Lecture 10 - Evolution of Eukaryotes

Meaning of endosymbiosis

- A living organism living within the cell of another living organism - they help each other out.

Origin of endomembrane system, nuclear membrane, ER, etc

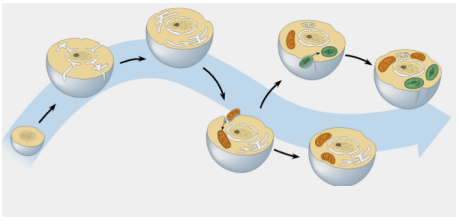
- The endomembrane system (nuclear envelope, ER) was thought to have been derived from infolding of the plasma membrane (supported by the face that ER is connected to the nuclear envelope.
- Having a nuclear envelope and a nucleus allowed for compartmentalization and was critical
- **Development of the n.e and ER is NOT part of endosymbiosis.

Origin of mitochondria and chloroplasts

- Mitochondria and chloroplasts are not derived from the same place the nuclear envelope was derived
- Mitochondria is said to have derived from an aerobic bacterium after a larger primitive cell brought it in through phagocytosis - it being the mitochondria.
- Chloroplast is said to have descended from a cyanobacteria
- 'Aerobic bacterium' that had mitochondria engulfed a cyanobacteria (lead to plants and such)

Bacteria that can undergo oxidative phosphorylation

May have been an anaerobic bacterium that had an endomembrane system



- Top branch gets you plants, algae, etc
- Bottom branch gets you animals, fungi, etc

- So, aerobic bacterium (free living bacteria) was engulfed and then evolved into the mitochondria over time
- A subset of those primitive cells that engulfed aerobic bacterium then engulfed cyannobacteria

Evidence supporting theory of endosymbiosis

1) Morphology - mitochondria look like bacteria (similar shape and what not) AND chloroplast kind of looks like cyannobacteria
2) Mitochondria/chlorplast divide within a living cell identical to the way bacteria divides
3) Mitochondria and chloroplasts are the only organelles within eukaryotic cells that have electron transport chains which means they probably did electron transport when they were free living cells as well
4) They have their own genomes (circular chromosomes)
5) They have transcription and translation machinery - have their own ribosomes and make their own proteins.
6) Mitochondria and chloroplast structure is very similar to prokaryotic cell structure

Factors driving development of early eukaryotic cells

- OXYGEN - about 2.2 bya cyanobacteria developed and they could do oxidative phosphorylation (split water - release oxygen into atmosphere)
- This paved the way for aerobic respiration --> provides MUCH more ATP than glycolysis and fermentation --> cell can now make more energy.

Why eukaryotic cells can be larger and more complex than prokaryotes

- Bacteria and Archaea have their centers of oxidative phosphorylation on their plasma membranes. So, if they try to get bigger, they need more centers of oxidative phosphorylation to give them more energy.
- Problem is that prokaryotes' plasma membrane's surface area increases as a function of radius SQUARED whereas the volume increases as a function of radius CUBED
- Volume gets bigger much quicker than surface area of plasma membrane and so the cell tries to compensate by putting more centers of oxidative phosphorylation on the plasma membrane BUT eventually it runs out of space on the p.m and that is what limits the size of prokaryotes (not being able to produce enough energy to support a larger cell)
- Prokaryotic Cells have a **HIGH** plasma membrane area to volume ratio

$A = 4\ \pi r^2$

$V = 4/3\ \pi r^3$

- However, eukaryotes can be much bigger because you have mitochondria and each mitochondria has many ox-phos centers and so you can produce MUCH more ATP and support a larger cell/genome
- **LOW** plasma membrane surface area to volume in eukaryotes.
- Eukaryotic cell has more energy to invest into protein synthesis and as a result can generate/express more proteins that are often more complex.

Evidence for lateral gene transfer from organelles to the nucleus

- Genes that code for proteins found in mito/chloro are not found and expressed in the nucleus (e.g - proteins that make up complex 1)

General idea about how lateral gene transfer is detected (Southern Blot)

- Isolate genomic DNA , run it on gel (make DNA single stranded)
- Add a single stranded DNA probe and see if hybridization occurs
- You can then see what genes you have in that genome.

- So you can isolate mitochondrial DNA, and nuclear DNA and add the probe of Oxidase 3 (mitochondrial gene)
- If the gene is found mtDNA but not the nDNA - Lateral gene transfer has not occurred
- If found in nDNA and not mtDNA - lateral gene transfer has occurred
- If found in both, LGT occurred and both genomes have a copy of this gene.

Hypothesis for why genes move to the nucleus from organelles (lateral gene transfer)

- Coordinated control / put them under tighter nuclear control
- Integrates the metabolism of the entire cell (see lecture 11)

Possible reasons why certain genes have NOT moved to the nucleus from organelles

- Lecture 11

Role of cpn60 in tracing endosymbiotic and lateral gene transfer event in eukaryotes

- cpn60 is a mitochondrial gene found in the nucleus (LGT)
- Giardia (creature with no mitochondria) has cpn60 in its nuclear genome so at one point it did have mitochondria but now cpn60 has no function in giardia.
- This means that the ancestor that gave rise to giardia and other eukaryotes had mitochondria with the cpn60 gene but giardia got rid of it to better adapt to its life style.

Lecture 11 - Intro to Prokaryotic Gene Structure

Relative sizes of typical mitochondrial, chloroplast, and nuclear genomes

Nuclear genomes(linear) > Chloroplasts Genome(circular) > Mitochondrial Genome(circular)

- If genome is 5000kb, and contains a total of 5000 genes, then every gene is made of 1000 base pairs (5,000,000/5000) and every amino acid is 3bp, so amount of amino acids in a gene of this organism is 333bp (1000/3)

Possible reasons why modern organelle genomes have become dramatically smaller over evolutionary time

- Got rid of genes that were useful for free living bacteria (genes coding for flagella)
- Got rid of redundant genes that were also in the nucleus - genes involved in glycolysis
- Certain genes have been deleted
- Hosts in which their mitochondria has suffered a mutation / deletion such that it has lost a bunch of genes, those genomes will be easier to replicate and those organisms will have a selective advantage
- Thus, it is in the organisms favor to get rid of all the genes they can in the mitochondria so long as they are redundant -- one way to get rid of these genes is by mutation and deletion, another way is to send them to the nucleus through *Lateral Gene Transfer*.

Possible reasons why genes have moved to the nucleus from organelles over evolutionary time

- Coordinated control
- Organelles, mitochondria and chloroplast are involved in electron transport and oxygen metabolism that generates reactive oxygen species. ROS's are very reactive and mutagenic and damage DNA and so it makes sense to get your DNA out of the organelles and into the nucleus -- get DNA away from ROS's
- To avoid any kind of cellular rejection system

Oxygen + electron = Reactive oxygen species

- However, **main reason** for LGT is *sexual recombination* - getting organelles out of the chloroplast and into the nucleus allows them to participate in sexual recombination and generate diversity.

Possible reasons why certain genes have NOT moved to the nucleus from organelles vs. eukaryotic nuclear environments

- They need to have **local control**
- **Transporting** proteins out of the cytosol into organelle may be to much of a hassle (too hard to move the protein - maybe too big -- **size**)
- **Structure** of genes in the organelle may not work in the nucleus - this is because the general environment in our organelles is essentially prokaryotic, whereas the environment in the nucleus is a eukaryotic environment - so it is a big change for the gene to go from one to the other.
- Could be due to chance - genetic drift. There hasn't been enough time yet for all of them to move - evolution is STILL occurring. Maybe in a million years there will be less DNA in mitochondria.

Rubisco structure and assembly from components coded by different genomes

- Complex protein components **(** proteins involved in Electron transport/Calvin Cycle [Rubisco] **)** are coded by DIFFERENT genomes.
- Some parts of the protein will be coded by genes in the nucleus, whereas others are coded from genes in the chloroplast.

Basic structure and function of RNA polymerase and ribosome

- Ribosomes are primarily RNA machines - it is the RNA that is catalytic
  The protein is structural
- In prokaryotes, RNA polymerase understands the information of the promoter - as a result, it binds to it and begins to transcribe
  RNA polymerase reads **3' to 5'** and synthesizes RNA **5' to 3'**
  - initial end of the RNA transcript is 5' and the other end is 3'

Examples of complimentary base pairing in gene expression

- Once mRNA is formed from transcription, it base pairs with itself - folding is critical for its structure.
- tRNA complimentary base pairs with itself to obtain proper 3D shape AND it also pairs with mRNA (codon with anticodon)
- rRNA in ribosome complimentary base pairs with itself to achieve proper 3D structure

Note:

- Mitochondria and chlorplast have their own gene expression machinery - they both have ribosomes
- Stop codons stop TRANSLATION , not transcription

Lecture 12 - Prokaryotic Gene Function (ISO)

Identify the sequence of standard “start” and “stop” codons

- Start codon : **AUG**
- Stop codons : **UAA, UAG,** & **UGA**

Identify the function of “start” or “stop” codons

- AUG codes for the amino acid methionine and is the first codon **translated** in ANY mRNA in both prokaryotic and eukaryotic cells.
- Stop codons (do not code for amino acids) act as “periods” indicating the the end of a polypeptide-encoding sentence. When a ribosome reaches one of the stop codons, polypeptide synthesis stops and the new polypeptide chain is released from the ribosome.

Compare the overall gene expression of prokaryotic vs eukaryotic cells

Eukaryotic Gene Expression	**Similarities**	**Prokaryotic Gene Expression**
• No SD box - instead, small subunit of ribosome recognizes 5‘cap, binds to it and moves along the mRNA (scanning) • Cleaving of polyadenylation signal signified termination of transcription	• Start codon is the same in translation • TATA boxes • Same termination ‘signal’ in translation.	• SD box helps initiate translation • Termination sequence is transcribed - forms loop and ends transcription • no micro RNA genes.

Lecture 12 - Prokaryotic Gene Function

Relative location of such DNA sequence “signals” as promoter, 5’ and 3’ UTR, “SD BOX”, start codon, stop codon, transcription terminator etc.(PROKARYOTIC)

- Gene begins with area called the promoter
- In between the promoter is the **TATA** box (TATA box is upstream of the transcription start point -- transcription start point is essentially where the promoter ends)
- The Transcription unit goes from the transcription start point to the transcription stop point
- The Gene includes the promoter AND the transcription stop point.
- **5’UTR - untranslated region** - stretch of DNA following the Promoter that is transcribed into mRNA but is **NOT** translated (downstream of the promoter, upstream of the start codon)
- Down stream of the 5‘UTR is the **start codon** , at the end of the gene is the **stop codon.**
- Downstream of the stop codon is the **3‘UTR -** transcribed - not translated
- At the end is the **terminator -** specific DNA sequence for a gene that signals the end of transcription of a gene.
- **SD box** is upstream of the start codon (found within the 5‘UTR)

In a DNA strand where promoter is on the left side, upstream is to the left of, and downstream is to the right of.

Note: Transcription starts here, not at the start codon.

Acts after it is transcribed

Region of DNA that once transcribed into mRNA, base pairs with ribosomal RNA to help the initiation of translation

Mechanism by which each signal is interpreted, or understood by cell (prokaryotes)

Promoter

- Sequence of promoters can be very attractive and so polymerase binds very stably or promoters can have sequences that less attractive and thus less efficient at transcription.
  - Promoter = signal , Interpreted by: RNA polymerase

Terminator

- In prokaryotes, terminator sequence in the DNA gets transcribed and ends up in the mRNA - this mRNA then complimentary base pairs with itself to form a loop (a hairpin) structure and the loop is what signals RNA polymerase to stop.

- Loop leads to a signal that causes mRNA to dissociate away from the DNA template - loop destabilizes mRNA thats bound to the DNA and so mRNA falls off and termination of transcription occurs
- **SIGNAL IS IN DNA** (Terminator DNA sequence) **, BUT IS UNDERSTOOD AS RNA** (RNA polymerase understands the 'loop' means stop)

SD Box

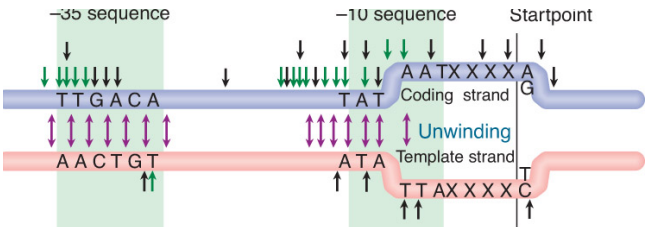
- SD box is found in DNA inside the 5'UTR, once transcribed, its RNA base pairs try to bond, however tRNA will not bond with them because translation starts at the start codon (downstream of SD box) and ribosome moves right (away from SD box)
- However, rRNA in the small subunit of the ribosome will pair with the mRNA to help initiate translation - SD box found in DNA but is understood/significant as RNA.

- The release factor is ALWAYS trying to get into the "A" site of the ribosome but is always outcompeted by tRNA. However, no tRNA binds to the stop codon and so the release factor has time to get in and BIND (**It does not base pair - it is a protein)** -- as a result of the binding of the release factor, translation stops.

Relationship between DNA sequence of signals and their function (how would low efficiency promoters be different than high efficiency promoters?)

Promoters

- Transcription starts at the startpoint (0) , and negative sequences indicate where the sequences are relative to the startpoint
- At the -10 sequence (upstream of startpoint) their is a lot of AT sequences - this is where the bubble begins (bubble is formed when Promotor attracts RNA polymerase and RNA polymerase interacts with the DNA)
- -35 sequence is the beginning of the promoter (upstream of -10 sequence)

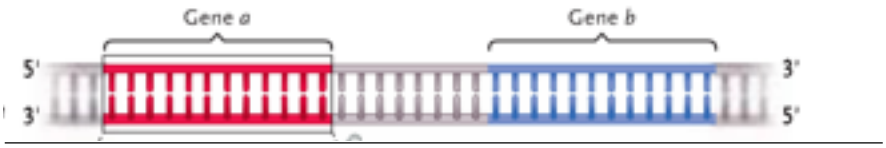


- Some promoters are very attractive - their sequence is such that polymerase makes a very stable bind and initiates transcription very frequently.
- Other promoters have different sequences and are less attractive and thus less efficient at transcription
  - Promoters have a general common sequence and structure but are also quite variable and can drive transcription at different rates.

Terminator

- If one terminator stops transcription 60% of the time and another only stops 40% of the time - this **could** be due to the length of the terminator sequence, longer = more stable and therefore higher termination rate
- One terminator sequence can have more G's and C's (3 hydrogen bonds b/w them) and so the loop will be more stable and as a result you have a more efficient terminator.

<u>Characteristics of promoters that require a particular position and direction</u>



- If the promoter of *gene b* is on the right, then polymerase will bind to to the right side of the blue gene and read 3' - 5' on the top strand

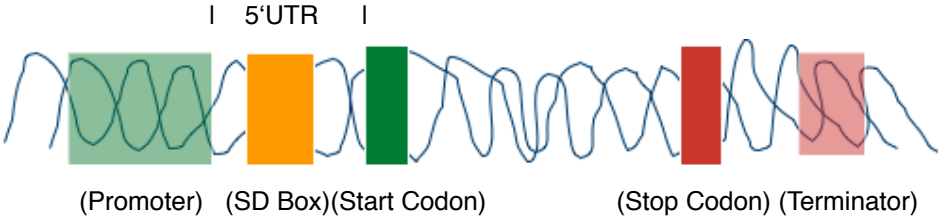
- If the promoter is on the **left** side of blue gene (gene b) then bottom strand will be the template strand that is read 3' to 5'
- **The strand that becomes the template strand depends on the position of the promoter.

Base sequence of start and stop codons as mRNA and DNA

- Start codon in DNA : 3' "TAC" 5' ,
  in mRNA : 5' "AUG" 3'

- Stop codon in DNA: 3' "ATT" 5' , 3' ATC 5' 3' ACT 5'
  in mRNA : 5' "UAA" 3' 5' UAG 3' 5' UGA 3'

The location of various signals given a diagram of gene expression

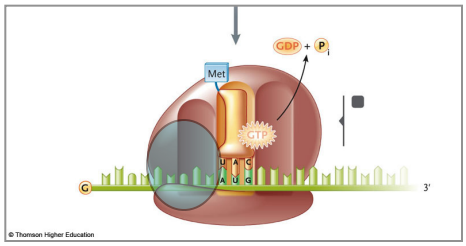


Note:

- Start codons are in DNA but are understood as RNA
- Start codons are **NOT** the start of transcription, they are the start of translation
- Start codon is NOT the first three bases transcribed
- Terminator is downstream of stop codon (terminate AFTER your transcribe stop codon)
- +1 nucleotide is the first nucleotide transcribed (right after promoter)
- **5'UTR** - transcribed but not translated - upstream/to the left of of start codon
- Some amino acids have MULTIPLE codons that code for them.

Mechanism by which each signal is interpreted, or understood by cell (prokaryotes) - continued

- Start codon (mRNA) attracts the first initiator tRNA (that codes for methianine) into the P-site of the ribosome and that initiates the process of translation



- To the left of the start codon is the UTR (mRNA to the left of start codon is UTR)
- These bases want to base pair, so in BACTERIA (prokaryotic) the ribosomal RNA in the small subunit of the ribosome pairs with those base pairs (in the blue circle) - this sequence is called the SD box
- **SD Box -** region on DNA that once transcribed into mRNA, complimentary base pairs with ribosomal RNA to help the initiation of translation
- SD box is found in DNA but not 'understood' as DNA

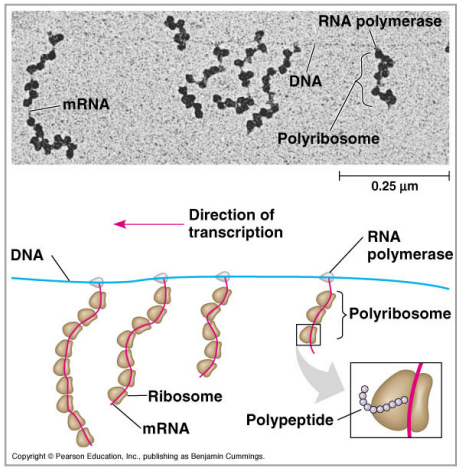
- Ribosome moves from the 5' to the 3' end of the mRNA and

- Ribosome eventually reaches the stop codon - no tRNA for stop codon, so release factor(a protein) has a chance to get in and BIND in the A-site (it does not base pair) - translation stops. (stop codon in DNA, understood as RNA)

Prokaryotic cell VS Eukaryotic cell

- In eukaryotes - transcription and translation are separated by the nuclear membrane which is NOT present in prokaryotes
- In prokaryotes, translation begins as soon as there is any mRNA available - **as soon as the SD box is transcribed, then ribosomes can jump on there and begin translating**
- Polymerase is making the message and at the same time, ribosome is translating this message.
- Multiple polymerases can be transcribing the SAME gene at the same time, as fast as they can, as fast as the promoter will allow.

- Multiple ribosomes translating every message all at the same time.
- Ribosomes are slower/bigger in eukaryotes and use method of scanning. (no SD box)



In this picture Polymerase moves from right to left (3' to 5') creating mRNA where the 5' end is furthest away from the DNA (blue line)

Ribosome translates from 5' to 3', therefore ribosome closest to the strand has been translating the longest and will have the longest polypeptide chain coming out of it.

Lecture 13 - Prokaryotic Gene Regulation (ISO)

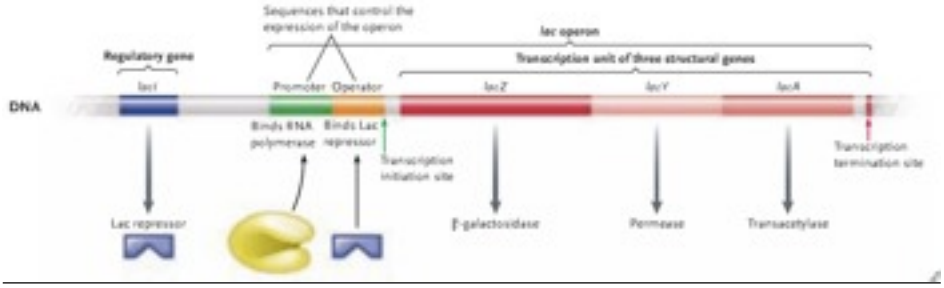
Identify the main features of bacterial operons

- **Operon -** cluster of prokaryotic genes and the DNA sequences involved in their regulation
- **Operator -** short segment that is a binding sequence for a **regulator protein** (repressor)
- When a **repressor** (regulatory protein) binds to the DNA, the likelihood that genes will be transcribed is GREATLY reduced.
- Other operons are controlled by regulatory proteins called **activators** - when bound to DNA, they increase likelihood of transcription.
- Each operon is transcribed as a unit from the promoter into a single mRNA - so the mRNA contains codes for several proteins
- A cluster of genes transcribed into a SINGLE mRNA is called a **transcription unit**

Identify the function of repressor proteins

- Binds to to the operator and as a result RNA polymerase is blocked from binding to the promoter (lac repressor acts as a road block)

Identify location of various components of the lac operon

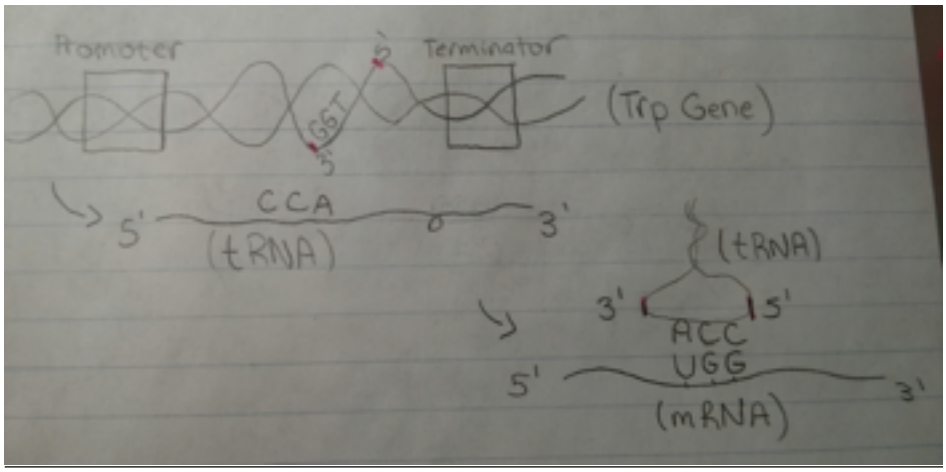


Lecture 13 - Prokaryotic Gene Regulation

DNA signals in RNA-coding genes

- RNA-coding genes are **NOT** translated - so the only DNA signals they need are a promoter and a terminator.
- No stop codon in RNA-coding genes - basically have no signals that would be understood during translation.
- tRNAS code for the anticodon that is used during translation.

DNA sequence of anticodon in tRNA gene, given the codon



UGG

- Tryptophan (Trp) gene (top drawing) has only a promoter and a terminator
- On the DNA strand (3' -> 5') , RNA polymerase transcribed the gene and makes tRNA (5' -> 3') - CCA
- Base pairing **MUST BE** complimentary and anti-parallel so when tRNA pairs with mRNA (5' -> 3') , the tRNA must be 3' -> 5' (ACC)
- Off the template strand of DNA read 3' -> 5' , the DNA base pairs of the anti codon were GGT if tRNA is 5' CCA 3' (before folding)
  Then DNA was 3' GGT 5'

- Trp mRNA codon is 5' UGG 3'
  So tRNA is 3' ACC 5' (in folded form - when pairing with mRNA)

Likely effect of base sequence substitutions in various DNA signals

- If you mutate the promoter you could inhibit polymerase binding or decrease the efficiency of the promoter OR the mutation could increase the efficiency of the promoter and make it more attractive to polymerase.

- If you mutate SD box you could make the SD box more functional - more attractive to rRNA or less functional - less attractive to rRNA

- Mutation to the start codon just screws it up - you cannot make the start codon 'work better' or more efficient (it is a lethal mutation)

- If you mutate the stop codon into another stop codon, nothing happens, BUT if you turn it into a different amino acid, then translation won't know to stop at that point. The ribosomal machinery will continue to read through and it won't stop until it finds the next in frame stop codon.

Stop codon is understood by TRANSLATION

As a result, over evolutionary time, many genes have got 2 or 3 stop codons in frame at their ends.

- If mutation occurs in the terminator, the effect will depend on whether the loop (hair pin) is strengthened and as a result the terminator becomes more efficient. OR the mutation could destabilize the loop - terminator becomes less efficient.

Change in amino acid coded, given a change in the DNA sequence (and Genetic Code table)

- **Silent mutation -** base-pair substation mutation that destroys one codon but creates another codon that codes for the SAME amino acid (no effect)
- **Missense mutation -** base pair substitution results in a different amino acid being coded - nothing serious happens if new a.a is similar enough to the old one, but for e.g if an amino acid that is hydrophobic is replaced with one that is hydrophilic - this could seriously affect protein function. (more severe than silent)

Or a negatively charged A.A for a positively charged one.

- **Nonsense mutation -** base-pair substitution results in the formation of a stop codon - this could result in a protein that will PROBABLY be too short and will PROBABLY be non-functional (more server than above two)

- **In/Del mutations -** addition or loss of a base pair. If a base pair is added, the entire reading frame shifts right (shifts downstream) and so a whole new set of amino acids are made (very severe)

Note

- Start codon determines which three amino acids end up being paired together - start codon sets the frame.
  - There can be three possible reading frames on a strand of DNA , e.g :

1)CAA ATG ACC.. --> 2)AAA TGA CC --> 3)AAT GAC --> 4)ATG ACC

In 4) your are back to the original reading frame

  - Therefore there are 3 possible reading frames on a strand of DNA

Could possibly occur as a result of three additions or three deletions

The location of various signals given a diagram of gene expression

(Other codons for other a.a's)

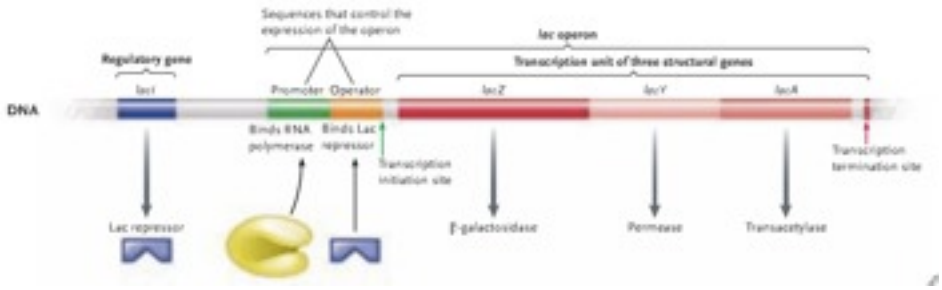
(Promoter) (SD Box)(Start Codon) (Stop Codon) (Terminator)

- Promoters are NOT transcribed // SD box IS transcribed /NOT translated
- Start codon IS transcribed and translated
- Stop codon is transcribed but NOT translated
- Terminator is transcribed but NOT translated

Basic structure of lac operon

- Going from upstream to down stream:

- LacI - regulatory gene, codes for Lac repressor protein
- Promoter - Where RNA polymerase binds
- Operator - Where Lac repressor binds
- Lac Z - encodes the enzyme *B-galactosidase*
- Lac Y - encodes the enzyme *Permease*
- Lac A - encodes the enzyme *Transacetylase*



Mechanism of action of lac repressor

- Lac repressor is encoded by the regulatory gene : ***lacI***
- LacI is upstream of the promoter/operator
- When there is no lactose in the medium , the lac repressor binds to the operator and prevents RNA polymerase from binding to the promoter

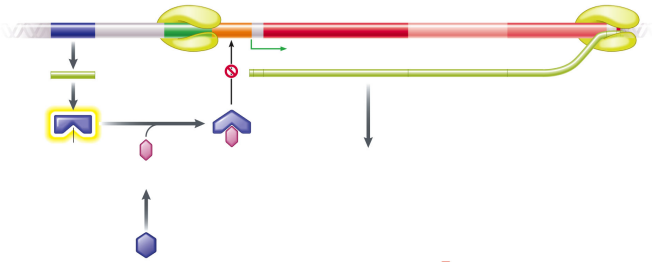
Function of lac operon in the presence, and absence, of lactose

Absence:

- lacI is transcribed and translated forming an *active* **lac repressor** (a protein)
- Lac repressor binds to operator and blocks transcription
- Transcription of structural genes occurs rarely (repressor occasionally falls of of operator temporarily) and so a few molecules of each enzyme is being made

Presence:

- Beta-galactosidase molecules (already present in the cell) convert some lactose into the inducer - **allolactose** (isomer of lactose)
- Allolactose binds to the repressor, inactivating it by altering its shape so that it cannot bind to the operator
- RNA polymerase binds to the promoter
- Transcription of *lac* operon (Z,Y,A) structural genes occurs
- Translation produces the three lactose catabolism enzymes



purple hexagon - lactose -->
Pink - allolactose

Note

- Our mitochondria have some operons (otherwise operons are only in prokaryotes)
- The lac operon is an inducible operon - its normal state is off
- lac I is independent of Z,Y,A and has its own promoter/terminator..etc
- Lac I repressor protein is a dimer
- Lac repressor creates a loop of DNA that prevents polymerase from transcribing the operon

Possible location of mutations in lac operon
Phenotype that would arise from a given mutation in lac operon under given conditions

Lecture 14 - Eukaryotic Genes

Basic structure of eukaryotic vs prokaryotic cell with respect to gene expression

- Eukaryotic cells are compartmentalized - they can keep their ribosomes away from their mRNA until the mRNA leaves the nucleus. This period of shelter allows for the mRNA to be enhanced in many ways.
- Prokaryotic cells have no nucleus and thus transcription and translation occur simultaneously

Structure of eukaryotic endomembrane system with respect go gene expression

- Eukaryotes have a nucleus that is surrounded by a nuclear envelope
- The nuclear envelope is semi permeable and so transcription products can leave the nucleus and protein products can enter the nucleus through the nuclear pores.

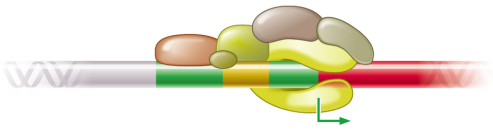
Structure of eukaryotic promoters/enhancers

Promoter:

- Contains TATA box (about 30 base pairs upstream of the transcription start point)
- Promoters usually have **proximal regions** upstream of them. Regulatory proteins that bind to promoter proximal elements may stimulate or inhibit the rate of transcription initiation.
- When a TATA binding protein binds to the TATA box, it makes the promoter MORE attractive to RNA polymerase II.
- TATA binding protein attracts polymerase AND several other activators - transcription factors that bind onto DNA to affect regulation
- *Promoter is position dependent (MUST be directly upstream of gene)

Regulatory sequences found in the promoter proximal region

TATA binding protein is a transcription factor.



Enhancers:

- Very far away from promoter (usually upstream)
    - Has **activators** bound to it, such that when the enhancer folds over, the activators on the enhancer interacts with a coactivator that interacts with activators/general transcription factors found on the promoter proximal region.
- This entire complex makes polymerase most attractive to the promoter - this leads transcription occurring at its maximal rate
- *Enhancer is position independent.

Protein motifs common in DNA binding proteins

- Proteins bind to DNA through electrostatic attraction - negatively charged DNA backbone with positively charged areas in protein - but this only works when proteins have a particular shape that fits into the helix - **motifs -** specific shapes that fit into the helix - there are 3 motifs :

A highly specialized region in a protein produced by the 3-D arrangement of amino acid chains within and between domains

1) **Helix-turn-helix -** alpha helix binds to base pairs in major groove of the DNA. A looped region of the protein (the turn) connects to a second alpha helix that helps hold the first helix in place
2) **Zinc finger -** series of amino acids that are able to associate with zinc cofactors, forming a shape that can bind to specific base pairs in the grooves of DNA.
3) **Leucine Zipper region -** is a dimer - hydrophobic interactions between leucine residues hold the monomers of the dimer together - Other alpha helices bind to DNA base pairs in the major grooves.
    - These 3 motifs are found in the DNA-binding domains of regulatory proteins -- binding occurs on/around the promoter and at the promoter proximal region.

Mechanism of transcription termination in eukaryotes

- During transcription, polymerase transcribes right through the **polyadenylation signal** (and well past the end of the gene) - turning it into mRNA
- This signal is in DNA, but understood in mRNA by an RNase - an enzyme that binds to the polyadenylation signal and cleaves the transcript just downstream of it.
- Once the signal is cleaved, polymerase 'realizes' its gone too far and so it stops.

Which gene expression components cross the nuclear membrane to get from where they are made to where they function

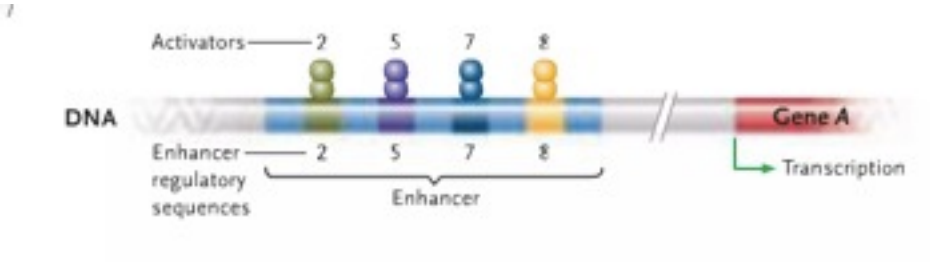
Gene expression component	Where they're made	Where they function
Ribosomal RNA	Nucleus (by transcription)	Cytoplasm
tRNA	Nucleus (by transcription)	Cytoplasm
snRNA	Nucleus (by transcription)	Nucleus (doesn't cross n.m)
Poly a polymerase (enzyme)	In the cytoplasm (by translation)	Nucleus

Various stages of gene expression subject to regulation

- Transcriptional regulation (chromatin remodelling, promoter affinity, transcriptional factors being added), enhancer activity, different activators binding to regulatory sequences
- Post transcriptional regulation (creation of pre-miRNA)
- Translational regulation
- Post-translational regulation (proteosome activity)

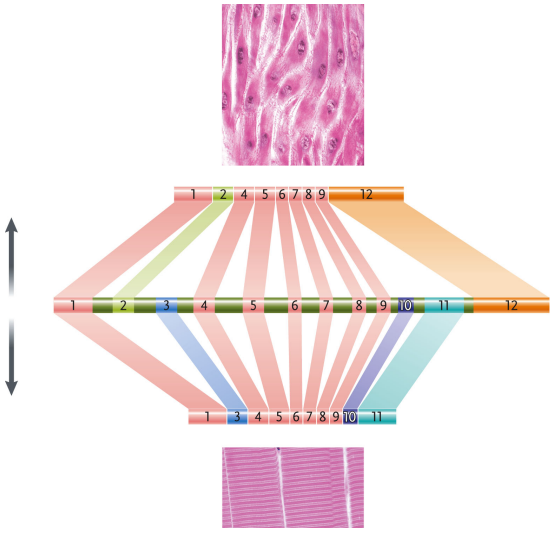
How organisms express different genes in each different tissue

- Different genes in different tissue are expressed differently by producing tissue specific activator proteins.
- Certain activators (or repressors) bind to different enhancer regulatory sequences in the enhancer and thus transcribe a different set of genes.
- Some activator proteins would be expressed only in the lens of your eye and therefore would only enhance certain sequences leading to certain genes that you would want expressed in the lens of your eye.
- In the liver, there would be a different set of activators that would be binding to a different enhancer and transcribing a different set of genes.



- For example gene A is controlled by activators 2,5,7, and 8 binding to regulatory sequences in its enhancer.
- Genes with multiple regulatory elements are under complex regulatory control (other genes may have only one regulatory element)

The advantages to alternative splicing



• In some tissue, “3” (blue) is expressed - it is an exon
•In other tissue (top) , it is removed - it is an intron

• Alternative splicing great;y increases the number and variety of proteins encoded in the cell nucleus without increasing the size of the genome.

• Allows you to generate many proteins from one single message - one gene can give rise to many different proteins.
• Alternative splicing is one of the reasons eukaryotic cells can be very complex without having excessive genes.

Mechanism of action of miRNA

• Micro RNA is coded for by a gene, gets transcribed, base pairs with itself and forms a stem-loop structure - this is precursor miRNA with two sides of the hairpin base-paired together, and a loop of unpaired bases.
• Pre-miRNA is exported to cytoplasm
• Encounters dicer enzyme that removes loop and leaves a double stranded RNA
• A protein complex then binds to the double stranded RNA
• An enzyme in that protein degrades one of the RNA strands, leaving the miRNA (single stranded) (dicer and protein complex attached now)
• miRNA binds to a target mRNA that has a complementary (or close to) base sequence in their 3‘UTRs - this ultimately blocks translation. (protein complex attached to miRNA gets in the way of translation.

Dicer is still attached at the head

- This is translational control and different cells (liver vs eye) have different miRNAs that shut off different genes.
- Cancer cells also have different miRNA than normal cells.

*This is an example of postranscriptional regulation that impacts translation

Mechanism of targeting proteins to cellular organelles

- Proteins are transported to their appropriate organelles and the cells knows where to transport them because they have a tag - a targeting tag.
- The tag is a peptide tag (amino acids) that tells the cells to put/move the protein to its appropriate location.
- Tags specific to chlorplast, nucleus , etc
- Some proteins are needed for the endomembrane system or outside of the cell - these proteins are targeted to the E.R (by a tag) and then sent to the appropriate location. (ER -> Golgi -> destination)
- This tag is coded for in the DNA and is found in the coding region of a gene. The tag is transcribed and translated into the appropriate amino acids that will function as a tag.

Mechanisms to regulate protein function after they are made

Mechanisms of ubiquitin/proteosome protein degradation

- Ubiquitin is added to the protein (tagged)
- The proteosome recognizes the ubiquitin-tagged protein , unfolds it and then digests the protein into small peptides
- Released peptides can be recycled and reused in protein synthesis
  - You can regulate the amount of protein around by regulating how quickly it is degraded through the proteosome.

Capping and Tailing and Intron removal

- 5' Cap is added soon after transcription begins : G-P-P-P(5'UTR)

- Once the polyadenylation signal is cut - the poly A tail is added
  A's are added by an enzyme called poly A polymerase (NO BASE PAIRING OCCURS) : (3'UTR)AAAAAAAAAAA...3'

- mRNA with introns attract the attention of snRNPs (small nuclear ribonucleoproteins - complex of protein and RNA)
- the RNA in snRNPs is called snRNA and it base pairs with itself
- snRNPs that bind have snRNAs that base pair with RNA sequences at the intron-exon junction - a bunch of snRNPs associating with the intron make a **spliceosome**
- The signals at each end of the intron (the intron-exon junctions) are found in the DNA and are understood as RNA by the snRNPs as a signal for cutting out the intron.
- Specifically, these signals on the ends of the intron are understood by complimentary base pairing with the snRNAs
- Spliceosome loops out the intron, bringing the exons very close , spliceosome then cleaves the pre-mRNA so that all that is left are the exons and they join together.
- Looped out intron that is cut out on both sides loops and bonds with itself -- gets degraded.

RNA is catalytic, protein is structural

All RNA's base pair with themselves.

Note:

- Synthesis and functioning of a typical tRNA requires complimentary base pairing with itself, DNA, and other RNA (mRNA)
- Lac I,Z,Y,A all have their own stop codons, they are individual genes.
- Bacteria only has one RNA polymerase. Eukaryotes have many.

Lecture 15 - Identical Twins are Not (ISO)

Identify various mechanisms for regulation of gene expression as summarized in Figure 14.6

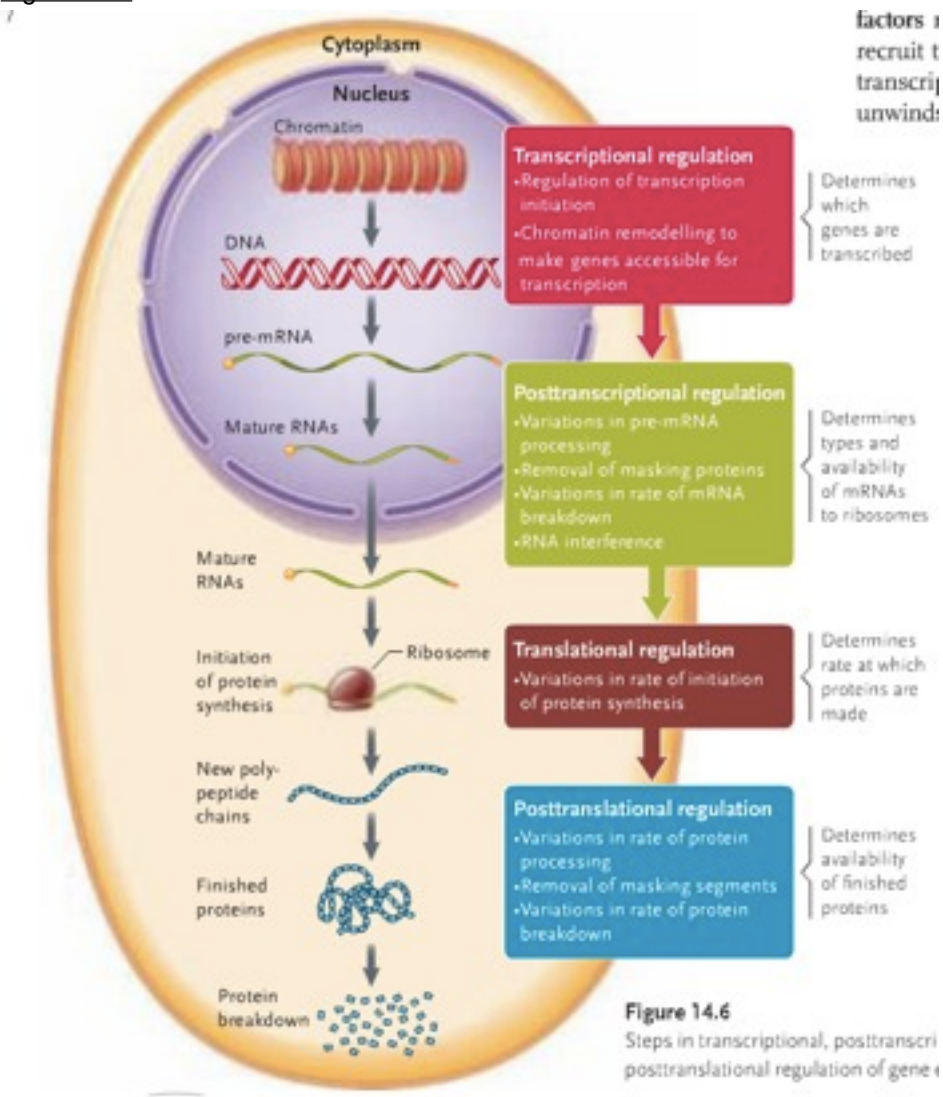


Figure 14.6
Steps in transcriptional, posttranscri
posttranslational regulation of gene

Identify characteristics of genetic mosaics

- Half of a females cells have the paternal X active and half of the cells have the parental X chromosome active. (No noticeable different usually in humans)
- In cats, different X's will be active and so different alleles will code for different colors of fur and this will result in patches of colors.

Identify role of histones in DNA packaging and expression

- **Histones -** small positively charges proteins that are complexed with DNA in the chromosomes of eukaryotes.
- Link to DNA by attraction with the negatively charged phosphate groups of DNA
- Histones pack DNA molecules into the narrow confines of the cell nucleus.

- DNA that is tightly packed is going to be hard to express - this is because their promoters are less accessible to the proteins that initiate transcription.
- DNA that is more open/loose is more available for transcription - easier for proteins that initiate transcription to bind.

Lecture 15 - Identical Twins Are Not

Reasons why, if identical twin women have sons with identical twin men, the sons will not be identical

- The kids will not be identical because each of these people got DNA from there mom and there dad, and the DNA they got from their mom and dad was different (mom and dads DNA is different but same between DNA b/w the two sons)
- So, during meiosis their going to combine their mom and dads DNA to make gametes that will be different.

Definition or explanation of “heritability” with respect to human disease risk

- Percentage of variation due to genetic variability
- Over 50% heritability for a human disease means that 50% of all the different expressions of that disease (schizophrenia) are due to genetic variability.
- Basically, the proportion of observable differences in a trait (disease) between individuals within a population that is due to genetic variability.

Characteristics that would, and would not be different between monozygotic ("identical") twins.

Similarities	Differences
• Same genes • Same alleles	• Gene expression (same genes, but different genes expressed) • Make different gametes • Environmental mutagens they are exposed too • Random mutations that occur during replication (mitosis) • X inactivation (females only) • They have inactivated different X's in different tissues • Random X inactivation by non coding RNA's (one twin could have only paternal X chromosomes expressed in her blood)

Process of random X inactivation leading to genetic mosaicism

- Both X chromosomes are active until about the 100 cell stage during early embryogenesis
- At this stage each cell will RANDOMLY inactive one of its X chromosomes (maternal or paternal) - There will be close to a 50/50 split between paternal and maternal X's that are activate.
- From the point of inactivation - the descendants of these cells will keep the same X inactive.
- This leads to genetic mosaicism - two different genetic expressions/cell lines from the same zygote.

All females are mosaic

How a mosaic is different from a heterozygote

- Being heterozygous at a gene locus means you have two different alleles of a gene. One allele is usually dominant and is thus expressed

- In mosaic, the allele on the active X chromosome will be expressed, regardless of whether it is dominate or recessive. So, if in one cell the paternal x chromosome is active - the fathers allele is expressed but if in another cell, the maternal X chromosome is active - then the mothers allele will be expressed.

Role of Xist RNA in X inactivation

- Xist is transcribed (but not translated) only in the inactive X chromosome
- Xist transcript (non coding RNA) coats the chromosome and keeps the DNA away from transcription factories - the coated X chromosome becomes so tightly condensed and essentially 'turns off' the chromosome

RNA polymerase and associated proteins

Role of Tsix RNA in regulation of Xist RNA expression

- The expression(transcription) of Tsix - (just the RNA) shuts off the expression of Xist
- The message of Tsix is antisense to Xist and they are antagonistic.
- Tsix expressed in active X chromosomes

Meaning of "antisense"

- A gene that when transcribed inhibits the expression of its counter gene (Xist vs Tsix)

Structure of nucleosomes

- **Nucleosome:** two molecules of each 4 different types of histones combine to form a beadlike, eight-protein nucleosome core particle around which the DNA winds around twice.
- Packing DNA acts as regulation - you can regulate gene expression by packing DNA more tightly or more loosely - this is because tightly packed DNA is harder to express than loosely packed/open DNA.

This is because their promoters are less accessible to the proteins that initiate transcription.

Function of nucleosomes in chromatic structure

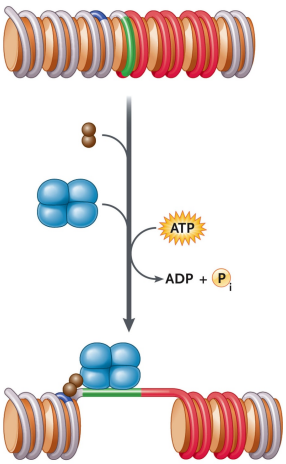
- DNA is wrapped around the nucleosome twice
- Compacts DNA by a factor of 7

1st type of chromatin remodelling

- Activator are the two small circles

Role of nucleosomes in chromatin remodeling

Chromatin remodeling - process in which the state of the chromatin is changes so that the proteins that initiate transcription can bind to their promotors.



- Activators bind to a regulatory sequence upstream of the genes promoter and recruits a **nucleosome remodelling complex** (blue clover)
- Using energy from ATP, the multi-protein complex to slide the nucleosome along the DNA to expose the promoter
- Thus, you can pack the nucleosomes tighter to shut genes off or move them around to open genes up. -- *Chromatin remodelling is an epigenetic change

Relationship between acetylation/deacetylation of histone tails on chromatin structure or protein binding (second type of chromatin remodelling)

- Nucleosomes have tails (each histone has a tail) so 8 tails in a nucleosome
- Tails can be chemically modified - you methalate them, phosphorylate them, or acetalate them.
- When you acetylate the lysine amino acids in the tails of histones in the nucleosome, you remove the positively charged amino acid group of lysine. As a result, the histones are less attractive to the negatively charged DNA.
- The histones then loosen their association with DNA and the promoter becomes accessible.
- Acetylation of histones tends to increase gene expression
- DEacetylation tends to shut down gene expression
- Thus, expression can be regulated by regulating acetylation of histones.

Relationship between methylation/demethylation of cytosine on chromatin structure or protein binding

- By adding methyl groups (CH3) to cytosine bases in the DNA, you can shut off genes.
- DNA methylation prevents the binding of transcription factors and as a result, shut the gene off
- DNA methylation might also attract deacetylase enzymes that would then compact the chromatin, which would also result in the gene shutting off.
    - If twins have different DNA/genes methylated as they age, then they will have different gene expression.
- Demethylation would allow for the gene to be expressed again

Factors that may influence epigenetic methylation patterns in DNA

- Diet - food you eat could affect methylation
- Chemical insults in the environment
- Maternal Care
- Smoking/exposure to harmful substances
- Childhood trauma/abuse - more methylation (genes not expressed)
- Social environment
- Heritability (inherited through genes)

Difference between “genetic” vs “epigenetic” changes that affect gene expression

1) Epigenetic changes affect gene expression by mechanisms that do not involve changing the DNA sequence (cannot be due to a mutation)
2) Epigenetic changes MUST be persistent (lac repressor constantly binding and falling off is not persistent --> not epigenetic)
    - Epigenetic mechanisms often involve shutting down/inhibiting genes,

- An example of epigenetics affecting gene expression:
- X chromosomes become inactive
- Chromosome remodelling.
- DNA methylation due to factors discussed above

Epigenetic changes can be stable over generations (heritable)

--> Genetic changes that affect gene expression are primarily due to mutations and CHANGE the gene sequence

Note

- Operator, polyadenylation signal and enhancers are understood by the cell through protein binding to nucleic acid.
- Packing of DNA is a kind of regulation.

Lecture 16 - Evolution of Multicellularity

Identify the function of caspases

- Intracellular proteases (proteins) that destroy the cell when trigger for programmed cell death is activated
- “Executioner” caspases are proteases that cleave essential proteins, leading to a controlled, but irreversible, biochemical cascade causing cell shrinkage, chromatin fragmentation and cell death.

Likelihood that modern multi-cellular life forms are monophyletic.

- Most multicellular life forms are NOT monophyletic
- Multicellularity has evolved MULTIPLE times
- Some taxonomic eukaryotic groups are all multicellular, some have some multicellular and some unicellular, and others have primarily multicellular
- Volvox genus is not monophyletic (evolved communal lifestyle multiple times independently)

Characteristics of Volvocine algae that make them a useful model system for studying the transition to multi-cellularity

- The divergence of volvox and chlamy is relatively recent - they last shared a common ancestor 50 mya - Would be easier to learn about this common ancestor than it would of the CA of plants and land animals (1200 mya)
- Volvocine algae are very accessible and can be grown in the lab
- Genetics of volvocine can be studied in depth
- Volvocine consists of somatic cells and reproductive cells (differentiated)

Relative structure/function of Chlamydomonas vs Volvox cells

- Volvox is a transparent sphere of extracellular matrix with thousands of somatic cells at the surface and much larger gonidial cells inside the sphere.
- Chlamy is unicellular and cells are undifferentiated

General process of Volvox asexual reproduction; role of somatic vs gonidial cells

- Gonidial cell divides until there are about 16 of them (all same size at this point)

- Then asymmetric division begins to occur that gives rise to small cells and larger cells.
- Large cells become gonidial (reproductive) and the small cells become somatic. As more cell division occurs, you end up with tons of somatic cells and large gonidial cells on the outside -- problem is that the gonidial cells are on the outside and the flagella are facing inwards, so the volvox flips all its cells thereby inverting the entire structure (reproductive cells are now inside and flagella is point outwards) - eyespots are looking out and can respond to light.
- Gonidial cells on the inside eventually 'pop out' and reproduce on their own to make a new organism/colony
- Somatic cell on the outside are left behind and eventually die.

<u>Genetic approaches to identifying genes relevant in rise/maintenance of multi-cellularity</u>

- Maybe you can use restriction endonucleases to cut out certain genes and then see what effect it has on Volvox.
- Mutate genes and once you have mutant baby volvox's , you can see which one's cannot invert or colonize and then see which gene was affected.
- Use southern blot to compare genomes of chlamy and volvox - see which genes are different and possible effect they can have
- Study gene expression in somatic vs reproductive cells.

- Use mutagensis : mutate genes and find volvox's that cannot do various 'steps'
- gls genes cause asymmetric division, so if volvox is mutated and all the cells end up being the same size, we know the gls gene is off
- If asymmetric division occurs, then the bigger cells turn on the lag gene and this represses somatic genes (flagella, eyespot, etc)
- In small cells If regA gene is on, then gonidial genes are repressed and they function as somatic cells instead.

Types of data or insight revealed by comparative genomic studies in Chlamydomonas vs Volvox

- Studies revealed that their genomes aren't much different.
- Volvox genome is only a bit bigger, amount of protein coding genes are almost identical.
- They have similar protein coding capacity, introns per gene, intron length... etc
- With the exception of ECM and cyclins (proteins that control cell cycling) , the developmental innovations in the volvox lineage did not involve major changes in the ancestral protein repertoire.
- Volvox has more genes regulating cell cycling and it has more genes to do with the ECM

Table 1. Comparison of the *Volvox* and *Chlamydomonas* genomes.

Species	Genome size (Mbp)	Number of chromosomes	% G and C	Protein-coding loci	% coding	% of genes with introns	Introns per gene	Median intron length (bp)
V. carteri	138	14*	56	14,520	18.0	92	7.05	358
C. reinhardtii	118	17	64	14,516	16.3	91	7.4	174

Types of data or insight revealed by mutation/rescue experiments in Volvox

- **gls gene** shifts the metaphase plate and produces cells of unequal size.
  - If you mutate volvox such that it cannot divide asymmetrically -- you can take a gene from chlamy and put it into the volvox mutant and RESCUE the mutant. (chlamy orthologue can rescue volvox mutants)
- You do not have to evolve a new gene in order to go from chlamy to volvox, you just have to evolve a new function for it. (repurpose it)

- **regA gene** in volvox functions in somatic cells by repressing transcription of nuclear chlorplast genes - by doing this, chlorplast is inefficient and does not make enough energy and so the cell does not have enough energy to reproduce.

- 50 mya in the common ancestor this gene probably caused cells to divide less in low light but this gene then duplicated and so the gene is found in chlamy now and does the same thing --> shuts/turns down photosynthesis in response to low light to stop unnecessary photosynthesis from taking place.
- In volvox this paralogous gene has the same function BUT a different regulation --> shuts/turns down chlorplast genes (photosynthesis) to inhibit reproduction
- Gene duplication 50 mya lead this gene to serve an environmental role in chlamy but a developmental role in volvox -- in chlamy the gene shuts down photosynthesis in response to low light but a related gene in volvox shuts down photosynthesis in response to what kind of cell you are. (regA is a chlamy paralogue)

- Inversionless (invA) - this gene codes for kinesin which is a microtubule motor - used in volvox to flip themselves.
- If there is a problem with kinesin (mutants), volvox cannot flip itself inside out.
- However, there is a gene in chlamy that when moved into the volvox, can rescue these mutants, it can provide the kinesin activity that has been lost in these mutants. (invA is an orthologue)

<u>Difference between orthologous and paralogous genes</u>

- Orthologue genes have been derived from a common ancestor and it is the SAME gene in both species but the gene has been repurposed and has a different effect in one of the species (in volvox for e.g) -- HOWEVER, if an orthologous gene is placed in the other species it will be able to take over for the function the mutant gene should be doing. ‘rescues mutants’
    - Orthologue genes can be taken from chlamy to rescue volvox mutants.

- Paralogous gene is a gene in the common ancestor and duplicates to form two genes in two different species with the same function but has a different REGULATION
- Gene is put under different regulation but essentially does the same job.

Note

- If you're going to participate in a multicellular organism, most cells have to give up reproducing
- Most cells have to give up getting their DNA into the next generation.
- *Maybe evolution of multicellularity isn't such a big deal - maybe we can just take genes we already have and repurpose(orthologue) them for different functions or put them under different regulation (paralogue) as we go.