Gene Control In Development

• Cellular differentiation—Concepts
• Themes of differentiation
• Example: Myogenesis in mammals
• Use of transgenic animals to identify requirement for gene in mechanism.
• Use of two-hybrid methods to identify protein-protein interactions.

The Problem

• A single fertilized cell gives rise to a complex, multicellular mass with groups of specialized cells.
• How does this single cell give rise to different cell types?
• We will use a simplified view of muscle development to demonstrate this issue.

The Assumptions

• The difference between different developed tissues is that different cell types express different sets of genes
• Many of the gene regulatory events used for development permanently switch off sets of genes (by mechanisms such as methylation or chromosome condensation).

More Assumptions

• The majority of gene expression mechanisms are transcriptional, and are therefore controlled by transcription factors
• Most gene regulatory events are assumed to be controlled by receptor signaling (some developmental receptors & ligands are both cell-surface proteins)

Hypothesis:

• Cell-cell signaling causes induction of “master” transcription factors, which in turn induce networks of cell type-specific genes, causing differentiation.

Basic Muscle Development

• Mammalian skeletal muscle develops from masses of cells called somites.
• Some somite cells get committed to muscle development by becoming myoblasts.
• Since myoblasts are committed to becoming muscle tissue, there are called "determined."
• Myoblasts can divide to produce more myoblasts, or they may differentiate further towards muscle, but they now can never become, e.g., liver cells.
Migrating Myoblasts

- Myoblasts migrate to regions where developing muscle is formed such as the limbs. (This allows cell-cell contact along path of migration.)
- They then stop cell division, and fuse into a syncytium which is a larger cell composed of many nuclei but a single cytoplasmic space.
- The syncytium further differentiates into a myotube which is the fundamental component of muscle.

Clues from the Laboratory

- A cell line, C3H10T1/2, is fibroblastic in culture.
- When 5-azacytidine is added to dividing cells, a small number of cells can be converted into muscle cells.
- 5-azacytidine is an analog of cytidine that cannot be methylated.

Aza group blocks methylation

- 5-azacytidine can be substituted for cytidine during cell division.
- After incorporation into DNA, this position can no longer be methylated.
- Therefore, we assume that changes in cell properties are due to loss of methylation.
- \[\text{Methylation inactivates genes during development.}\]
Hypothesis:

- If demethylation can lead to re-differentiation of these fibroblasts, could we instead add demethylated versions of genes to these cells to induce muscle development?
- More specifically, a small number of genes must be activated to induce muscle development.

Screen for demethylated gene

- Using cDNA libraries prepared from untreated and 5-azacytidine-treated C3H 10T1/2 cells, scientists compared the mRNA populations expressed in the two cell types. The method used was “subtractive hybridization.”
- The result was four cDNA clones that were unique to the 5-azacytidine-treated cells.

Screen yielded four cDNAs

Development by “replacement”

- If any of these four cDNAs satisfied the hypothesis, then expression of the cDNA into mRNA in a normal C3H 10T1/2 cell should drive it toward muscle development.
- One cDNA, named myoD, satisfied this requirement.
- Later, three other genes were identified in similar experiments—myogenin, myf5 and mrf4.
- Labeled muscle-regulatory factors (MRF).

How do MRF genes promote muscle development?
They’re transcription factors!

- The myogenic proteins are helix-loop-helix (HLH) transcription factors. (HLH is a modified form of the leucine zipper family).
- The HLH family of transcription factors generally forms homo- and heterodimers, and they interact with a common DNA sequence element called the "E-box" with the sequence CANNTG.

MyoD & E-boxes

- This sequence element is found in multiple copies in muscle-specific structural genes.
- MyoD binds only a single E-box...
- But transcriptional activation requires multiple myoD-E-box complexes interacting together in a cooperative fashion.

E2A

- Another protein plays a role in this mechanism. The E2A protein (also an HLH transcription factor) forms heterodimers with myoD.
- E2A not muscle-specific; required for B cell development.
- This heterodimer has a 10-fold higher affinity for the E-box than a homodimer of myoD alone.
- MRF+E2A also binds muscle enhancer binding factors (MEF), which themselves are transcription factors.
- Proposed that MyoD induces transcription, not E2A

MEFs function in concert with MRFs to confer myogenic specificity

How do we demonstrate the hypothesized roles of myogenic proteins?

Creation of “Knockout” Mice

- Specific genes are inactivated by homologous recombination.
- This is done in embryonic stem cells in culture.
- Modified ES cells are combined with a fertilized egg and implanted into a pseudopregnant female.
Gene replacement

• “Flanking regions” from gene direct site of homologous recombination
• neo<sup>R</sup> selects for inserted gene
• tk<supHSV</sup> provides selection against non-homologous recombination
• Successful replacement should be both neo<sup>R</sup> and gancyclovir resistant.

Construction of Mouse

• “Knockout” ES cells are injected into the blastocoeal of an early embryo, where the cells fuse with the existing embryo and undergo normal development
• Coat colors are used to trace cells in progeny mice.
  – ES cells come from a black mouse strain.
  – Use white mouse as “host embryo.”
Finding our Knockout

- Most F1 progeny are “mosaic” (contain cells from two different genetic sources).
- Mate mosaic F1s with white mice to create heterozygous knockouts (completely black).
- Inter-mate F2s to obtain homozygous knockout
- Confirm by DNA analysis -- all cells are missing gene that is “knocked out.”

<table>
<thead>
<tr>
<th>Gene Knocked Out</th>
<th>Viable</th>
<th>Myoblasts</th>
<th>Muscle</th>
<th>Role of Myogenic Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>myoD</td>
<td>Yes</td>
<td>+</td>
<td>+</td>
<td>Required for myoblast formation or survival</td>
</tr>
<tr>
<td>myf5</td>
<td>Yes</td>
<td>+</td>
<td>+</td>
<td>Required for myoblast formation or survival</td>
</tr>
<tr>
<td>myoD; myf5</td>
<td>No</td>
<td>+</td>
<td>+</td>
<td>Required for myoblast formation or survival</td>
</tr>
</tbody>
</table>

Either myoD or myf5 is dispensible, but knockout out of both is lethal. Therefore myoD and myf5 are parallel signals of development.

Myogenin knockouts are lethal--there is no other gene that can substitute for its function. But they can produce intermediate myoblasts.

Model based on knockout results

MyoD and the Cell Cycle

- Strong expression of exogenous MyoD induces growth arrest.
- Muscle differentiation requires cell cycle arrest.
- Requires Rb protein—phosphorylation state of Rb directly correlates with cycling state of cell.
- Studies showed a direct binding of MyoD and under-phosphorylated Rb in vitro.
- Is MyoD acting via a MyoD:Rb complex to arrest the cell cycle?

More effects of MyoD

- Overexpression of cyclin D1 overcomes MyoD arrest of cell cycle.
- Cyclin D1 also promotes nuclear translocation of cdk4, binds MyoD and prevents transcriptional activity.
- Recent work (Zhang et al.) shows that MyoD also inhibits cdk4 kinase activity.
- This inhibits Rb phosphorylation by cdk4 in vitro and in cells.
- Does MyoD bind cdk4 and inhibit it to arrest the cell cycle?
MyoD interacts directly either Rb or cdk4 (or both) in cells to arrest the cell cycle

- But does it? How do we test the binding of one protein to a target in living cells?
- A “two-hybrid system” is commonly used to assess protein-protein interactions in vivo.

**The Yeast Two-Hybrid System**

**How’s it work?**

- A transcription factor is split into its DNA Binding Domain and its Transcriptional Effector Domain.
- One half (“Bait”) is linked to a known coding sequence by cDNA cloning methods.
- A cDNA library is randomly linked to the other half (“Prey”).
- Plating of transfected yeast on selective medium allows the identification of yeast harboring a prey fusion protein that can bind bait by expression of target reporter gene.

**Mammalian Two-Hybrid Luciferase Assay**

- Same idea as yeast, performed in mammalian cells with two selected test proteins.
- Rb was cloned into a gal4 DNA binding domain vector.
- MyoD was cloned into a vp16 activation domain plasmid.
- A third plasmid includes gal4 binding sequences in a promoter for firefly luciferase protein (reporter).
- Quantity of luciferase activity produced in cells is a measure of bait-prey interaction.

**Close-up of protein-protein interaction**

Two proteins must interact to drive expression from the reporter plasmid.
Does MyoD bind Rb in cells?

Conclusions

- Transcription factors are important developmental mediators.
- Cells are sometimes committed to differentiation by inactivation of sets of genes, such as by methylation.
- Replacement of methylated genes by a few, key, unmethylated genes can drive differentiation.
- “Master” differentiation factors control expression of additional transcription factors.
- Parallel pathways allow for errors.
- Differentiation mechanisms may overlap with cell cycle control.

Remaining Questions

- What turns on myoD and the other myogenesis proteins?
- We believe that receptors are important for these events and these receptors may be activated by either diffusible factors or by cell-cell contact. What receptors?
- What genes are regulated by myogenesis proteins? Are they also transcription factors?