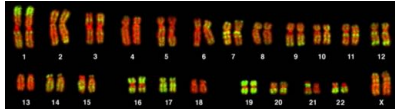




ANBI 139

Evolution of Human Disease

Lecture 10: Domesticated disease? Endogenous retroviruses, Transposons etc.



Pascal Gagneux

Winter 2021

Transposons: “Jumping Genes”



MOBILE DNA: A false-color transmission electron micrograph of a transposon, a segment of DNA that can move around chromosomes and genomes.
© PROFESSOR STANLEY N. COHEN/
SCIENCE SOURCE

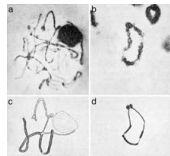
More than half of our genome!
These types of DNA elements come:
in different sizes
with different capacities to copy and paste themselves

Transposons, originally discovered in plants (Maize) are mobile genetic elements (chunks of DNA of various sizes, hundreds to thousands of bases long) that can “jump around the genome” and also make additional copies of themselves, which then insert themselves in the genome.

Transposon discovered in Maize in the 1950s



By Biologist
Barbara Mc Clintock



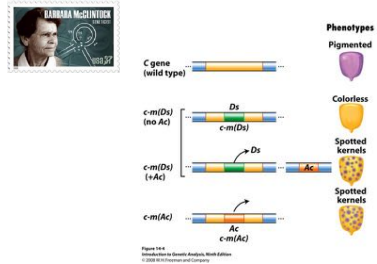
seed color genetics,
maize chromosomes



When McClintock first reported her findings in the 1960s, most people believed that this was something unique to corn. But later, as transposable elements were discovered in *E. coli*, yeast, and higher organisms, it became apparent that she had been the first to describe a phenomenon that was far more universal, suggesting that genomes were far more dynamic than first supposed. In 1983, she was awarded the Nobel Prize in Physiology or Medicine for her early work on corn transposons.

What species were transposons discovered in?
Maize

Transposons: Jumping Genes



Some jumping genes become dependent on others for jumping: Ds and Ac elements

Top row: Wild type pigmented kernel.

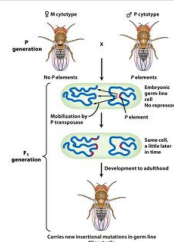
Second row: Ds element is inserted into pigment gene (C) permanently, disabling it. By itself, it can't move. It's stuck. Ds is a non-autonomous element.

Third row: Ds and Ac both present, Ds can now excise from the C gene in some cells (i.e., it can transpose) during development, creating developmental fields that can produce pigment. This is because Ac has provided the elements needed for Ds to transpose.

Fourth row: Ac is inserted into pigment gene, but not permanently, as it can provide the elements that allow its removal from the gene. Ac is an autonomous element.

And the kicker: Rarely, an Ac type was sometimes found to transform into the Ds type, apparently because the Ac element spontaneously turned into a Ds element. (This could mean that Ds is simply a mutant version of Ac that has lost the ability to encode the elements that allow it to jump around.)

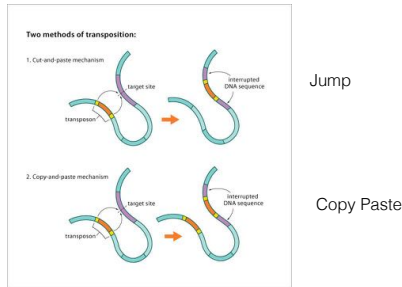
Transposons in fruit flies



Transposons in fruitflies p (passenger) elements

Such transposable elements can cause infertility in fruit flies.

Transposon

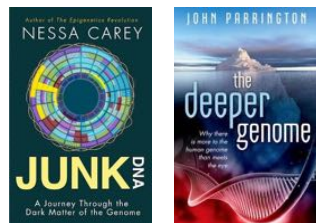


Two principal ways DNA transposons can jump around the genome

Practice Question:

Why could the activity of a transposon be dangerous for organism in which it happen?
genomic instability/mutation/ disruption of gene function

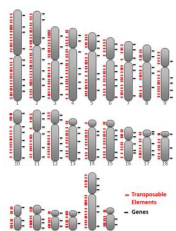
Junk DNA or Deeper Genome???



1.5 % protein coding DNA
20% intron DNA
~50% Transposon DNA!!!

Two recent books discussing the possible functions of “junk DNA”

Transposons = Transposable Elements



Transposable elements vastly outnumber genes.

Transposable elements and genes are depicted as horizontal black or red lines respectively along chromosomes. Figure not to scale, actual TE and gene counts are ~ 4.5 million and ~25,000 respectively.

TE can be thought of as genomic parasites or molecular parasites, copying themselves at the cost of the genome. Genomes have evolved protective mechanisms against the spread of TE: including methylation of DNA to “shut down” DNA encoding TEs.

Biological definition of a parasite: an organisms that lives on or in another organism and can potentially harm it.

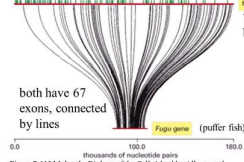
Practice question: Does the human genome contain more genes or more transposable elements?

More TEs.

Transposons, some species can get rid of many

Comparison of Fugu and human huntingtin gene:

(green indicates transposons prevalent in human version)



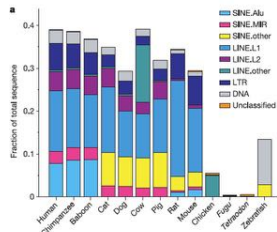
Fugu (鰐) puffer fish/ Botete, has lost most of its transposons and has a much smaller genome than most vertebrates.

Practice question:

What is the genome of the pufferfish so much smaller than that of most other vertebrates?

The ancestors of the pufferfish purged transposable elements from their genome.

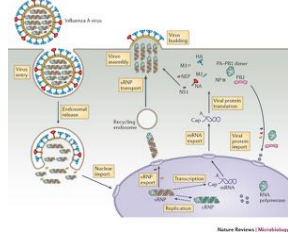
Transposable elements gone WILD!!



Different types of transposable elements have spread across the genomes of different species of animals.

VIRUSES AND RETROVIRUSES

Virus: **Influenza A**
negative strand RNA
RNA → Protein



Retrovirus:

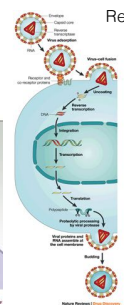
HIV

RNA

→ DNA

→ RNA

protein



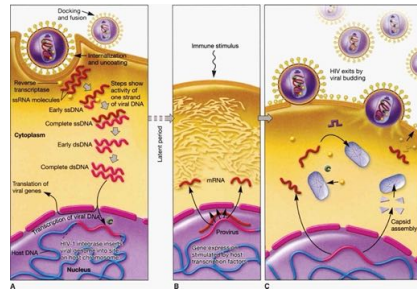
Viruses hijack the cell to make more viruses (they are sub-cellular parasites). Unlike transposons (jumping genes) viruses can jump from cell to cell and also from individual host to individual host. Retroviruses have reverse transcriptase and can generate DNA copies that integrate into the host genome

Practice question:

What makes a retrovirus a retrovirus?

Its genome encodes an enzyme that can “back transcribe” RNA into DNA.

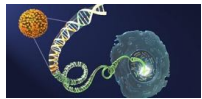
RETROVIRUSES can insert in the host genome



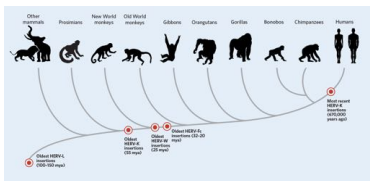
Retroviral (HIV-1) life cycle.

- A. Viral entry and post entry (reverse transcription, DNA synthesis, and integration) events;
- B. B. Viral gene expression (transcription and protein synthesis); C. Virus assembly and release.

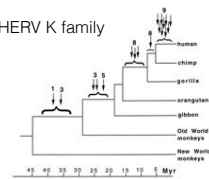
Endogenous retroviruses (ERV)



Around 8 percent of our genetic code stems from HERVs, the bulk of which integrated during primate evolution.



HERV K family



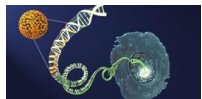
These genomic parasites came from outside and then “went germ line” in their host. When a virus integrates in the genome of its host, it can become part of the host genome. Such viruses are known as endogenous viruses.

HERV are human endogenous retroviruses, these are only found in humans and infected our germ line after the common ancestor with chimpanzees and bonobos. They have now become “us”.

What is an endogenous retrovirus?

A retrovirus that has been incorporated in the host genome.

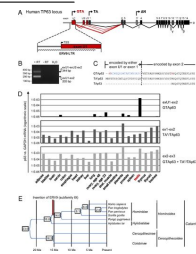
Domesticated ERV: Tp63



Control of male germ cell apoptosis by Tp63, a tumor suppressor gene that became regulated by an endogenous retrovirus (HERV9). guardian function was greatly enhanced by integration of an endogenous retrovirus upstream of the TP63 locus that occurred 15 million years ago. By providing increased germ-line stability, this event may have contributed to the evolution of hominids and enabled their long reproductive periods.

Endogenous retrovirus drives hitherto unknown pro-apoptotic p63 isoforms in the male germ line of humans and great apes

Beyera et al. 2011 *PNAS*



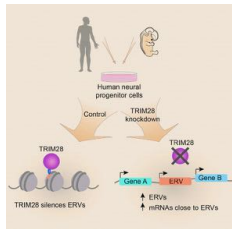
Apoptosis: controlled cell death

pro-apoptotic: favoring cell death

Give an example of a “domesticated endogenous retrovirus”.

HERV9 in Tp63!

Genomic Defenses against ERV



TRIM 28 acts as a transcriptional corepressor of the viral genome.

The protein binds to the histones of the viral chromatin, then recruits a histone methyltransferase that induces heterochromatin formation.

This heterochromatin formation prevents the transcription of the viral genome.

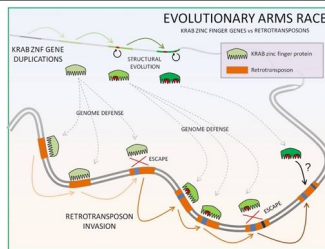
Transcriptional co-repressor: protein that binds to DNA and/or other proteins that bind to DNA and help suppress gene expression (of among other things molecular parasites).

Transcription (literally: “writing across”) is the process by which DNA is transcribed into RNA. Our cells have evolved special proteins that help control the potentially catastrophic spread of selfish DNA elements.

Practice question: How can our genomes defend themselves against molecular parasites such as transposons?

By inactivating the chromatin in which the transposons are located.

Arms Race between selfish DNA & genome!



Transcriptional regulators in mammals. These proteins bind nucleic acids and play important roles in various cellular functions, including cell proliferation, differentiation and apoptosis, and in regulating viral replication and transcription.

There is a host-parasite runaway coevolution (evolutionary arms race) nested within each of our cells.

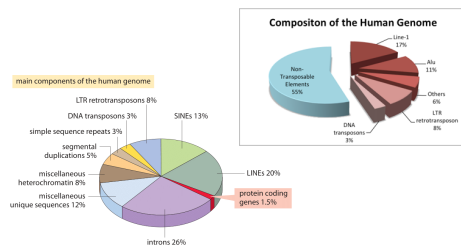
Genomes have to defend themselves against danger of massive increase in copy number and disruption of genes by transposable elements.

Practice Question:

How could there be an “arms race” between selfish DNA and our genomes?

Genomes have to evolve defenses against rampant and disruptive transposon activity but successful transposons evolve ways around these defenses.

8% of Genome comprised of ERV



Repetitive DNA, initially excluded from DNA comparisons, thus only 1.3 % difference between human and chimp DNA.

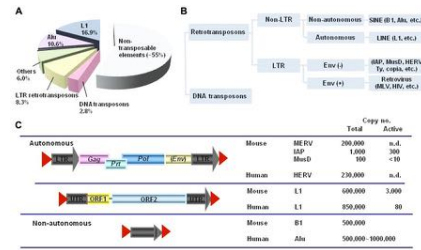
If one includes repetitive DNA (mostly transposable elements) the difference is ~5%.

Practice question:

Why did the figure for genetic difference between humans and chimpanzees recently go from ~1 % to ~5%?

Earlier figure of 1% did not include comparisons of highly repetitive DNA.

8% of Genome comprised of ERV



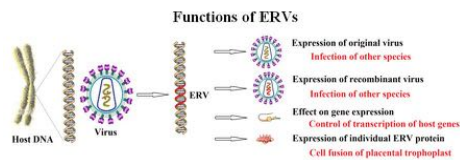
Anatomy of different transposons:

autonomous ones have all the tricks to copy and paste themselves

non-autonomous ones piggy-back on autonomous ones:

Molecular parasites have their own parasites....” babushka doll of parasites”

Endogenous Retrovirus (ERV) function



There are many ways in which ERV can affect gene function of their host cells.

Think of endogenous retroviruses as “mobile tuning buttons” for gene expression. They can often affect gene expression near the places where they insert themselves.

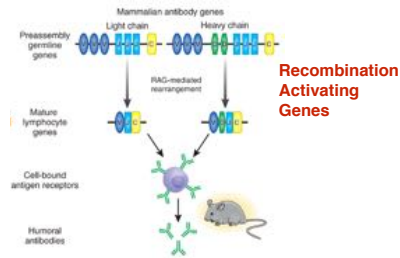
Evolution of immunity coopted transposable elements



Transposons set to work for G.O.D.

The cover focuses on some of the immunities of various animals, such as DSCAM in insects, spRAG1I and spRAG2I in echinoderms, VLR in jawless vertebrates, and various immunoglobulin subtypes in animals that possess the adaptive immune system. The white on grey background displays receptors that play key roles within each of the immune systems, such as TLR for the innate immune system, and TCR, MHCI, and MHCII in the adaptive immune system. Transposons set to work for **Generation Of Diversity!** (G.O.D.)

RAG to Riches
from selfish DNA transposon to basis of our adaptive immune system



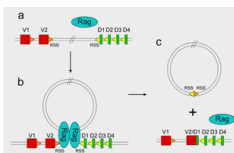
Jawed vertebrate antibody genes are assembled via recombination activating gene (RAG)-mediated joining of immunoglobulin gene fragments consisting of variable (V), diversity (D) and joining (J) elements, as well as constant (C) exons. The mature antigen receptors in both cases are expressed on the surfaces of lymphocytes and can be secreted to the plasma. RAG are encoding recombinase enzymes and appear to have evolved from transposon enzymes (transposases).

Practice Question:

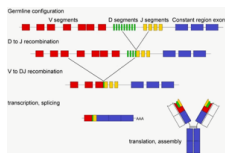
Which cell types in the human body activate recombination activating genes (RAGs)?
B-cells and T-cells of the immune system.

RAG to Riches
from selfish DNA transposon to basis of our adaptive immune system

Mechanism of recombination

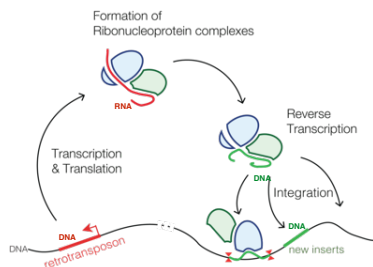


V(D)J Recombination



Jawed vertebrate antibody genes are assembled via recombination activating gene (RAG)-mediated joining of immunoglobulin gene fragments consisting of variable (V), diversity (D) and joining (J) elements, as well as constant (C) exons. The mature antigen receptors in both cases are expressed on the surfaces of lymphocytes and can be secreted to the plasma.

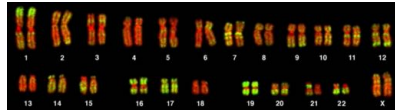
Retrotransposon



Transposons, transposable elements that go from DNA to RNA and back again, they encode their own reverse transcriptase.

These make up the bulk of TE in the human genome.

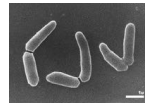
Alu elements in primates



Karyotype from a female human lymphocyte (46, XX).

Chromosomes were hybridized with a probe for Alu elements (green) and counterstained with TOPRO-3 (red).

Alu elements were used as a marker for chromosomes and chromosome bands rich in genes.



Arthrobacter luteus bacteria

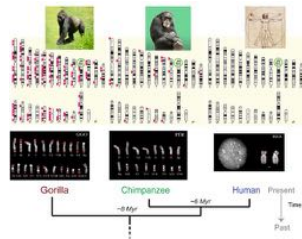
Alu elements are named after *Arthrobacter luteus* bacteria. An enzyme from this bacteria cuts DNA at a sequence carried by all these million of copies of a ~ 300 basepair element.

Practice question:

Why would a bacterial enzyme that recognizes a certain DNA sequence be able to cut the genome of primates into thousands of chunks?

If the recognized sequence is identical to sequence of an Alu element, the enzyme would cut DNA wherever one of these thousands of transposons are across the genome.

Other apes have their own expansions of transposable elements



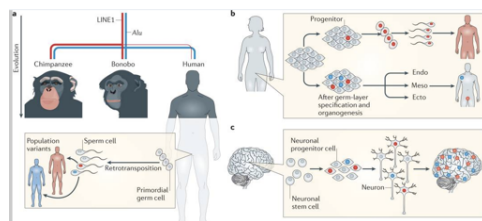
Marques-Bonet et al., Nature, Feb. 12, 2009

Apes have their own expanded elements, that have remained single copy in the human genome. African Great ape genomes are slightly larger as a result!

Practice question: are human and chimpanzee genomes the same size?

No, chimpanzee genomes are slightly larger due to accumulated repetitive sequences.

The necessary junk: new functions for transposable elements.

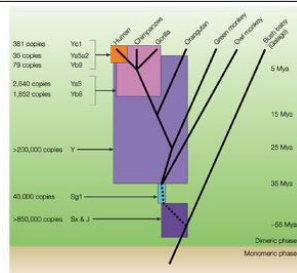


Muotri, et al. Hum. Mol. Genet. (2007)

As humans, chimpanzees and bonobos evolved from a common ancestor, retrotransposons actively mobilized in the ancestral germ lines, which resulted in the generation of genomic variation that natural selection then acted upon. Alu retrotransposition rates (represented by the thickness of the blue line) remained relatively similar between the three species; however, long interspersed nuclear element 1 (LINE1) retrotransposition rates (represented by the thickness of the red line) were suppressed in the human lineage. Retrotransposition of LINE, Alu and SINE-VNTR-Alu (SVA) elements continues to occur in the human germ line, which creates population variants that are present in every cell of an individual's body and are also passed on to future generations. Whether LINE1 or Alu somatic retrotransposition rates differ between human and non-human primates is unknown. b | Somatic retrotransposition can happen at any time during embryogenesis. Retrotransposition events that occur in early pluripotent progenitor cells will result in somatic mosaicism: these unique cells will contribute to all tissues of the body of the

individual, including the germ line. Somatic retrotransposition that happens after germ-layer specification and organogenesis, however, results in germ-layer- or tissue-specific insertions. These will not contribute to the germ line. c | Somatic retrotransposition increases as neural stem cells differentiate into neurons and results in neurons with unique genomes. Variability exists between the rates of retrotransposition and regions in which it occurs between individuals. High rates of retrotransposition events seem to occur in the hippocampus in some individuals

Expansion of *Alu* elements in primates



Batzer & Deiniger 2001 *Nat Rev genet.*

The expansion of *Alu* elements in primates. The expansion of *Alu* subfamilies (Yc1, Ya5a2, Yb9, Yb8, Y, Sg1, Sx and J) is superimposed on a tree of primate evolution. The expansion of the various *Alu* subfamilies is colour coded to denote the times of peak amplification. The approximate copy numbers of each *Alu* subfamily are also noted. Mya, million years ago.

Practice question:

Do we humans have transposable elements that do not exist in the apes?

Yes, a subset of young *Alu* elements for example, have only replicated and inserted after our last common ancestor with apes.

Retrotransposons and Disease

Neurofibromatosis: *de novo* *Alu* insertion

The following human diseases have been linked with *Alu* insertions:

- Breast cancer
- Ewing's sarcoma
- Familial hypercholesterolemia
- Hemophilia
- Neurofibromatosis
- Diabetes mellitus type II

And the following diseases have been associated with single-nucleotide DNA variations in *Alu* elements impacting transcription levels.

- Alzheimer's disease
- Lung cancer
- Gastric cancer

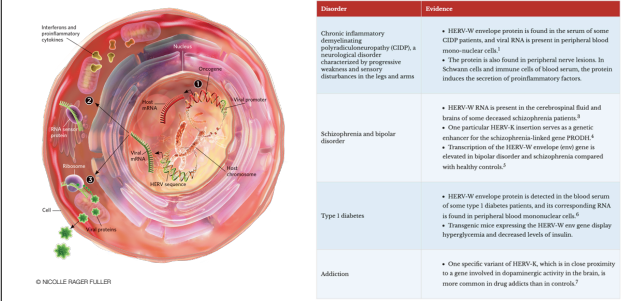


Singer et al Gage Lab Salk Trends in Neurosciences Vol.33 No.8

When a new copy of a transposable element inserts itself in the genome, this often disrupts gene function and can lead to disease.

Give an example of a human disease caused by an *Alu* element: Neurofibromatosis.

Activation of HERV and Disease



Endogenous retrovirus activity and disease:

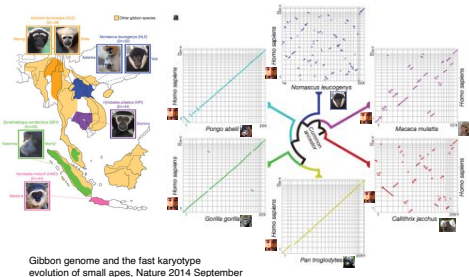
(1) Activation of viral promoters: Ancient retroviral infections have left viral promoters throughout the human genome. Although our bodies have coopted many of them to drive the expression of our own genes, a lot of those promoters are kept silenced through epigenetic repression. Reactivation of these elements can result in abnormal expression of nearby oncogenes or tumor-suppressor genes.

(2) Expression of viral genes: Under some circumstances, such as cancer, many regions of the genome that are normally silenced can awaken. This can activate the transcription of HERVs, causing viral RNA to accumulate in the cytoplasm. According to the “viral mimicry” theory, these molecules alert cellular RNA-sensing pathways to the viral material, triggering an immune response.

(3) Translation of viral proteins: Some viral RNAs are translated into proteins, which can be secreted and travel to other cells. It's unclear what effects these proteins have, but some researchers hypothesize they activate surface receptors and ultimately initiate immune reactions.

1. Mult Scler J, 18:1721–36, 2012; 2. EBioMedicine, 6:190–98, 2016; 3. PNAS, 98:4634–39, 2001; 4. PNAS, 110:19472–77, 2013; 5. Transl Psychiatry, 2: e201, 2012; 6. JCI Insight, 2:e94387, 2017; 7. PNAS, 115:10434–39, 2018

Gibbon Karyotype (chromosome) Evolution driven by transposable elements



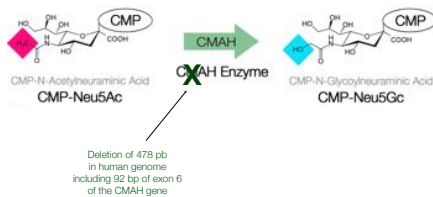
Gibbon chromosomes have been profoundly altered due to millions of copies of a new transposon LAVA, only so far found in gibbons.

Practice question:

How could transposable elements affect ape evolution?

TE such as the LAVA TE in gibbons causes massive chromosome rearrangements. This can potentially cause reproductive incompatibility and with it speciation.

Watershed event:
Loss of **Neu5Gc** and excess of **Neu5Ac**



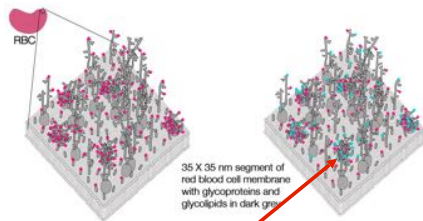
A small and big difference due to an Alu insertion.

Humans for example cannot make a certain form of sialic acid (Neu5Gc) that is very common in other mammals. Humans fixed a mutation that inactivates the enzyme that can modify the precursor sugar Neu5Ac into the derived form Neu5Gc by adding a single oxygen atom.

Practice question: What is the difference between Neu5Ac sialic acid found in all vertebrate and the animal Neu5Gc sialic acid lost in the human lineage?

Answer: the presence of one additional oxygen atom in the non-human Neu5Gc.

Watershed event:
Loss of **Neu5Gc** and excess of **Neu5Ac**



Antibodies against Neu5Gc (= Xenoglycan)

Modified from Vitala & Järnfeldt, 1985

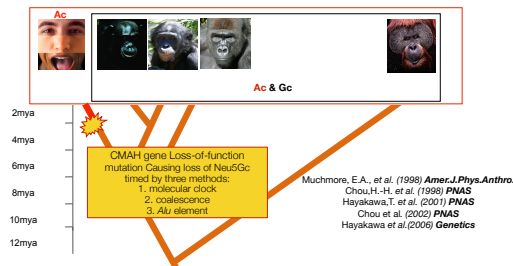
The molecular “flavor” of human cell surface has massively changed, as each cell in our bodies carries hundreds of millions of sialic acid molecules.

Practice question:

How can a mutation caused by an Alu element in one enzyme result in an over-all change of most cell surfaces?

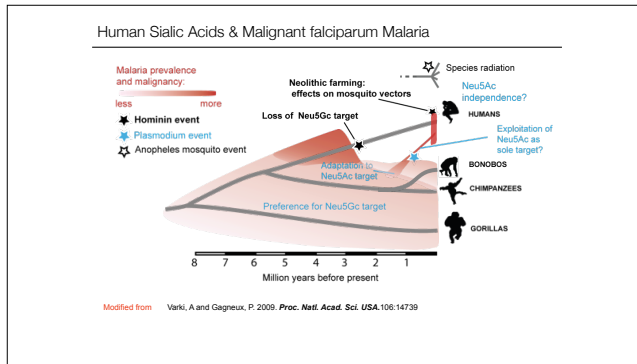
If the enzyme is a glycan modifying enzyme, it would affect millions of glycan chains on most cells.

Human-specific loss of the sialic acid Neu5Gc ~2-3mya



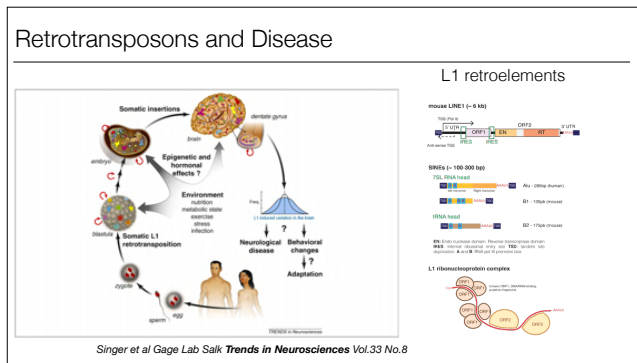
Humans are “genetic knock-outs” for Neu5Gc

Time line for human specific Alu mediated mutation.



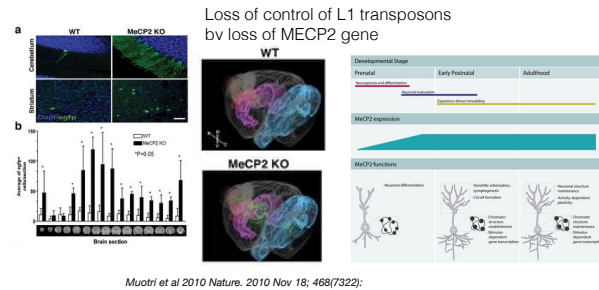
Scenario for the role of ancient malaria in the loss of CMAH function and with it any Neu5Gc sialic acids on cell surfaces targeted by malaria.

Forces internal to the cells: genomic (molecular) parasites such as Alu elements, interact with pathogens and parasites from outside the organism to shape susceptibility to disease!



L1 elements are a common retrotransposon in humans and other apes. Putative implications of L1-mediated somatic mosaicism in the brain. In a reversal of the commonly held belief that retrotransposition occurs primarily in the germline it became clear that L1 elements are expressed in many somatic tissues, including the brain. Recent evidence shows that L1 retrotransposition (curved red arrows) does not occur in the parental germline but in the soma during early embryonic development (colored dots), resulting in individuals that are genetically mosaic with respect to L1 composition. It has been suggested, however, that L1 RNA could be transcribed in the parental germline and carried over in both male and female germ cells in the form of RNPs (black line with red dots) and integrated into the genome at the preimplantation stage (colored spots); however, these events are probably rare because retrotransposons are effectively silenced in the germline through a small-RNA-induced mechanism. Somatic L1 retrotransposition events that occur during embryogenesis would result in clonal sectors of cells (colored patches) that carry the same insertion event. The size of the clonal sector depends on the developmental stage when the insertion occurred and the number of subsequent cell divisions. L1 insertion events that take place during embryonic brain development will be found in different brain regions (colored patches and dots), whereas events that take place during adult neurogenesis will be restricted to specific areas such as the dentate gyrus (insert). According to our hypothesis, L1-induced mosaicism could increase variability in the brain (blue curve), and this could have implications for behavioral phenotypes. The environment could influence regulation of somatic L1 retrotransposition in the brain and this influence could be mediated by epigenetic or hormonal mechanisms. Depending on its impact upon the brain and the consequences, L1-induced somatic variability could either increase the risk of neurological disease or induce behavioral changes that could help the organism to adapt better to changing environments.

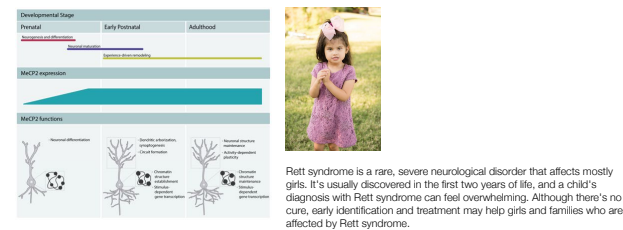
Retrotransposons and Disease: Rett's Syndrome



MECP2 gene encodes a protein that can act as “brake” against L1 transposable elements. Genetically altered mice lacking MeCP2 have much more retrotransposition (visualized with green fluorescent protein engineered into their L1 transposons).

Retrotransposons and Disease: Rett's Syndrome

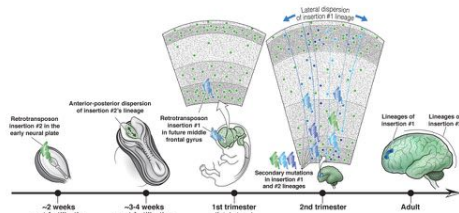
Loss of control of L1 transposons by loss of MECP2 gene (on X-chromosome)



Males do usually not survive to live births, as they only have one X-chromosome and lack of back-up causes pre-natal mortality.

Why is Rett's Syndrome only observed in females?
The X-linked mutation is lethal in males.

Retrotransposons and the brain: somatic variation!



Cell lineage analysis in human brain using endogenous retroelements.

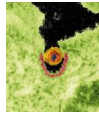
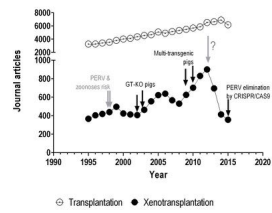
Line1 retrotransposon activity during brain development.

Practice question:

Is the human body a perfect clone of the fertilized egg cell?

No, certain tissue have somatic mutations (often associated with transposable elements), especially neurons.

Xenotransplantation and PERV

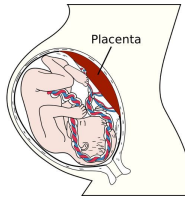
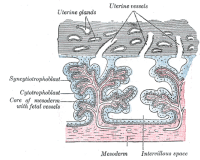


There is justified fear that this could open a Pandora's box of novel viruses. Pig endogenous retroviruses (PERV) could reactivate once a pig organ is transplanted into an immune-suppressed patient. Such events could potentially lead to the emergence of completely new pathogens.

How could xenotransplantation represent health risks for the general human population?
By allowing animal viruses to infect the immune suppressed human patient, to adapt and then to infect other healthy humans.

[illegible]

cell fusion: syncytium
and immune suppression by
viral glycoprotein



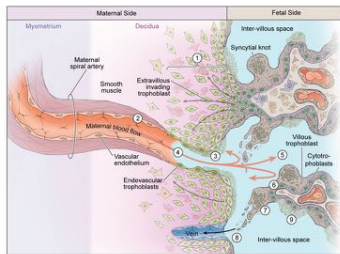
Placentation and Domesticated Retroviruses

Syncytin 1 and 2 are captive viruses:

cell fusion: syncytium
and immune suppression by
viral glycoprotein

The diagram illustrates the placental interface, showing the Maternal Side (Myometrium, Decidua) and the Fetal Side (Inter-villous space, Syncytial knot, Villous cytotrophoblasts, Chorion-phoblasts, Villous space). Key structures labeled include the Maternal spiral artery, Extravillous invasive trophoblasts, Maternal blood flow, Intervillous endothelium, Endovascular trophoblasts, and a Vein. The diagram also shows the Syncytial knot and Villous cytotrophoblasts. Numbered circles 1 through 6 indicate specific points of interaction and viral components.

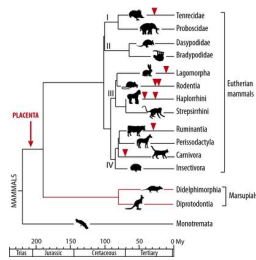
cell fusion: syncytium
and immune suppression by
viral glycoprotein



Invasive mammalian placentas have evolved in part by “co-opting” endogenous retroviruses for syncytium formation and immune evasion of maternal immunity.

Recurrent capture of viral proteins

Mammals repeatedly captured viral genes to make better placentas



Proc Natl Acad Sci U S A. 2015 Feb 3; 112(5):

Phylogeny of mammals and previously identified syncytin genes. Different endogenous virus genes have been co-opted many times independently for placentation. Mammals comprise the monotremes, the marsupials, and the eutherians, the latter comprising the four major clades: Afrotheria (I), Xenarthra (II), Euarchontoglires (III), and Laurasiatheria (IV) (data from ref. 1). Branch length is proportional to time (in My), and the time of insertion of the different syncytins identified to date is indicated with arrowheads.

Baby food from defensive enzyme:

Lactase synthase complex includes modified defensive protein: lysozyme.

Lysozymes are glycan-cleaving defensive enzymes that degrade bacterial cell walls.



Talking about co-option: Mammalian milk was “invented” by the cooption of an anti-bacterial glycan cleaving enzyme (lysozyme) into a novel glycan synthesis (lactase synthase, a complex between a lysozyme and a glycosyl transferase) enzyme.

Practice question:

Describe the evolutionary invention of milk:

Ancestral mammals evolved milk as a nutritious food for their young from modified sweat glands. They used a modified anti-bacterial enzyme to evolve a lactose synthesizing enzyme, making a sugar (disaccharide lactose) that is easy to share between mother and infant but difficult to “steal” by most microbes.

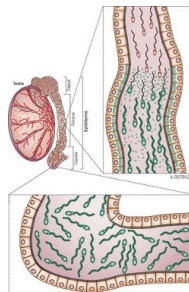
What is evolutionary bricolage (François Jacob)?

“Evolutionary tinkering”: novelty can be pieced together from different existing features.

Sperm invisibility cloak from defense

beta defensin 126 glycoprotein “hides” sperm from female immune system.

Defensin proteins are hole-punching defense tools against bacteria.



Another co-option: Defensin 126, that is related to anti bacterial “hole-punching proteins” is used to provide immunological invisibility to human sperm.

Summary

- DNA can duplicate independently of the genome, giving rise to molecular parasites!
- "selfish DNA" jumping genes, transposons and retrotransposons clutter the genome.
- Retroviruses can integrate and become part of the genome.
- Humans and other species have co-opted several transposons and endogenous retroviruses during their evolution (evolutionary tinkering!).
- Transposition can cause disease but also adaptation.
- Adaptive immunity became possible thanks to transposon genes: RAG 1 and RAG2.
- Invasive placentation became possible thanks to viral genes: Syncitin 1 and 2.
- Genomes have to defend themselves against molecular parasites.
- Male germ cells use viral genes for quality control (Tp63).

