Article

The Tree of Life: Tracing the Genetic Pathway from the Last Universal Common Ancestor to Homo Sapiens (Part III)

Chris King*

ABSTRACT

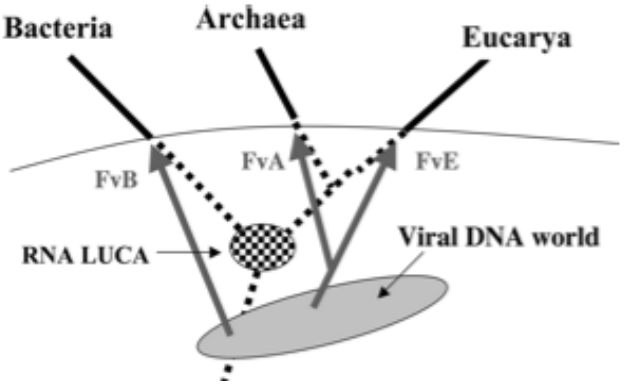
This series of articles are fully referenced research reviews to overview progress in unraveling the details of the evolutionary Tree of Life, from life's first occurrence in the hypothetical RNA-era, to humanity's own emergence and diversification, through migration and intermarriage, using research diagrams and brief discussion of the current state of the art. The Tree of Life, in biological terms, has come to be identified with the evolutionary tree of biological diversity. It is this tree which represents the climax fruitfulness of the biosphere and the genetic foundation of our existence, embracing not just higher Eucaryotes, plants, animals and fungi, but Protista, Eubacteria and Archaea, the realm, including the extreme heat and salt-loving organisms, which appears to lie almost at the root of life itself. To a certain extent the notion of a tree based on generational evolution has become complicated by a variety of compounding factors. Gene transfer is not just vertical carried down the generations. There is also evidence for promiscuous incidences of horizontal gene transfer, genetic symbiosis, hybridization and even the formation of chimeras. This review will cover all these aspects, from the first life on Earth to Homo sapiens.

Part III of this article includes: 7. Viral Influences on the Nuclear Genome; 8. The Symbiotic Face of Eukaryote Mobile Elements; 9. Endogenous Retroviruses and the Placenta ; 10. The Cambrian Radiation, Homeotic Genes, Metamorphosis and Hybridization; & 11. Mammalian Radiative Adaption: Traditional DNA versus Micro RNAs.

Key Words: tree of life, genetic pathway, common ancestor, Homo Sapiens, biological diversity.

7. Viral Influences on the Nuclear Genome

Fig 17: Proposed viral contribution of DNA polymerases FvA etc. founder viruses (Forterre). Other cellular and viral genealogies are possible and the scheme is merely representative.



* Correspondence: Chris King http://www.dhushara.com E-Mail: chris@sexualparadox.org

Forterre looks likewise to a three component origin, but his emphasis is on the idea that viruses have contributed major components to the genome of all three groups, possibly providing each of three RNA-based cell lineages with independent transitions to DNA-based genomes by contributing DNA-polymerases, thus radically improving the stability and competitiveness of these cell lines who became the eventual survivors. In addition to the ribosomal proteins and rRNAs having distinct qualitative features in each domain, many DNA informational proteins exist in different nonhomologous families (usually with several versions for one family). There are already six known nonhomologous families of cellular DNA polymerases. In the case of DNA polymerases of the B family, there is one version in Bacteria (only found in some proteobacteria), one in Archaea, and several in Eukarya. The distribution of the different versions and families of cellular DNA informational proteins among domains is erratic most of the time and does not fit with any of the models proposed for the universal tree, suggesting abrupt insertion into the cellular genomes by viral transfer.

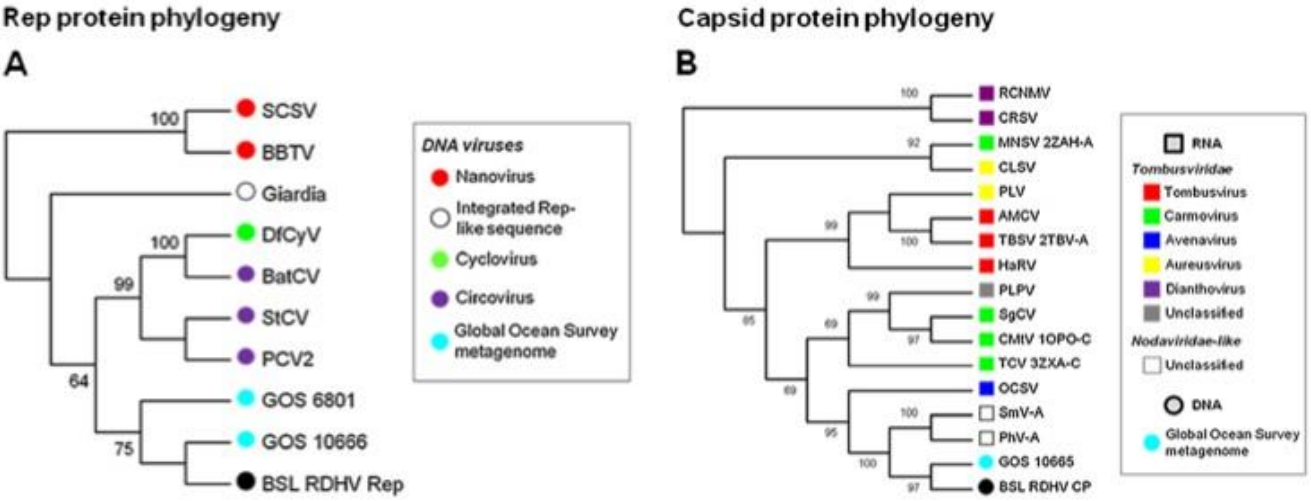


Fig 17b: The differing DNA and RNA viral taxomomy of the Rep and capsid genes of BSL RHDV (Diemer & Stedman 2012).

A very unusual virus has given an indication how the transfer from RNA to DNA could have meen mediated by viruses. Although viruses are very promiscuous, they generally only recombine with viruses of s similar type or at least the same mode of replication. Thus until recently no instances were known of viral recombination bridging the three major groups: RNA viruses, DNA viruses and retroviruses encoding DNA from RNA insktructions by reverse transcriptase. However a chimeric circular, putatively single-stranded DNA virus BSL RHDV encoding a major capsid protein similar to those found only in single-stranded RNA viruses was discovered in a hot acidic lake (Diemer & Stedman 2012). They also found that something very similar had turned up in samples of ocean water sequenced by a team led by Craig Venter. This gives the beginning of an explanation how DNA-based RNA-viral genes in endemic viruses, presumably via reverse transcriptase, have made multiple RNA to DNA transitions of other viral and cellular genes, probably when RNA, DNA and retroviruses cohabited cells.

Viruses such as phages appear to have no evolutionary tree, with genomes across widely diverse habitats consisting of cut and paste components, implying viral adaption has resulted almost entirely from

utilization of advantageous genes from horizontal transfer. Around 10% of all bacterial genes sequenced to date consist of ORFans that bear no resemblance to genes seen anywhere else, suggesting horizontal viral origin (Hamilton).

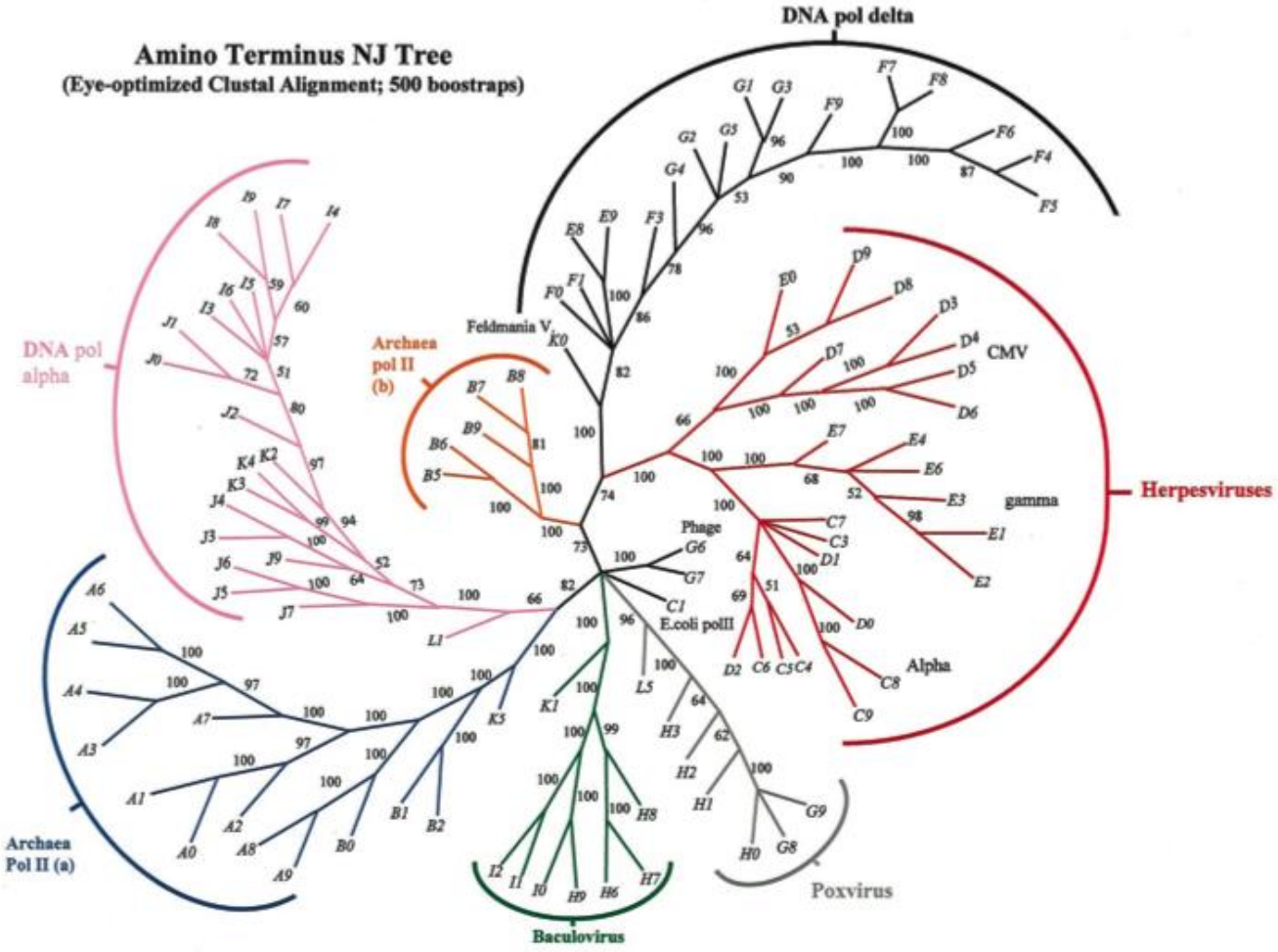


Fig 18: Evolutionary tree of DNA polymerase amino termini (Villareal and Defilippis)

Villareal and Defilippis have likewise investigated the idea that DNA viruses are the origin of DNA replication proteins, by investigating the amino terminus and constructing an evolutionary tree which shows DNA polymerases of DNA viruses, eukaryotes (&alpha,&delta), archaea, E coli and two phages rooted in a tree consistent with a viral origin.

This idea has a great deal of plausibility because viruses are now know to have a potentially primal origin, rather than being recent escapees from cellular genomes which have undergone reductive parasitic changes to their genome. Viruses clearly also have retained both RNA-RNA, DNA-DNA and retrotranscription DNA-RNA-DNA using both RNA and DNA stages in their capsid viral forms, so they retain all the transitional states between RNA and DNA-based replication.

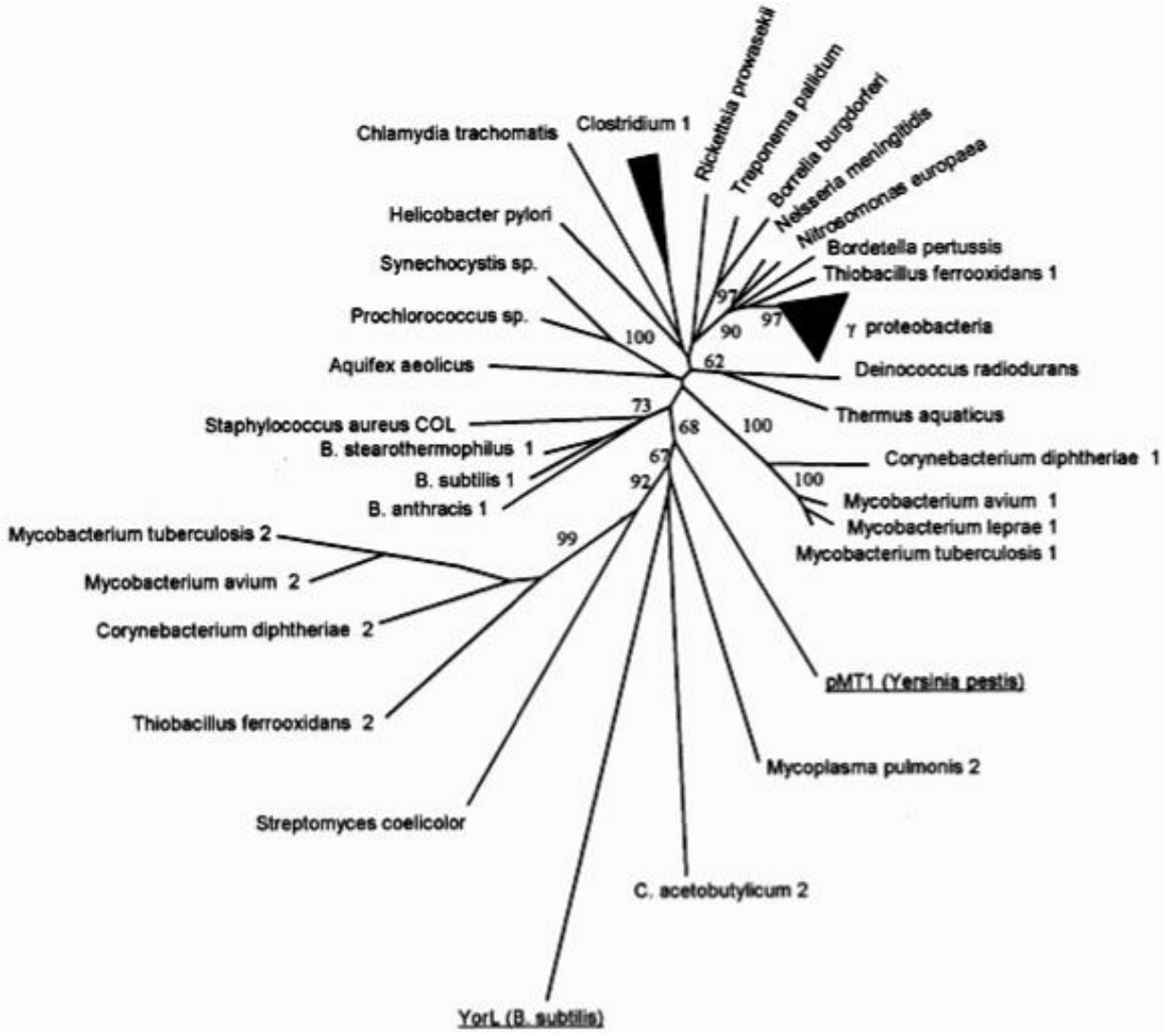


Fig 19: Bacterial DNA polymerases also show viral members (underlined) close to the root of the tree (File et. al.).

Furthermore the retroviruses and related mobile genetic elements have a common ancient evolutionary origin, which is related to telomerase, which itself uses an RNA primer to initiate chromosome duplication. There is thus a plausible case that telomerase is in fact a biological fossil of a retroviral conversion of the founding Eukaryote cell line to a DNA genome.

8. The Symbiotic Face of Eukaryote Mobile Elements

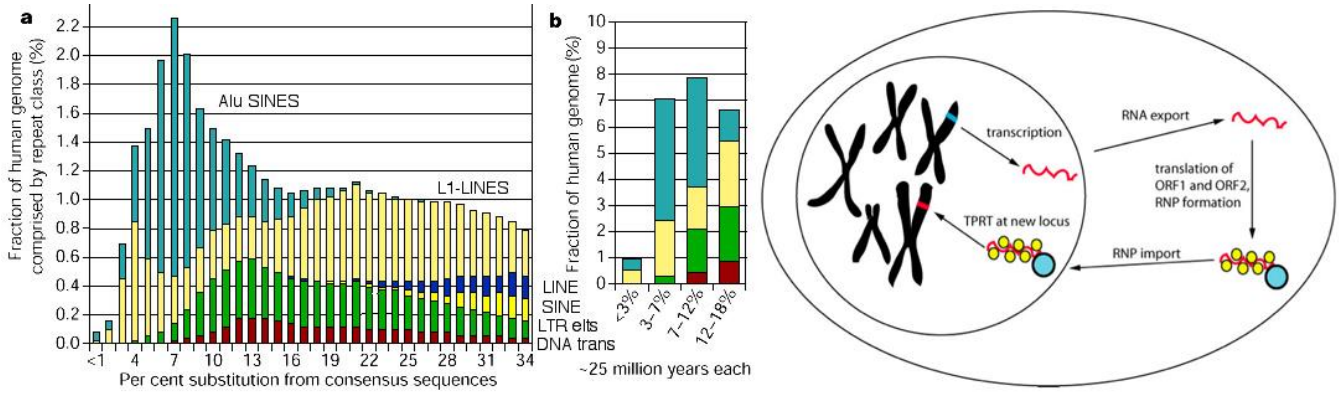
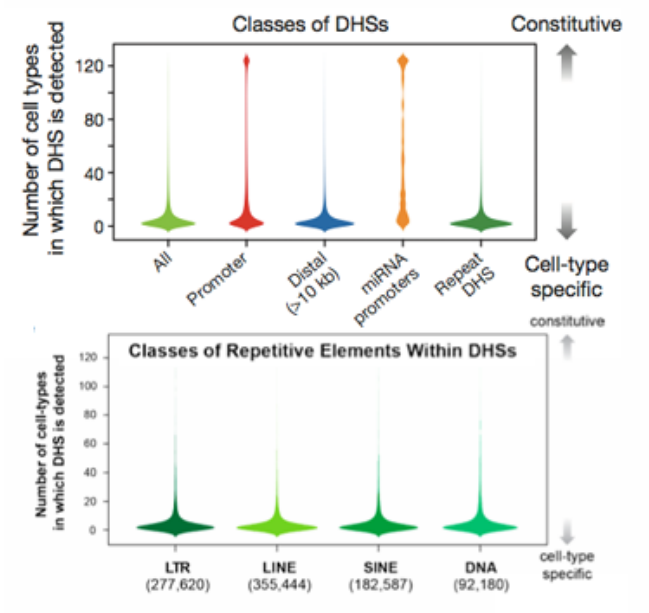


Fig 20: (Left) Human transposable element evolutionary history of L1-LINEs (cream), Alu elements (lt. blue), retrovirus-like LTR (long-terminal repeat) elements (green) and DNA transposons (dark brown). Older L2-LINEs and SINEs are in yellow and dark blue. This history extends back over 200 million years indicating the very ancient basis of this potentially symbiotic relationship (Human Genome Consortium). Comparison with the mouse genome can be accessed at Waterston et. al. (2002). (Right) L1 replication: L1 is transcribed in open reading frames ORF1, an RNA-binding protein, and ORF2 an endonuclease/reverse-transcriptase. The bound RNA-protein complex RNP is transported to the nucleus where target-primed reverse transcription to chromosomal DNA takes place (Han and Boeke).

In 1978, following the work of Darryl Reanny (1974-6), I proposed (1978, 1992) that viruses and transposable elements, far from just being selfish genes (Dawkins 1976), formed part of a dynamical system of genetic symbiosis between the hosts and the mobile genetic elements, because the mobile elements permitted forms of coordinated gene expression and the formation of new genes in a modular manner, which would otherwise be impossible, achieving in return perpetuation of their own genomes over evolutionary time scales. Most of the details of this proposal have proved to be realized. The ENCODE project has demonstrated involvement of all the major classes of human transposable element in regulatory enchancer activity, most specific to a single cell type (Thurman2012).

By some reckonings, 40 to 50 per cent of the human genome consists of DNA imported horizontally by viruses, some of which has taken on vital biological functions. Taken together, virus-like genes represent a staggering 90 per cent of the human genome (Hamilton). Coding sequences comprise less than 5% of the human genome, whereas repeat sequences account for at least 50% and probably much more. Transposable LINE or long-intermediate repeat retroelements, common to mammals (Han and Boeke), and insects (Jensen et al, Sheen et al) with a history running back to the Eukaryote origin are specifically activated in both sperms and eggs during meiosis (Branciforte and Martin, Tchénio et. al., Trelogan and Martin), although subjected to down regulation by interfering piRNAs (Aravin et al). They replicate from transcribed RNA copies of themselves thus using RNA to instruct DNA copies, indicating an origin in RNA-based life, as does the active RNA processing of our own Eukaryote cells. Their RNA-based reverse transcriptase shows homologies with the telomerase essential for maintaining immortality in our germ line, indicating a common and symbiotic origin. 100,000 partially defective LINEs, around 100 of which remain fully active in humans, and their 300,000 dependent smaller fellow

traveller Alu SINEs make up a significant portion of the human and mammalian genomes, along with pseudogenes, apparently defective copies of existing genes translocated by elements such as LINEs.



Chromosome	Repetitive element name	Repetitive element family	Repetitive element class
chr1	L2a	L2	LINE
chr1	L2c	L2	LINE
chr1	L2c	L2	LINE
chr2	MLT1F2	ERVL-MaLR	LTR
chr5	L2c	L2	LINE
chr6	(CA)n	Simple_repeat	Simple_repeat
chr7	MLT1N2	ERVL-MaLR	LTR
chr7	Tigger15a	TcMar-Tigger	DNA
chr7	Tigger15a	TcMar-Tigger	DNA
chr8	MLT1I	ERVL-MaLR	LTR
chr9	MIR3	MIR	SINE
chr9	FordPrefect	hAT-Tip100	DNA
chr13	L1PA15-16	L1	LINE
chr14	AluJr4	Alu	SINE
chr14	HERV3-int	ERV1	LTR
chr15	L2b	L2	LINE
chr21	MLT2A2	ERVL	LTR
chr21	MER34C2	ERV1	LTR
chr21	L1MB7	L1	LINE
chr21	AluJb	Alu	SINE
chr21	MLT1K	ERVL-MaLR	LTR

Fig 20b: ENCODE data showing involvement of the major classes of transposable element in enhancer activity (Thurman 2012).

These elements travel passively down the germ line with chromosomal DNA, so their specific activation during meiosis suggests they may perform a role of coordinated regulatory mutation. This suggests that the type of symbiotic sexuality embraced by bacteria and plasmids also continues to function in higher organisms in a form of sexual symbiosis between our chromosomes and transposable genetic elements. This is consistent with the 1.4% point mutation divergence between humans and chimps, being overshadowed by an additional 3.9% divergence to 5.4% overall (Britten), when insertions and deletions are accounted.

SINEs, such as human Alu, a free-rider on the LINE reverse transcriptase derived from the small cellular RNA used to insert nascent proteins through the membrane, are in turn implicated in active functional genes (Reynolds, Schmid) particularly some involved in cellular stress reactions, again suggesting genetic symbiosis. Humans have about 13 times as many RNA edits as non-primate species, including inosine insertions associated with Alu elements, as well as intron deletions (Holmes) and newly inserted exons (Ast), which may differentiate humans from other apes through alternative splicing of genes expressed in the brain. RNA editing is abundant in brain tissue, where editing defects have been linked to depression, epilepsy and motor neuron disease. There is a new Alu insert about every 100 births. As many as three quarters of all human genes are subject to alternative splice editing.

Recent explosion of the area of interfering miRNAs as regulatory elements in gametogenesis and development (Großhans) has provided an explanation of how pseudogenes, including those retrotransposed via LINE elements, can gain functional regulatory significance even though they do not produce translatable mRNAs.

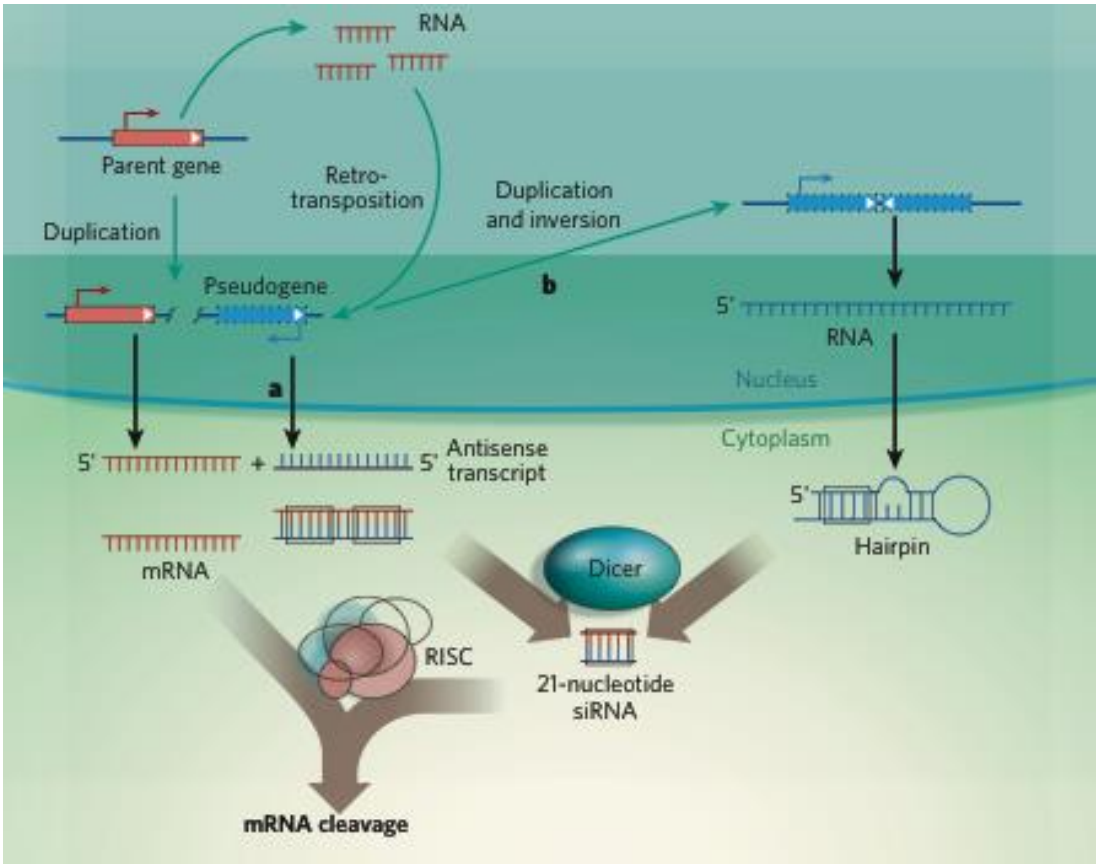


Fig 21: Pseudogene-mediated production of endogenous small interfering RNAs (endo-siRNAs). Pseudogenes can arise through the copying of a parent gene (by duplication or by retrotransposition). (a) An antisense transcript of the pseudogene and an mRNA transcript of its parent gene can then form a double-stranded RNA. (b) Pseudogenic endo-siRNAs can also arise through copying of the parent gene as in a and then nearby duplication and inversion of this copy. The subsequent transcription of both copies results in a long RNA, which folds into a hairpin, as one half of it is complementary to its other half. In both a and b, the double-stranded RNA is cut by Dicer into 21-nucleotide endo-siRNAs, which are guided by the RISC complex to interact with, and degrade, the parent gene's remaining mRNA transcripts. The mRNA from genes is in red and that from pseudogenes is in blue. Green arrows indicate DNA rearrangements (Sasidharan and Gerstein).

Although the data from the human genome project indicated that human LINEs are becoming less active as a group by comparison with the corresponding elements in the more rapidly evolving mouse genome, there remain about 60 active human LINE elements which are known to be responsible for mutations in humans. More recent investigation (Boissinot et. al.) shows that the most recent families are highly active. Around four million years ago shortly after the chimp-human split, a new family Ta-L1 LINE-1 emerged and is still active, with about half the Ta insertions being polymorphic, varying across human populations. Moreover 90% of Ta-1d, the most recent subfamily are polymorphic, showing highly active lines remain present. LINEs are more heavily distributed on the sex chromosomes with X chromosomes containing 3 times as many full length potentially active elements and the Y chromosome 9 times as many! This is consistent with a continuing mutational load on humans which is removed more slowly

from the sex chromosomes by crossing over in proportion to the degree to which crossing over is inhibited in each (i.e. totally on the Y and largely in males in the X but not in females). Sexual recombination is a protection from mutational error in a process called Muller's ratchet.

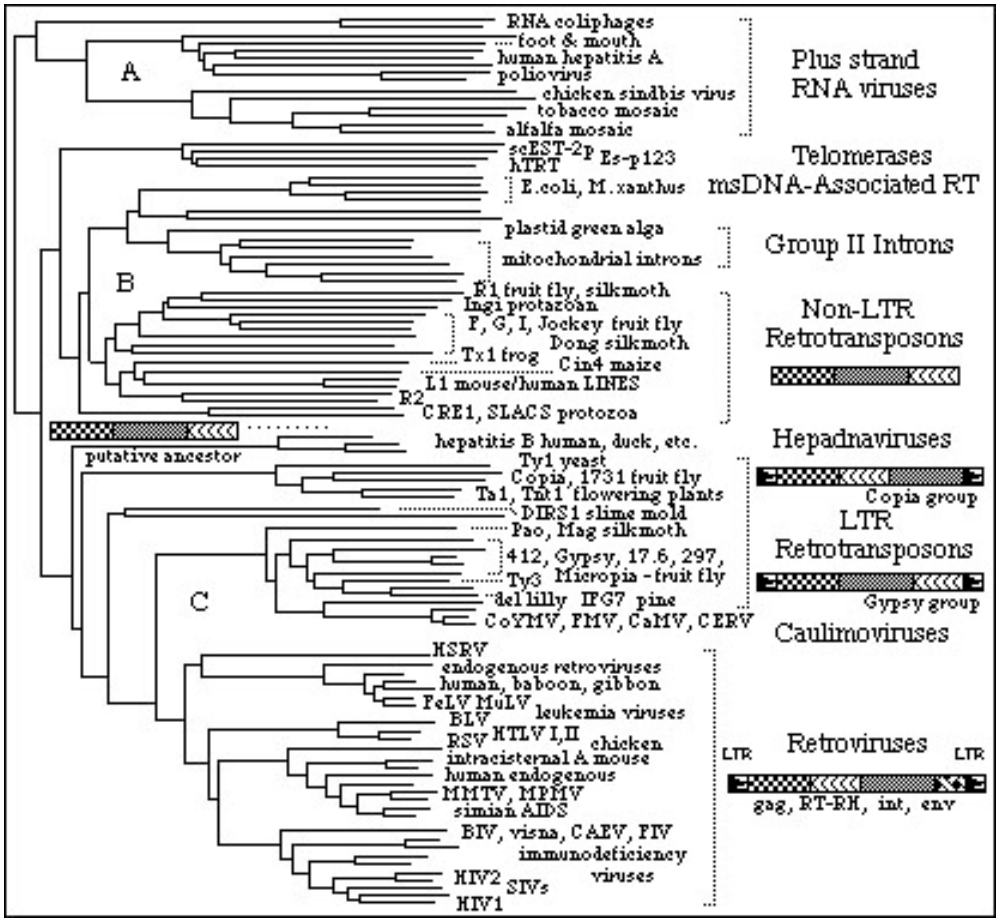


Fig 22: Evolution of reverse transcriptases from a common ancestor bearing a LINE archetype (Xiong and Eickbush,Nakamura et. al.). The root of their evolution goes back to the transfer from RNA to DNA at the beginning of life. They form a complementary evolutionary tree to that of cellular life as genetic symbionts of metazoa travelling down the germ line. Their group includes telomerases essential to the reproductive cycle.

LINEs are preferentially expressed in both steriodogenic and germ-line tissues in mice (Branciforte and Martin, Trelogan and Martin), suggesting stress could interact with meiosis. L1 expression occurs in embryogenesis, at several stages of spermatogenesis including leptotene, and in the primary oocytes of females poised at prophase 1. Conversely the SRY-group male determining gene SOX has been found to regulate LINE retrotransposition (Tchénio et. al.). Similarly LINE elements have been proposed to be 'boosters' in the inactivation of one X chromosome that happens in female embryogenesis (Lyon). This could enable somatic stress to have a potential effect on translocation in the germ-line which might enable form of genetic adaption in long-lived species such as humans. They have diverse means both to cause mutational damage and novel alleles (Han and Boeke).

Both L1 and Alu elements may be able to self-regulate rates of replication, through the existence of stealth drivers, viable elements which maintain a low transcription rate of active elements, with little genomic impact and hence little negative selection. These occasionally seed daughter master elements, which may replicate actively to form new families when conditions permit. This picture is consistent with long periods of quiescence, punctuated by bursts of 'saltatory' replication leading to large copy numbers (Han et. al.).

Further evidence of a symbiotic relationship comes from Drosophila telomeres, which are maintained by the non-LTR retrotransposons, the Line-like TART (Jensen et al, Sheen et al) and HeT-A (Biessmann et al). Likewise the recombination activating gene protein RAG1/RAG2, essential for the mutational variability of the vertebrate immune system, appears to have evolved from an ancient DNA transposon common to the metazoa (Agrawal et al,). Significant similarities exist in the catalytic proteins of Hermes hAT transposase in insects, the V(D)J recombinase RAG, and retroviral integrase superfamily transposases, thereby linking the movement of transposable elements and V(D)J recombination (Zhou et al).

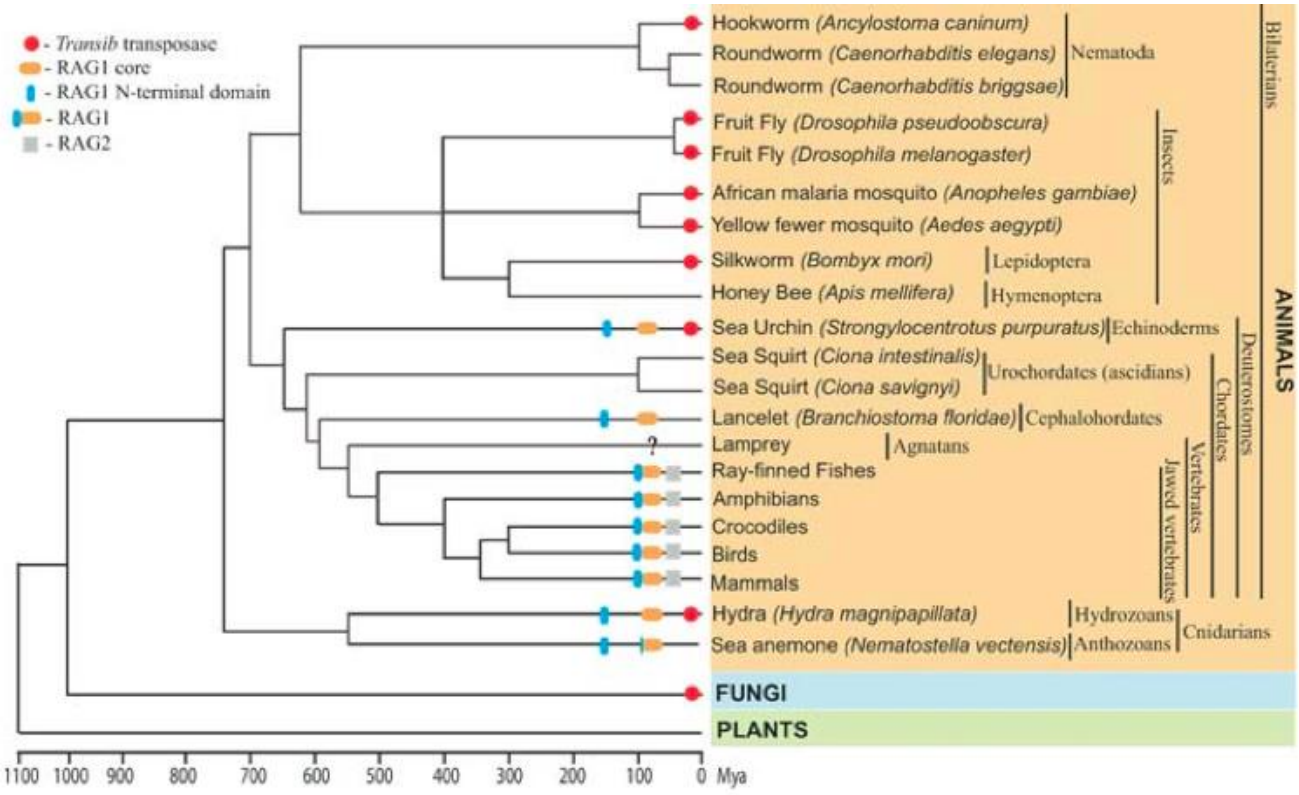


Fig 22b: Transib evolutionary tree spans the eucaryotes (Kapitonov and Jurka).

The approximately 600-amino acid ‘‘core’’ region of RAG1 required for its catalytic activity is significantly similar to the transposase encoded by DNA transposons of the Transib superfamily discovered recently based on computational analysis of the fruit fly and African malaria mosquito genomes. Transib transposons also are present in the genomes of sea urchin, yellow fever mosquito, silkworm, dog hookworm, hydra, and soybean rust (Kapitonov and Jurka).

Looking at the varying pace of evolutionary change in a very ancient gene family and one of the largest in the human genome, the G-protein linked receptor family has roots going back to the first eucaryotes, with two major types of serotonin receptor 5HT1 and 5HT2 diverging before the molluscs, arthropods and vertebrates diverged and originating between 750 million and 1 billion yeas ago. Consequently serotonin functions in mood and circadin rhythms in a similar manner in insects and humans. The diversity of neurotransmitters in humans particularly the amines serotonin, dopamine, norepinephrine and histamine and the amino acids glutamate, gamma-amino-butyric acid and glycine originate from the need of single celled eucaryotes to communicate major pathways of strategic survival from nutrition through aversion to reproduction and sporulation. The family also includes the opsins of visual perception, development receptors and a diverse array of olfactory receptors which are evolving far more rapidly. Serotonin appears with the first photosynthetic bacteria that used tryptophan to hold the porphyrin reactive center and continues, along with melatonin to play a crucial role in light and circadian cycles in humans, as well as mood and social responsiveness.

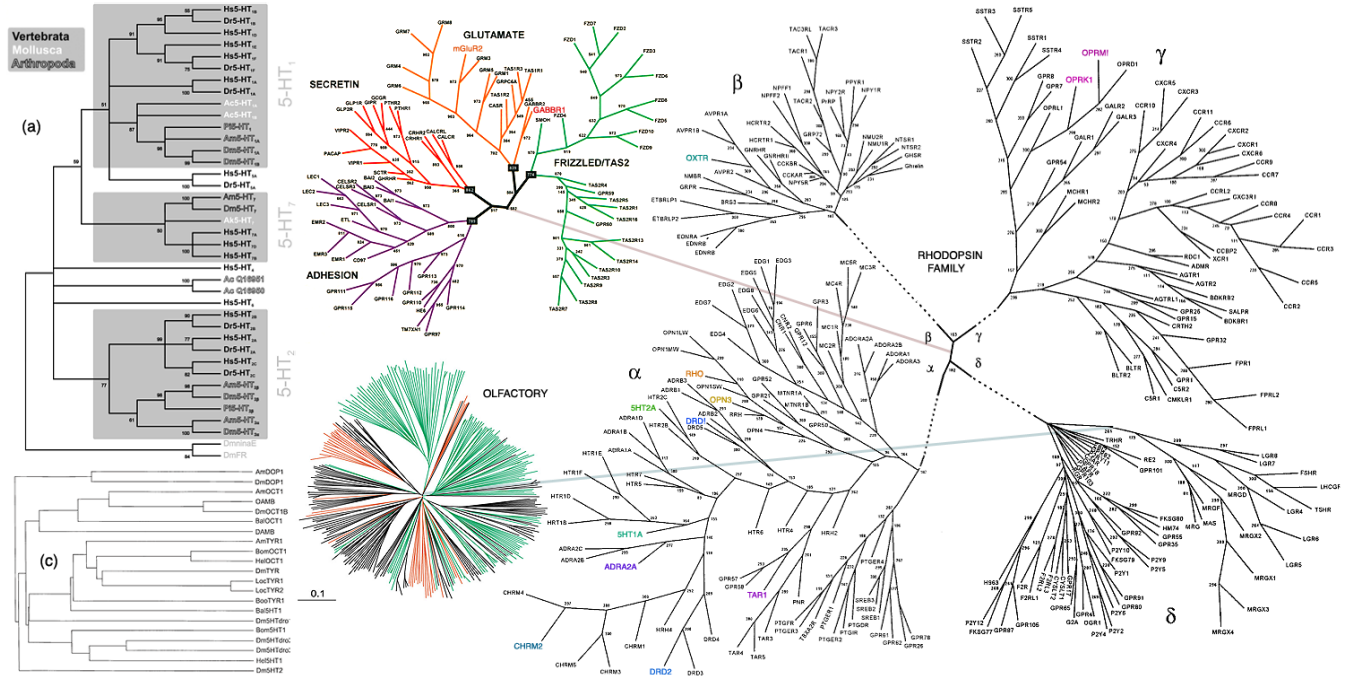


Fig 22c: (a) The two major serotonin receptor types 5HT1 and 5HT2 separated before the molluscs, arthropods and vertebrates diverged (Blenau & Thamm). (b) Evolutionary tree of the human G-protein linked receptors with examples highlighted incolor. On the α branch are amine receptors - serotonin 5HT1A and5HT2A, dopamine D1, and D2 (DRD1, DRD2), adrenergic α2a (ADRA2A), muscarinic acetylcholine (CHRM2), trace amine TAR1, as well as rhodopsin (RHO) and encephalopsin (OPM3). On the glutamate branch are metabotropic glutamate mGluR2 and GABA GABBR1. On the β branch is oxytocin (OXTR) surrounded by vasopressin receptors and Ghrelin. On the γ branch are opioid κ and μ (OPRK1, OPRM1). Olfactory and the non-rhodopsin receptors are linked to their respective points on the rhodopsin family tree. (Fredriksson R et al, Zozulya S. et al). (c) Insect tree of receptors for serotonin, dopamine, tyramine and octopamine neurotransmitters (Blenau & Baumann).

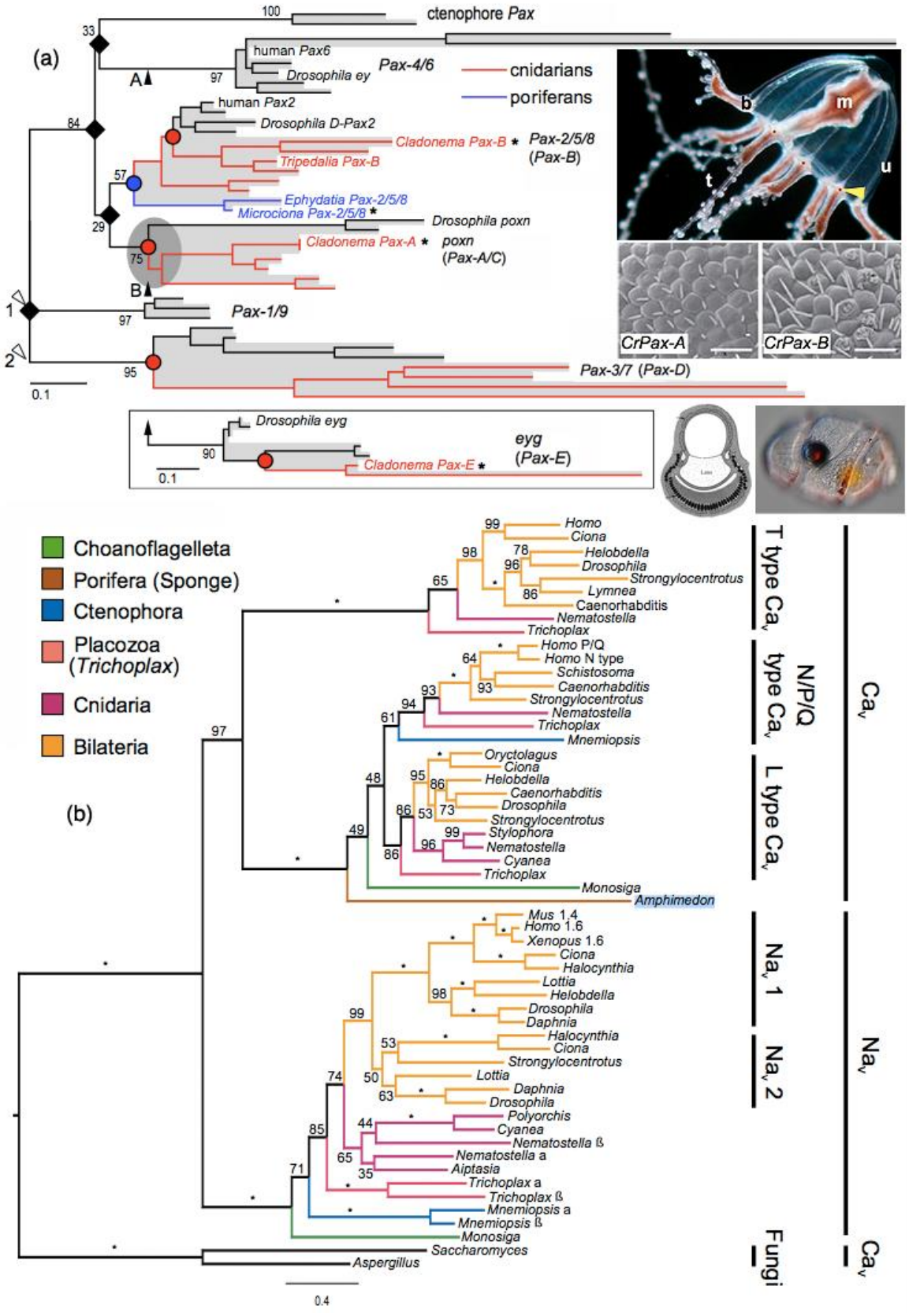

(a)
ctenophore Pax
human Pax6
Drosophila ey
Pax-4/6
cnidarians
poriferans
human Pax2
Drosophila D-Pax2
Cladonema Pax-B
Tripedalia Pax-B
Pax-2/5/8 (Pax-B)
Ephydatia Pax-2/5/8
Microciona Pax-2/5/8
Drosophila poxn
Cladonema Pax-A
poxn (Pax-A/C)
Pax-1/9
Pax-3/7 (Pax-D)
Drosophila eyg
Cladonema Pax-E
eyg (Pax-E)
b
m
t
u
CrPax-A
CrPax-B
Choanoflagelleta
Porifera (Sponge)
Ctenophora
Placozoa (Trichoplax)
Cnidaria
Bilateria
(b)
Homo
Ciona
Helobdella
Drosophila
Strongylocentrotus
Lymnea
Caenorhabditis
Nematostella
Trichoplax
T type Cav
Homo P/Q
Homo N type
Schistosoma
Caenorhabditis
Strongylocentrotus
Nematostella
Trichoplax
Mnemiopsis
N/P/Q type Cav
Oryctolagus
Ciona
Helobdella
Caenorhabditis
Drosophila
Strongylocentrotus
Stylophora
Nematostella
Cyanea
Trichoplax
Monosiga
Amphimedon
L type Cav
Cav
Mus 1.4
Homo 1.6
Xenopus 1.6
Ciona
Halocynthia
Lottia
Helobdella
Drosophila
Daphnia
Nav 1
Halocynthia
Ciona
Strongylocentrotus
Lottia
Daphnia
Drosophila
Nav 2
Polyorchis
Cyanea
Nematostella ß
Nematostella a
Aiptasia
Trichoplax a
Trichoplax ß
Mnemiopsis a
Mnemiopsis ß
Monosiga
Nav
Saccharomyces
Aspergillus
Fungi
Cav
0.4

Fig 22d (a) Common involvement of PAX genes in eye formation from jellyfish, insects and vertebrates suggests a single common origin despite the differing mechanisms. The jelly fish pax genes, like mouse pax-6, induce ectopic compound eyes in fruit fly (right) (Suga et al, Kozmik et al). The small camera eyes of a jellyfish are shown top right (yellow arrow). It has even been suggested that the jellyfish could have gained the eye development pathway through symbiosis with certian single-celled dinoflagellates which possess an eyespot ocelloid, complete with lens and retinoid organelle (lower right) and may have in turn inherited this functionality from cyanobacterial chloroplasts via red algae (Pennisi, Keim) (b) Evolutionary diversification of Na^{+} channels from Ca^{++} channels, essential for the action potential, appears to have occurred before the existence of nervous systems in founding single-celled eucaryotes leading to the metazoa before the choanoflagelates such as monosiga (Liebeskind et al).

The evolution of the metazoan sodium channel essential for the neuronal action potential, from the Calcium channel shared by fungi and animals ocurred in single celled eucaryotes before the metazoa evolved from choanoflagellate-like ancestors. Metazoan eyes also appear to have a common origin, as indicated by the capacity of both jellyfish and mouse pax genes to elicit ectopic compound eyes on fruit files.

9. Endogenous Retroviruses and the Placenta

Endogenous retroviruses, or ERVs, which also travel down the germ line as free-riders, although some may retain infectious capacity, may be essential for placental function, as every mammal tested has placental blooms of endogenous retroviruses which appear to both aid the formation of the syncytium, the super-cellular fused membrane that enables diffusion from the mother to the baby and the immunity suppression, which prevents rejection of the embryo, both characteristics of retroviruses such as HIV.

Mi and colleagues (2000) found a placental gene whose sequence was homologous to several retroviral envelope proteins. The sequence, now called syncytin, is identical to the envelope protein of the HERV-W retrovirus (Blond et. al.) which exists in around 40 apparently defective viral copies, including those in which the two syncytin viral *env* genes are fully functional (Mi et. al.). Syncytin is expressed at high levels in the syncytiotrophoblast (and at low levels in the testes) and nowhere else. Most of the other genes of the provirus have been mutated, suggesting that the envelope glycoprotein function was specifically selected. If cultured cells are made to express syncytin, they will fuse together, and this fusion can be blocked with antibodies against syncytin. HERV-W is only found in primates, but mice have similar retroviral blooms and ERV-related Syncytin genes have also been found in them (Dupressoir et. al.). The ability of mammals and thus ourselves to form a viable placenta and give birth to live young may thus depend on the mammals having harnessed a viral gene somewhere in our evolutionary lineage.

Magiorkinis G et. al. (2012) have verified that the loss of the Env gene which enables cell infection, is associated with super-amplification of germ-line retroviral elements by a factor of about 30, as exogenous retroviruses, switch to endogenous modes of selection. They have investigated the widespread occurence of retroviruses, including intracisternal A-particles, or IAPs, across the diversity mammal groups. Notably following Jern and Coffin (2008) the evolutionary tree of retroviruses both

spans the vertebrates and includes both endogenous and exogenous habits, rooted in exogenous viral types.

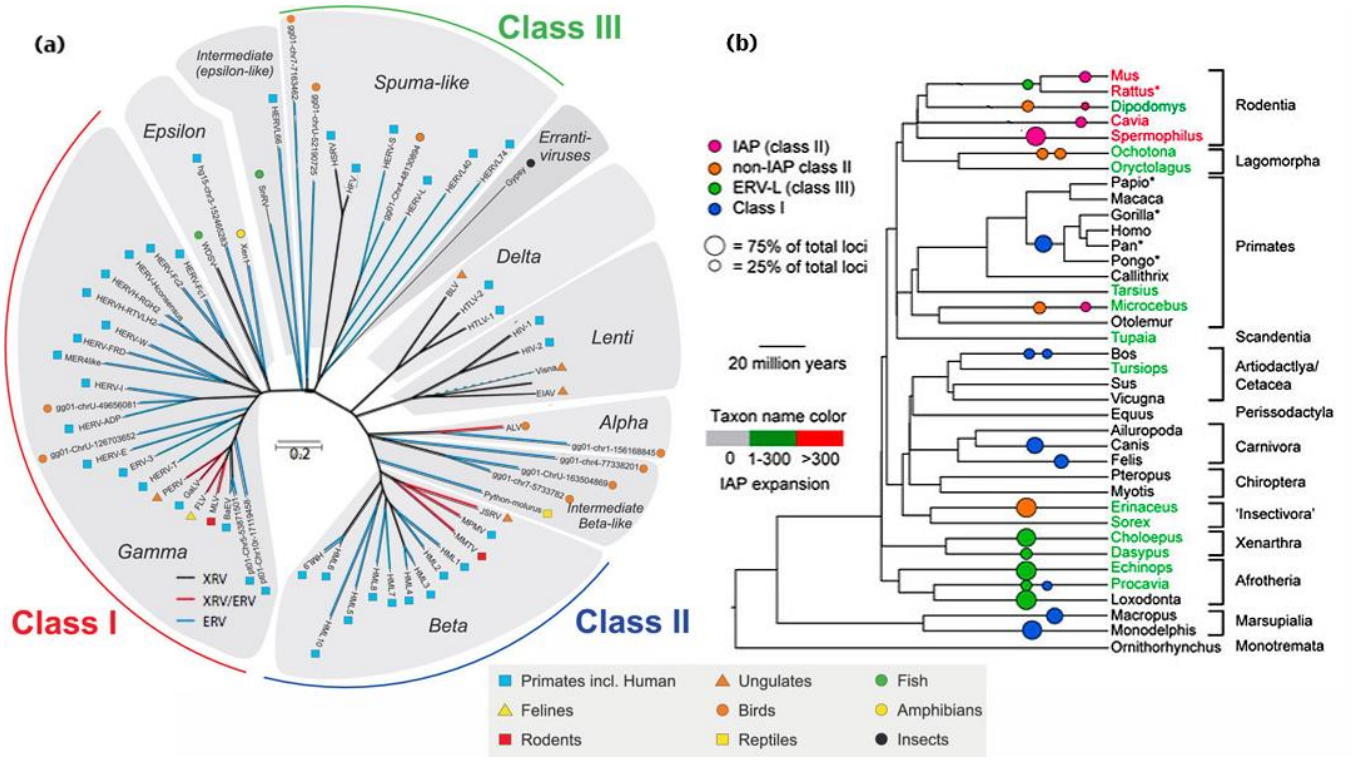


Fig 23: (a) The seven retroviral genera: alpha-, beta-, gamma-, delta-, epsilon-, lenti-, and spuma-like retroviruses and their intermediate groups based on Pol sequences. Black branches indicate viruses known only in exogenous infectious forms (XRV); redbranches indicate viruses present in both XRV andendogenous (ERV) forms; and blue branches indicate ERVs Jern and Coffin (2008) . (b) Phylogeny of mammals with ERV megafamilies shown as colored circles (area is proportional to the percentage of the ERV loci in the genome represented by that family Magiorkinis G et. al. (2012). When retrotranspons are included (fig 22) they extend to all Eukaryote realms.

The defective copies of endogenous retroviruses may also serve to protect the host against further infection by becoming transcribed and causing incorporation of defective elements into the replicating virus (Best et. al.).

10. The Cambrian Radiation, Homeotic Genes, Metamorphosis and Hybridization

One of the most stunning and puzzling aspects of evolution is the Cambrian radiation some 550 million years ago which over a very short period of geological time, gave rise to the major phylla of multicellular animals we see today. This radiation forms the core of the evolutionary tree of fig 1. The previous evolutionary epoch - the Ediacarian by contrast has far fewer and less elaborate fossil forms such as *Charnia*, *Dicksonia* and *Spriggina* in fig 1, and particularly few organisms with well-preserved mineral skeletons.

There have many proposals why such a rapid and abundant radiation could have occurred, including geological scenarios involving the ending of a snowball earth epoch in which the earth became frozen and thus reflected radiant heat, until rising CO_2 levels caused a rapid thawing, setting off a major expansion of cyanobacteria, filling the atmosphere with oxygen and changing the ocean from an acidic state with dissolved iron and litle oxygen to one capable of harbouring diverse forms of multicellular life.

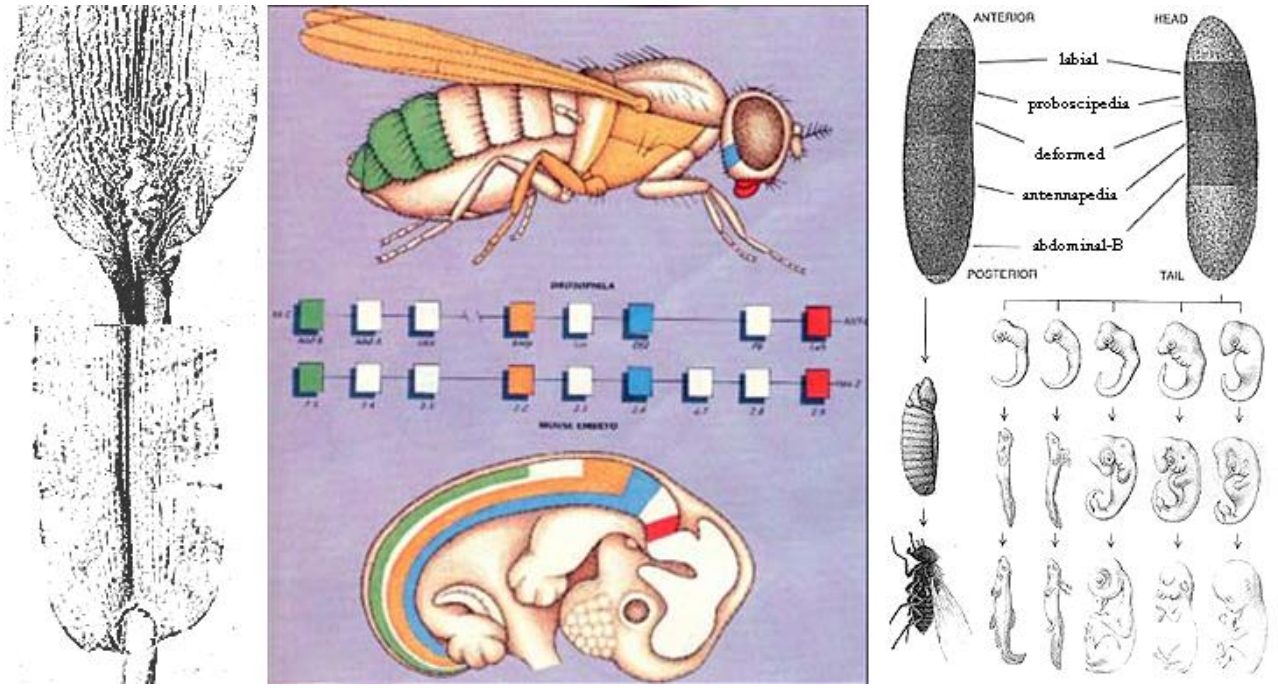


Fig 24: Homeotic genes specifying differentiation along the bodily axis have closely related sequences and are organized in a parallel scheme in arthropods and vertebrates. Mutations of related genes in maize cause disruption of leaf development.

However the underlying reasons may be genetic and more to do with the evolution of a pangenomic algorithm for generating the multicellular body plan based on homeotic and related developmental genes which are highly conserved and spread across the major animal and even plant kingdoms. Closely related schemes of homeotic genes drive the vertebrate and arthropod development along the bilateral axis, being involved in segmentation and notochord differentiation.

For example pax-6, a gene involved in eye formation in the mouse, will induce ectopic eyes on the fruit fly, intriguingly the compound eyes insects usually have, indicating a deep commonality between the genes organizing the body plan in these two pivotal phylla.

This suggests it may have taken evolution a considerable time to come up with such an algorithmic regulatory process, but that once it came into play, it permitted almost symphonic variations, leading to the diversity of the major animal phylla over a short geological time.

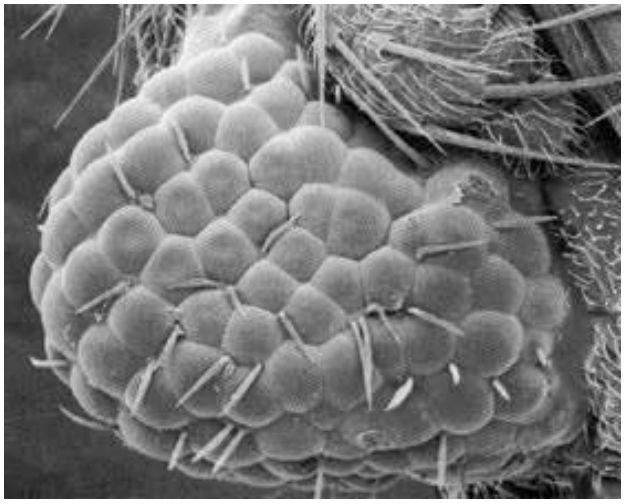


Fig 24b: Ectopic compound eye on the leg of a fruit fly induced by mouse pax-6.

Pivotally, given a context in which such organisms were still in the early stages of genetic diversification in their radiative adaption, it is possible that occasional wholesale horizontal transfer of whole genomes may have occurred as a result of adventitious hybridization.

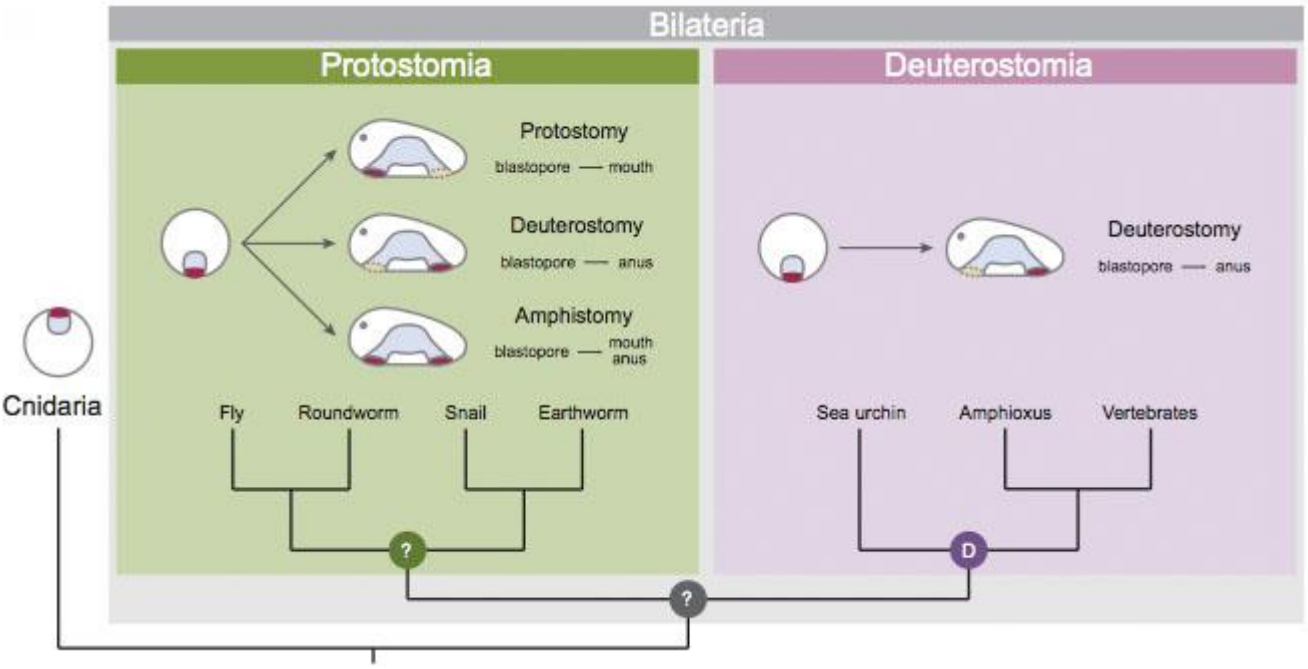


Fig 25: Body schemes of the bilateria (Martın-Duran et al. 2012).

A central division in the Cambrian radiation is the one that led to the bilateria - organisms with the left right symmetry possessed by both arthropods and vertebrates. The conventional argument is that the founding event differentiating bilateria from earlier organisms such as the cnidaria which have a mouth only is the symmetry introduced by the formation of an anus and an intestinal tube. Here things get complicated because embryonic development can go either from mouth first, based on the blastopore

originating in cnidaria and then to anus or vice versa. In the conventional division the deuterostomes including the vertebrates go "arse-first" i.e. anus > mouth but the proterostomes go mouth > anus. The difficulty is that key deuterostomes such as the priapulid worm*Priapulus caudatus*, which was abundant in the middle Cambrian, actually develops on a deuterstome plan although in terms of molecular evolution the expression of bra, cdx, foxA, gsc, and otx during early development is similar to nematodes and arthropods spanning the "skin-shedding" ecdysozoa. Given the fact that the Chaetognatha or arrow worms also have this pattern the deuterstome pattern appears to be ancestral with the proterostomes forming a diverse set of body plans (Martın-Duran et al. 2012).

The conventional theory of insect morphogenesis is the evolution from eggs giving rise to small adult form individuals to delayed maturation of embryonic forms in the form of larvae, which did not compete with the adults in food consumption and habit, thus leading to a two-stage life cycle, with non-competing foraging larval and reproductive adults forms (Jabr). A controversial theory (Ryan) posits that aspects of metamorphosis, which we also see in insects, but more pivotally in marine organisms such as echinoderms, may have resulted from early hybridization even between organisms which have now become distinct phylla, such as vertebrates and echinoderms. For example *Nectocaris pteryx* appears to have a body plan looking like a chimera of an arthropod head and the abdomen of an entirely different phyllum, although this may be a result of the way fossils are depicted in drawings and this species may simply be an early cephalopod (Smith & Caron). The validity of this idea, at least in insects, is highly controversial and hotly disputed (Hart & Grosberg, Williamson).

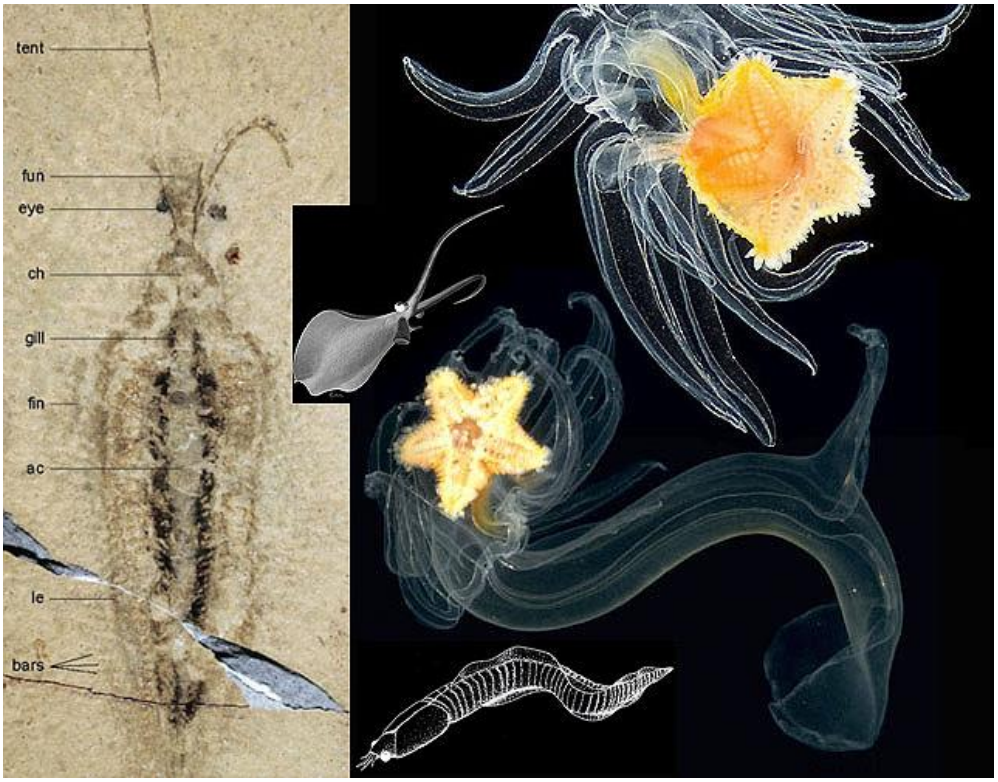


Fig 26: Fossil and two very different scientific illustrations (insets) of *Nectocaris ptaryx* and two views of the larva of *Luidia sarsi* emitting an echinodermal 'offspring'.

Studies attempting to trace the evolutionary tree depending on the simpler larval forms from which one would expect the adult form to have evolved have yielded contradictory results, suggesting the two might have independent genetic origins. Genetic analysis of sea squirts, which have vertebrate larval forms with a notochord and a primitive brain but metamorphose into fixed sea-floor feeders, suggests they have two genomic components, one coming from vertebrates and the other from and an unknown but now extinct non-vertebrate at a very early stage in the evolution of animals. Echinoderms themselves have larvae with a bilateral body plan which later becomes colonized by pluripotent cells in the abdominal cavity forming radially symmetric organisms which grow into adults. In the starfish *Luidia sarsi*, the embryonic form some 4 cm long survives for several months as a vegetarian living off phytoplankton after its starfish 'offspring' have burst out to their carnivorous habit of hunting other starfish.

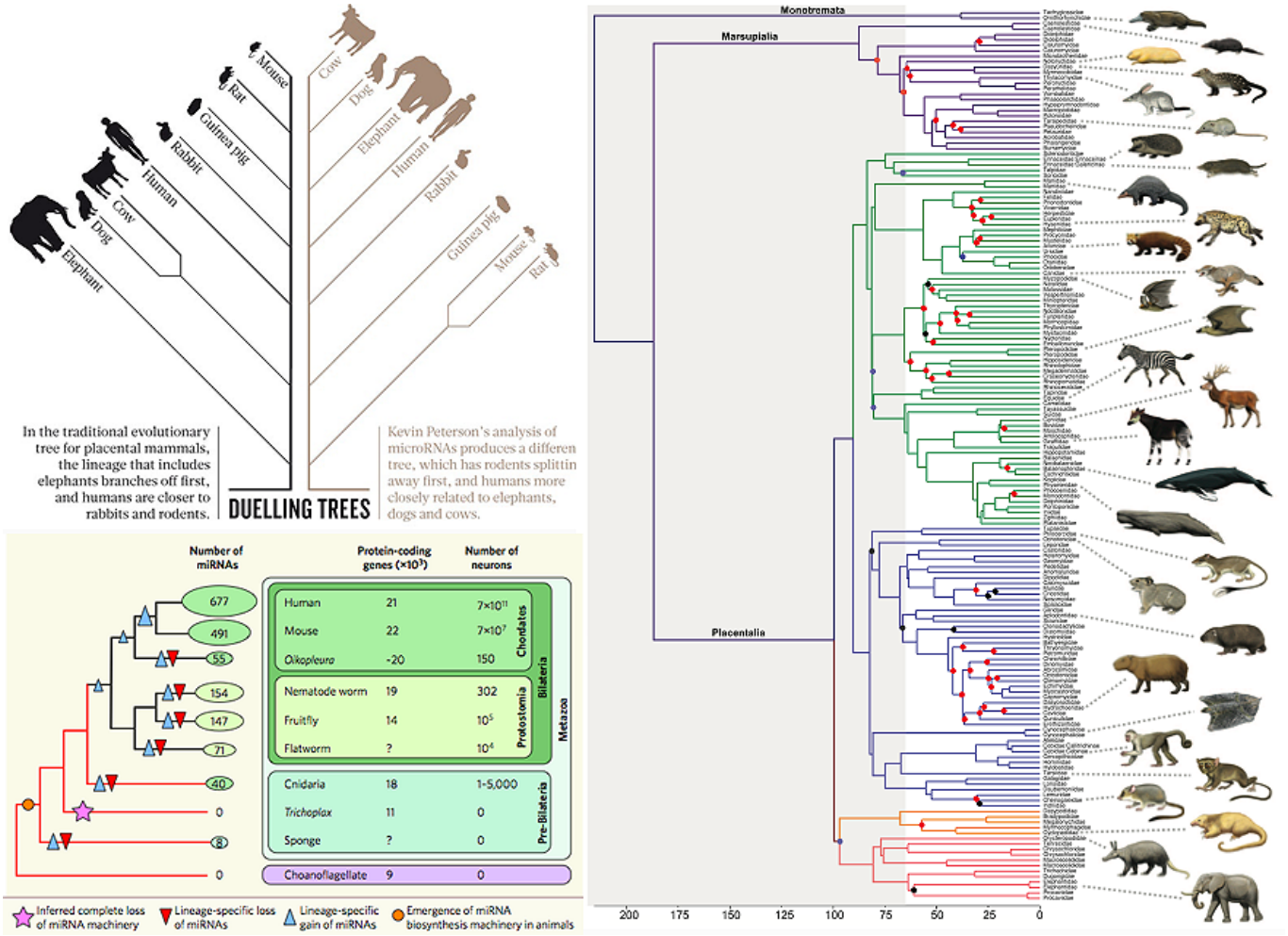


Fig 26b: Conflict in the tree of mammalian diversification. Detailed traditional DNA-based evolutionary tree of the mammals (right) (Meredith et al, dos Reis et al) tends to have a different order of diversification from one based on the number of new miRNAs appearing in successive branches (top left Dolgin). Micro miRNA numbers have also been suggested to be correlated with neural complexity (bottom left Technau). Larger image of the right figure. Major mammal groupings image.

Donald Williamson, who in the 1950s advanced the 'larval transfer' theory claimed to have successfully hybridized fertilised eggs from the sea squirt *Ascidia mentula* with sperm from the sea urchin Echinus esculentus. Then in 2002, in an unpublished study with Sebastian Holmes and Nic Boerboom, he did the reverse cross, using eggs from the urchin and sperm from the sea squirt. Both crosses resulted in large numbers of offspring, the majority of eggs developing into easel-shaped larvae - the '*pluteus*' form typical of sea urchins, rather than the tadpole larvae that are the hallmark of sea squirts. Most of these larvae subsequently metamorphosed to a rounded adult form, which Williamson called a 'spheroid'. The first cross created spheroids with a suction cup, that enabled them to attach to surfaces. Most intriguingly, the second produced spheroids that reproduced asexually through budding, the pinching off of a section of the body to create a clone. However these were never subjected to genetic analysis. A cross between two echinoderm species has however resulted in new developmental phenotypes with confirmed hybridized genomes.

11. Mammalian Radiative Adaption: Traditional DNA versus Micro RNAs

Different assay methods are shedding an intriguing light on radiative adaptation and diversification of animal species, from the Cambrian through to the present day. The detailed branching of the tree of life when calculated by traditional mutational DNA methods (Meredith et al, dos Reis et al), appears to differ significantly from a new technique developed by Kevin Peterson (Dolgin) that depends on the number of newly accrued micro miRNAs which modulate gene expression by selectively binding to specific messengers inhibiting their expression.

A single miRNA is thus able to modulate the expression of a diverse array of mRNAs to which it binds, thus providing for sophisticated forms of coordinated regulation conducive to phylogenetic complexity. It has also been suggested that neuronal complexity correlates with the number of miRNAs (Technau,Grimson et al) an interesting question in itself to do with how complex nervous systems are generated in development. Notice here that humans have fewer protein genes than a mouse, roughly 21,000 against 22,000 although we have a brain with 10,000 times as many neurons, so we need to have an idea how organismic complexity evolves in terms of sophisticated gene regulation, and miRNAs do just that.

Consitent with such a role in multi-celled evolution, the appearance of miRNAs goes back to the earliest multicelled animals. Sea anemones already carry up to 40. Metazoa, from sponges to bilateria, also share the two classes of piRNA, the second of which plays a role in suppressing transposable elements in gametogenesis, by containing a sequence complementary to a transposase mRNA. In fruit flies these are directed against DNA transposons, but in mammals they target LINE L1 and IAP transcription during meiosis in the germ line, by methylating L1 and IAP DNA sequences (Aravin et al).

While a traditional DNA-based tree places primates and humans much closer to rodents, as highly evolved branches, with the elephants diverging earliest, an miRNA analysis places rodents as branching out earliest, something which might seem to be consistent with their possibly closer correspondence to the founding shrew-like mammalian type. The critical question determining the fate of the miRNA perspective is what the rate of loss of these small RNA molecules is in evolution. A higher rate of loss would tend to remove the inconsistency. While the picture is consistent with retaining miRNAs in

mammalian diversification, in insects and a primitive chordate sudden losses have occurred. Ida's controversial place in the primate famiy tree.

For modern lineages of birds and mammals, few fossils have been found that predate the Cretaceous–Palaeogene (K-Pg) boundary, although molecular studies using fossil calibrations have shown that many of these lineages existed at that time. One intriguing way of checking the evolution of mammals and its timing is to examine the parasites of birds reptiles and mammals and to develop a "tree of lice". Smith et al. (2011) demonstrate that the major louse suborders began to radiate before the K-Pg boundary, lending support to an earlier Cretaceous diversification of many modern bird and mammal lineages.

(Continued on Part IV)