Gene Regulation

Sonja J. Prohaska

Computational EvoDevo, University of Leipzig
Santa Fe Institute, Santa Fe, NM, U.S.

WS 2017/18
Gene regulation: what is it?

- What is a gene?
- What is regulation?
- What is “gene regulation”?
- What does it mean “a gene is regulated”?
- How are genes regulated?
- How can we observe that a gene is regulated?
- Why is it important that genes are regulated?
What is a “gene”?

What is a gene?
(the physical basis of) the heritable unit of function
sloppy: a DNA sequence encoding a protein (or RNA)

What is it made of?
ribonucleic acid, specifically DNA (or RNA)

What does it do?
nothing, it is very passive
it holds, provides and stores information about the protein (or RNA)

What is gene expression?
it is the molecular process that turns DNA into RNA (transcription) and
(eventually) RNA into protein (translation)

How can we define a gene?
sloppy: the DNA sequence(s) giving rise to a (functional) gene product, such as
a protein or RNA
conceptually: the gene is a construct with a structural (i.e. DNA sequence) and
a functional (i.e. function of the gene product) component which are set into
relation by the process of gene expression

1 More detailed descriptions/definition will be given later.
2 The reason to mention RNAs explicitly is: RNA genes are almost as abundant as protein genes in the human genome.
What is “regulation”?

- Where do we observe regulation?
  in biological systems, societies, systems designed to do so

- How can we define regulation?
  principle or rule employed in controlling, directing, or managing an activity, organization, or system

- How is regulation established, changed, and maintained?
  e.g. by an authority, adaptation and reinforcement

---

1. There is no formal Theory of Regulation (in natural sciences).
2. If you are really interested in this read up on control theory.
What is “gene regulation”? 

- What is gene regulation?
  the regulation of gene expression

- What is regulated?
  - the **variant/isofom** of the gene product produced
  - the **amount** of gene product produced
  - the **timing** of gene product production
  - the spatial **localization** of gene product

- Who is regulating?
  cellular/organismal level: external and internal signals/stimuli
  molecular level: gene products of regulatory genes

- How is regulation of gene expression achieved?
  by a very large set of molecular mechanisms...

Regulatory genes are regulating regulated genes.

A regulated gene can be a regulatory gene.
Example: Lactose Intolerance in Humans

- all babies can drink milk
- some adults cannot drink milk
- milk contains lactose
- **lactose** is a disaccharid of galactose and glucose
- humans have a gene (names: LHP1, LAC, LCT) that encodes **lactase**
- lactase is an enzyme that cleaves lactose into galactose and glucose
- lactase is expressed only in enterocytes until 3-4 years of age
- when we drink milk lactose gets into our gut
- if lactase is expressed in enterocytes (cells in the gut) lactose is cleaved (this is good)
- else lactose is fermented by gut microbes gases are produced → bloating (this is bad)
Example: Lactose Intolerance

Video:
https://en.wikipedia.org/wiki/Lactose_intolerance
Example: Lactose Intolerance

Video:
https://en.wikipedia.org/wiki/Lactose_intolerance
How can we observe that a gene is regulated?

applying a strategy from math we ask the inverse question first: How can we observe that a gene is NOT regulated?

The gene is expressed/the gene (product) function is present

- at the same amount
- in every cell/location
- at every point in an organisms lifetime
  with a variation of $\epsilon$.

What should we be looking for to observe regulation?

We should be looking for non-random/statistically significant changes

- in the amount
- the cell or organismal location
- and the time
  of gene expression

So, how do we do that? Compare (enough) samples.
Let’s measure gene regulation!

To measure the amount of

- about 23,000 protein coding and 17,000 RNA genes
- in about 270 cell types
- at $x$ time points (How dense do we want to sample?)
- on RNA and protein level $^1$
- for 3 biological replicates $^2$

we need $x \times 64.8 \times 10^6$ measurements.

BIG DATA!

$^1$ because we don’t know how well they are correlated
$^2$ because this is the absolut minimum for doing statistics
Expression of the JAK1 gene across multiple tissues.

- measured on the level of RNA

JAK1 gene expression in 79 human tissues taken from the bioGPS database.
Relative amount of gene expression in 6 samples under two conditions.

- measured on the level of RNA

red – overexpression, green – underexpression
RNA expression of a genomic region at level of nucleotide resolution.

- from transcriptome sequencing
- measured on the level of RNA

5 samples; 200kb genomic region; many genes
Gene expression at 9 development time points.

- measured on the level of RNA

38 genes; 9 time points; 2 independent methods for verification

Identification of expressed/present proteins.

- measured on the level of protein

2D-gel electrophoresis; separation by size (x axis) and charge (y axis)
Protein localization in the body.

- measured on the level of protein

antibody staining of hox genes (lab to Abd-B) in the crustacea Parhyale sp.

taken from work by Dr. Heather Bruce from the lab of Prof. Patel.

http://www.patellab.net/portfolio-view/heather-bruce/
RNA localization in the cell.

- measured on the **level of RNA**

RNA-FISH, mRNA of beta-actin (red) in the cytoplasm (green) of MEF cells; DNA (blue)