

Genes, Proteomes and Genetic Determinism

Can we model a cell?

Karen F. Greif, Ph.D.

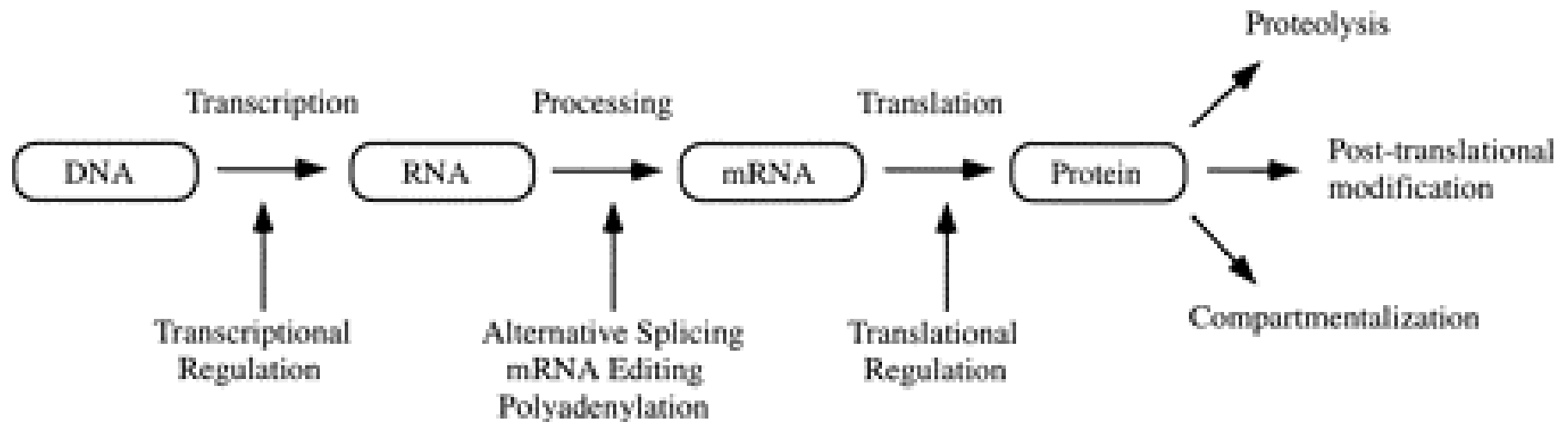
A molecular biology glossary

- Genes: units of information in DNA encoding functional molecules; divided into structural and regulatory components
- Genome : all the nuclear DNA for a given organism (i.e., all its genes plus “non-coding” DNA)
- Proteome: all the proteins in a cell (not one-to-one with genes!)
- Cis-regulatory sequences : sequences of DNA generally “upstream” of the gene’s structural sequence that regulate its expression. Also referred to as Promoters and Enhancers
- Transcription factors: proteins that bind to cis-regulatory sequences to control gene expression. These are encoded by other genes!

The “-omics”

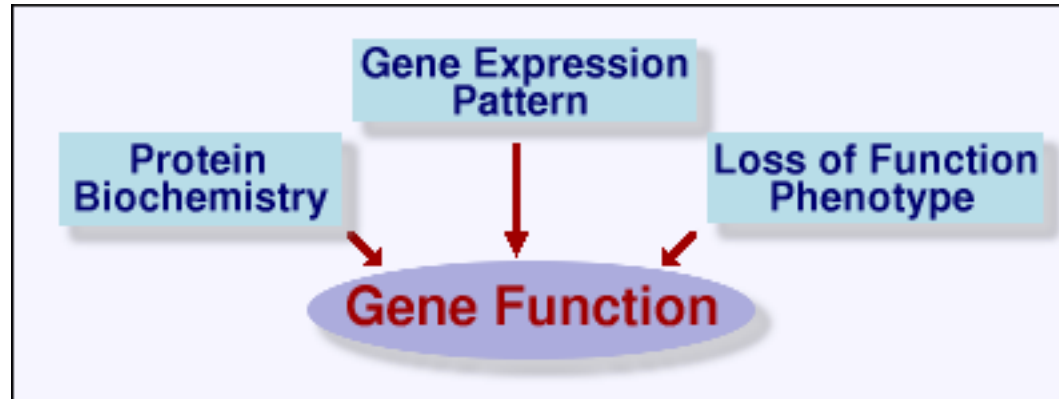
- Genomics : identification and characterization of all the genes in an organism
 - Reverse genomics: determination of the minimal genome required for life
- Proteomics: characterization of all the proteins in an organism
 - Functional proteomics: characterization of all interactions between proteins in space and time

Information flows (generally) from DNA ->RNA->protein



Information content may be strongly influenced by external factors at all stages! Different proteins may be produced as a result.

Understanding a gene's function requires more than just analysis of its expression



Also required: understanding the interactions with other proteins (many proteins operate as part of larger complexes). Researchers are attempting to analyze all interactions between proteins in cells (Functional proteomics)

Some questions

- Given complete information about all the genes in an organism (its genome), can we model it?
- Can we sort out all the influences on gene expression?
- Can we determine all the interactions between proteins?

One group's approach to understanding gene interactions



Eric Davidson

The sea urchin, *S. purpuratus*

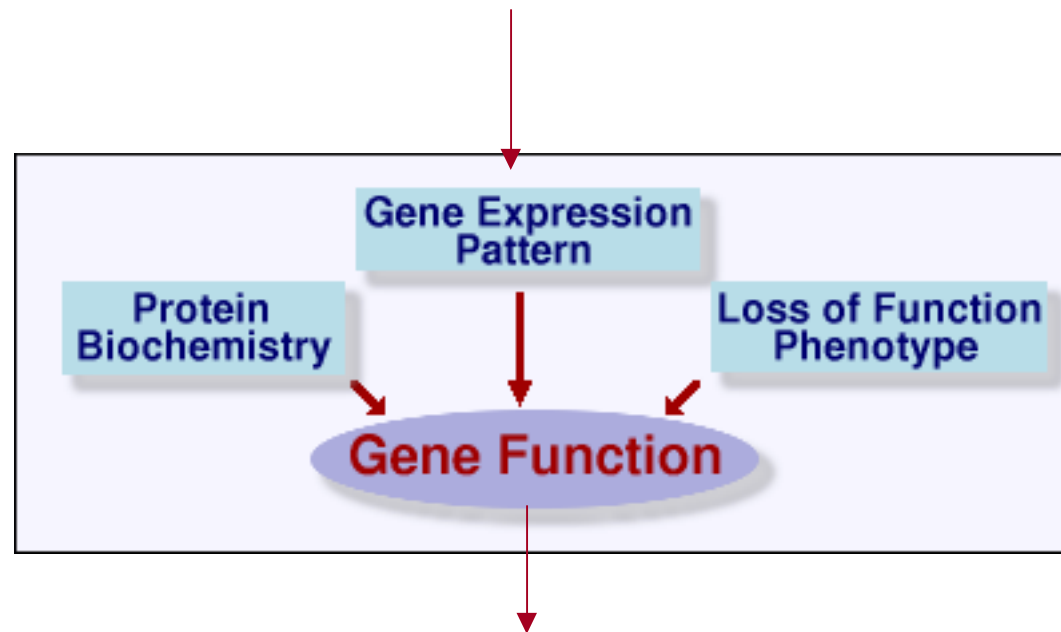


<http://sugp.caltech.edu:7000/>

Study the genes that control embryonic development of the organism

The Davidson strategy: What controls gene expression in time and space?

Regulatory elements that control gene expression

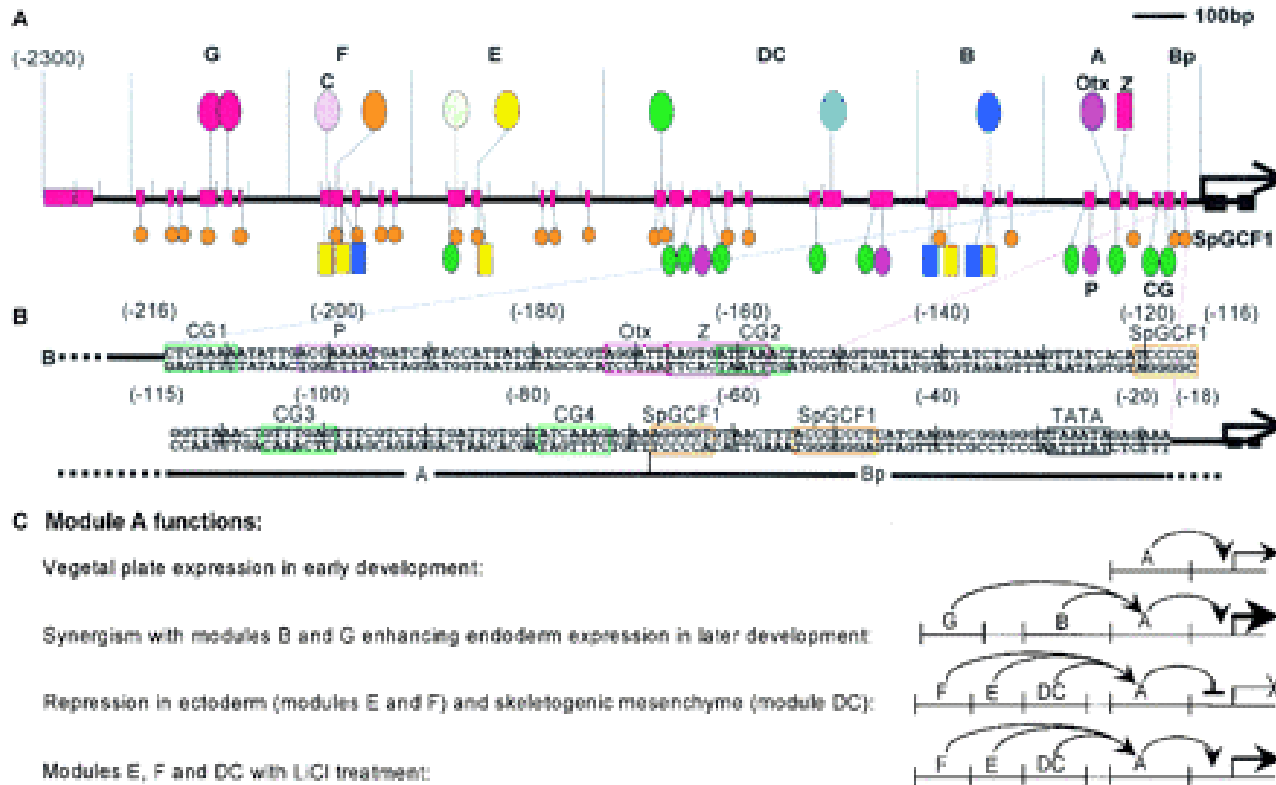


Cascades of regulatory genes that control other genes

Strategy

- Most traditional molecular biology approaches can only examine the function of a few genes at a time; not conducive for studying networks
- By focusing only on regulatory elements (transcription factors and their cis-regulatory targets on other regulatory genes) the regulation of genes may be studied and then added into the network.
- This is possible because determination (the progressive narrowing of ultimate cell fate) of cell types in early sea urchin embryos is separable from differentiation (expression of cell-type specific proteins), but only by a few steps

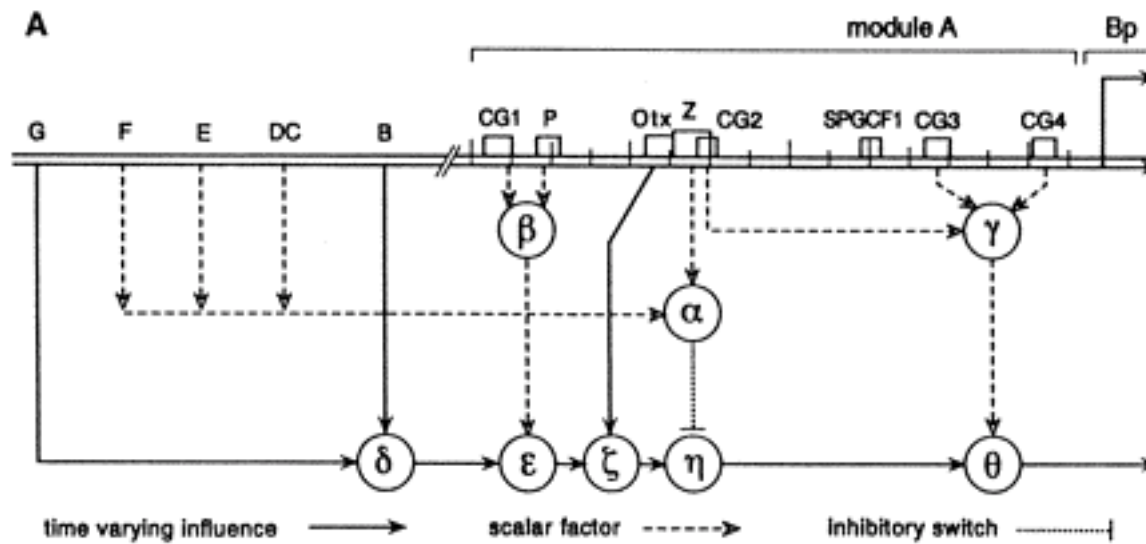
The cis-regulatory network of a single gene (Endo 16)



The promoter is divided into “modules” that have different functional outcomes

Yuh et al (1998) Science 279: 1896.

The model for functional interactions between modules



The regulatory “program” derived from empirical data

B

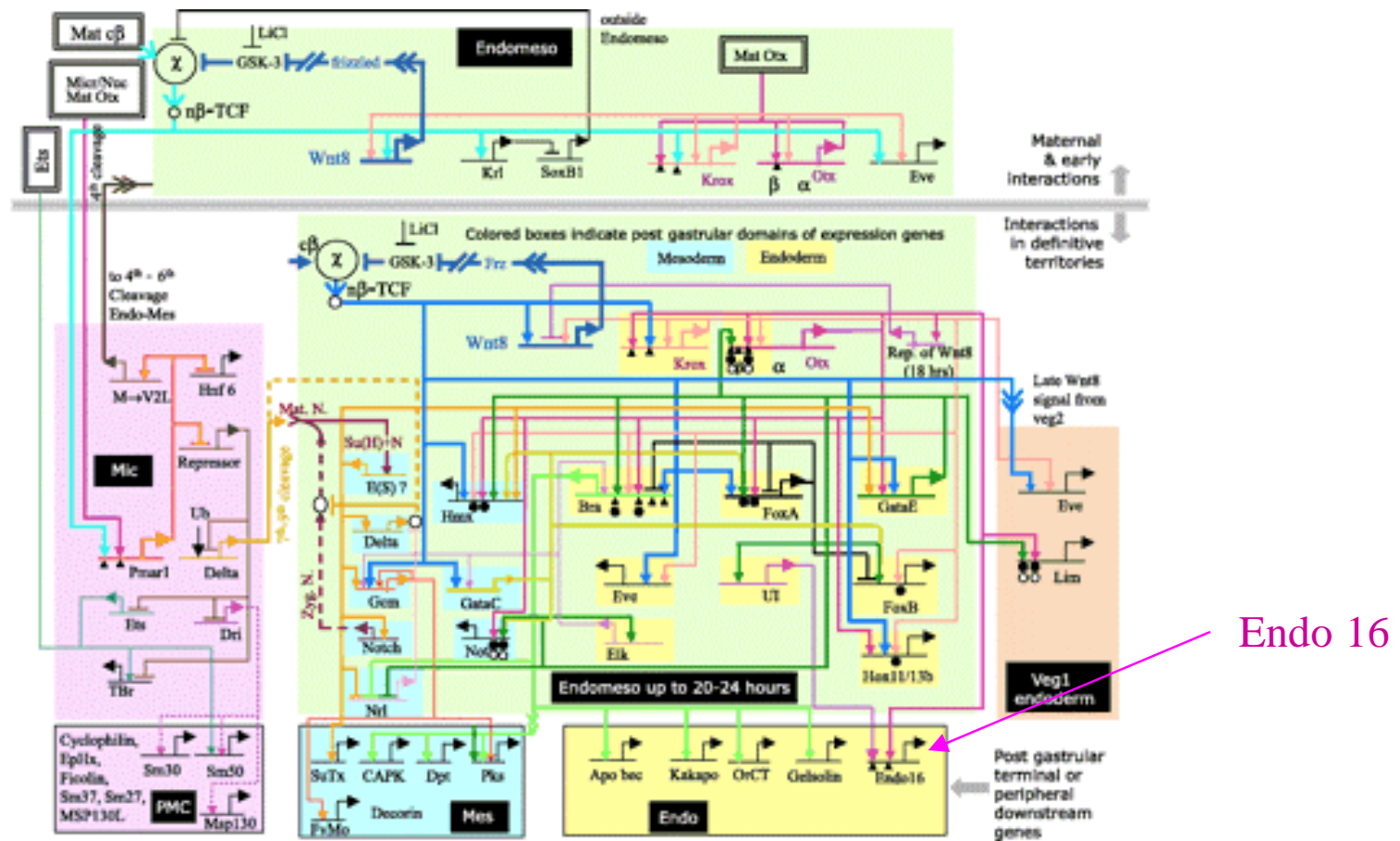
if (F = 1 or E = 1 or CD = 1) and (Z = 1)	Repression functions of modules F, E, and DC mediated by Z site
$\alpha = 1$	
else $\alpha = 0$	
if (P = 1 and CG ₁ = 1)	Both P and CG ₁ needed for synergistic link with module B
$\beta = 2$	
else $\beta = 0$	
if (CG ₂ = 1 and CG ₃ = 1 and CG ₄ = 1)	Final step up of system output
$\gamma = 2$	
else $\gamma = 1$	
$\delta(t) = B(t) + G(t)$	Positive input from modules B and G
$\varepsilon(t) = \beta \cdot \delta(t)$	Synergistic amplification of module B output by CG ₁ -P subsystem
if ($\varepsilon(t) = 0$)	Switch determining whether Otx site in module A, or upstream modules (i.e., mainly module B), will control level of activity
$\xi(t) = Otx(t)$	
else $\xi(t) = \varepsilon(t)$	
if ($\alpha = 1$)	Repression function inoperative in endoderm but blocks activity elsewhere
$\eta(t) = 0$	
else $\eta(t) = \xi(t)$	
$\Theta(t) = \gamma \cdot \eta(t)$	Final output communicated to BTA

Two ways to view a network model

The “view from the genome”: all relevant inputs into each regulatory element that may occur in all cells at any time

The “view from the nucleus”: only those interactions occurring in given nuclei in the particular time frame (not all genes are expressed by all cells and/or at all times)

The “view from the genome” of regulation of one stage of embryonic development in sea urchins (Just follow the arrows!)



Davidson, E et al (2002) Science 295: 1669.

(n.b.: only about 50 of the approximately 8500 genes expressed in the embryo are shown...)

How far can we go using this strategy?

Developmental regulatory network analysis can be done in any organism where the necessary genomics, a high-throughput method of gene transfer, and the ancillary molecular methods are available. But it requires a new mix of technologies and a new level of close interactions between system-minded biologists and computational scientists...**The model thus represents an outline of the heritable developmental program, but the program is not the machine.** The DNA regulatory network coexists with many other multicomponent systems that constitute the machine. These systems execute biochemical functions, produce signal transduction pathways and cause cell biological changes to occur. They sum to the majority of the working parts of the cell. Their mobilization is controlled by the transcriptional switches that hook them into the genomic regulatory control system. (p. 1677)

Davidson, E. et al (2002) Science 295: 1669.

The “big” questions

- Can we model the networks in complex organisms? (or, for that matter, in simple ones?)
- If we can, does that allow us to predict behavior reliably?
- How much of behavior (in all contexts) is genetically determined?