Horizontal Gene Transfer

Genetic Variation:
The genetic substrate for natural selection

Dr. Carol E. Lee, University of Wisconsin

What about organisms that do not have sexual reproduction?

In prokaryotes:

**Horizontal gene transfer (HGT):** Also termed Lateral Gene Transfer - the lateral transmission of genes between individual cells, either directly or indirectly.

This transfer of genes between organisms occurs in a manner distinct from the vertical transmission of genes from parent to offspring via sexual reproduction, which could include transformation, transduction, and conjugation.

These mechanisms not only generate new gene assortments, they also help move genes throughout populations and from species to species. HGT has been shown to be an important factor in the evolution of many organisms.

From some basic background on prokaryotic genome architecture

General Principles

- **Most conserved feature of Prokaryotes is the operon**
- **Gene Order:** Prokaryotic gene order is not conserved (aside from order within the operon), whereas in Eukaryotes gene order tends to be conserved across taxa
- **Intron-exon genomic organization:** The distinctive feature of eukaryotic genomes that sharply separates them from prokaryotic genomes is the presence of spliceosomal introns that interrupt protein-coding genes

Figure 1: Approximate distribution of evolutionary constraints across genomes with different architectures. The fractions of different classes of sequences that are subject to constraints of varying strengths are shown as rough approximations of the values that are typical of the respective class of genomes. The data are from [Rogers 2005](Rogers 2005), as discussed in the main text.
Small vs. Large Genomes

1. **Compact**, relatively small genomes of viruses, archaea, bacteria (typically, <10Mb), and many unicellular eukaryotes (typically, <20 Mb). In these genomes, protein-coding and RNA-coding sequences occupy most of the genomic sequence.

2. **Expansive**, large genomes of multicellular and some unicellular eukaryotes (typically, >100 Mb). In these genomes, the majority of the nucleotide sequence is non-coding.

**Prokaryotic Genomes**

- Even though bacteria and archaea are not closely related, they share certain features in Genome architecture

**Prokaryotic Genomes**

- Prokaryotes (archaea and bacteria) have compact genomes, though with larger intergenic regions than viruses

- Many prokaryotic genes are organized into co-transcribed groups, or **operons** (Miller and Reznikoff, 1978; Salgado et al., 2000)

  - **Operons**: https://www.youtube.com/watch?v=10YWgqmAEsQ

**Prokaryotic Operons**

- A key distinctive feature of prokaryotic genomes are that they are organized into **operons**, clusters of co-regulated genes

  - **Definition**: groups of adjacent, co-expressed and co-regulated genes that encode functionally interacting proteins

  - Genes within operons are close together in the genome and co-transcribed and co-regulated

  - Grouping related genes under a common control mechanism allows bacteria to rapidly adapt to changes in the environment

  - Operons occur primarily on prokaryotes, but have been found in some eukaryotes (nematodes, Drosophila)
The organization of genes into an operon allows for simultaneous expression of all the genes that are located in cis (i.e., on the same contiguous piece of DNA) in the operon. It also allows the set of genes to undergo horizontal gene transfer as a unit.

Some operons belong to complex, interconnected neighborhoods: “super operons” or überoperon (Lathe et al., 2000) are large arrays of genes that include several operons with a complex pattern of regulation, such as the ribosomal superoperon. The majority of genes in the überoperons encode proteins involved in the same process and/or complex but highly conserved arrangements including genes with seemingly unrelated functions exist, as well.

Operons with identical or similar gene organization are often found in highly diverse organisms, and across different functional systems. Examples are numerous metabolite transport operons, such as transmembrane, ATPase, and periplasmic subunits of diverse permeases. The persistence of such common operons in diverse bacteria and archaea has been interpreted within the framework of the selfish operon concept, where the selfish character of these compact genetic elements make them prone to horizontal spread among prokaryotes (Lawrence, 1997, 1999; Lawrence and Roth, 1996).

An operon contains one or more structural genes which are generally transcribed into one polycistronic mRNA (a single mRNA molecule that codes for more than one protein). Most prokaryotic mRNA is polycistronic, having a single mRNA that encodes for multiple different polypeptides.

Most of the mRNA found in bacteria and archaea is polycistronic, having a single mRNA that encodes for multiple different polypeptides. Bacterial Operons Produce Polycistronic mRNAs: Polycistronic mRNA carries the information of several genes, which are translated into several proteins. These proteins usually have a related function and are grouped and regulated together in an operon.

So, Most Prokaryotic mRNA is Polycistronic:
• Most of the mRNA found in bacteria and archaea is polycistronic, having a single mRNA that encodes for multiple different polypeptides.
• Bacterial Operons Produce Polycistronic mRNAs: Polycistronic mRNA carries the information of several genes, which are translated into several proteins. These proteins usually have a related function and are grouped and regulated together in an operon.
• (most eukaryotic mRNA is monocistronic)
Lack of Gene Order Conservation in Prokaryotes

- One of the unexpected findings of the first comparisons of complete bacterial genomes has been the near lack of gene order conservation, beyond the level of operons, even between relatively close species such as *Escherichia coli* and *Haemophilus influenzae* (Koonin et al. 1996; Tatusov et al. 1996)

- This breakdown in "synteny" (gene order) in prokaryotes is thought to be caused by horizontal gene transfer and also inversions around the origin of replication

Syntaxeny

- Definition 1: Classical Genetics: Physical co-localization of genetic loci on the same chromosome
  (but bacteria often have one just one chromosome…)

- Definition 2: Genomics: syntenic regions refer to the case where both sequence and gene order are conserved between two closely related species

Evolution in Prokaryotes

- Mutation rate: high within a given amount of time because of short generation time
- Also, often high mutation rate per base pair per replication

- Horizontal Gene Transfer (= Lateral Gene Transfer)

Horizontal Gene Transfer

Consequences:
- Phylogenetic relationships are sometimes difficult to discern (as genetic material is being swapped around)
- Rapid transfer of functional genes: pathogenicity genes, rapid evolution of drug resistance
- Bacteria effectively have a HUGE genome size (Pan-Genome), a large genome to draw from, as individual cells can share genes with other individuals
Core and Pan Genomes

• The **Core Genome** consists of genes shared by all the strains studied and probably encode functions related to the basic biology and phenotypes of the species

• The **Pan-Genome** is the sum of the above core genome and the dispensable genome
  – The dispensable genome contributes to the species’ diversity and probably provides functions that are not essential to its basic lifestyle but confer selective advantages including niche adaptation, antibiotic resistance, and the ability to colonize new hosts.

• The Pan-Genome tends to be much much larger than the Core Genome of a prokaryotic “species”

The Pan-Genome

Each gene can be classified into one of three groups:

**Core Genome** (blue): the extended core genes, which include those that control translation, replication and energy homeostasis. ~250 gene families.

‘**Character**’ genes (red): involved in adaptation to particular environmental niches, such as those that control photosynthesis or endosymbiosis. ~7,900 character gene families.

**Accessory genes** (green): nearly limitless in size. These genes are often specific for a strain or serotype, and in many cases have no known function.

Transformation and Transduction

• **Transformation**: when a prokaryotic cell takes up and incorporates foreign DNA from the surrounding environment

• **Transduction**: movement of genes between bacteria by bacteriophages (viruses that infect bacteria)

Conjugation

**Conjugation** is the process where genetic material is transferred between bacterial cells

• Sex pili allow cells to connect and pull together for DNA transfer

• A piece of DNA called the **F factor** is required for the production of sex pili

• The **F factor** can exist as a separate plasmid or as DNA within the bacterial chromosome
The F factor can exist as a separate plasmid or as DNA integrated within the bacterial chromosome.

(a) Conjugation and transfer of an F plasmid

(b) Conjugation and transfer of part of an Hfr bacterial chromosome

The F Factor as a Plasmid

- Cells containing the F plasmid function as DNA donors during conjugation: they are able to construct the sex pilus
- Cells without the F factor function as DNA recipients during conjugation
- The F factor is transferable during conjugation

The F Factor in the Chromosome

- A cell with the F factor built into its chromosomes functions as a donor during conjugation
- The recipient becomes a recombinant bacterium, with DNA from two different cells

The F Factor in the Chromosome

- A cell with the F factor built into its chromosomes functions as a donor during conjugation
- The recipient becomes a recombinant bacterium, with DNA from two different cells

R Plasmids and Antibiotic Resistance

- R plasmids: plasmids that carry genes that encode antibiotic resistance
- R-plasmids are typically transferred through conjugation, less so via transduction; difficult to treat bacteria with drugs as resistance alleles can be readily transferred
- Antibiotics select for bacteria with genes that are resistant to the antibiotics, and antibiotic resistant strains of bacteria are becoming more common

R Plasmids and Antibiotic Resistance

- R plasmids: plasmids that carry genes that encode antibiotic resistance
- R-plasmids are typically transferred through conjugation, less so via transduction; difficult to treat bacteria with drugs as resistance alleles can be readily transferred
- Antibiotics select for bacteria with genes that are resistant to the antibiotics, and antibiotic resistant strains of bacteria are becoming more common

Hot spots for HGT

- Certain environments are conducive to promoting HGT; for instance, the animal gut environment thought to promote HGT
- Bacteria in the gut might cause inflammation which increases HGT
- HGT is assumed to be important in archaea, but less well studied

Hot spots for HGT

- Certain environments are conducive to promoting HGT; for instance, the animal gut environment thought to promote HGT
- Bacteria in the gut might cause inflammation which increases HGT
- HGT is assumed to be important in archaea, but less well studied

Horizontal Gene Transfer Across Domains of Life

- Analysis of 40 animal genomes (e.g. Drosophila, C. elegans, primates) found that all the animal genomes acquired ~10s-100s active genes via HGT
- In the case of humans, 145 genes seemed to be acquired from other species, most from bacteria and protists, likely via HGT
- We do not know the mechanism of HGT of these events

Horizontal Gene Transfer Across Domains of Life

- Analysis of 40 animal genomes (e.g. Drosophila, C. elegans, primates) found that all the animal genomes acquired ~10s-100s active genes via HGT
- In the case of humans, 145 genes seemed to be acquired from other species, most from bacteria and protists, likely via HGT
- We do not know the mechanism of HGT of these events

Crisp et al. 2015. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. Genome Biology. 16:50
Questions

(1) What is horizontal gene transfer?
(2) What are the different mechanisms of horizontal gene transfer?
(3) What is the typical unit of horizontal gene transfer?
(4) What is an operon?
(5) What are the key differences between genome architectures of prokaryotes and eukaryotes?
(6) What is the Pan Genome of prokaryote “species?”

1. Which of the following does NOT contribute to rapid adaptation in prokaryotes?
   (a) High mutation rate
   (b) Short generation times
   (c) The huge Pan-genome, serving as a reservoir of potentially adaptive genes
   (d) Horizontal gene transfer
   (e) Unequal crossing over

2. What is the following is FALSE regarding prokaryotic genome evolution?
   (a) The typical unit of horizontal gene transfer is the operon
   (b) The genetic material that could be exchanged in a bacterial “species” tends to be much larger than the portion that is not exchanged
   (c) Transduction involved transfer of genetic material through a viral vector
   (d) Horizontal gene transfer leads to increases in heterozygosity
   (e) Horizontal gene transfer leads to decreases in synteny in bacterial genomes

3. Which of the following is TRUE regarding prokaryotic genomes?
   (a) The core genome usually contains the genes that aid in environmental adaptation and antibiotic resistance
   (b) The Pan Genome of a bacterial “species” tends to be very large, often larger than animal genomes
   (c) Genetic recombination (chromosomal crossing over) in prokaryotes reduces linkage disequilibrium
   (d) Horizontal gene transfer leads to increases in genotypic diversity
   (e) Horizontal gene transfer leads to increases in synteny (gene order) of bacterial genomes

4. Which of the following is TRUE regarding prokaryotes?
   (a) The Pan genome of a bacterial species includes all the genes except those of the core genome
   (b) For conjugation to occur, an operon is required for the production of sex pili
   (c) The operon is typically part of the core microbiome
   (d) Transformation is the genetic exchange between bacteria that is mediated by viruses
   (e) The genes expressed within an operon are typically functionally related
• Answers
• 1E
• 2D
• 3B
• 4E