

Week 7 & 8

The Nucleus: Structure, nuclear transport

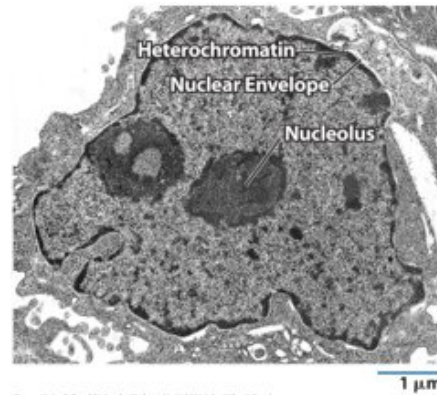


Figure 12-1a Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)

- Reading in Chapter 12 (Karp 9th Edition)
- 12.3 The Nucleus of a eukaryotic cell

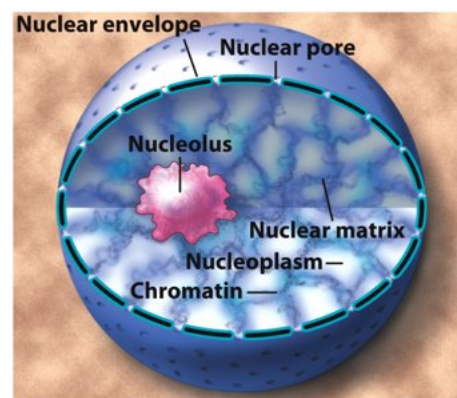
http://www.youtube.com/watch?v=NO_djGYepE0

Structural overview of the nucleus

The contents of the nucleus are enclosed by the *nuclear envelope*.

A typical non-dividing nucleus includes:

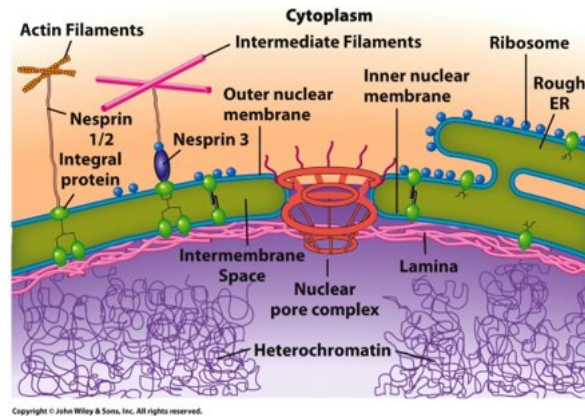
- DNA - chromosomes associating with (and organized by) proteins to form *chromatin*.
- *Nucleoli* for rRNA and ribosome synthesis.
- *Nucleoplasm* as the fluid where solutes are dissolved.
- The *nuclear matrix*, which is the protein-containing fibrillar network.



Major components of the nucleus

The nuclear envelope separates the nucleus from the cytoplasm

- It consists of two membranes separated by a nuclear space.
- Approx. 60 integral membrane proteins
- The two membranes are fused at sites forming a nuclear pore.
- The inner surface of the nuclear envelope is lined by the **nuclear lamina**, an intermediate filament made from lamin

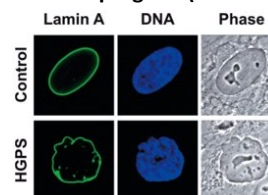


The nuclear lamina is essential for nucleus function

- Mutations in nuclear laminar genes sometimes lead to a number of diverse human diseases
- e.g. **Hutchinson-Gilford progeria syndrome (HGPS)**
- occurs from a single base pair mutation in Lamin A which causes an incorrect splice site on the mRNA leading to a truncated protein
- causes premature aging and death as a teenager from heart attack and stroke



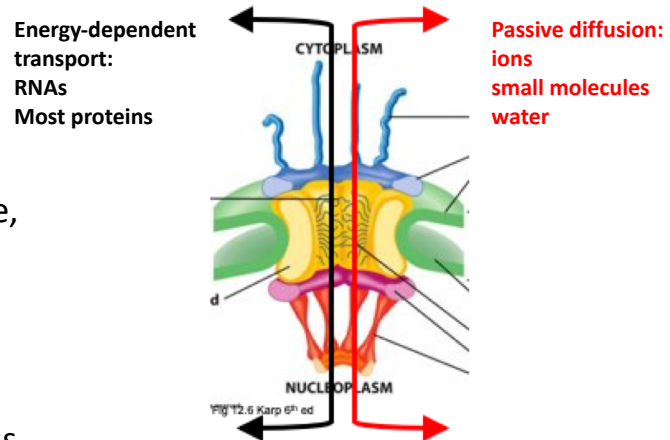
Leon Botha: South African hip hop artist had progeria (died at 26)



HGPS nuclei are mis-shapen due to truncated lamin A protein

Nuclear pores mediate transport across the nuclear envelope

- Regulate transport of RNA and proteins into and out of nucleus
- Traffic is FAST!
- A typical human cell has to import 560,000 ribosomal proteins/minute, export 14,000 assembled ribosomes/minute
- Thus, each nucleus has approx. 3000 nuclear pores
- Small molecules and proteins of less than 9 nm can pass freely
- Large proteins must be assisted



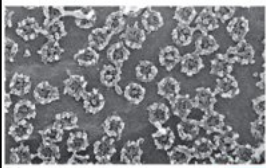
Nuclear pores regulate the movement of proteins and nucleic acids across the nuclear envelope

Cute video with link below. Incorrect link in handouts for in class video !!

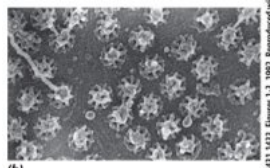
- <https://www.youtube.com/watch?v=UyhqLpjicZg>

Nuclear pore structure

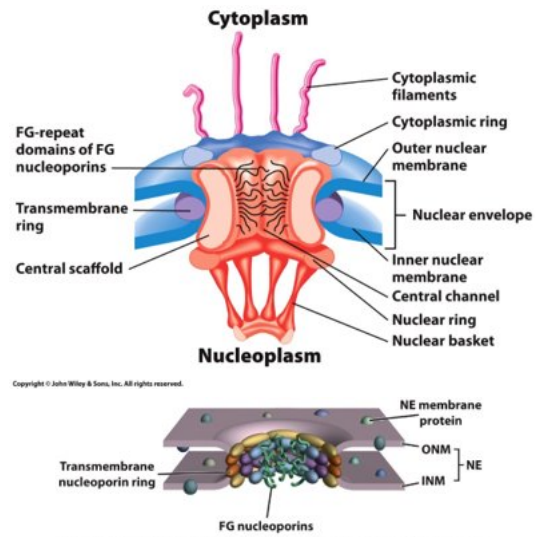
- Big! 15 to 30x times the size of a ribosome.
- Opening is lined with proteins to form the **nuclear pore complex**
 - at least 30 different **nucleoporin** proteins
 - in octagonal symmetry
 - cytoplasmic filaments extend into the cytoplasm
 - FG nucleoproteins extend into the pore space to form a barrier



(a) Cytoplasmic side of nuclear envelope showing octagonal ring symmetry



(b) Nuclear side of nuclear envelope showing the nuclear 'basket'



FG proteins are repeats of phenylalanine (F) and glycine (G)
 Very *disordered*, no secondary structure
 Thus wiggly snakes as seen in the video Block anything over 40 kDa in size

How is the import of large proteins regulated?

Proteins that are to be imported into the nucleus have a **Nuclear Localization Signal (NLS)** as part of their amino acid sequence

a series of positively charged amino acids in specific sequence:

- pro - lys - lys - lys - arg - lys - val -

⊕ ⊕ ⊕ ⊕ ⊕

 **NLS in the protein**

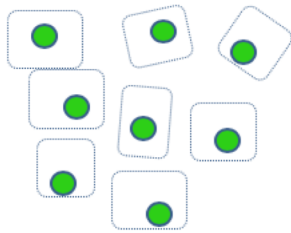
Proteins with an NLS stretch bind to a NLS receptor (*importin*).

Conformation of the nuclear pore complex changes as the protein passes through.

What happens when the NLS is altered?

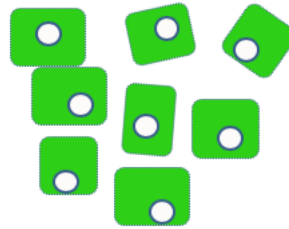
In 1984, Kalderson & colleagues (Cell 39:p499) mutated a single amino acid of a protein's NLS:

- Wild-type NLS:
- Pro – Pro – **Lys – Lys – Lys – Arg** – Lys – Val –



- The green-labeled protein is in nucleus
- (and not in cytoplasm)

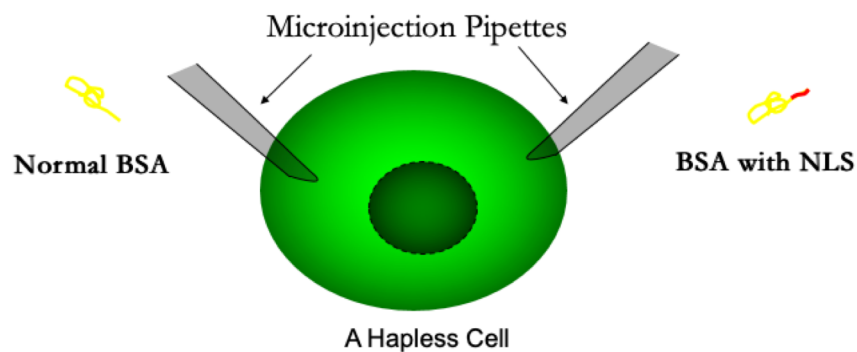
- Mutated NLS
- Pro – Pro – **Lys – Thr – Lys – Arg** – Lys – Val –



- The green-labeled protein is in cytoplasm
- (and not in nucleus)

Targeting non-nuclear proteins to the nucleus

- What happens when we use recombinant DNA techniques to add the **NLS** to a dummy protein?
- Normal or modified Bovine Serum Albumin (NLS added) and injected to the cytoplasm



Targeting non-nuclear proteins to the nucleus

This provided evidence of **nuclear transport receptors**

- family of proteins associated with the nuclear pore complex

This highlights two concepts in cell biology:

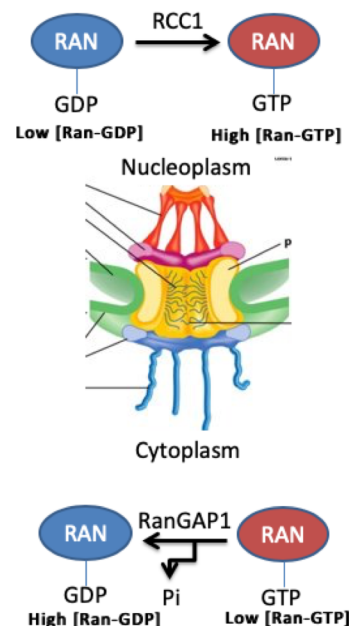
- **Necessary**: the NLS was shown to be necessary for transport, because if it was 'broken' the 1st protein wasn't imported
- **Sufficient**: adding the NLS was sufficient to transport the albumin into the nucleus

Info for next set of slides:

- **Importins** recognize the NLS and bring proteins in
- Another set of proteins, the **exportins**, work in the opposite direction
 - These recognize other signals called nuclear export signals

A Ran-GTP gradient is crucial in driving import and export through the nuclear pore

- The pore is a gate that allows small molecules to pass, but prevents large proteins from passing without recognition
- Ran is a GTP binding protein (like Sar1) that is active when GTP is bound and inactive when GDP is bound
- The cell keeps high levels of Ran-GTP *inside* the nucleus through the activity of RCC1, which exchanges GDP for GTP
- In the cytoplasm, RanGAP1 promotes the hydrolysis of Ran-GTP into Ran-GDP
- Thus, there's a **concentration gradient** of Ran-GTP across the nuclear envelope

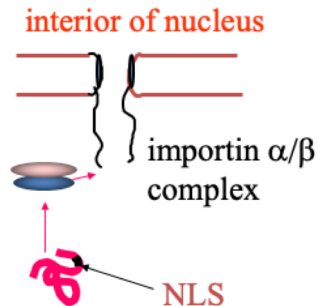


Mechanism of protein import through the nuclear pore complex

Step 1

Protein binds to a two-protein complex (importin α and importin β) |
 Importin α is a receptor for the NLS portion of the protein to be imported
 i.e. the NLS recognizes and sticks to importin α .

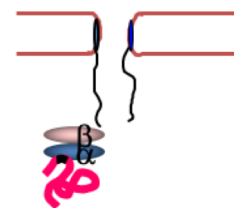
Step 1



Step 2

Importin complex and protein stick to cytoplasmic filament
 - mediated by importin β

Step 2



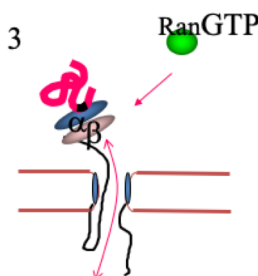
Mechanism of protein import through the nuclear pore complex

Step 3

Complex moves into nucleoplasm

Not an energy consuming step, it can go back at this point unless captured by the RanGTP in next step:

Step 3

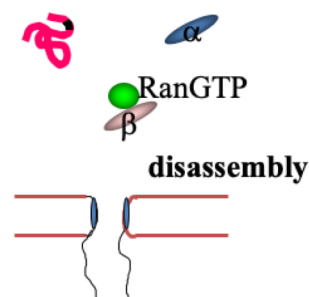


Step 4

Complex binds Ran-GTP; after binding Ran-GTP, complex dissociates

- importin β stays on the Ran-GTP

Step 4



Mechanism of protein import through the nuclear pore complex

Step 5

The Ran-GTP - importin β complex moves back to the cytoplasm, down the RanGTP concentration gradient

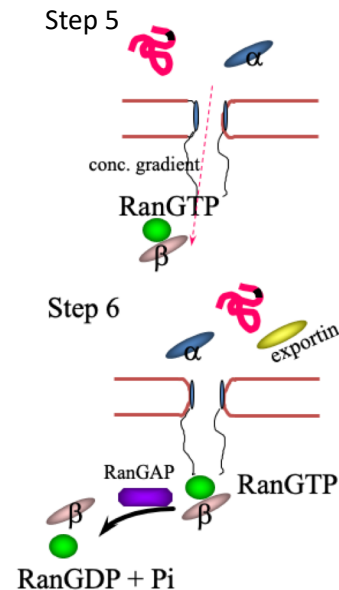
Step 6

Two things happen now:

the Ran-GTP is converted to Ran-GDP plus phosphate by the enzyme **RanGAP**. This causes Ran-GTP to release importin β

an **exportin** molecule binds to importin α , setting it up for transport out of the nucleus

Note that it's the hydrolysis of RanGTP to RanGDP that's the energy consuming step in this process!



Mechanism of protein import through the nuclear pore complex

Step 7

Ran-GDP diffuses back to the nucleus (1)
down its concentration gradient, i.e. from high to low concentration.

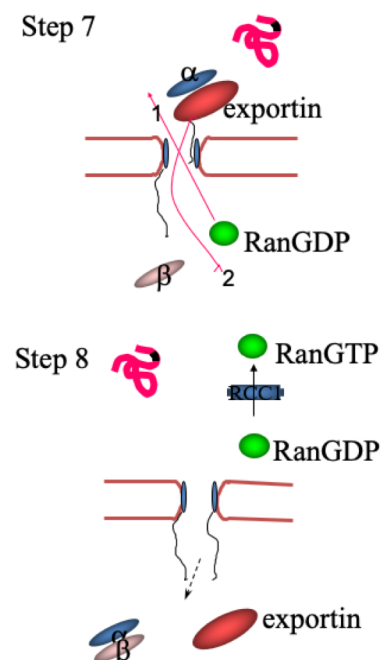
Exportin carries importin α out of the nucleus (2)

Step 8

Restoration to initial state

The importin α and importin β complex re-forms.

Enzyme **RCC1** re-forms Ran-GDP to Ran-GTP



pre-mRNAs vs hnRNAs

- **hnRNA** includes *all* RNAs in the nucleus: pre-mRNAs, pre-rRNAs, pre-tRNAs etc
- **pre-mRNA** is unprocessed mRNA (not yet spliced)

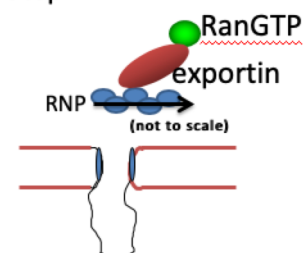
RNA export

RNAs move through the NPC as ribonucleoproteins (RNPs)

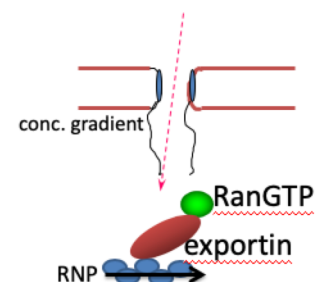
- Some of the RNP proteins have nuclear export signals that bind to exportins and RanGTP
- Once in the cytoplasm, GTP is hydrolyzed and the mRNA released and the associated proteins shuttled back into the nucleus.
- There's a functional link between splicing and mRNA export... only mRNAs with fully spliced introns are exported!

In class note: snRNA is carried through the nuclear pore *twice*.

Step 1



Step 2

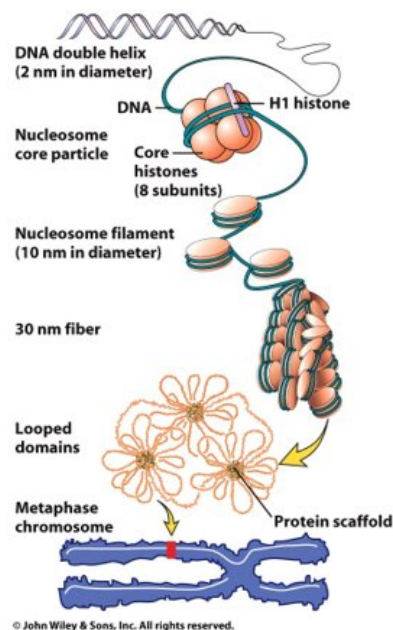


What you should have learned from this section

- Key traits about nuclei
- Nuclear pore structure, function (what they regulate and what they don't), abundance
- What is a NLS and how do we know how it works?
- Steps involved in importing a protein into nucleus
- Steps involved in exporting a mRNA

Week 07- DNA packaging and (some) epigenetics

Please read: Karp, 9th ed.
Chromosomes and
chromatin 12.4

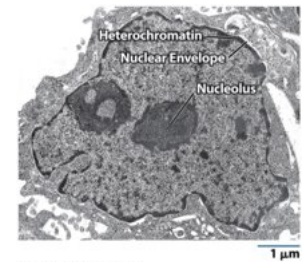


Chromosomes and chromatin

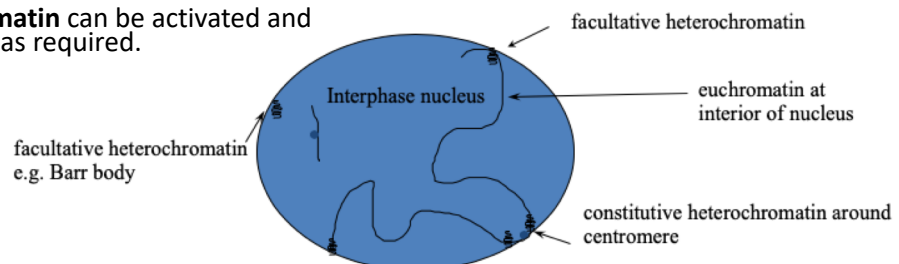
- 6.4 billion bps of DNA in the average human cell!
- 46 chromosomes in humans ($n = 23$), each a continuous strand of DNA
- Together our DNA is 2 m long ($1 \text{ bp} = 0.34 \text{ nm}$)!
- How does this fit into a nucleus that's $10 \mu\text{m}$ in diameter??? How do the separate chromosomes not get hopelessly tangled??
- **Answer:**
 - Chromosomes consist of **chromatin fibres**, composed of DNA and associated proteins.
 - Each chromosome contains a single, continuous DNA molecule that's packaged

Types of chromatin

- **Euchromatin** is in a dispersed state during interphase. Where genes are mostly found, transcriptionally active.
- **Heterochromatin** is condensed during interphase (10% of chromatin)
 - **Constitutive heterochromatin** remains condensed all the time. Consists of repeat DNA, especially in the telomeres and centromere (α satellite DNA).
 - **Facultative heterochromatin** can be activated and inactivated dynamically as required.



heterochromatin is located around the nucleus periphery; active genes, in the euchromatin, are located toward the interior



DNA packaging video

- <https://www.dnalc.org/resources/3d/07-how-dna-is-packaged-basic.html>

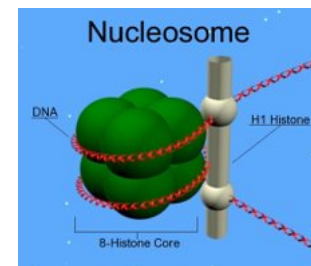
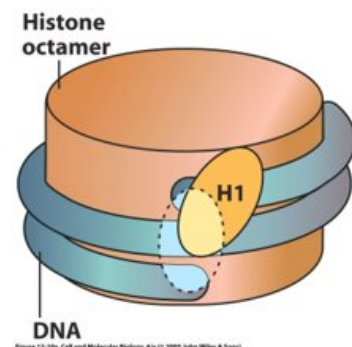
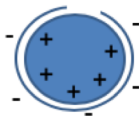
Note: some of the terminology is wrong in this video. The chromosomes are always present!

The condensed *mitotic* chromosomes uncondense and are no longer visible by light microscopy

DNA organization level 1: nucleosomes

DNA winds around **nucleosomes**, which are composed of a group of positively charged, highly conserved proteins – histones

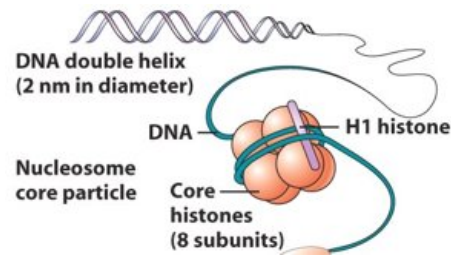
DNA is held by the histones due to non-covalent bonds, especially ionic bonds between negatively charged phosphate backbone of DNA and positive charge of histones (from lysine and arginine amino acids)



Nucleosome structure

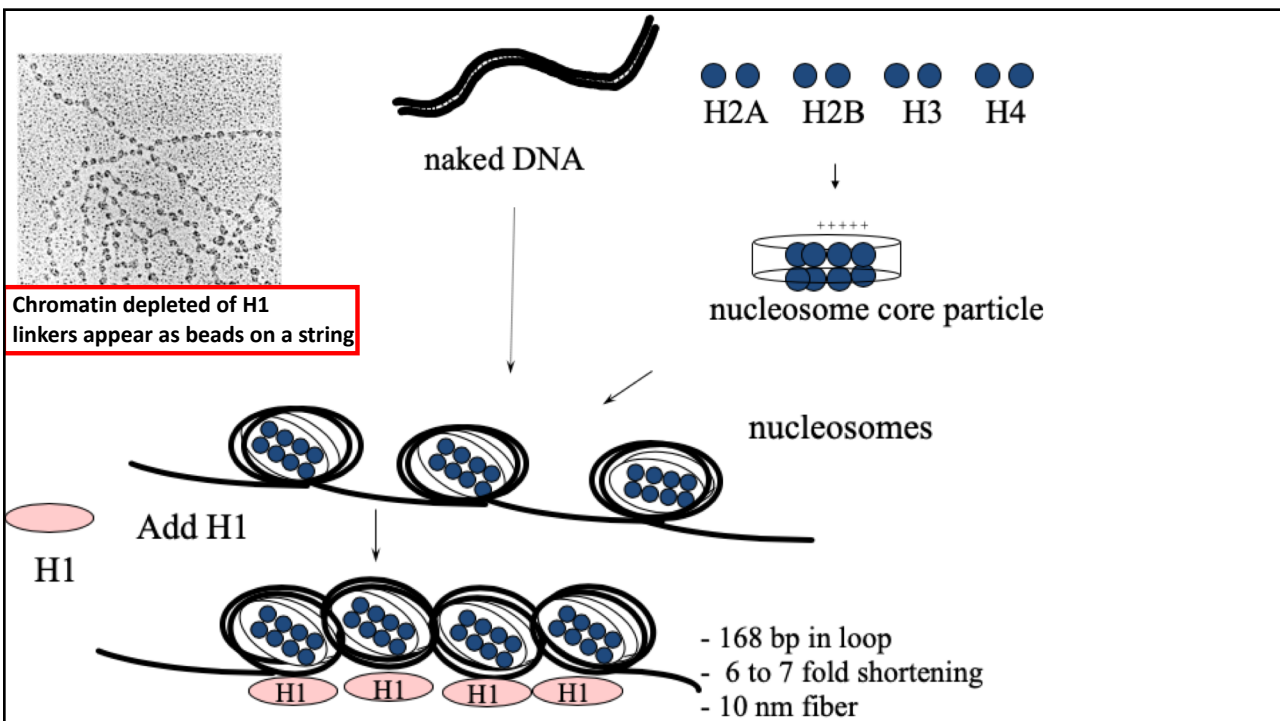
Made up of four types of histones proteins

- two of each:
 - Histones H2A, H2B, H3 and H4
 - DNA wraps around the **nucleosomes** 1.8 turns or 146 nucleotide bases per nucleosome
- Histone H1 (fifth type) – the linker histone
 - links adjacent nucleosome core particles.
 - A total of 168 nucleotide bases per unit
 - 7:1 packing ratio, the fiber is 10 nm thick



Histones are remarkably conserved: only 2 amino acid differences (out of 102) between pea and cow H4

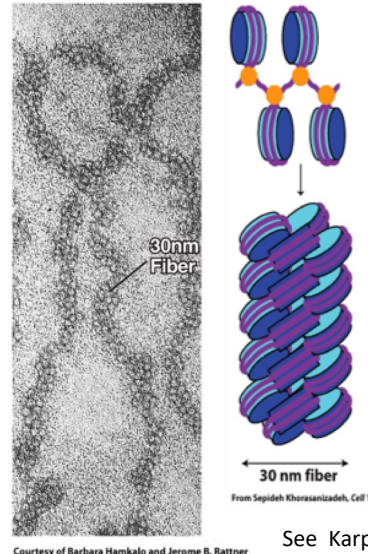
Why? Because DNA is DNA, and the function of histones is exactly the same in all eukaryotes



DNA organization level 2: 30 nm fibers

Spontaneous assembly of adjacent nucleosomes results in a 40:1 condensation in length, with 30 nm thick fibers

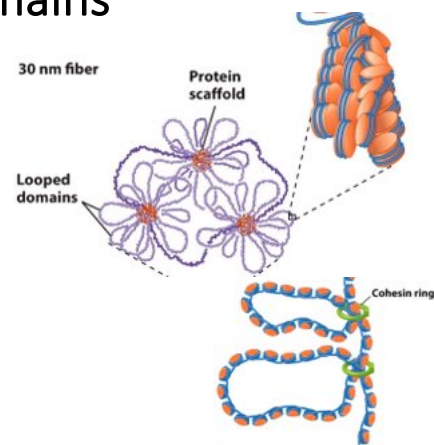
- H1 histone may bind adjacent nucleosomes
- N-terminal tails of H4 may extend far enough to reach other nucleosome histones



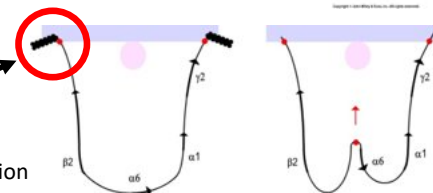
See Karp Fig 12.14

DNA organization level 3: looped domains

- The 30 nm fibers gather into supercoiled loops that are tethered to protein nuclear scaffold
- loops are 20,000 to 100,000 bases
- Cohesin – ring-shaped protein may maintain the loops
- DNA strands associate with nuclear matrix
 - AT-rich domains on DNA form **MARs (matrix associated regions)**
- Includes a satellite (centromere repetitive DNA) and other non-coding DNA



MARs – DNA sequence that has affinity for nuclear matrix; note that DNA loop can be moved to facilitate transcription



Scaffold proteins

Nuclear matrix

- a protein fiber framework
- major organizing structure for RNA polymerase (makes messenger RNA), RNA processing, DNA replication.

• Topoisomerase II

- an “un-tangling” protein (for unwinding DNA)

• Insulator Protein

- keeps loops separate

• Nuclear lamina (lamin IFs)

- lamins bind telomeres and a satellite DNA of the centromere

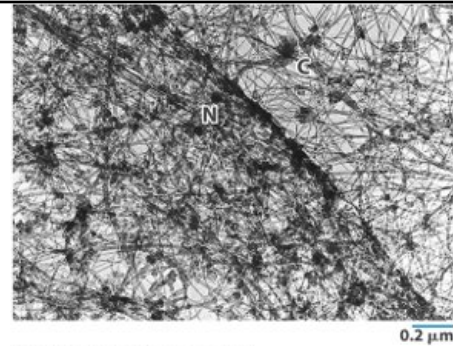
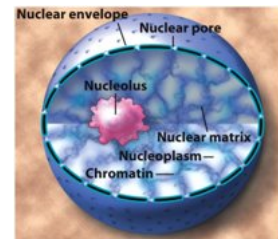


Figure 13-26a Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)

Nuclear scaffold can be observed by digesting away all DNA and removing all histones



DNA organization level 4: Mitotic chromosomes

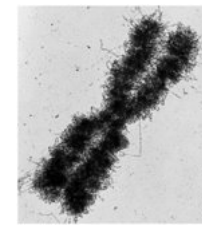
The ‘ultimate’ in packaging, only occurs during mitosis!! A further 10000 fold decrease in length.

Also have looped domain structure

- No nucleus is present when we see chromosomes in this state
- DNA loops on **Condensin** protein
- More compact than interphase chromatin
- Condensin is activated by phosphorylation via MPF

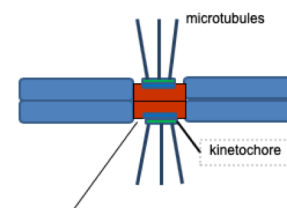
Centromere is a site of extreme condensation

- contains α-satellite DNA
- binds to proteins, including the kinetochore



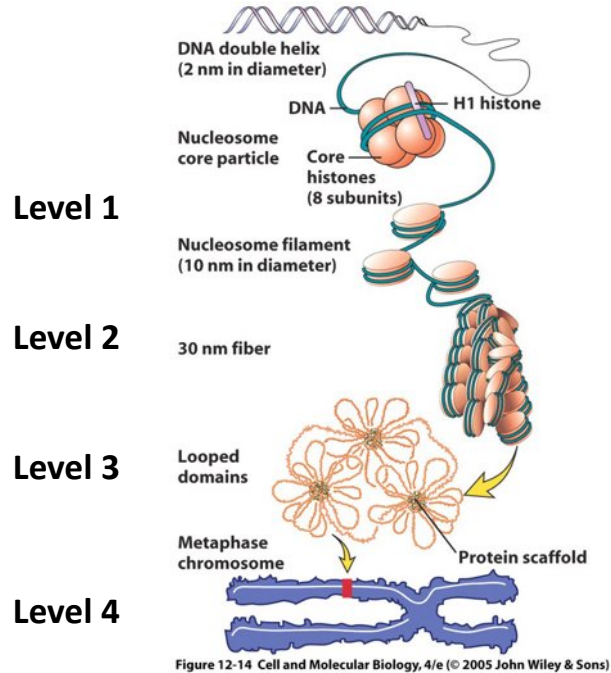
See Karp Fig 14.13

α-Satellite DNA:
comprised of 171 nucleotide repeats



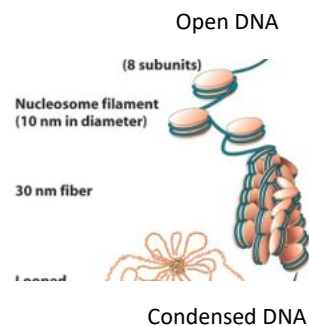
Centromeric DNA with α-satellite DNA

All together now!



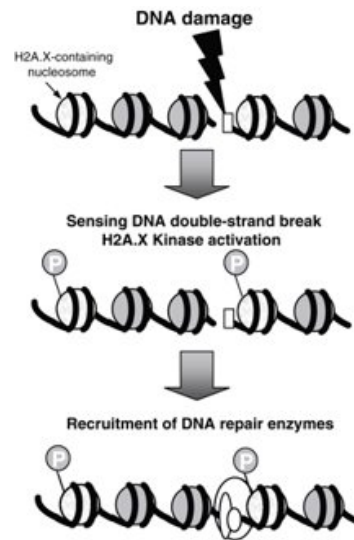
Histone tails have a 'code'

- DNA is frequently being transcribed and duplicated
- Necessary machinery has to work around nucleosomes
 - The cell makes DNA more or less accessible to other proteins by modifying histones
- Mechanisms of Histone modification
 - Replacement with modified types of histone
 - Chemical modification of the histones

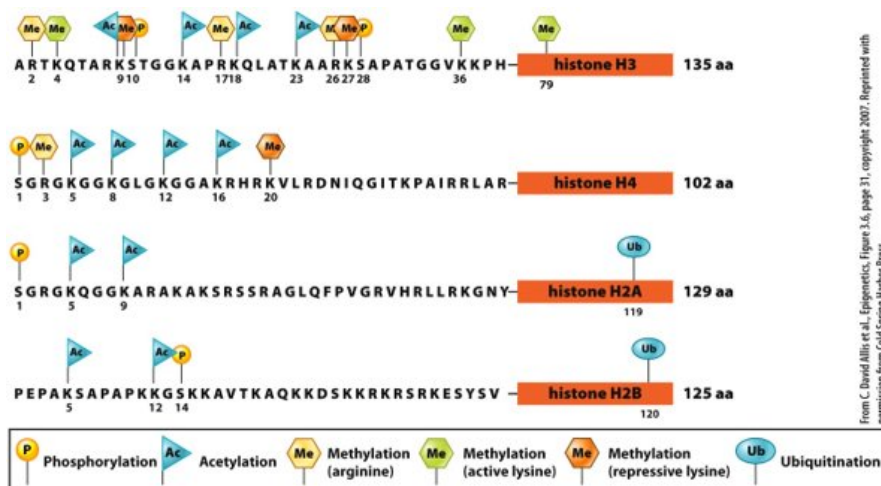


Histone protein variants

- **H2A.X**
 - DNA being repaired
- **H2AZ** and **H3.3**
 - DNA transcribed by RNA polymerase II
- **CNEP-A** (a version of H3)
 - on the centromere, kinetochore assembly
- **macroH2A**
 - Included in heterochromatin (eg Barr bodies... coming soon!)
 - keeps chromatin condensed and inactive



Histone N-termini can be chemically modified

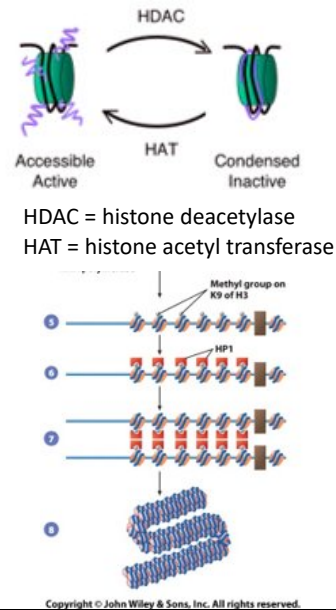


Don't memorize this!

Types of histone chemical modifications

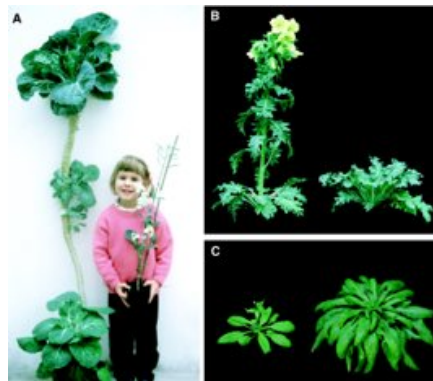
- Acetylated histones -looser DNA binding -> transcriptionally active
-
- Phosphorylated histones – can increase or decrease DNA winding on histones
-
- Methylated histones
- Alter the accessibility of DNA to processing enzymes
- Alter binding to organizational structures such as nuclear matrix
- often tighter DNA binding
- e.g. methylation of histone H3 leads to HP1 binding, which helps form heterochromatin

HP1 = heterochromatic protein 1



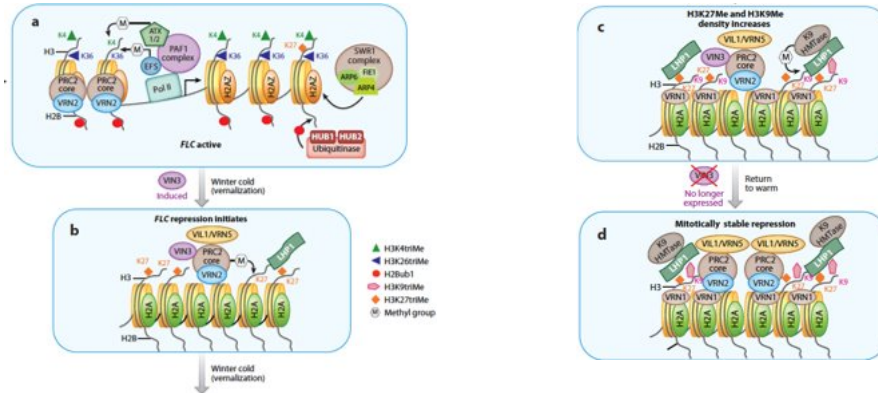
Flowering plants use various environmental cues to regulate flowering time and are linked to reproductive success

- Photoperiod – long days (LD) vs short days (SD)
- Defined by daylength
- Temperature cues - extended periods of certain temperatures
- Vernalization period of cold to prime for flowering in the spring
- Long day plants often require vernalization
- These two combined prevent flowering at the incorrect time
- Match with pollinators and other individual plants if outcrossing



Amasino 2004 Plant Cell

Vernalization results in epigenetic changes and happens over multiple seasons



Kim et. al. *Annu. Rev. Cell Dev. Biol.* 2009.25:277-299.

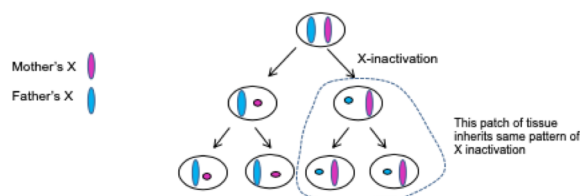
X chromosome inactivation

- In humans, gender is determined by the sex chromosomes:
 - females have two X chromosomes, males have one X and one Y chromosome
- The X and Y chromosomes look very different: almost all the genes on the X have no counterpart on the Y
- To ensure that males and females have equal levels of gene expression derived from the X chromosome, one X in every female cell is inactivated
 - the choice of which X is inactivated is apparently random

Female: XX

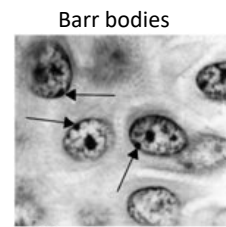
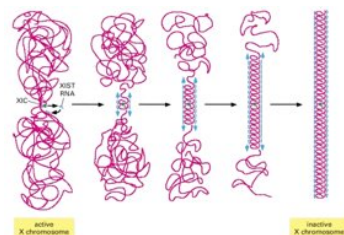


Male: XY



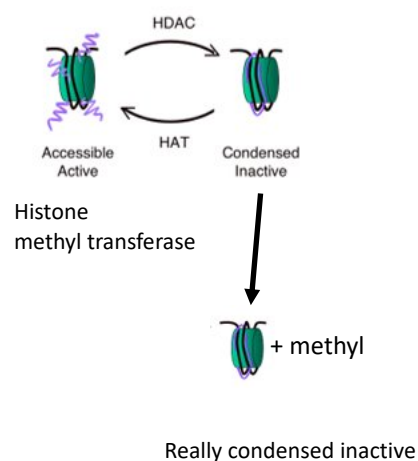
Mechanism of X inactivation

- Inactivation of an X chromosome requires a gene on that chromosome called ***XIST***.
 - *XIST* encodes a large non-coding RNA.
- *XIST* RNA accumulates along the X chromosome containing the active *XIST* gene and proceeds to inactivate almost all of the other genes on that chromosome.
- *XIST* RNA does not travel over to the other X chromosome in the nucleus.
- **Barr bodies** are inactive X chromosomes "painted" with *XIST* RNA.



How is the Barr body condensed??

- *XIST* RNA induces several later events (by recruiting various enzymes):
 - Deacetylation of H3
 - Methylation of H3
 - Ubiquitination of H2A
- H2A is replaced with macroH2A
- All these changes are undone in egg cells, so a random X chromosome can be passed onto the next generation



Fun fact: this is how calico cats work

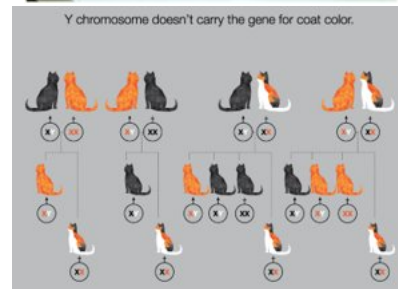
Random inactivation of X results in mosaic patterns of gene expression

Calico cats: always female (except in rare XXY males)

- X chromosome has gene for coat colour that encodes either black or orange pigment
- Since males have only one X, they can be either black or orange, but not both

The X chromosome inactivation in females is random, so different patches of epidermal cells can have different colour genes

(the gene for white coat is on another c'some)



<http://catjaa.com/wp-content/uploads/2015/07/Calico-Cat-genetic-diagram.png>

Thought question – how do you make a male calico cat?

- This majestic cat is Odd Sock, my grad student's cat.
- Odd Sock is male but has a calico paw.
- How could this have happened?
- Odd Sock is not XXY (see previous slide).

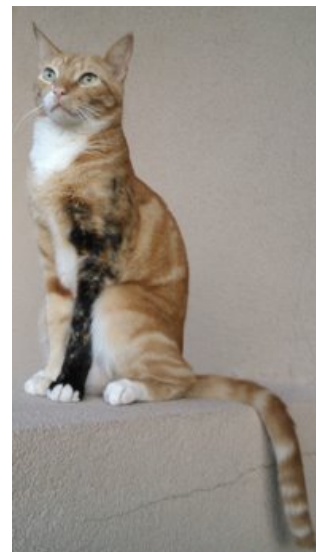


Photo: M. Balogh

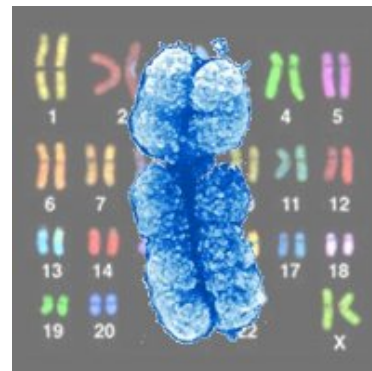
What you should know from this section

- Positioning of chromosomes in nucleus and relevance to gene expression
- Difference between heterochromatin and euchromatin
- Different histones and how they interact with each other and with DNA (including concept of nucleosome)
- Different levels of DNA condensation/packaging and what proteins mediate each level
- Some regions of DNA that interact directly with proteins (e.g. MARs, satellite DNA, kinetochore)
- Histone code - different histones, or chemical modifications of the standard histones – consequences of changing histones
- X chromosome inactivation process and consequences

Week 07 Stem cells and chromosomes

Read Chapter 12.4

- The structure of the mitotic chromosome
- Chromosomal aberrations
- Telomeres
- Centromeres
- Nucleus as an organized organelle



Stem cells and 'potency'

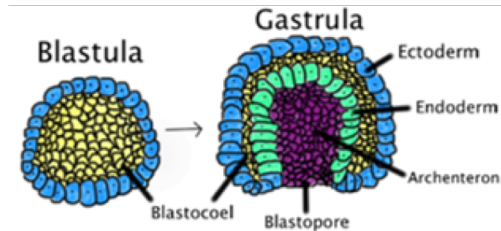


The initial stages of our life cycle consist of cells that can differentiate into *any* of our adult cell types. These cells are called **totipotent**. Most plant cells are totipotent throughout development!

At a certain early stage of embryo development, these cells differentiate into progenitor cell types that will ultimately give rise to the major three germ layers:

endoderm (interior stomach lining, gastrointestinal tract, the lungs),
mesoderm (muscle, bone, blood, urogenital)
ectoderm (epidermal tissues and nervous system)

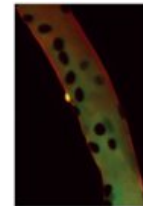
These are called **pluripotent** at this stage



Potency = a cell's ability to differentiate into other cell types

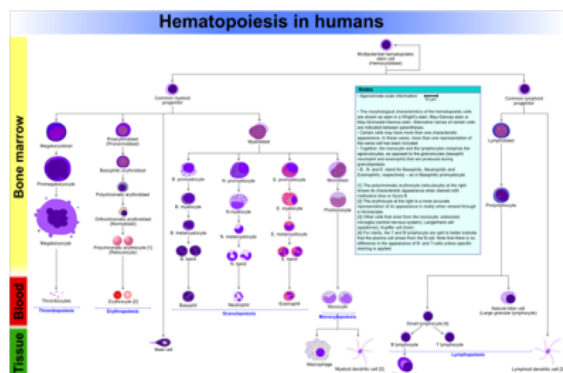
Adults have adult stem cells

- Different cell types have different longevity in the body:
 - Intestinal epithelial cells last days
 - Blood cells last months
 - Many neurons are made during development and last until death



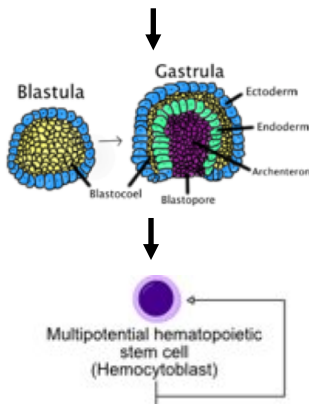
That wee yellow dot is a muscle stem cell

- **Stem cells** are defined as undifferentiated cells that are capable of:
 - self-renewal (production of more cells like themselves)
 - differentiation into 2 or more mature cell types.
- Most, if not all, of the organs in human adult contain stem cells capable of replacing the particular cells of the tissue in which they are found.



Because these stem cells can only differentiate into a few cell types, they're called **multipotent**

One factor in decreasing potency is the establishment of facultative heterochromatin



Totipotent cells have most of their DNA (except constitutive heterochromatin) in a euchromatin state

In the pluripotent cells, the DNA/genes involved in the pathway that the cells will become are kept as euchromatin, the DNA containing the genes in the other two pathways will be condensed into heterochromatin

As the adult stem cells develop, more genes will be switched into a heterochromatic state. The genes necessary for the 'downstream' potential cells are kept euchromatic.

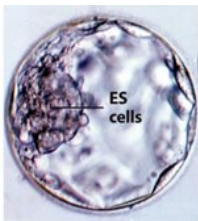
Stem cells may cure many diseases

Many diseases arise from dead cells which cannot be replaced:

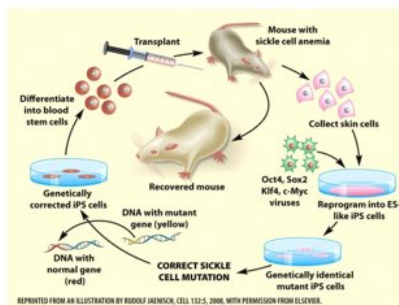
- Parkinson's from dead dopamine-generating neurons
- Type I diabetes from dead β pancreatic cells

The idea of stem cell therapy is take stem cells, get them to differentiate into the missing cell types (which we're already good at), and inject them into the body.

Possible sources of stem cells:



Embryonic stem cells have some major ethical concerns!!

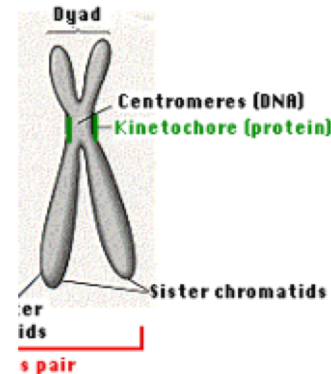


Induced pluripotent stem cells: take a differentiated cells and make them de-differentiate back to pluripotent cells, then induce them to form the cell type we wish

Chromosomes are the carriers of genetic information

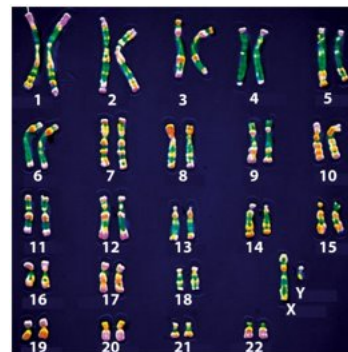
Chromosomes

- term coined in 1880s: “coloured bodies”
- during S phase (DNA replication phase), chromosomes duplicate – form **dyads**
- the attached duplicated chromosomes are called **sister chromatids**
- **ploidy** indicates the number of sets of chromosomes
 - Most animals are **diploid** – have two copies of each chromosome, one from the mother & one from father
 - Gametes are **haploid** – have only one set – arise through meiosis
- Plants alternate between diploid and haploid life phases – alternation of generations. Sporophytes are **diploid** and make spores. Spores become gametophytes that are **haploid** and make gametes.



Diploid vs haploid

- Diploid number ($2n$) – number of chromosomes per somatic (body) cell
- Haploid number (n) is number of chromosomes in gametes
 - Humans: $2n = 46$
 - Mouse: $2n = 40$
 - Drosophila $2n = 8$
- The complete set of chromosomes in the cells of an organism is its **karyotype**
- Human $2n$: 22 pairs of autosomes
- 2 sex chromosomes
- Autosome = not a sex chromosome



Karyotypes can be observed by isolating mitotic cells and staining the chromosomes

Aneuploidy: abnormal number of chromosomes

e.g. trisomy = one extra c'some; monosomy = missing chromosome

Some stats:

In 10,000 conceptions, 800 chromosomally abnormal → 750 spont. abort

> 140 are 45, X. They lack an X or a Y chromosome → 139 abort

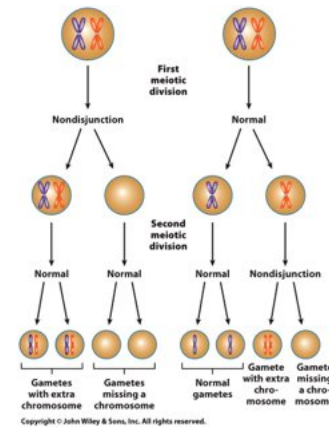
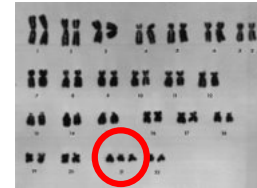
> 110 have an extra chromosome 16 → all abort

> 20 have an extra chromosome 18 → 19 abort

> 40 have an extra chromosome 21 → 35 abort, rest have Down's

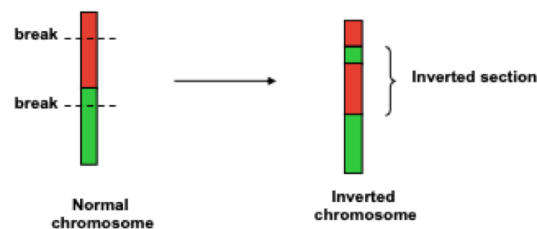
The rest have various different chromosomal abnormalities

This occurs through **nondisjunction** during meiosis, where one of the gametes receives an extra chromosome



Chromosomes are not indestructible

- Chromosomes can be broken by DNA damage, and when repaired, they may get scrambled:
- Translocations – movement to new chromosome
- Inversions – portion of chromosome is “flipped”
- Deletions – section of DNA excised
- Insertions – section of DNA inserted



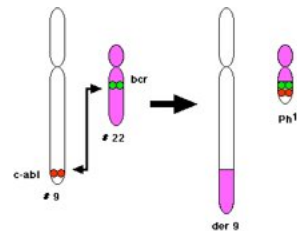
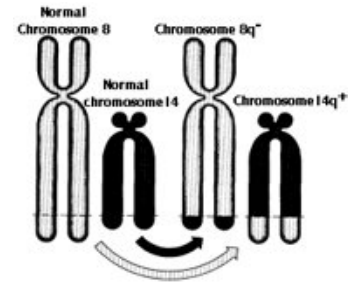
Translocations

Translocations are the transfer of a piece of one chromosome to a **nonhomologous chromosome**

- can result in changes in phenotype by:
 - Breaking a gene
 - Moving gene to a region where it can be controlled by another regulatory sequence
 - Creating hybrid gene

Some cancers are due to translocations (e.g. chronic myelogenous leukemia = CML)

- The *c-abl* gene encodes a kinase that regulates cell proliferation;
- the translocation alters the controls
- of this gene, placing it under the control of another gene's (*bcr*) promoter



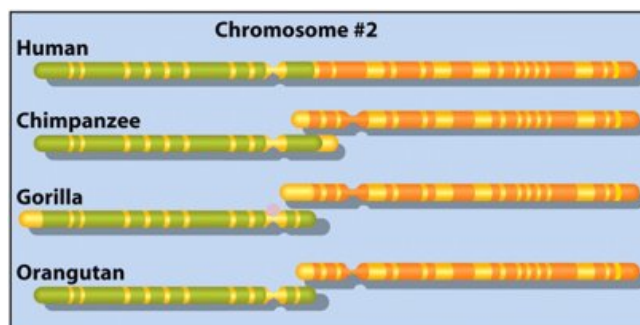
Reciprocal translocation between c'some 9 & 22, producing extra long c'some 9

A translocation distinguishes human and ape karyotypes

- Apes have 48 chromosomes, humans have 46
- The genes in humans and chimps differ by <1%
- Looking at human and chimp chromosomes, we share all, *except* the human chromosome 2 is a fusion of two ape chromosomes (but it is not the only difference)



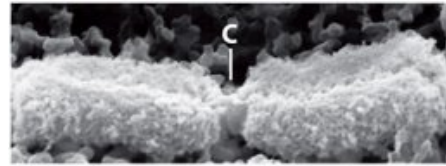
Microsoft Clip Art



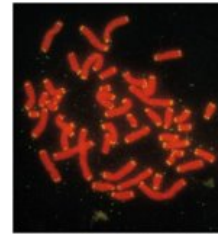
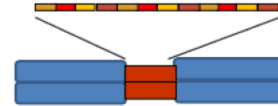
© John Wiley & Sons, Inc. All rights reserved.

• Centromeres

- the constricted (central) portion of each chromosome,
- the DNA contains **a-satellite DNA**; made of non-transcribed 171 base repetitive sequences.
- Repeated many thousands of times
- Attach to the kinetochores during M phase



α -satellite DNA:
comprised of 171 nucleotide repeats



Yellow dots are fluorescently-labeled probes that bind telomere sequences

• Telomeres

- non-coding regions at the ends of the chromosomes.
- short repeated sequence: repeated 500-5000X
- include specialized proteins
- form a "capped" end structure

- TTAGGGTTAGGGTTAGGG

Importance of Telomeres

In humans: ...TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG...



Sequence conserved in vertebrates, and similar in most organisms, so function must be important.

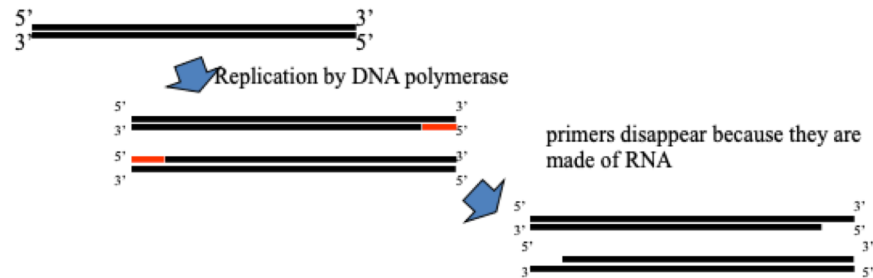
Telomere functions:

- protect the ends of chromosomes from nucleases
- allow cells to distinguish chromosome ends from broken DNA
- prevent chromosomes from fusing with each other
- attachment to nuclear scaffold

Problem – when chromosomes are replicated in most cells, the telomeres get progressively shorter

End-replication problem: the shortening of DNA molecules with each cell division

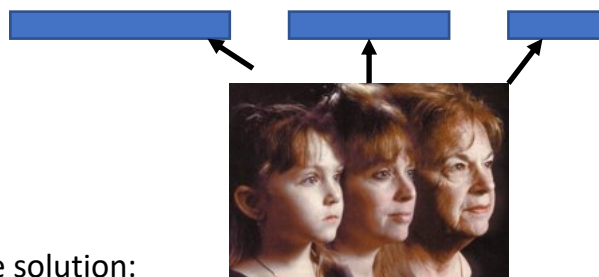
DNA polymerase - an enzyme that replicates DNA strands.



- DNA polymerase can't start on it's own... it needs to add onto an existing DNA or RNA
- In replication an RNA primer (in red) is used to get started
- DNA polymerase builds in a 5' to 3' direction to make a new DNA strand
- RNA primer is removed, leaving a small 3' overhang
- the overhang folds back to form a loop (safer from nucleases)
- Every time DNA is replicated it gets a bit shorter – eventually this would be a problem.

End replication problem - the DNA shortens over time

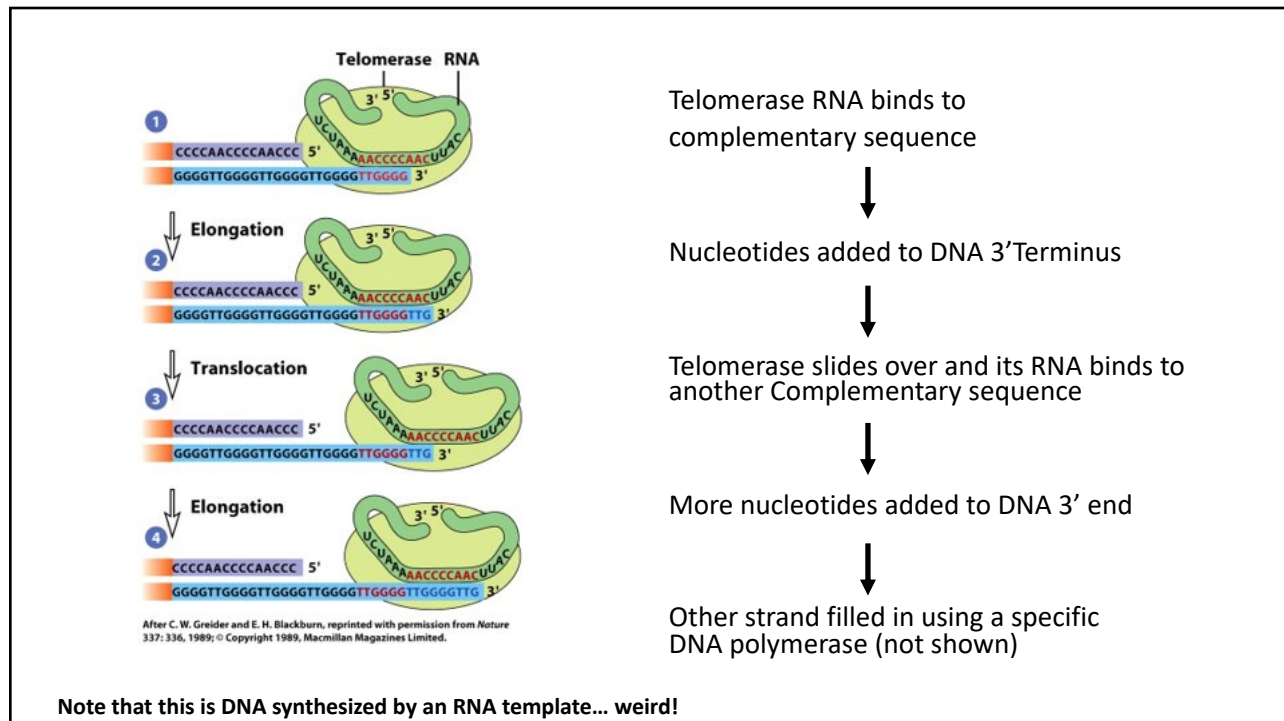
- each round of DNA replication leaves 50-200 bp DNA unreplicated at the 3' end
- cells with telomeres that are 10-12 kb in length (average) divide 50-60 X
- **cellular senescence** (cells will no longer divide and will die) is triggered when telomeres are about 4-6 kb long



Telomerase is the solution:

Telomerase: an RNA-containing enzyme that adds more

- nucleotides to the 3' end of the telomere DNA strands

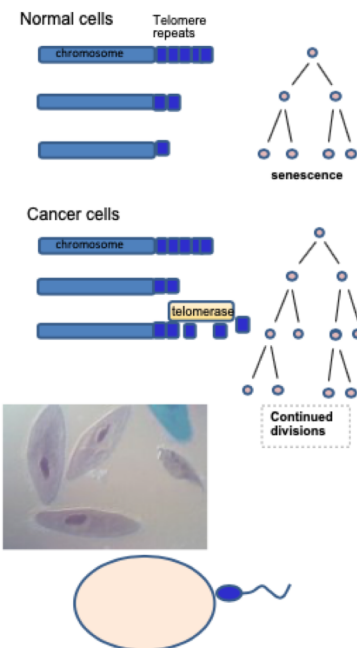


A. The steady loss of telomeres results in a limited life-span for most cells

- in culture, most of our cells cycle a set number of times, then die
- same seems to happen in eukaryotic organisms
- due to lack of telomerase enzyme in our cells

B. Telomerase is not present in all cells

- **Absent** in cultured cells from normal tissue - 50 divisions
- **Present** in many tumor cells - unlimited divisions
- **Present** in one celled organisms - unlimited divisions
- **Present** in primordial germ cells - unlimited divisions
- **Present** in tissue stem cells
- **Experimentally added** to normal cells with added telomerase gene - more cycles observed



Could loss of telomeres cause the aging process?

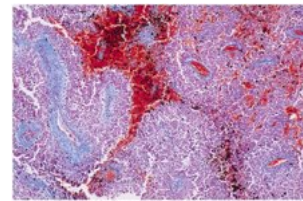
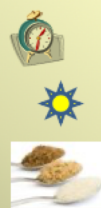
- perhaps, but there are other factors too.
- Loss of telomeres certainly provide a maximal limit to the number of cell divisions
- Could we provide telomerase therapy to restore telomere length?

Many cancer cells have reactivated their telomerase

- some anticancer therapies are directed at inhibiting telomerase
- so if you think you could extend life by adding telomerase, perhaps you will increase the risk of cancer

Some factors in cell aging:

- Telomere shortening
- Accumulated errors
- Chronic risk exposures such as oxidants, UV
- Glycation – sugar binding DNA, proteins, lipids



Cancer cells (red) proliferating in tissue

The problem with Dolly

- 1996: Dolly the sheep was the first cloned mammal from an adult cell
- achieved by taking a nucleus from the udder of an adult ewe, and transplanting it into an de-nucleated embryo

Dolly died at age 6 (most sheep live to about 12 years)

- Dolly's telomeres were 20% shorter than normal sheep by age 3
- she started developing serious arthritis at age 5
- premature aging occurred

Questions (to which we still don't have answers to!):

1. How do shortened telomeres cause aging?
2. Why didn't the de-nucleated embryonic cell restore the length of the telomeres in the 6 year old nucleus?



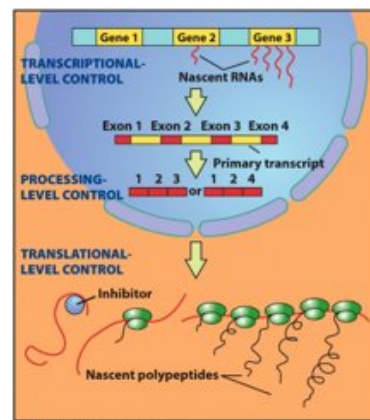
Aaaaand... they stuffed her and put her in a museum

<http://www.youtube.com/watch?v=6wFtyGtph2w&feature=related>

What you have learned in this section

- Concept of potency, how stem cells become differentiated cells, stem cell therapy
- Concept of chromosomes and ploidy
- Different types of aneuploidy
- Defects or rearrangements in chromosomes
- Examples of translocations
- Defining telomeres and centromeres
- How telomeres shorten, and how telomerase can restore telomeres in certain cells
- Why Dolly the sheep died prematurely

Control of Gene Expression



Readings: Chapter 12.7

Overview of gene expression in eukaryotes

The role of transcription factors in regulating gene expression

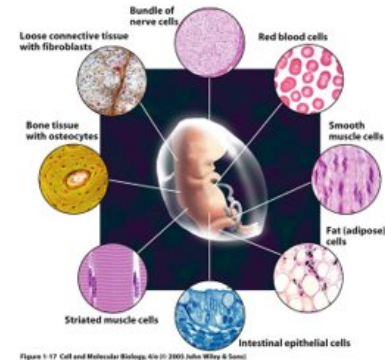
The structure of transcription factors

Different cells express different sets of genes

- The human genome encodes approx. 20k genes coding for proteins

Different *expression patterns* of different genes occur in different cell types:

- Genes encoding basic cell function proteins (e.g. ATP synthase, actin) are on in all cell types. We thus call these 'housekeeping genes'
- Some genes are only expressed when stem cells differentiate and are then turned off
- Some genes are only expressed in cell types with specialized functions
- Some genes are only expressed in response to certain stimuli (e.g. hormone induction)



Five main 'levels' of regulation that accounts for how much protein will be present in a cell

1. Epigenetic level

- Whether the gene is in a heterochromatin or euchromatin region

2. Transcription level

- if and how often a gene is transcribed into mRNA. This interplays with the epigenetic level

3. Processing level

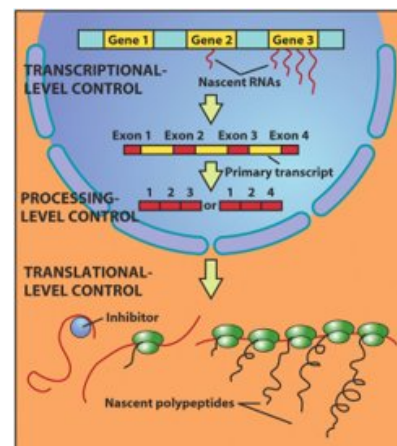
- different messenger RNAs made from a given gene (alternative splicing)

4. Translational level

- How much of the mRNA is made into protein
- Control of mRNA lifetime

5. Post-translation

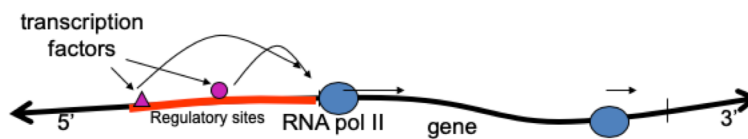
- Protein lifetime (e.g. ubiquitination)



Transcriptional regulation in a nutshell

Differential transcription is the most important mechanism by which eukaryotic cells determine which proteins are synthesized (more so than epigenetics!)

1. There are regulatory sequences of DNA that are 'upstream' (or 5') of the protein-coding portion of a gene. Together, these sequences are called the **promoter**.
2. Regulatory proteins called **transcription factors** bind to specific nucleotide codes in the promoter. These modify how easily and how often the initiation complex comes together and whether it can initiate transcription.
 - transcriptional activators and will promote transcription of the gene
 - transcriptional repressors and will inhibit the transcription of the gene

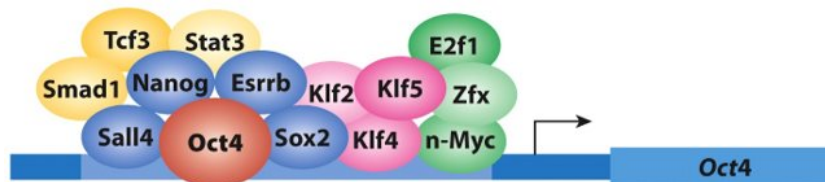


There are a *lot* of transcription factors

Between 5% and 10% of all the genes in the human genome encode for transcription factors!

- These regulate the transcription levels of all the other genes, and each other.
- One single transcription factor can regulate many genes
- One gene can be regulated by many transcription factors

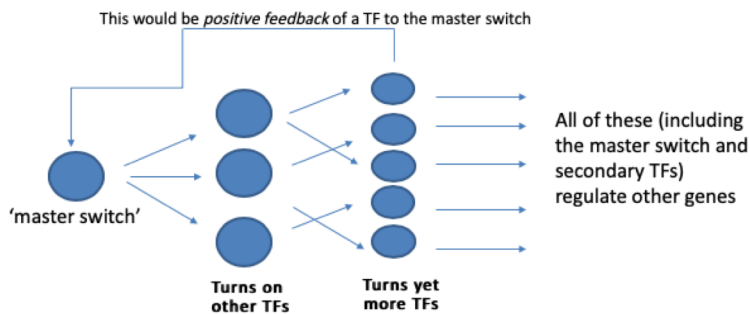
Different combinations of transcription factors can have different effects on how much a gene is transcribed. This gets very complex!



From Ng, Huck-Hui; Surani, M. Azim, Nature Cell Biology 13:490, 2011. Nature Cell Biology by Nature Publishing Group. Reproduced with permission of Nature Publishing Group in the format reuse in a book/textbook via Copyright Clearance Center.

Some transcription factors are 'master switches' in cell development

Because transcription factors can regulate the transcription of *other* transcription factors, some are more 'upstream' in a developmental pathway than others.



Examples:

- Expressing *myoD* TF is sufficient to get a fibroblast cell to turn into a muscle cell
- Expressing a single TF gene in the wrong part of a fly is sufficient to grow an eye on a leg.



The structure of transcription factors

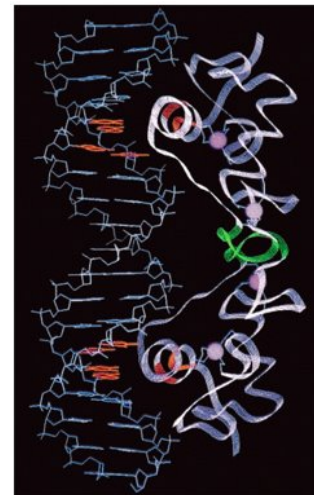
TFs work by having different domains:

1. A DNA binding domain

- Amino acids fit into the major groove of DNA and can directly interact with bases
- the structure of the protein forms tighter interactions with specific combinations of nucleotides (how these can recognize specific DNA regulatory sequences)

2. Regulatory domain

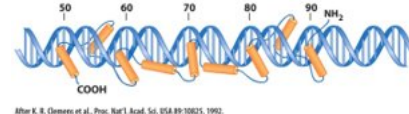
- Parts of the protein that interact with other proteins (other TFs, components of the initiation complex, etc)



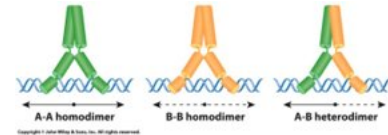
Most TFs operate as dimers (both heterodimers and homodimers occur). These come together in dimerization domains (area #3 in the diagram)

TF are classified based upon the 'motif' by which they bind DNA

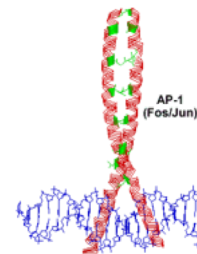
Zinc-finger motif – multi-“finger-shaped” projections of protein fit into major grooves of DNA



Helix-loop-helix (HLH) motif – forms homo- and heterodimers; humans have ~120 HLH genes – can produce 1000s of different transcription factors



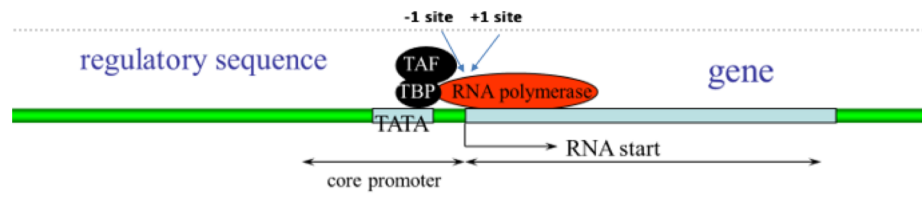
Leucine-zipper motif – two helices “zipped” together to form a coiled-coil structure as dimers



Components of the promoter 1: Core promoter

DNA sequence -1 to -40 bases from the start of the coding DNA.

Mostly defines the position of the start site.



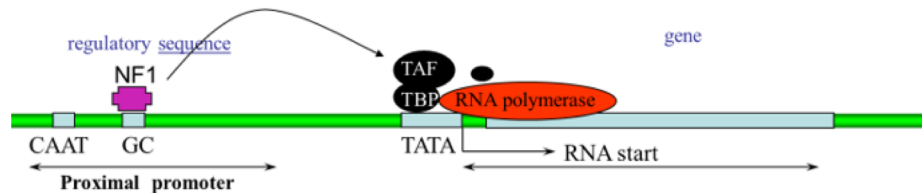
It is recognized by a series of DNA-binding proteins: **general transcription factors**, comprising the pre-initiation complex, including:

- **TBP (TATA binding protein)** recognizes the TATA sequence about 30 bases upstream from the start of the gene.
- **TAFs (TBP-associated Factors)** a group of General Transcription factors accessory proteins necessary for RNA polymerase II
- **RNA polymerase II**, which produces the **RNA**

Note that these are *generalized* parts of the promoter. Not all genes have a TATA box, for example, but are still transcribed.

Components of the promoter 2: Proximal promoter

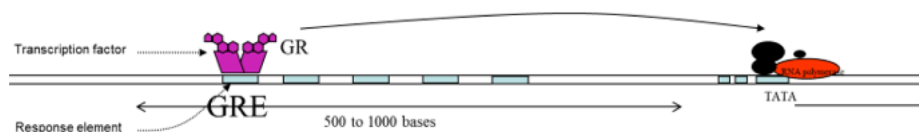
approximately -40 to -150 bases away from -1



- CAAT and GC boxes are bound to transcription factors such as NF1 & SP1
- NF1 recruits a co-activator needed for RNA polymerase to work (more later)
- While the core promoter determines whether or not transcription can take place, the proximal elements regulate **frequency** of transcription.
- When Cs within GC box are methylated by the cell, the gene is transcriptionally **inactive** (again, more later)
- Again, these elements are common, but not found in all genes

Components of the promoter 3: Distal promoter

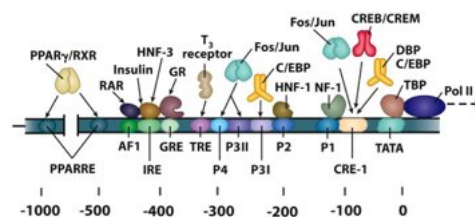
Contains response elements (-500 to -1000 bp).



Response elements = DNA sequences to which **specific transcription factors** bind

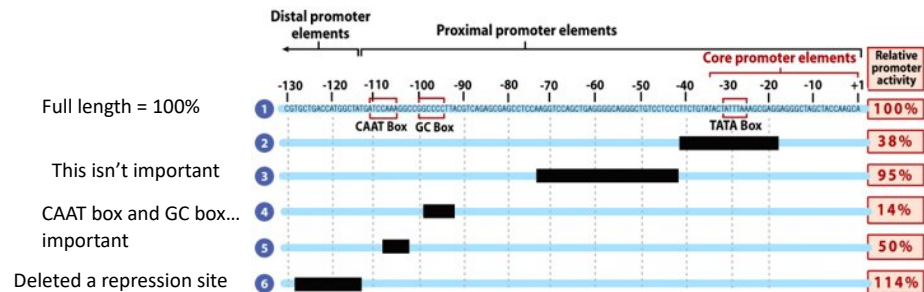
- These specific TFs are usually specific to one gene or a group of related genes
- These TFs can activate or repress the transcription of the gene
- An example is the activation of *PEPCK* which encodes phosphoenolpyruvate carboxykinase (active in gluconeogenesis)
- *PEPCK* is expressed mostly in the liver, and expression is increased by hunger and stress, decreased by eating

- This is a cartoon of the *PEPCK* promoter.
- Note that there are many different specific TFs that bind to different response elements. Different combinations of these TFs are active in response to different stimuli to modulate the expression of *PEPCK*



From S. E. Nisizaki et al., J. Nutrition 126:2099, 1996; © American Society for Nutritional Sciences.

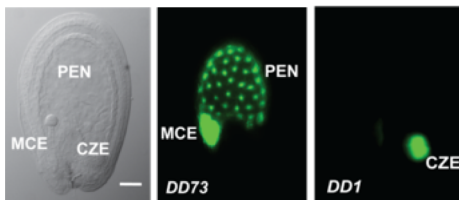
Use of 'deletion mapping' to identify response elements



1. Take a promoter and fuse it 5' to the coding sequence of a reporter gene (GFP will work)
2. Make a series of DNA deletions (black boxes in diagram) of parts of the promoter
3. Introduce these DNA constructs into a cell (one at a time!), and measure how much GFP fluorescence there is *relative* to the full-length promoter
4. If the fluorescence is low, you've deleted an important response element for the activation of transcription
5. If the fluorescence is higher, you've deleted a response element to which a transcriptional repressor binds to the promoter

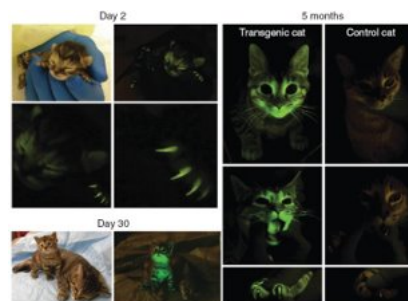
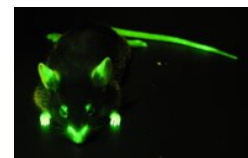
More fun with promoters

- Promoters specify which cells or tissues a gene is expressed in.
- to see where a gene may be active, make a promoter: GFP fusion construct and put it into a multicellular animal.
- Fluorescence = where and when that gene will be made!

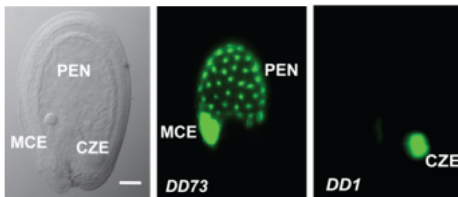
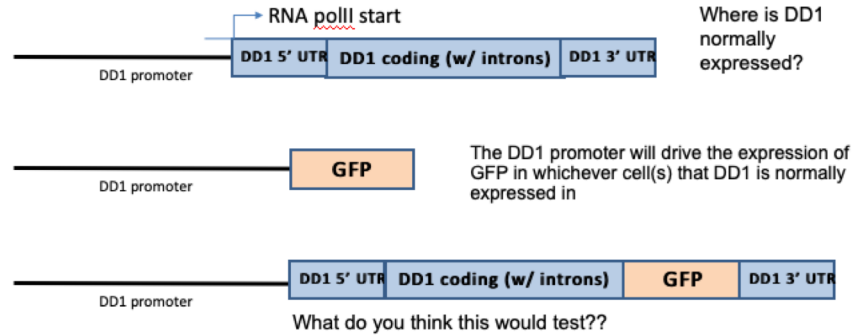


2 genes that will be expressed in different parts of a plant embryo

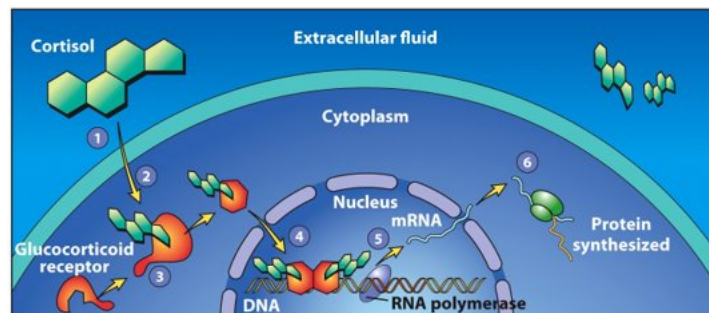
Rats! Kitties! Bunnies!



Characterizing a promoter.....



Example of a TF activating gene transcription



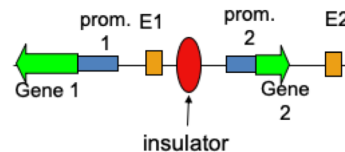
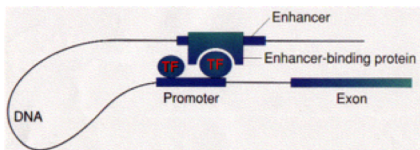
1. Glucocorticoid (cortisol) is released from adrenal gland following tissue damage
2. Cortisol passes through liver cell membrane
3. Cortisol binds to glucocorticoid receptor – receptor changes shape
4. A newly exposed NLS allows entry into nucleus
5. Glucocorticoid receptor (with bound hormone) can bind to a glucocorticoid response element (GRE) – binds as a dimer - PEPCK gene transcription begins – newly made RNA leaves the nucleus
6. PEPCK protein made – gluconeogenesis begins (make glucose to help in stress)

Note: It's not just PEPCK that's regulated by cortisol. Any gene that has a GRE in the promoter can be potentially activated cortisol

Components of the promoter 4: Enhancers

Enhancers are specific DNA sequences, often found far from the target gene, which bind specific **transcriptional activators** (or enhancer binding proteins), and activate gene expression

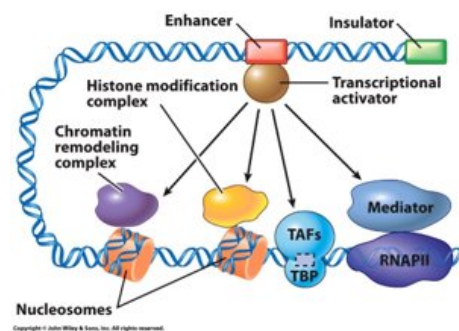
- Tens of thousands of bases away;
- can be in any orientation
- One enhancer, when activated, can activate a number of genes.
- Enhancers are separated from each other by **insulator sequences** that bind to proteins of the nuclear matrix that separate loops.



Note that enhancers can be on either side of a gene

So how do TFs and enhancer binding proteins *really* modulate transcription?

- **TFs Bind** to the response elements in the distal promoter region (recognizes the nucleotide base sequence), enhancer-bound activators fold around and interact with the promoter region
- Together, they **recruit** proteins which help the pre-initiation complex work. These are called **Coactivators**.
- **The coactivators** enhance the RNA polymerase activity



In this figure, only the general TFs are shown. The chromatin and histone modification complexes and Mediator are examples of coactivator proteins.

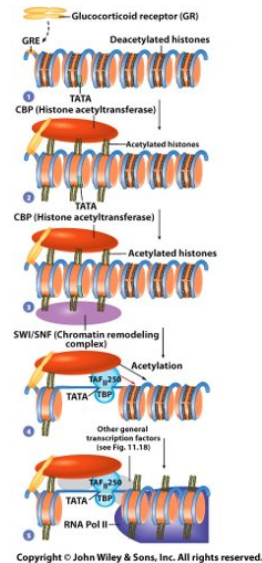
Co-activator function

The coactivators enhance the RNA polymerase activity in two ways:

- They assist in the assembly of the preinitiation complex and the start of transcription
- They act on the histone chromatin state around the gene, making the DNA more accessible for translation

In the GR example, CBP is a coactivator that is recruited.

- In the GR example, GR binds to the distal GRE and recruits CBP
- CBP is a histone acyltransferase that acylates the histones in the proximal and core promoter, causes slight loosening of DNA and histone associations
- The acylated histones recruit SWI/SNF (another coactivator!) which makes the promoter DNA really accessible (for the purposes of our discussion, it just removes the histones)
- This allows for the assembly of the preinitiation complex



Preinitiation complex has its own histone acetyltransferase activity: subunit TAFII250

Transcriptional repression

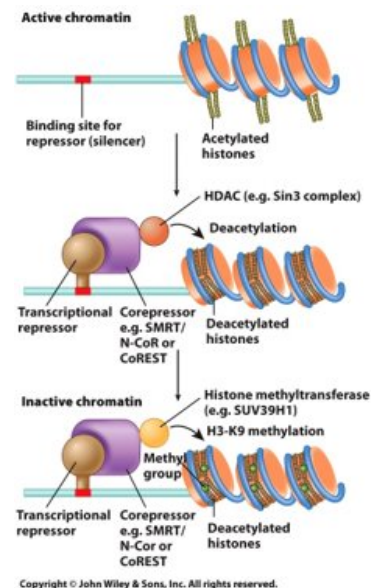
Gene transcription can also be turned *off* by factors that work at the promoter

Histone Deacetylases (HDACs) remove acetyl groups, and DNA winds up more tightly – gene activity is silenced

Histone methyltransferases – methyl groups added to histones causes tighter DNA binding – gene activity is silenced

DNA methyl transferases add methyl groups to DNA - always at carbon 5 of cytosine

- This essentially 'tags' regions of DNA so that they are utilized (transcribed) differently – usually a sign of inactivation
- This is a reversible process, but DNA methylation is "passed-on" through cell division
- This methylation is another form of *epigenetic* regulation!



What you should know from this section

- Identify different ways by which gene expression can be regulated
- Identify the types of proteins involved in regulating gene expression and how their structure dictate function
- Identify the different component of the promoter, where they are located, and how each regulates gene expression
- How a steroid hormone can regulate gene expression
- How histones and DNA packaging changes in transcriptional activation and repression