Differential Gene Expression

June 18, 2009
Differential Gene Expression

Overview
Chromatin structure
Gene anatomy
RNA processing and protein production
Initiating transcription: promoters and basal transcription factors
Regulation of gene transcription
  enhancers
  specific transcription factors
Epigenetic processes (modifications of DNA structure)
  methylation
  imprinting
  dosage compensation
  non-coding RNAs
Alternative RNA splicing
Differential Gene Expression

Concepts

Development requires precise control of gene expression in both location and time.

Gene transcription is regulated at multiple levels - regulatory elements found within and outside the DNA sequence.

Variation in internal or external regulatory elements increases the range and complexity of possible developmental outcomes.

Transcriptional regulation often involves modification of chromatin; specifically, changes in histone acetylation and methylation.

Gene messages are extensively processed and modified prior to translation.
Chromatin Structure

Histone lysines are protonated = net positive charge
DNA PO₄ deprotonated = net negative charge
Electrostatic attraction ensures tight binding
- DNA inaccessible to polymerases, etc.

~140 bp

~60 bp
**Nucleosome**

DNA bound to an octameric complex composed of two each of histones H2A, H2B, H3, and H4.

Histone H1 – associates nucleosomes into higher-order folded complex

Chromatin Acetylation

Transcription requires access to DNA for transcriptional complex (polymerase, etc.)

1. Tight packing of chromatin prevents access = no transcription

2. Histone acetylation \( \left[ \text{\text{CH}_3} \right] \) controls DNA binding
   - deacetylated histones (**histone deacetylases**) stabilizes nucleosome = transcription repressed
   - acetylated histone (**histone acetyltransferase**) destabilizes nucleosome = transcription allowed

3. Histone methylation (\( \text{CH}_3 \)) \( \rightarrow \) further repression
Chromatin Configuration

Euchromatin
- “active” chromatin

Heterochromatin
- remains tightly condensed throughout most of the cell cycle
- replicates later
Anatomy of the Gene

Transcription produces heterogeneous nuclear RNA (hnRNA)
= mature mRNA, ready to be transported to the cytoplasm and translation
Production of β-globin

m^7 GpppAC ("Cap")

Processing

Exon 1 - Intron 1 - Exon 2 - Intron 2 - Exon 3

m^7 GpppAC ("Cap")

Leader

Exon 1 AUG Exon 2 Exon 3 UAA

Translation

H_2N COOH β-GLOBIN PROTEIN

Posttranslational modification

β-globin α-globin Heme

HEMOGLOBIN
Differential Gene Expression

Overview

Chromatin structure
Gene anatomy
RNA processing and protein production
Initiating transcription: promoters and basal transcription factors
Regulation of gene transcription
  enhancers
  specific transcription factors
Epigenetic processes (modifications of DNA structure)
  methylation
  imprinting
  dosage compensation
  non-coding RNAs
Alternative RNA splicing
Promoters and Transcription Factors

RNA polymerase II binds to the *promoter* region at the TATA box
- however, Pol II cannot initiate transcription alone

Various proteins bind to regulatory sequences upstream and downstream of transcription initiation site……

These proteins make up the **Basal Transcription Factors**
- facilitate Pol II binding and activity
Basal Transcription Factors

aka General Transcription Factors:

- small nuclear proteins
  [TFIIB, TFIIC, etc.]

- constitutive, ubiquitous

- bind to promoter *sequentially*

- binding of individual TFs is mediated by other small proteins:
  - e.g. TBP-associated factors (TAFs)
  - these exist in a mediator complex (~25 proteins)

With RNA Pol II the TFs, etc. form the

Transcription Initiation Complex (a.k.a. Preinitiation Complex)
Transcription Initiation Complex

1. TFIID complex binds to TATA box through TATA Binding Protein (TBP) subunit

2. TFIID is stabilized by TFIIA

3. TFIIB and TFIIH join the complex on the TATA box; TFIIIE and TFIIIF associate with RNA polymerase II
4. RNA polymerase II is positioned by TFIIB, and its carboxy-terminal domain (CTD) is bound by TFIID.

5. The CTD is phosphorylated by TFIIH and is released by TFIID; RNA polymerase II can now transcribe mRNA.
Transcriptional Initiation Complex Stabilized by TAFs

TAF(s) – TBP-associated factor(s)
TBP – TATA binding protein

(A) A minimal complex of TBP and a TAF fails to activate transcription

(B) Addition of the p110 TAF and the p150 TAF allows stabilization of the TBP by both NTF and Sp1

(250)

Sp1
NTF-1

(110)

Sp1
NTF-1

250

150

Transcription initiation

+1

(SP1, NTF-1 = Transcription factors)
Histone acetylation = opening/access to DNA
Differential Gene Expression

Overview

Chromatin structure
Gene anatomy
RNA processing and protein production
Initiating transcription: promoters and basal transcription factors
Regulation of gene transcription
  - enhancers
  - specific transcription factors
Epigenetic processes (modifications of DNA structure)
  - methylation
  - imprinting
  - dosage compensation
Non-coding RNAs
Alternative RNA splicing
Enhancers

*Cis*-acting DNA sequences that regulate gene expression by interacting with the transcription initiation complex on the promoter.

- **cis** – (same or same side); elements that reside on the same DNA strand; e.g. DNA sequences
- **trans** – (other side); elements that originate from another DNA strand, e.g. regulatory proteins
Enhancers

*Cis*-acting DNA sequences that regulate gene expression by interacting with the transcription initiation complex on the promoter.

Function - bind specific regulatory proteins
  (i.e. *specific transcription factors*)
  - activate or repress transcription via promoter region

Location in DNA - highly variable:
  - upstream (5’), downstream (3’), or within transcribed region
  - in close proximity to gene or as many as $10^6$ bp away

Enhancers and promoters are both DNA regulatory sequences, but enhancers:
  1) need a promoter to work
  2) can work at a distance
  3) can work in reverse orientation
Enhancer Generalizations

1. Most gene transcription requires enhancers.

2. Enhancers are the major determinants of differential transcription in cell types and through developmental stages.

3. There can be multiple signals (e.g. multiple enhancer sites) for a given gene, and each enhancer can be bound by more than one transcription factor (though, not at the same time).

4. Transcription is regulated by the interaction of transcription factors bound to enhancers and the transcription initiation complex assembled at the promoter.

5. Enhancers are modular. A gene can have several enhancer elements, each of which may turn it on in different sets of cells.
Enhancer Modules

Enhancers are *modular*: e.g. Pax6 gene expressed in the eye, pancreas, and nervous system.
- also expressed in different cells within these tissues
Enhancer Generalizations

6. Enhancers are *combinatorial*. Various DNA sequences regulate temporal and spatial gene expression; these can be mixed and matched.
Enhancer Modules

Within modules, transcription factors work in a *combinatorial fashion*. Transcription factors often operate in *cascades*:
- activation of one enhancer stimulates the production of several other TFs
Enhancer Generalizations

6. Enhancers are *combinatorial*. Various DNA sequences regulate temporal and spatial gene expression; these can be mixed and matched.

7. Enhancers (transcription factors) often operate in *cascades*.

8. Enhancers generally activate transcription by remodeling chromatin to expose the promoter, or by facilitating the binding of RNA polymerase to the promoter by stabilizing TAFs.

9. Enhancers can also inhibit transcription (aka *Silencers*).
Method – Reporter Genes

**Reporter genes** - fuse enhancer (regulatory) elements for a target gene to a gene that will produce a detectable protein.

- e.g. attach Pax6 enhancers to a gene for β-galactosidase (blue)
- insert the reporter gene into a cell or embryo
- wherever (or whenever) Pax6 is transcribed, detect by blue color

Pax6

- mouse – Pax6
- β-galactosidase

Reporter protein gene, e.g. β-galactosidase (blue) green fluorescent protein luciferase (produces light)
Transcription Factors

*Trans*-acting regulatory elements (proteins)

- *cis* – (same or same side); elements that reside on the same DNA strand; e.g. DNA sequences
- *trans* – (other side); elements that originate from another DNA strand, e.g. regulatory proteins

TFs - proteins that bind to enhancer or promoter regions
- activate or repress transcription

Most bind to specific DNA sequences (e.g. enhancers)

Transcription factors are grouped together in families, based on structural similarities
- families share common framework in DNA binding sites
- slight differences in binding sites cause differences in recognition
<table>
<thead>
<tr>
<th>Family</th>
<th>Representative transcription factors</th>
<th>Some functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homeodomain:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hox</td>
<td>Hoxa1, Hoxb2, etc.</td>
<td>Axis formation</td>
</tr>
<tr>
<td>POU</td>
<td>Pit1, Unc-86, Oct-2</td>
<td>Pituitary development; neural fate</td>
</tr>
<tr>
<td>LIM</td>
<td>Lim1, Forkhead</td>
<td>Head development</td>
</tr>
<tr>
<td>Pax</td>
<td>Pax1, 2, 3, 6, etc.</td>
<td>Neural specification; eye development</td>
</tr>
<tr>
<td>Basic helix-loop-helix (bHLH)</td>
<td>MyoD, MITF, daughterless</td>
<td>Muscle and nerve specification; <em>Drosophila</em> sex determination; pigmentation</td>
</tr>
<tr>
<td>Basic leucine zipper (bZip)</td>
<td>C/EBP, AP1</td>
<td>Liver differentiation; fat cell specification</td>
</tr>
<tr>
<td>Zinc finger:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>WT1, Krüppel, Engrailed</td>
<td>Kidney, gonad, and macrophage development; <em>Drosophila</em> segmentation</td>
</tr>
<tr>
<td>Nuclear hormone receptors</td>
<td>Glucocorticoid receptor, estrogen receptor, testosterone receptor, retinoic acid receptors</td>
<td>Secondary sex determination; craniofacial development; limb development</td>
</tr>
<tr>
<td>Sry-Sox</td>
<td>Sry, SoxD, Sox2</td>
<td>Bend DNA; mammalian primary sex determination; ectoderm differentiation</td>
</tr>
</tbody>
</table>
estrogen receptor zinc finger domain
Transcription Factor Domains

Three major domains:

1. DNA-binding
   - recognizes particular DNA sequence
Three major domains:

1. DNA-binding  
   - recognizes particular DNA sequence

2. trans-activation – activates or represses transcription  
   - often interacts with proteins that bind RNA polymerase II;  
     e.g. TFIIB, TFIIE  
   - often involved with enzymes that modify histones

3. protein-protein interaction domain  
   - promotes dimerization  
   - allows it to be modulated by TAFs or other transcription factors
Transcription Factor Domains

Transcription factor MITF
- basic helix-loop-helix
- homodimer is the functional protein

The *trans*-activating domain is contained in the center of the protein.
- when bound to a promoter or enhancer, the protein is able to bind a TAF (p300/CBP)
- TAF p300/CBP is a *histone acetyltranferase*
Differential Gene Expression

Overview

Chromatin structure
Gene anatomy
RNA processing and protein production
Initiating transcription: promoters and basal transcription factors
Regulation of gene transcription
  enhancers
  specific transcription factors
Epigenetic processes (modifications of DNA structure)
  methylation
  imprinting
  dosage compensation
  non-coding RNAs
Alternative RNA splicing
DNA Methylation

- found in vertebrates; not *Drosophila*, nematodes, inverts
- methylation stabilizes nucleosome; stable nucleosome = transcriptional repression
- degree of methylation is proportional to degree of transcription
DNA Methylation

Absence of methylation correlates with tissue-specific expression

- methylation patterns maintained throughout cell division by DNA (cytosine-5)-methyltransferase
Genomic Imprinting

Special case of DNA methylation: alleles from maternal and paternal genome - differentially methylated
- activity inhibited in different tissues

Males:
- UBE3A imprinted
- ubiquitin pathway

Females:
- 7 genes imprinted; (SNRP, necdin, etc)

Chromosome 15

Wild-type

Angelman syndrome

Prader-Willi syndrome

Lethal
Dosage Compensation

Mammals – female = XX, male = XY
   - excess X chromosome products has negative consequences
   e.g. XXY males (Klinefelter’s syndrome)
     - infertility, some degree of language impairment, etc.

X chromosome dosage is equalized in females by inactivation of a single X chromosome in mammalian XX cells

*Drosophila* – transcription rate of male X is doubled

*C. elegans* – both Xs partially repressed (♀ = hermaphrodite)
X Chromosome Inactivation

XX cell

Barr bodies
(heterochromatin)

early embryo – both active

Inner cell mass

Trophoblast

late – only one active

Extraembryonic tissue of placenta

Embryonic cells

Extraembryonic yolk sac precursors

Note – the inactive X chromosome remains inactivated throughout the life of the organism
**X Chromosome Inactivation**

**Lyon hypothesis:**
1. Both X chromosomes active in very early female development.
2. One X is inactivated in each cell.
3. Inactivation is random.
4. The process is irreversible. All progeny cells will retain the same inactivation pattern.

- heterozygous for coat color
- genes contained on X chromosome

**Calico cats**

*early*  
*late*  
*X chromosome inactivation*
Non-Coding RNA

The majority of the genome is transcribed into non-coding RNA.

**Short ncRNA**
- function depends on sequence complementarity
- processed by Dicer and RISC (Argonaut)
- miRNA, siRNA – silence target genes
- piwi-interacting RNAs (piRNA) – roles in silencing repeated gene elements during spermatogenesis

**Developmental function:**
- Dicer-deficient mice die early in embryogenesis
  - also have defective embryonic stem cells
- miRNAs repress expression of certain Hox genes (patterning genes) along the antero-posterior axis
Non-Coding RNA

**Long ncRNA** (300 bp – many kb)
- transcribed, spliced, polyadenylated, developmentally regulated

**Mechanism:**
- sequence-independent
- *cis*-functioning:
  - transcribed on sense or anti-sense strand – interferes with transcription of neighboring gene
- *trans*-functioning:
  - processed ncRNAs binds to genes or chromosomes
  - recruit transcription repressor proteins

**Developmental function:** mostly unknown
- may repress *Drosophila Hox* genes in *cis* (transcriptional interference)
- may regulate mammalian *Hox* genes in *trans*
- *Xist* RNA coats X chromosome targeted for inactivation (*trans*)
  - recruits proteins that methylate histone H3 (inactivation)
Alternative Splicing

\( \alpha \)-Tropomyosin

\( \alpha \)-TM EXON GENE ORGANIZATION

\( \alpha \)-TM mRNA TRANSCRIPTS

- striated muscle
- striated muscle
- myoblast
- smooth muscle
- neuroblast/muscle
- hepatoma
- brain