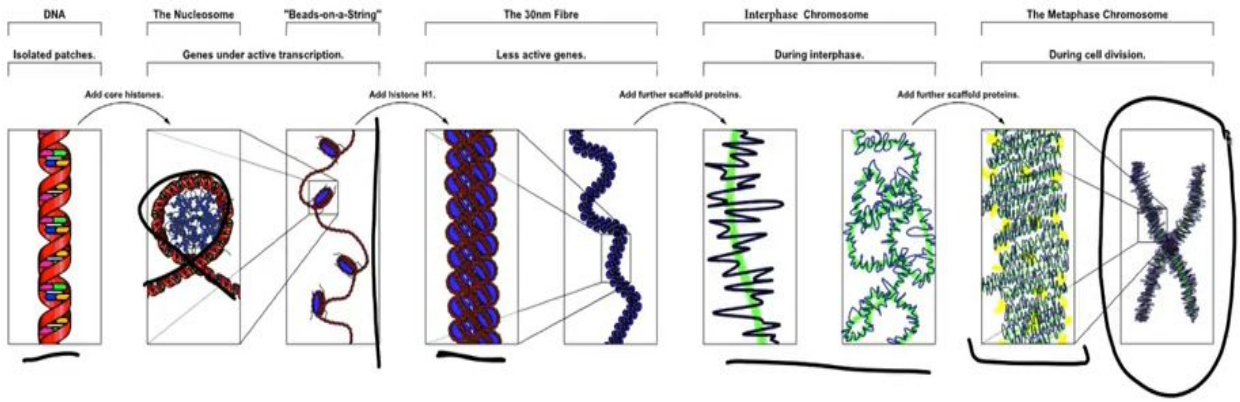
EPIGENETICS COURSERA CLASS: LECTURE WEEK 1

~25,000 genes in humans but not all are used at same time in all cells. Epigenetics is concerned with this regulation. It is the comments and import statements in the code base of DNA.

- ~1000 genes are involved in epigenetic control
- Terminology: **DNA** wrapped around **histones** forms **chromatin**
- more than 2 meters of DNA in each nucleus => packed into 10 micrometer of nucleus
- **nucleosome** is the smallest unit of chromatin, made up of ~8 histones
- positively charged histones + negatively charged DNA => attract
- "nucleosomes are like beads on the string"
- chromatin can be tightly (**Hetrochromatin**) or less tightly (**Euchromatin**) packaged => controls accessibility of DNA to transcription factors, RNA polymerase
- chromatin packaging forms higher order structure in **30nm fibre**. These form even higher order structure around scaffold proteins



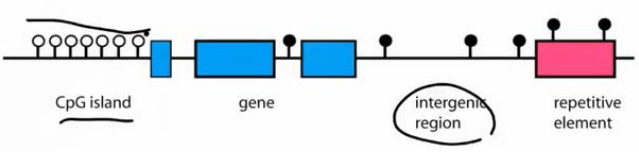
- chromosomes rarely look like X's. Usually only so tightly packed during **mitosis** (cell division)
- **DNA methylation**: can happen at C (when usually followed by G in mammals) (called CpG dinucleotides). It is laid down de novo by **DNMT3(a|b**) - simply appends a CH3 to C (this is a pretty strong carbon carbon bond). After cell division, new protein **DNMT1** recognizes hemi-methylated DNA (only one side is methylated) and adds methylation to the daughter strand. Many CpG islands (more than by chance) tend to be found in promoter regions upstream from genes (where transcription factors bind). If methylation occurs at these, it leads to **gene silencing**.

- **X inactivation is a good process to study as an example of epigenetic control**:

in females, one X gets deactivated (or otherwise female would have twice the X gene product) Occurs at ~100 cells stage early in embryo. This is done by DNA methylation of all CpG islands on one of the X chromosomes.

- CpG islands – usually unmethylated
- Intergenic regions – usually methylated
- Repetitive elements – usually methylated

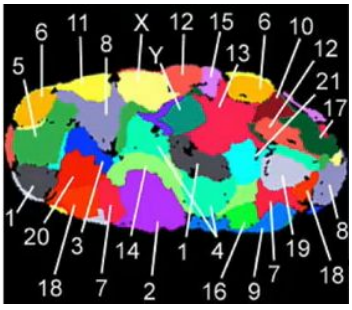
- Left: filled circles mean methylation.
- At intergenic regions methylation usually adds to genomic stability



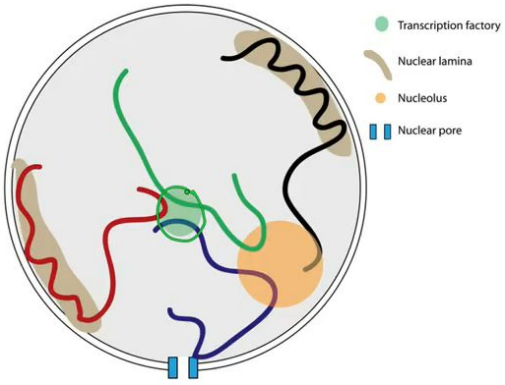
- At repetitive elements: Also to add to stability.
- DNA methylation can be added/removed actively in the cell

EPIGENETICS COURSERA CLASS: LECTURE WEEK 2

- Acetylation or Methylation (among other things) can happen at N-terminal tails of histones.
- Various molecules can bind to histones, some suggest there is a “histone code”, as these all have effects on transcription at those locations
- **Histone Acetylation**: can be made, unmade by 18 different molecules. Acetylation is correlated with gene activity (when it binds, it reduces positive charge of histones and loosens the DNA to be transcribed). **Non-heritable** (as far as we know)
- **Histone Methylation:** can correlate with either transcriptional activity or inactivity. Can be laid down/removed with many types of molecules too. They act as docking sites for other chromatin proteins that can act on the chromatin structure or even lay their own methylation marks.
- There are **many types of Histone variants -** some occur at centromeres, some occur where the DNA is broken, and they recruit repair proteins etc.
- **DNA repair** is intricate mechanism: DNA must be wound out, repaired, and wound back in!
- **MicroRNA (miRNA)** (19-24 nucleotides) - does post-transcriptional gene silencing! Binds to a complex that finds specific mRNA and snaps it! (Also commonly used for experimental knockdown of genes)
- **Piwi-interacting RNA (piRNA)** (24-35 nucleotides) - involved in silencing transposable elements in the DNA and improve genomic stability
- **Long noncoding RNAs (lncRNA)** (>200 nt in length) - constrained to nucleus, ~10,000-200,000 found in mammalian genome. Incriminated in X inactivation, Hox gene silencing, Genomic imprinting, DNA damage response. They function by acting as Guides, Scaffolds, etc. Mainly, the length of these allows for very precise targeting of a particular place in genome and allows epigenetic complexes to be very specific.

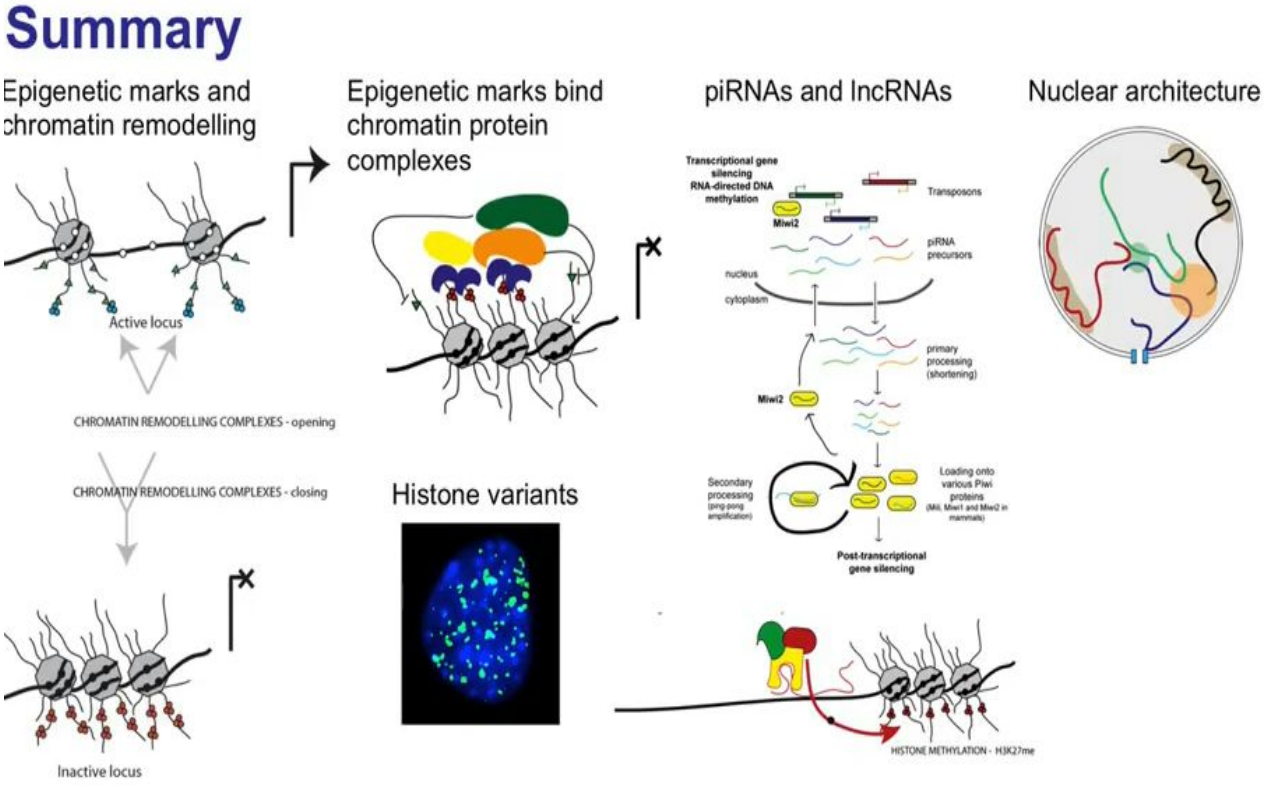


- example lncRNA: **Xist:** critical determinant in X inactivation. 17kb long. Transcription of Xist => silencing of that chromosome (in cis).
- example 2 lncRNA: **HOTAIR**: lives around HOX C cluster, helps silence genes in HOX D cluster, in trans! (but this is controversial)
- What?? Chromosomes have their own territories inside nucleus! Their own “place” where they stay clumped up. These regions vary across cell-type, developmental time. It’s very unlikely to see translocations across chromosomes that are far away physically



- **Nucleus** is organized:
- **Nuclear lamina (periphery)** - associated with silencing of transcription, Hetrochromatin
- **Nuclear pore** - associated with euchromatic sites, probably for fast export out of nucleus?
- **Nucleolus** - ribosomal DNA (rDNA) repeats are here, get transcribed here as well. Inside there is RNA Pol I which transcribes **rDNA**. Nucleolus is surrounded with a region that has a lot of RNA Pol III. Which transcribes

transfer RNA (**tRNA**)
- **Transcription factories** - associated with Euchromatin sites. These are several hot zones where transcription happens. Lots and lots of RNA Polymerase II, transcription factors (the RNA Pol that transcribes ordinary genes). DNA is fed through these sites.
- There are only **3 RNA Polymerase: I, II, III**
- many more types of compartments inside nucleus

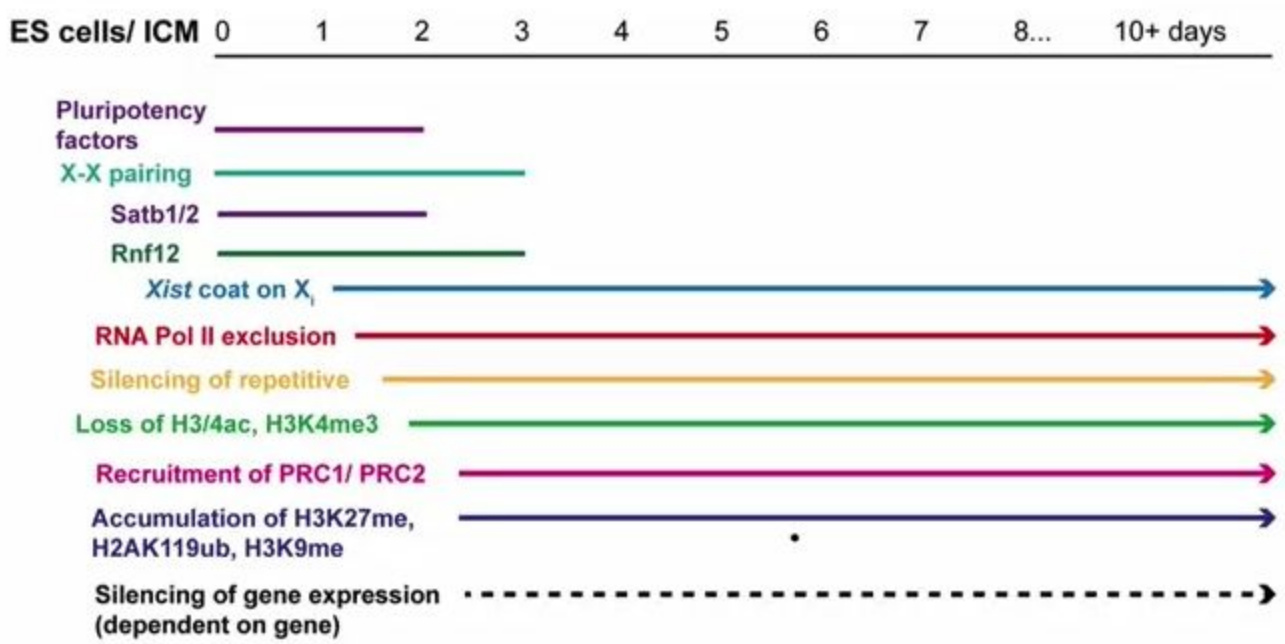


EPIGENETICS COURSERA CLASS: LECTURE WEEK 3
X inactivation history: 1949: Barr and Bertram found structure at nuclear periphery only in females. Later identified as hetrochromatin, and in fact X chromosome. 1961 was proposed this was a mechanism for dosage compensation
- occurs early in development, one X chromosome randomly inactivated per cell.
- Another type: Imprinted form of X inactivation: the (paternal or maternal X) is inactivated for all cells. Only occurs in placenta, in mice, not known to occur in humans.
- Few crucial states in X inactivation:
    - **Counting** of X chromosomes: The X:A ratio (A = autosomal) to determine if inactivate
    - **Choice** of X chromosome to inactivate
    - **Initiation**: **Xist** expression from the XIC (X Inactivation Centre)
    - Progressive **spreading** of inactivation, in cis along the chromosome
    - **Establishment** of inactivation, turning Xist signal into silence
    - **Maintenance** of the inactive X

- Xist recap: 17kb noncoding lncRNA, in Nucleus. Acts only in cis -- inactivates chromosome that it comes from.

- There are several other lncRNA segments around Xist that either enhance or suppress Xist.
- In addition, a gene few hundred kb upstream called **Rnf12** (actually coding). Involved in doing the counting. Threshold levels of Rnf12 are needed to initiate X inactivation. When one of the X is inactivated, Rnf12 will go down in concentration so another X inactivation doesn't occur. Male cells don't have enough Rnf12 to activate Xist. But this is sketchy, incomplete understanding
- Other set of factors control Xist: "**Pluripotency factors**". (these also maintain pluripotent state of embryonic stem cells). These proteins bind to Xist to repress it, and also suppress Rnf12, and also suppress lncRNA that activate Xist. These are found in stem cells. => X inactivation will not occur in embryonic stem cells.
    - Eventually later in development these go away => Rnf12 will come on, accumulate => activate Xist => X inactivation => Rnf12 levels drop to male levels. Nice.
- How is the **choice** made? Turns out the two X's pair up physically (still in embryonic stem cell) in space and "fight". Called "chromosome kissing". One gets inactivated.
    - This **choice can be _SKEWED_,** all the way up to 95-5%. Not necessarily 50-50. There exist stronger/weaker "choice alleles"
- Xist creates a silencing cluster by coating first the repetitive elements on X, then other genes progressively, slowly over several days.
- A small handful of **"escapee" genes** on the X chromosome remain active! They never get drawn into the Xist inactivation cloud
- The process: Loss of active histone marks, accumulation of repressive histone marks. Protein complexes (**PRC1, PRC2** for example) bind to the Xist RNA to enact these changes.

TIMELINE OF THE PROCESS RECAP:



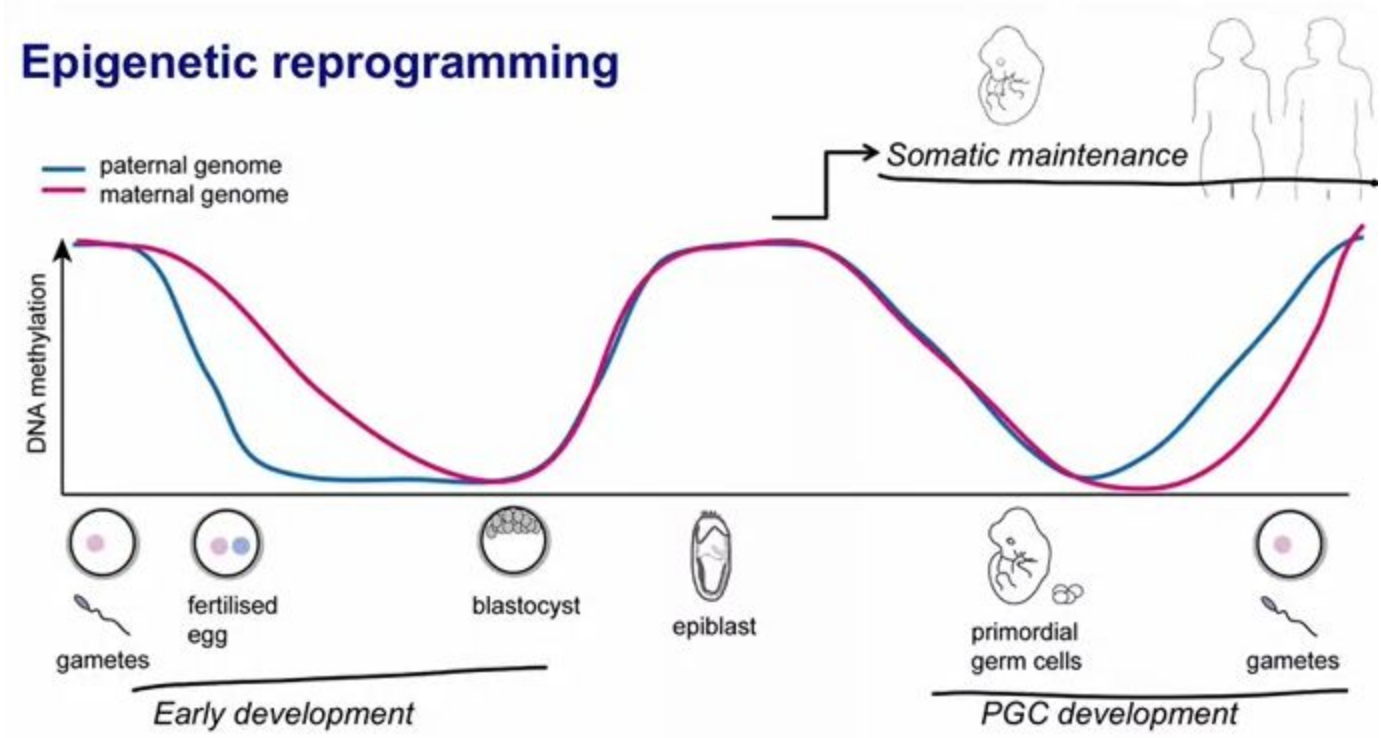
- **Maintenance**: Two factors implicated: Dnmt1, Smchd1
- **Dnmt1:** If inactive X is already established and removed, X inactivation is not maintained! So necessary in maintenance. Recognizes hemi-methylated DNA (after cell division) and methylates the other half.

- **Schmd1**: co-localizes with inactive X chromosome. Also involved in maintenance
- There are several layers of redundancy involved in X inactivation
- Wow **interesting**!:
    - in worms, instead, females (XX) have both X chromosomes downregulated to 50%!
    - in flies: the male X chromosome (Xy) gets UPREGULATED to 200%.
    - both are ways of doing dosage compensation, achieving same ratio of X product to autosomal product.

-Q: I don't understand why keep the inactive X around the entire lifetime... hm. Is it because not all genes are completely shut off, and those are somehow important?

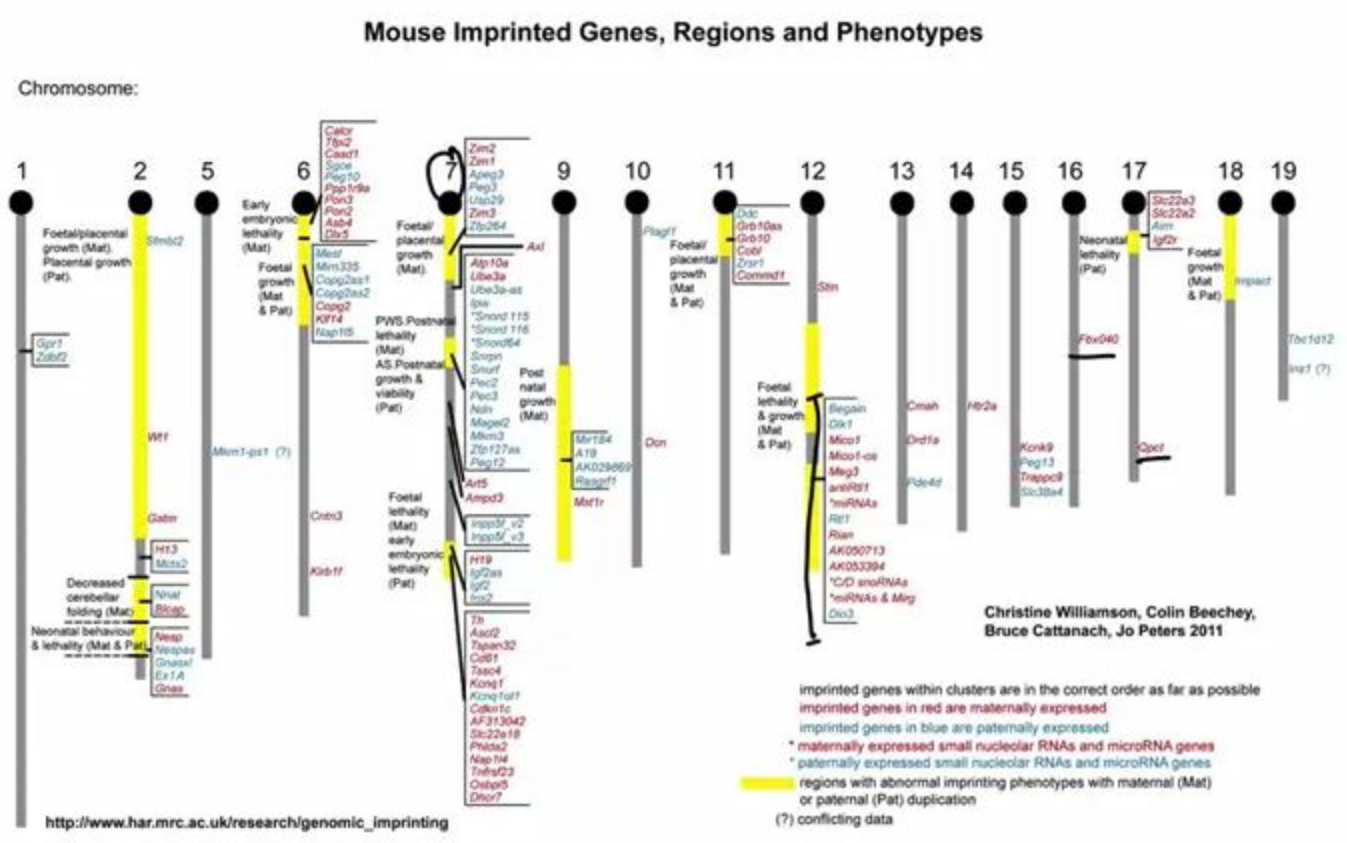
- Associated video: http://www.wehi.edu.au/x_inactivation_and_epigenetics/
    - Human genome: ~3 billion base pairs, ~30,000 genes
- Position effect variegation: Position of a gene relative to heterochromatin results alters expression of the gene. (closer to hetrochromatin => more likely to be silenced). These genes can be sometimes on, sometimes off depending on the distances
- Hetrochromatin can spread from telomeres, centromeres, but is usually limited by DNA elements called "boundary elements" that insulate other regions from the spread.

EPIGENETICS COURSERA CLASS: LECTURE WEEK 4

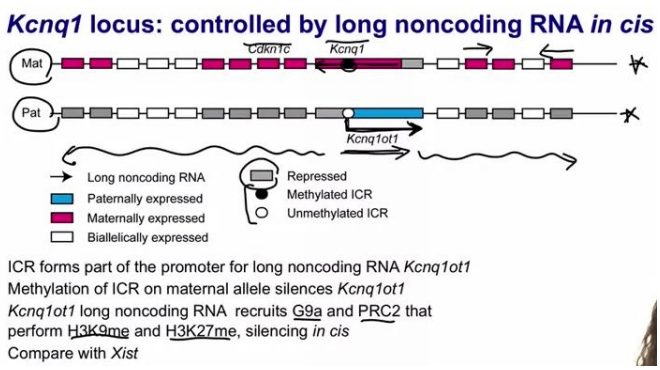


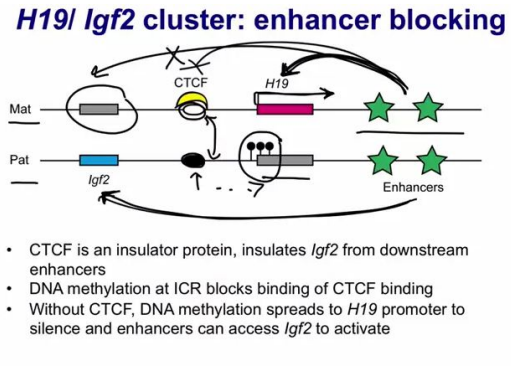
- Right after egg is fertilized, the paternal genome is heavily de-methylated very rapidly (in ~6 hours). The maternal genome a little slower. This has to occur because the DNA had particular methylation patterns required for proper cell functioning of the sperm and the egg.
- In fertilized egg, the two parts that exist for a while (from mom, dad) are referred to as "pro-nuclea".
- **Genomic imprinting**: Monoallelic gene expression based on parent-of-origin of the allele.

Known to be critical for embryo viability
- Expression of imprinted genes is controlled by **Imprint Control Regions** (**ICR**). Imprint is associated with DNA methylation at the ICR. The way ICR brings about imprinted gene expression differs by locus: lncRNA, enhancer/insulator blocking, etc
- ICR methylation is established in primordial germ cells to achieve parent-of-origin specific marks, maintained in the embryo
- This ICR methylation is actively protected during the vast reprogramming that happens in the embryo by specific proteins that are found in the egg. Holy shit
- When new germ cells are formed, these ICR regions are erased and re-methylated according to the gender
- There are ~150 genes imprinted in humans that we know of right now.
- They exist in clusters and can share ICRs
- Gene expression of these imprinted genes is **tissue specific**. Most imprinted expression is seen in placenta and the brain.



- To measure methylation: **Sulfide sequencing.** The bisulfite conversion replaces all C's with Uracil (U), but the methylated C's don't undergo this change. Then amplify the DNA and sequence. SNPs can disambiguate between paternal/maternal strands
- Examples of mechanisms of regulation: (Methylation not necessarily simply silencing!)
- **Kcnq1** locus ICR: controlled by lncRNA in cis
- **H19 / Igf2** cluster: works by enhancer blocking. The enhancers switch to enhancing a different gene? :S





- Example of disorder: Beckwith Wiedemann syndrome. Maternal transmission from carrier mother to affected offspring (50% of cases since one mutant allele). Problem is with a mutation that makes the chromosome look like it came from Father.
    - Parent-of-origin specific inheritance, not sex-specific
- There is some evidence that epigenetic reprogramming may be disrupted with use of **assisted reproductive technologies** (ART) (in particular IVF or IntraCytoplasmic Sperm Injection (ICSI)) => Leads to 3-4x higher chance of epigenetic diseases
- In **cloning**, very few animals make it through (<1%). And even ones that do have very different gene expression later in life. Seems the issue may be with imprinted genes:
    - Somatic nucleus (that is taken from adult donor cell) didn't go through the necessary premordial germ cell (PGC) development epigenetic reprogramming that normally goes on in production of germ cells. The ICRs are not reset, and the packaging is all wrong.
- In transforming somatic cells (like skin cells) to **induced pluripotent cells** (iPS cells), we are essentially trying to reverse a cell to look like one of the ones found in blastocyst very early in development. Need to:
    - remove all lineage-specific epigenetic marks
    - restore pluripotent epigenetic marks
    - restore open pluripotent chromatin state
    - retain imprints at ICR
    - remove XCI marks (in females) so that there are again two active X chromosomes
- this is easier in some somatic cells than others

EPIGENETICS COURSERA CLASS: LECTURE WEEK 5

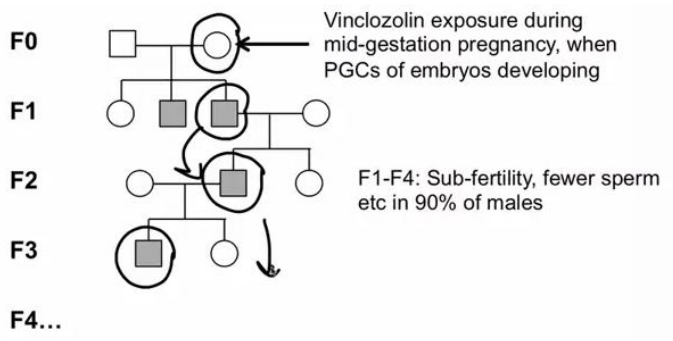
- environmental effects on epigenetic control: Diet, Paternal/Maternal effects
- wow, **Turtles**: males and females are genetically identical. Sex-determination is temperature dependent during egg phase.(=> some epigenetic influence)
- **Tulips** as well: vernalisation cold needed to allow flower
- Worker **bees** and Queen bees are genetically identical, but Queen bee is much bigger/different. Caused by consumption of "royal jelly". i.e. clearly diet influence!
- **Transgenerational epigenetics**: inheritance of phenotypes or gene expression patterns from parent to offspring by passage through gametes, not explained by genetic differences.

- Example of environmental effects on epigenetic control #1: **Dutch Famine**

- In 1944-1945, dutch families isolated due to war/winter, calorie intake => 580 cals/day (bread, potatoes) for many months
- Slightly controversial results, but exposure to famine seemed to have effects on peri-conceptional period, lead to increased mental/metabolic disorders in offspring
- Possible that there was interference during epigenetic reprogramming during embryogenesis/gametogenesis. Some evidence that there were changes in DNA methylation of small number of (imprinted) genes.
- But, some evidence that the next generation seems to have normal offspring again.

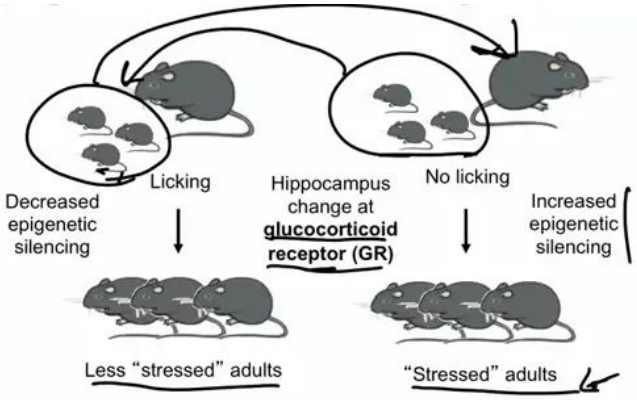
- Example of environmental effects on epigenetic control #2: **Overkalix**
  - Overkalix is an isolated town in Sweden, have good records of famine/feast
  - Suspicious link between availability of food in grandparents and longevity in grandchildren, but only paternal grandfather to grandsons or paternal grandmother to granddaughters.
  - Very suspicious study. Very weak evidence. Lack of good controls.

- Example in **mice**: exposure to **Vinclozolin** (used as fungicide) in mother during mid-gestation pregnancy (when PGC of embryos are developing) leads to subfertility in her sons (with 90% penetrance!). More distressingly, breeding further to F2,F3,F4,etc. males still display subfertility. Also there are effects in females. Substituting Vinclozolin with Methyoxychlor similar results. While females are affected, effect propagates only down the male line. Also, some evidence that Vinclozolin also has genetic changes (copy number alteration).



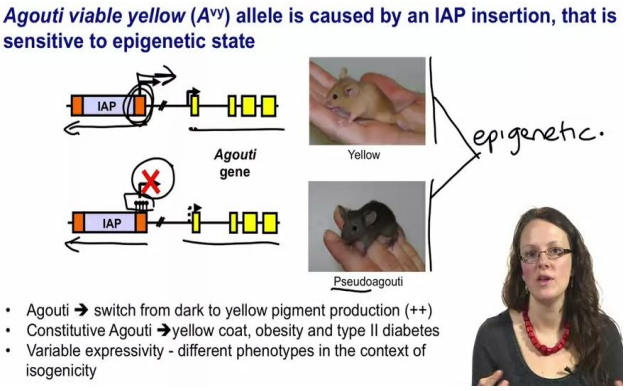
- Also more confoundingly, doing this study in two different inbred mice populations, get different results. So the sensitivity to these compounds is affected by genetics. Well that's just great.

- Another example in mice: amount of licking leads to phenotypic differences in adults. Can control for genetics by swapping offspring. (confusion note: can this not be some simple neural/learning effect? how to control?) (also, what about F2?)

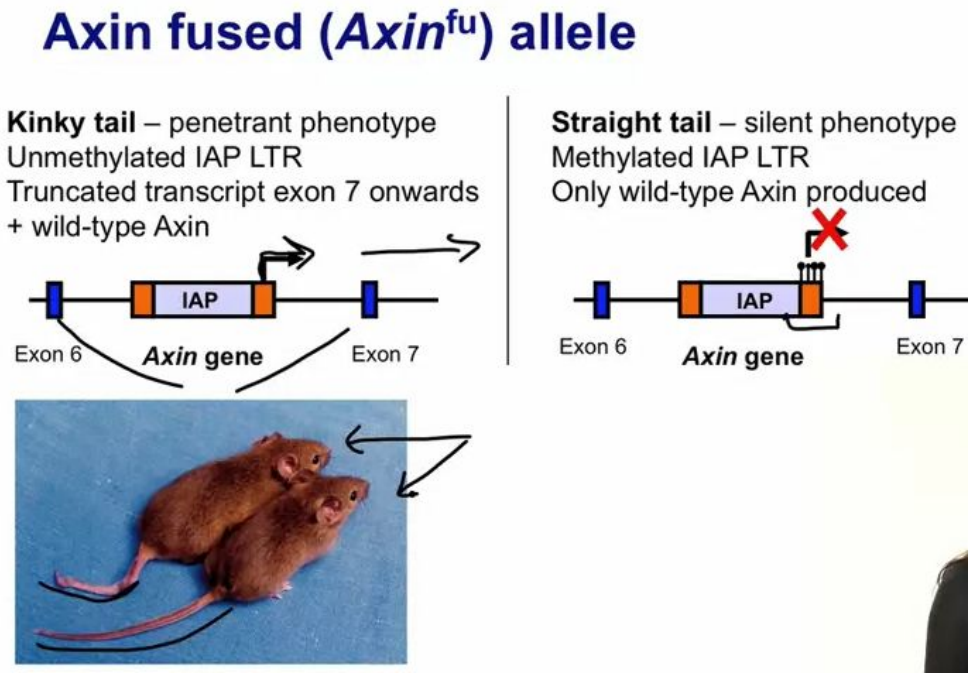


- Example in plants: **Linaria vulgaris** has a heritable DNA methylation that affects its phenotype in how the plant looks.

- Studying transgenerational epigenetic inheritance is best done in mice. Best studied example: **Agouti** (Avy) viable yellow allele, **Axin** (Axinfu) fused allele.



- The Avy allele is an example of **metastable epiallele:**
- **epiallele** -> epigenetic (mitotically heritable methylation marks)
- **metastable** -> they remain fixed during lifetime of organism, but may get inherited in offspring
- The Avy allele is involved with other things than simply signaling yellow pigment to be produced, so when there is too much of it around and always, mice become obese, diabetic. This can be inherited in mice, but only through the mother. The marks are cleared in paternal copy.
- **<u>Fascinating</u>: If you feed mother folate, choline, betaine, vitamin B12, or ethanol (these are all methyl donors so there is more "substrate" to create methyl marks) during a critical period (in this case during early development in utero or even slightly before conception), it results in more pseudoagouti mice! As if there were more methyl marks to go around so they simply get directly applied. Folate in particular, is recommended to be consumed for pregnant mothers.**
    - **Also, feeding "genistein" (soy protein) also leads to many more dark mice. Molecular mechanism not known.**
    - **And feeding bisphenol A (a common plastic) actually produces more yellow mice!**



- Axinfu: Also **metastable epiallele**. There is an IAP in intron region of the Axin gene, which controls development and makes sure the tail is straight. The Axin gene drives expression starting from Exon 7 and Axin isn't produced correctly. Methylation silences the IAP and all is good.

- **Molecular mechanisms** for Transgenerational epigenetics. Studies looked at Avy allele:
    - Appears to be completely cleared during CPG reprogramming in early embryonic development! So that's not it.
    - Could be histones, since even though most histones are replaced with promines(?) in DNA carried by sperm, some remain.
    - Could be messenger RNA that carries the mark and then causes re-establishment of DNA methylation or histone modification

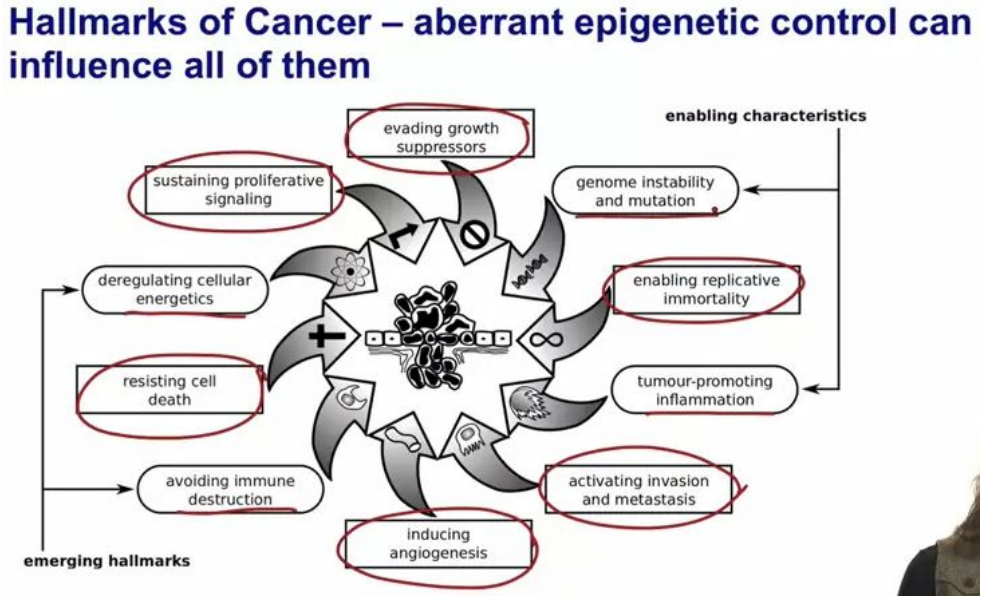
- **Paramutation**: can an RNA molecule influence transgenerational epigenetic inheritance?
- what can happen in mice is that breeding a m/+ with +/+ wild type, you don't get 50% mutant phenotype. Can get as much as 90%, and double check that indeed you have genetically correct ratios, but somehow offspring is epigenetically different. How is this inherited?
- Some suggestions that miRNA or piRNA is involved, as we know that there is RNA in both sperm and egg. These could be passing on some of the epigenetic state.
- Transgenerational epigenetics in humans example:
    - evidence that MLH1 epimutation (due to hypermethylation at CpG island) can run in families, just like mutations, increases chances of cancer in patient (despite no genetic mutation). But more careful study revealed that this is probably not the case, with more suitable controls. Controversy! :p

Required reading:
We are what our mums ate
http://www.theguardian.com/mending-broken-hearts/british-heart-foundation-we-are-what-our-mums-ate

EPIGENETICS COURSERA CLASS: LECTURE WEEK 6

- Cancer Epigenetics

- generally involves activation of oncogenes (genes for growth promotion), and inactivation of tumour suppressors (both can be achieved genetically and epigenetically)

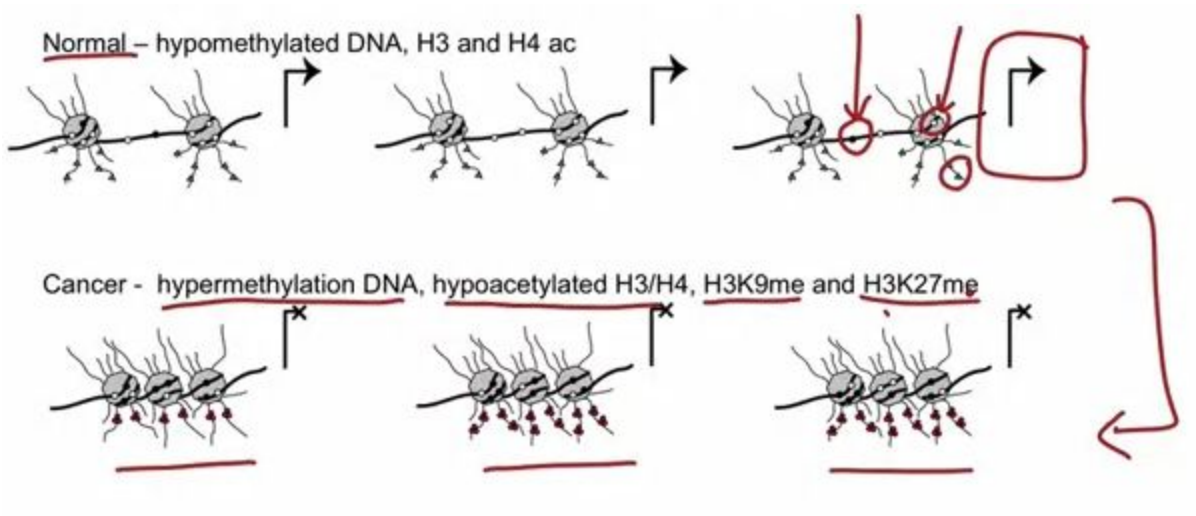


- Knudson

hypothesis: usually many errors have to build up and work in concert for cancer to develop
- DNA methylation: Is a strong candidate for introducing errors in the system as it is mitotically heritable. CpG islands on tumor suppressor genes can become hypermethylated which inactivates the genes (similar effect can be achieved more traditionally through a mutation in that gene).
- good news is that this is a reversible effect unlike mutations => potential treatments
- CpG island hypermethylation can be used as single **biomarkers,** but usually can get better results looking at panels of markers at the same time
    - For example: sending skin tissue to lab can test for skin cancer as certain regions of the DNA are known to be hypermethylated in case of skin cancer.
    - We have technology to relatively reliably tell how heavily a region in DNA is methylated
- the methylation biomarkers can be used for Diagnosis, Prognosis, and informing treatment
- In addition in cancer as a general rule of thumb, various intron, intergenic, and esp. repetitive element regions of the DNA tend to become **hypomethylated** (i.e. less than usual methylation). This leads to all kinds of trouble due to genomic instability (deletion, insertions, reciprocal translocations)
- Experiment: Take adult mouse, suppress DNMT1 in some cell type selectively (which leads to hypomethylation is tissue) => increased genomic instability => animal develops cancer.
- However, deletion of DNMT1 can both enhance or suppress tumorigenesis since different tumours have different dependencies:
    - If driven by tumour suppressor hypermethylation, depletion of DNA methylation appears to suppress tumorigenesis
    - If driven by chromosomal instability, depletion of DNA methylation enhances tumorigenesis

- **Aberrations in terms of Histone Modifications:**
- Histones can also look different in cancer, can be mutated, screw up genomic stability, bad things happen. Much less is known though.
- **Carcinogens** (compounds that increase incidence of cancer) can be **mutagenic** (i.e. DNA screws up) or **non-mutagenic** (DNA seems okay). An example of non-mutagenic carcinogen are some heavy metals. It seems these can alter activity of histone modifier enzymes and may therefore act epigenetically.

- Long-range epigenetic alterations in cancer and alterations to nuclear architecture
- Groups of genes. Megabase-region spans that are abnormal
- Large regions of active genes <-> Large regions of inactive genes in cancer
- As an example of what can happen:



- Probably what's going on is that there are more global changes to the entire nuclear architecture in cancer cells.
- Indeed, these global changes are discriminative for identifying cancer, for example:
    - Nuclear size, Nuclear shape, Ploidy (#copies of autosomes), chromatin organisation

- Altered expression on piRNAs and long noncoding RNAs in cancer

- btw micro RNA (miRNA) is also globally screwed up in cancer, but we don't consider this an epigenetic effect because these act post-transcriptionally
- **piRNA** (reminder: expressed from transposons, in bi-directional manner), can have effects on post-transcriptional gene silencing, or can be transported back to nucleus for RNA-directed DNA methylation.
- **lncRNA** (reminder: acts as molecular scaffold that chromatin regulatory complexes bind to)
    - As an example, HOTAIR (a lncRNA) is overexpressed in breast cancer
    - Another example, PCA3 is indicative of prostate cancer, can be detected in urine (yay, non-invasive!)

Q: How do all these epigenetic mistakes come about?
A: Well, genetic changes => Epigenetic machinery/regulators mutations => bad things happen

- Therapy + Epigenetics

- One of the easier things to do is to inhibit epigenetic regulatory enzymes (such as DNMT, HDAC, HMT, HAT, HDM, TET, Chromatin remodelers).
- DNMT inhibitors work by attracting DNMT and then trapping it indefinitely.
- HDAC inhibitors
- Chromatin readers

Economist article on "Cancer's epicentre" reading http://www.economist.com/node/21552168

- Many of the genes whose breakage leads to cancer are themselves involved in epigenetics. Interestingly, while damage in DNA is not easily treatable, epigenetic damage might be as it is more easily reversible.

- Example: enzyme EZH2 inhibitors. Normally, enzyme attaches methyl groups to histones, but in lymphomas (cancers of immune system) mutations occur that make EZH2 overactive => more methylation that should be => silences surrounding genes (perhaps including tumor-suppressor genes). Preliminary tests with EZH2i encouraging on cell cultures and lab animals
- Example 2: a substance JQ1 (which blocks epigenetic regulator BRD4) blocks activity of gene Myc. Myc codes for a transcription factor that is involved with expression of almost 15% of DNA. Researchers had hard time finding way to down-regulate Myc directly, but going through BRD4 using JQ1 does the job indirectly.

“”

The complexity of the mammalian genome is regulated by heritable epigenetic mechanisms, which provide the basis for differentiation, development and cellular homeostasis. These mechanisms act on the level of chromatin, by modifying DNA, histone proteins and nucleosome density/composition.

“”

More given reading: (nice review!)
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3480634/