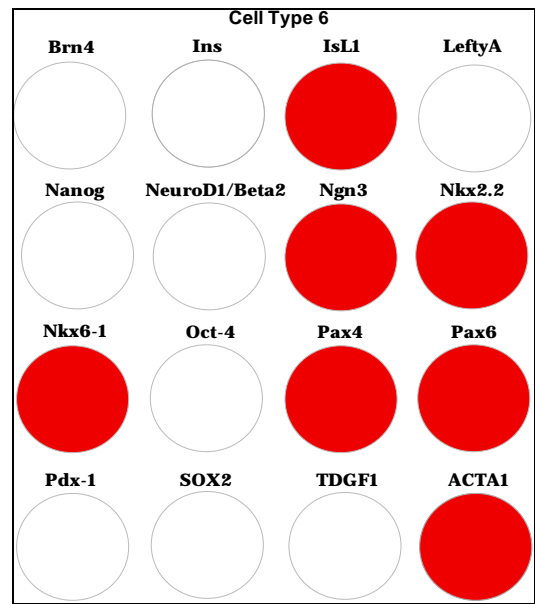
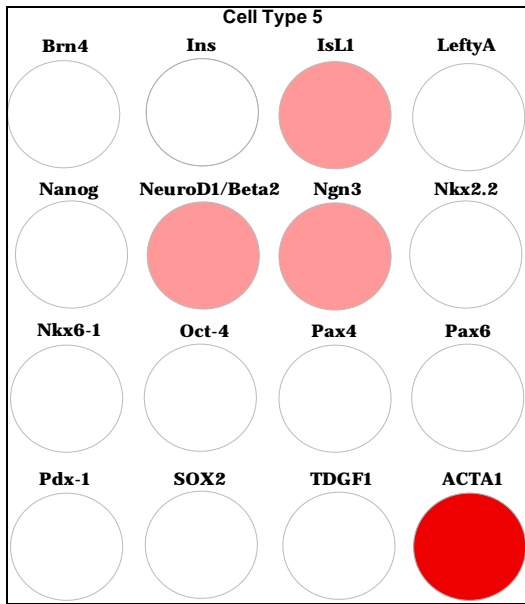
Cell Type 1			
Brn4	Ins	Isl1	LeftyA
			
Nanog	NeuroD1/Beta2	Ngn3	Nkx2.2
			
Nkx6-1	Oct-4	Pax4	Pax6
			
Pdx-1	SOX2	TDGF1	ACTA1
			

Cell Type 2			
Brn4	Ins	Isl1	LeftyA
			
Nanog	NeuroD1/Beta2	Ngn3	Nkx2.2
			
Nkx6-1	Oct-4	Pax4	Pax6
			
Pdx-1	SOX2	TDGF1	ACTA1
			

Cell Type 3			
Brn4	Ins	Isl1	LeftyA
			
Nanog	NeuroD1/Beta2	Ngn3	Nkx2.2
			
Nkx6-1	Oct-4	Pax4	Pax6
			
Pdx-1	SOX2	TDGF1	ACTA1
			

Cell Type 4			
Brn4	Ins	Isl1	LeftyA
			
Nanog	NeuroD1/Beta2	Ngn3	Nkx2.2
			
Nkx6-1	Oct-4	Pax4	Pax6
			
Pdx-1	SOX2	TDGF1	ACTA1
			

Cell Types 5 - 8 Microarray Sheet



Microarray Summary Sheet

The microarray results for each cell type in this study provide you with information about to what degree specific genes are being expressed and which are not. Using this information, you will also be able to determine the source of 6 of the 8 cell types used in this activity.

For each cell type, record the names of the genes that are being expressed.

Cell Type	Summary	Cell Type	Summary
1	(a) LeftyA (b) Nanog (c) Oct-4 (d) SOX2 (e) TDGF1 (f) ACTA1	2	(a) _____ (b) _____ (c) _____
3	(a) _____ (b) _____	4	(a) _____ (b) _____ (c) _____ (d) _____ (e) _____ (f) _____ (g) _____ (h) _____
5	(a) _____ (b) _____ (c) _____ (d) _____	6	(a) _____ (b) _____ (c) _____ (d) _____ (e) _____ (f) _____ (g) _____
7	(a) _____ (b) _____ (c) _____ (d) _____	8	(a) _____ (b) _____ (c) _____ (d) _____

Gene Reference Sheet

Each microarray analysis done in this study provides you with information about which genes are being expressed, to what degree they are expressing, and which genes are not being expressed. Using this information, you will be able to determine the source of six of the eight cell types represented on the *Stem Cell Differentiation Flowchart*. For example, the results of one analysis will indicate that the source of the cells being analyzed on the microarray is the pancreas, specifically pancreatic beta cells. Another analysis will indicate that the cells are embryonic stem cells

When using the *Gene Reference Sheet to identify the source of the cell types*, note that some genes are expressed during the differentiation of stem cells into cells of more than one germ layer. Timing, neighboring cells, growth factors and which genes are already being expressed all influence other genes and ultimately the fate of a cell.

Some Genes and What They Do

- **Brn4** (brain 4 gene) This gene codes for a transcription factor. It is highly expressed in glucagon-producing islet cell lines
Partial sequence: cgc tga gca gcg
- **Ins** (Insulin gene) In adults, insulin is expressed in the β -cells of the pancreatic islets. The name "insulin" comes from the Latin word *insula* that means, "island." Complex mechanisms control insulin expression at the correct time and place during embryonic development.
Partial sequence: cag cta caa tca
- **Isl1** (Islet-1 gene) The transcription factor it encodes plays an important role in the formation of islets of Langerhans in the pancreas.
Partial sequence: ttg gag taa gag
- **LeftyA** (Left-right determination factor) This gene establishes a signaling network that allows for the temporal and spatial control of other genes involved in the differentiation of stem cells into cells of endoderm and mesoderm. As the cells differentiate and become more specialized, LeftyA is expressed less.
Partial sequence: ctg aga ccc tcc
- **Nanog** (Nanog Homeobox gene) This gene is expressed in embryonic stem cells and codes for a transcription factor that is important in maintaining pluripotency.
Partial sequence: att ata aat cta
- **NeuroD1/Beta2** (Neurogenic differentiation gene) This gene helps regulate the formation of pancreatic islets. It also codes for a transcription factor that regulates the expression of the insulin gene in islet beta cells.
Partial sequence: ctc ttg ccc ctc
- **Ngn3** (Neurogenin3) During embryonic development, this gene specifies which cells will differentiate into the cells that make up the endocrine pancreas. It is thought to be absent after birth. If expression of the Ngn3 gene is not allowed to take place in cells grown in culture, new insulin producing cells are not generated. This could be important in understanding beta cell regeneration.
Partial sequence: cct cgg acc cca
- **Nkx2.2** This is a homeobox gene that encodes a transcription factor needed to initiate beta cell differentiation. It also affects insulin gene expression.
Partial sequence: gcg gcc gcc gga

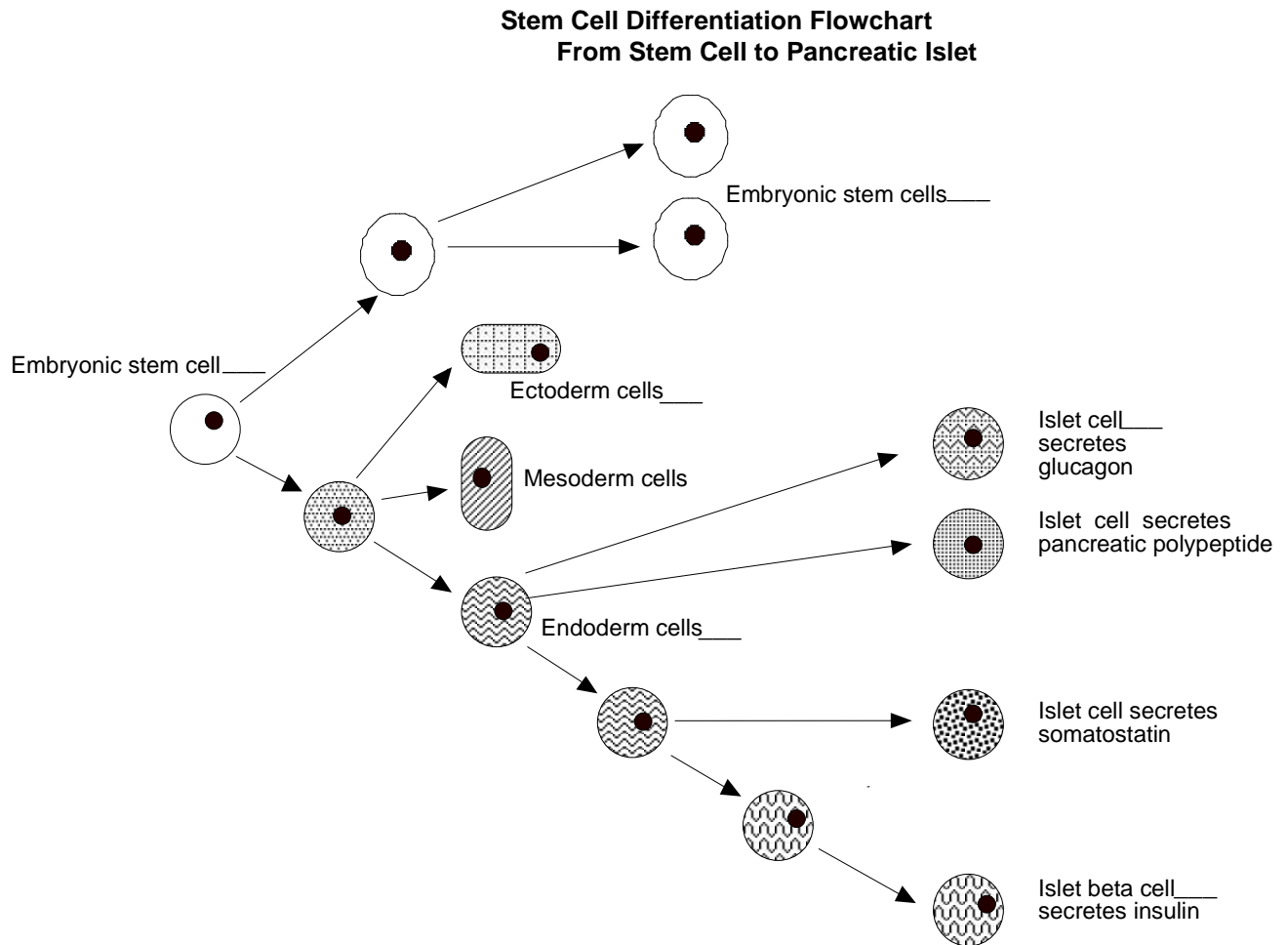
- **Nkx6-1** (NK6 Homeobox 1) This gene codes for the most beta cell specific transcription factor known in the pancreas. It is necessary for the formation of insulin-producing beta cells.
Partial sequence: ccg ccg gga gag
- **Oct-4** (Octamer-4) Oct-4 is a commonly used synonym for **POU5F1** (POU class 5 homeobox 1). The Oct-4 transcription factor is active in embryos throughout the pre-implantation period of development. The expression of the Oct-4 gene is associated with maintaining an undifferentiated phenotype and with tumors.
Partial sequence: cct tcg caa gcc
- **Pax4** (Paired box gene 4) This gene is important during fetal development. It is involved in pancreatic islet development; specifically in the differentiation of insulin-producing beta cells
Partial sequence: ggg cag caa gga
- **Pax6** (Paired box gene 6) This is the most important of the Pax genes. It codes for a transcription factor that is important in the development of both the central nervous system and the pancreas. It is required for the differentiation of pancreatic islet alpha cells.
Partial sequence: cgc gtg ggc gcg
- **Pdx1** (Pancreatic and duodenal homeobox 1. It is also known as insulin promoter factor 1) It codes for a transcription factor necessary for pancreatic development and beta cell formation.
Partial sequence: gaa ctg tca aag
- **SOX2** (also known as sex determining region Y (SRY)-box 2) It is a transcription factor that is important in the self-renewal of undifferentiated embryonic stem cells.
Partial sequence: cta tta act tgt
- **TDGF1** (teratocarcinoma-derived growth factor 1) It is one of several genes important in maintaining pluripotency.
Partial sequence: tcc ccg ccc cga

Housekeeping gene

- **ACTA1** (Actin) It is found in almost all eukaryotic cells and is involved in cell motility, cell division and cytokinesis, organelle movement, signaling, and maintaining cell shape.
Partial sequence: cca ccg cag cgg

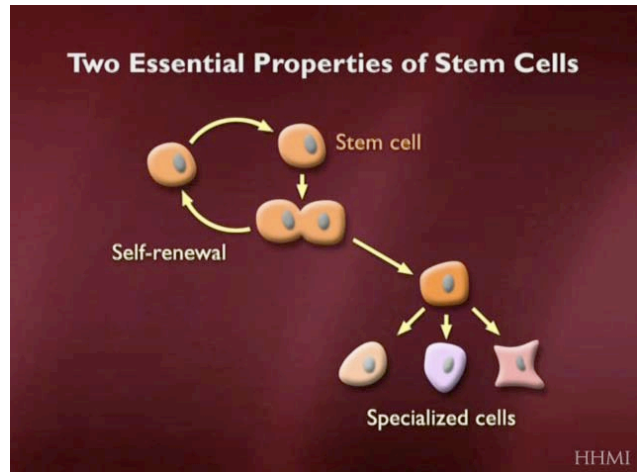
Analysis

- Record the number of each cell type by the appropriate cell on the *Stem Cell Differentiation Flowchart*. *Note:* This will be possible for only 6 of the 8 cell types. Two cell types do not readily correlate to any particular cell type shown of the flowchart. Use the *Gene Reference Sheet* and the *Microarray Summary Sheet*.



- Two of the cell types you analyzed were impossible to accurately place on the *Stem Cell Differentiation Flowchart*. Explain what additional steps researchers would need to take in order to know where these cells should be positioned on the flowchart.

3. Pluripotent stem cells sometimes divide to form more pluripotent stem cells. Other times they divide to form one or two cells that differentiate to become specialized tissue. Identify at least two factors that influence whether or not a stem cell differentiates after dividing.

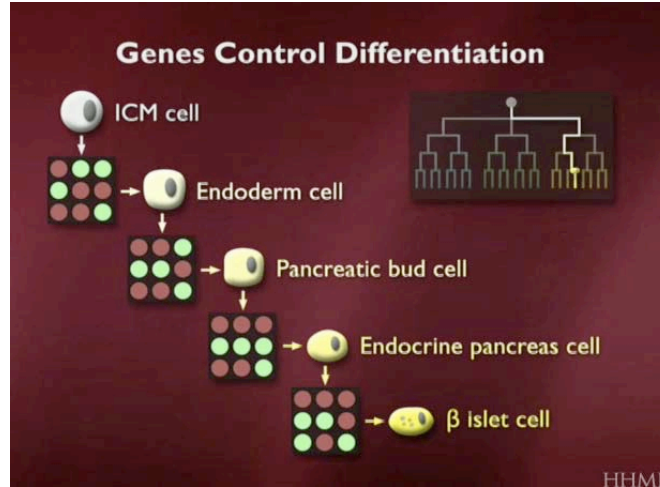


4. If scientists were to manipulate cultures of pluripotent stem cells to differentiate to form insulin-producing β cells, explain what factors would need to be regulated and why.

5. Stem cells initially differentiate into cells that form the early ectoderm, mesoderm, and endoderm. Identify two or more tissues that develop from each of these germ layers.

6. Individuals receiving either islet or pancreas transplants must take immunosuppressant drugs. Would this be necessary if the source of the transplanted β cells was pluripotent stem cells manipulated in culture? Support your answer.

7. The microarrays associated with each step provide information that can be used to eventually direct the differentiation process in culture.



Describe what type of information the microarrays provide and how it can be used to direct embryonic stem cells to differentiate to become β cells.

8. Researchers are actively investigating how to deprogram and reprogram already differentiated cells so that they would become pluripotent stem cells. Such cells are referred to as induced pluripotent stem cells (iPS cells). If the technique is perfected, iPS cells could be created from a patient's skin cells and then used to replace diseased or damaged tissue. Explain why this would be better than using embryonic stem cells obtained from a different source.
