Biological Action in Read-Write Genome Evolution

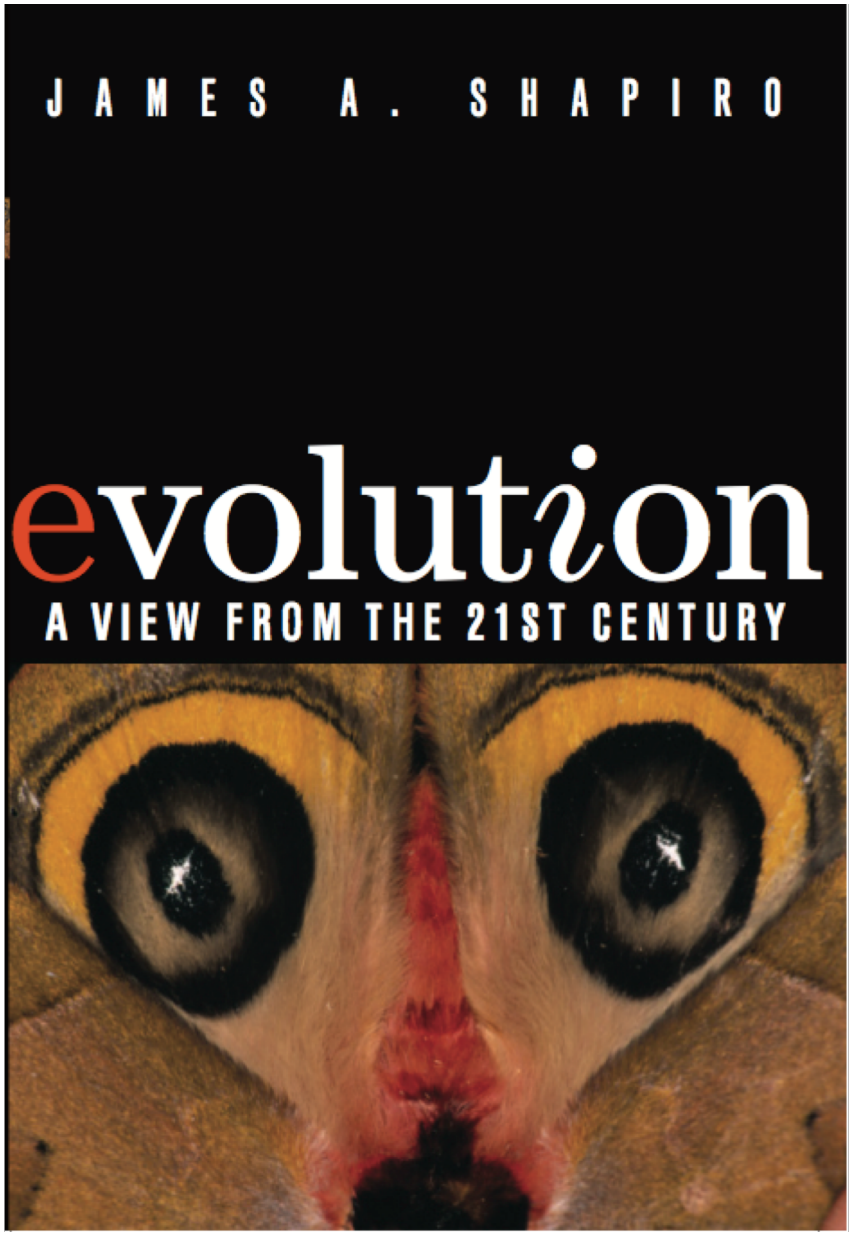
Royal Society-British Academy,
"New trends in evolutionary biology,"
November 7-9, 2016

James A. Shapiro

University of Chicago

jsha@uchicago.edu

www.shapiro.bsd.uchicago.edu



Living Organisms Regularly Use Generic Activities to Facilitate Evolution

- Taxonomic divergences and adaptive innovations often result from abrupt, widespread cell-mediated and biochemical modifications of RW genomes rather than gradual accumulation of independent localized copying errors in ROM genomes.
- Abrupt genome rewriting processes include symbiogenesis, horizontal DNA transfers, interspecific hybridizations, and mobilization of repetitive DNA elements to rewire genome networks and regulatory ncRNAs.
- Ecological conditions and population interactions stimulate evolutionary variation.
- Conclusion: living organisms have core biological/molecular tools to rewrite their genomes actively when challenged.

Well-documented classes of active RW genome inscription lead repeatedly to adaptive and taxonomic novelties:

- Hereditary variation, adaptations and taxonomic originations by cell mergers (symbiogenesis);
- Acquisition of evolved functional adaptations from other taxaa by horizontal DNA transfers;
- Taxonomic originations and adaptive radiations by interspecific hybridizations and whole genome doublings;
- Protein evolution by exon shuffling;
- Amplification of so-called “noncoding” repetitive components in the genome as evolution produces increasingly complex organisms;
- Mobile DNA elements rewiring taxonomically-specific regulatory networks (*e.g.,* evolution of mammalian reproductive and nervous system biology).

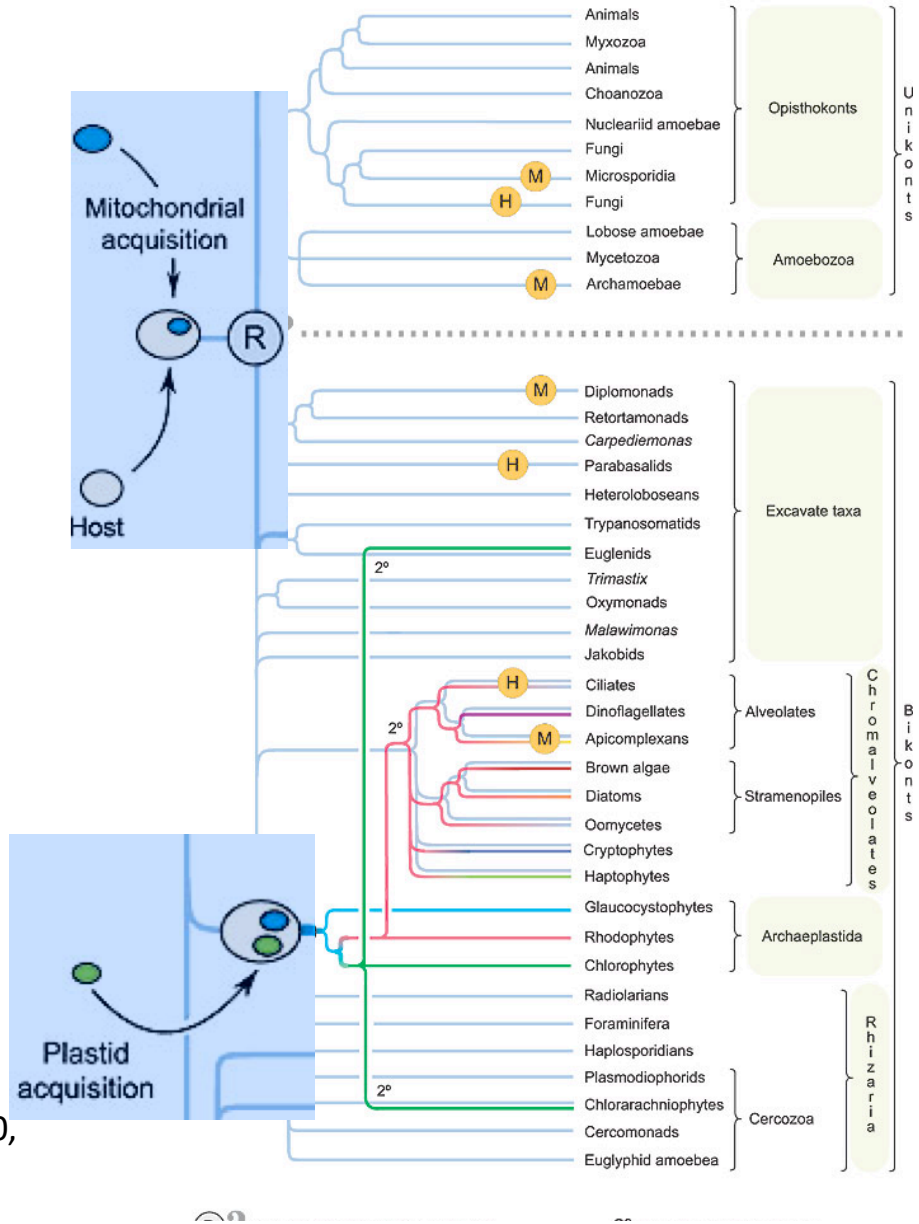
Biomath: One + One = One
(Margulis, 2010; Archibald, 2014)

Ubiquitous Cell Mergers in Reproduction and Evolution

- Infection and phagocytosis
  - Pathogenesis
  - Endosymbiosis, heritable symbionts and holobionts
  - Symbiogenesis (Margulis and Sagan, 2002)

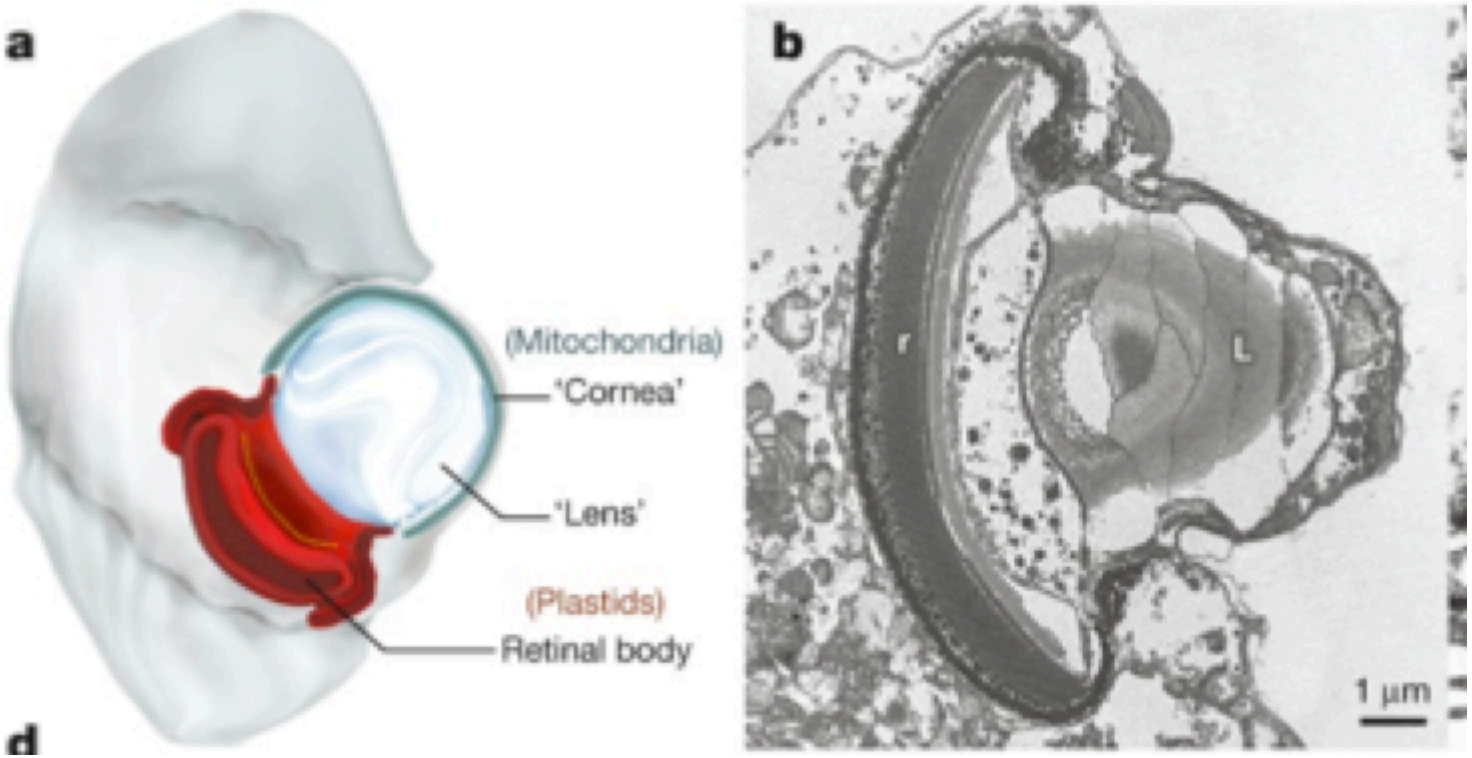
- Sexual reproduction
  - Reassorting of alleles
  - Introgression
  - Hybridization, genome restructuring and hybrid speciation

Evolutionary genome writing by cell fusions – symbiogenetic origins of the eukaryotic cell and descendant photosynthetic taxa



T. M. Embley and W. Martin. 2006. Eukaryotic evolution, changes and challenges. Nature 440, 623-630.

Evolution of an “eye” in a single-celled protist (a *warnowiid dinoflagellate*) by serial symbiogeneses



GS Gavelis *et al.* **Eye-like ocelloids are built from different endosymbiotically acquired components.** *Nature* 523, 204–207 (2015) doi:10.1038/nature14593

Acquiring Adaptive Innovations Through Horizontal DNA Transfers Rather Than Reinventing the Wheel

- Hundreds of *eubacteria-archaea* DNA transfers in the independent evolution of four distinct mesophilic *archaeal* taxa (Lopez-Garcia, P., Y. Zivanovic, et al. (2015). "Bacterial gene import and mesophilic adaptation in *archaea*." Nat Rev Microbiol **13**(7): 447-456.)
- 128 distinct DNA transfers from *eubacteria*, *archaea*, and viruses encoding activities involved in xylem formation, plant defense, nitrogen recycling, and biosynthesis of starch, polyamines, hormones and glutathione into the nuclear genome of a primitive land plant (the moss *Physcomitrella patens*) (Yue, J., X. Hu, et al. (2012). "Widespread impact of horizontal gene transfer on plant colonization of land." Nat Commun **3**: 1152.)
- Dozens of DNA transfers from *eubacteria* and fungi conferring the ability to digest phytopolymers to distinct taxa of plant parasitic nematodes (Haegeman, A., J. T. Jones, et al. (2011). "Horizontal gene transfer in nematodes: a catalyst for plant parasitism?" Mol Plant Microbe Interact **24**(8): 879-887.)

Synthetic species origination in plant domestication: major phenotypic changes after interspecific hybridization and allopolyploid speciation with whole genome doubling

G. Ledyard Stebbins.
Cataclysmic Evolution.
Scientific American 184, 54-59 (April 1951)

Cataclysmic Evolution

Many plants (e.g., wheat, cotton, tobacco) evolved suddenly by a process involving the doubling of chromosomes. The same process is artificially induced to create useful new species

by G. Ledyard Stebbins, Jr.

TWO GRASSES, blue wild rye (*left*) and squirrel-tail grass (*right*), were crossed to produce a hybrid (*center*). The hybrid was sterile, but when its chromosomes had been doubled with colchicine, it became fertile.

Examples of Speciation and Adaptive Radiations by Interspecific Hybridization and Genome Doubling

- **Fungi**: *Saccharomyces* species (Greig et al., 2002; Muller & McCusker, 2009; Sipiczki, 2008)

- **Plants** (Soltis & Soltis, 2009):
  - *Pinus densata* (Mao, 2011)
  - *Primula* (Guggisberg, 2008)
  - Sunflowers (Ungerer, 1998)
  - Irises (Arnold, 1991)
  - *Orchidaceae* (Vega, 2013)
  - *Brassica napus* (Albertin, 2007; Chalhoub, 2014)
  - *Arabidopsis* (Schmickl, 2011)
  - *Nicotiana* (Kelly, 2010; Fuentes, 2014)
  - Potatoes (Pendinen, 2008)
  - Wheats (Stebbins, 1951; Ozkan, 2001, 2003; Qi et al., 2012)
  - Cultivated oats, cotton and sugar cane (Stebbins, 1951; Li, 2015)
  - *Triticale* (Hulse & Spurgeon, 1974)

- **Animals** (Dowling & Secor, 1997; Schwenk, 2008):
  - Tephritid fruitflies (Schwarz, 2007)
  - Swallowtail Butterflies (Kunte,2011)
  - *Heliconus* butterflies (Mallet, 2007; Pardo-Diaz, 2012, Heliconus Genome Consortium, 2012)
  - Army ants (Kronauer, 2011)
  - Sculpins (*Cottus* sp., *Teleostei*)(Renaut, 2011)
  - Sailfin silversides (*Teleostei*) (Herder, 2006)
  - **E. Africa *Cichlids*** (Keller, 2013; Brawand, 2014; Selz, 2014; Svensson, 2016)
  - Sparrows (Hermansen, 2011)
  - Audubon's warbler (Jacobsen,2011)
  - **Galapagos finches** (Grant & Grant, 2014 etc. ; Lamichhaney, 2015; Palmer, 2015)
  - Clymene dolphin (Amaral, 2014)
  - Bats (*Noctilio*) (Khan, 2014)
  - Cats (*Felidae*) (Li, 2016)

Genomic Consequences of Interspecific Hybridization and Genome Doubling

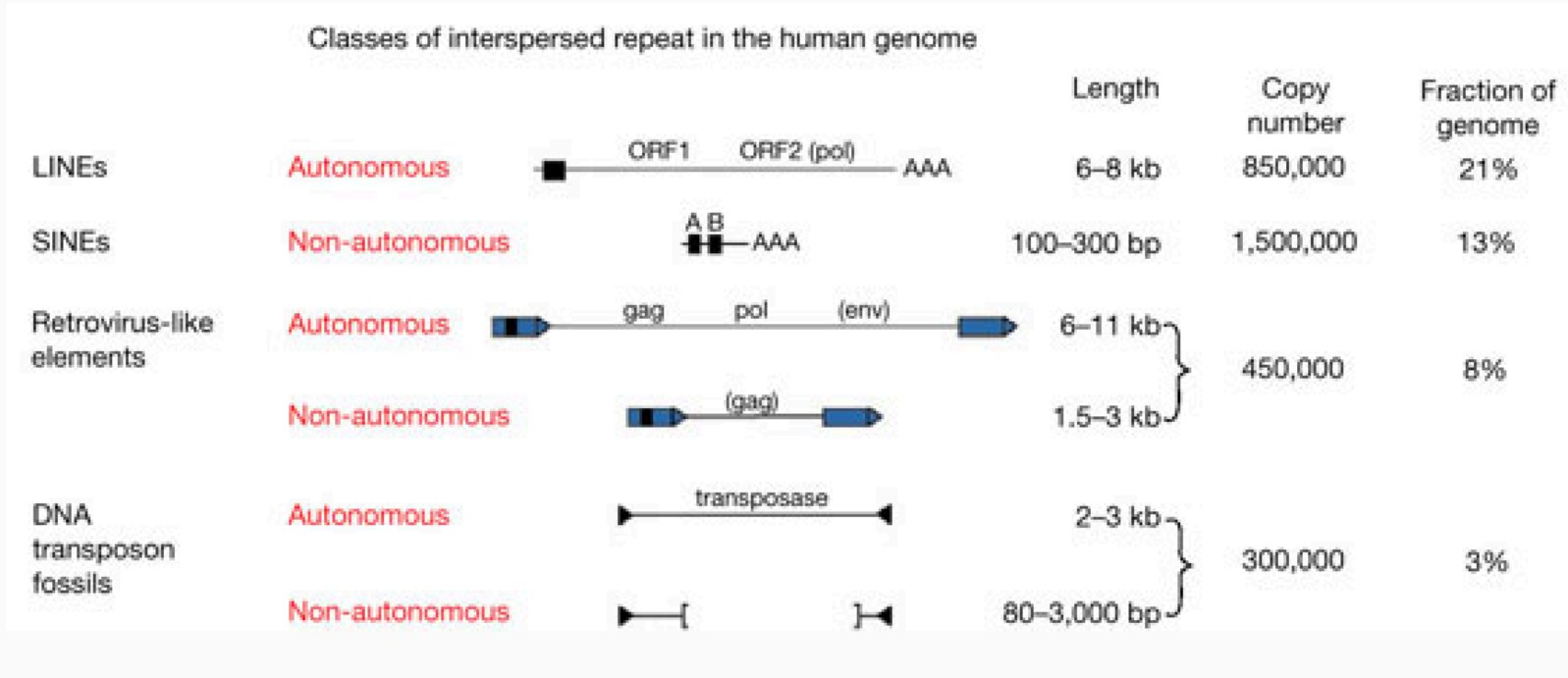
- Allopolyploidy: “All data stress allopolyploidization as a shock associated with drastic genome reorganization” in plant hybrids
- Epigenetic changes affecting mobile DNA elements and other natural genetic engineering activities
- Massive and repeated patterns of repeat amplifications and chromosome rearrangements
- Whole genome duplications followed by selective loss of duplicated regions

Parisod, C., et al., *Impact of transposable elements on the organization and function of allopolyploid genomes.* New Phytol, 2010. **186**(1): p. 37-45.

The Special Impacts of Interspecific Hybridization on Evolutionary Innovation

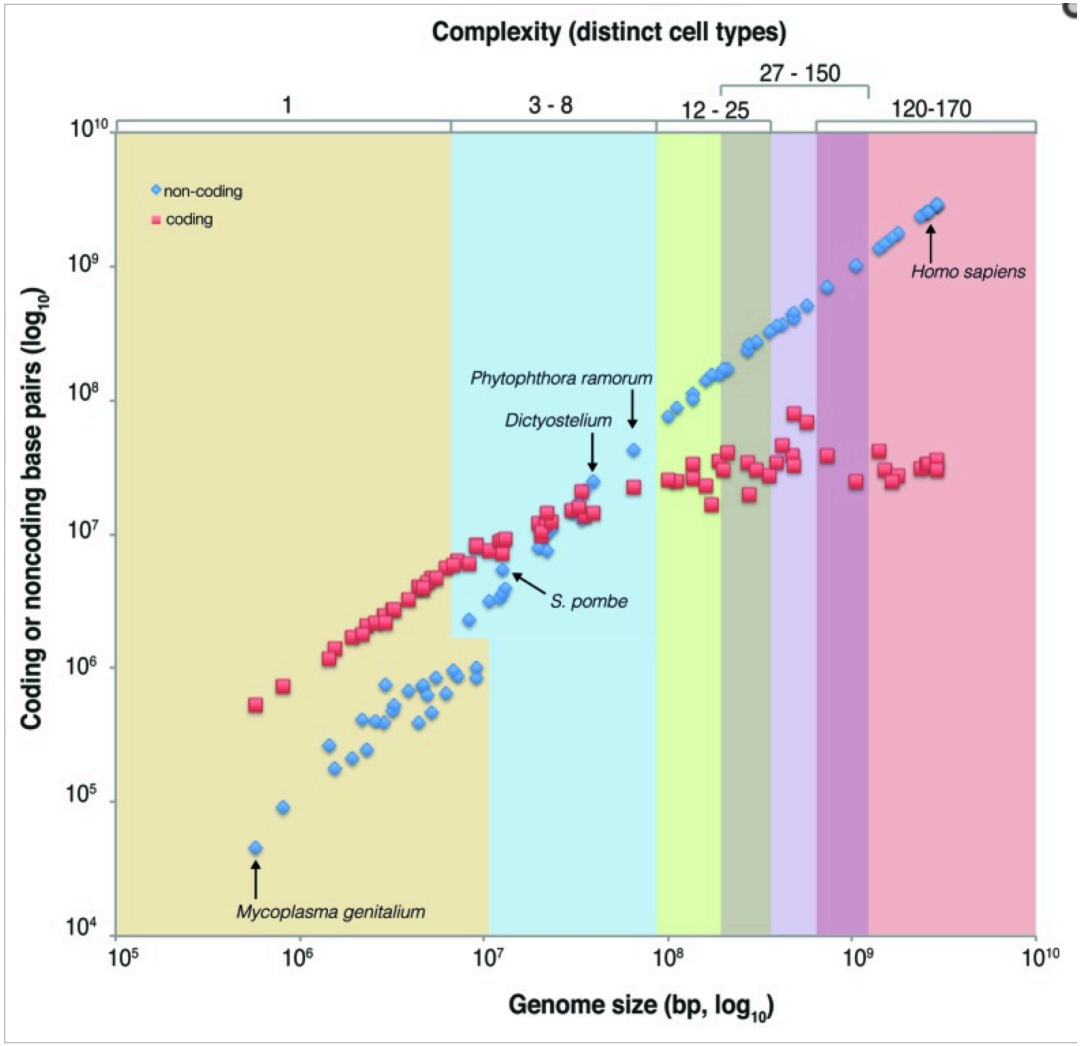
- The initiating event involves the entire genomes of each parent species – so all organismal traits are affected in a single evolutionary episode.
- Since interspecific hybridization serves as a major stimulus to genome variability (“genome shock”), including chromosome rearrangements and activation of mobile DNA elements, hybrid organisms have a markedly elevated potential for generating novel DNA sequence configurations that were not present in the genome of either parent species.

Evolutionary time:
Genome writing by natural genetic engineering – massive amplification of dispersed repetitive mobile DNA elements



International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409, 860 - 921 (2001)

Growing importance of “non-coding” DNA with organismal complexity



Liu G[1], Mattick JS, Taft RJ. **A meta-analysis of the genomic and transcriptomic composition of complex life.** Cell Cycle. 2013 Jul 1;12(13):2061-72.

Mammalian Reproduction Network Wiring by Mobile DNA: Both Sides of the Fetal-Maternal Interface

PLACENTA

- **Endogenous retroviruses (ERVs) provide Promoters for Human Placental-Specific Transcripts (Macaulay et al, 2011).**
- **ERV family MER41 contributes hundreds of human-specific enhancers for development of placental trophoblast stem cells (Chuong et al., 2013).**
- **ERV family RLTR13D5 contributes hundreds of mouse-specific enhancers for development of placental trophoblast stem cells (Chuong et al., 2013).**

UTERUS

- **Eutherian-specific DNA transposon MER20 elements as enhancers for progesterone- and cAMP-stimulated expression of 200 uterine functions (Lynch et al, 2011).**
- **MER20-MER39 transposons provide Prolactin promoter (Emera & Wagner, 2012).**
- **1,721 uterine-expressed Progesterone Receptor binding sites found within Mammalian- or Eutherian-specific transposable elements (Lynch et al., 2015).**

Mammalian Reproduction Network Wiring by Mobile DNA and lncRNA – Stem Cells, Nervous System, Innate Immunity

PLURIPOTENT STEM CELLS

- **Human endogenous retrovirus HERVH present in promoters of more than 100 lncRNAs providing transcriptional regulatory signals for pluripotent stem cell-specific expression (Kelley & Rinn, 2012; Robbez-Masson, 2015; Durruthy-Durruthy, 2016).**
- **25% human and mouse binding sites for stem cell pluripotency transcription factors Oct4 plus Nanog located in species-specific transposable elements (Kunarso et al., 2010).**
- **>98% human-specific individual pluripotency transcription factor binding sites in mobile DNA (Glinsky, 2015)**

BRAIN

- **MER130 retrotransposon dispersed repeat family provides active enhancers in development of the mouse dorsal neocortex (Notwell et al., 2015).**
- **"…the brain is the organ where lncRNAs have the most peculiar features including the highest number of lncRNAs that are expressed, proportion of tissue-specific lncRNAs and highest signals of evolutionary conservation" (Aprea and Calegari, 2015).**

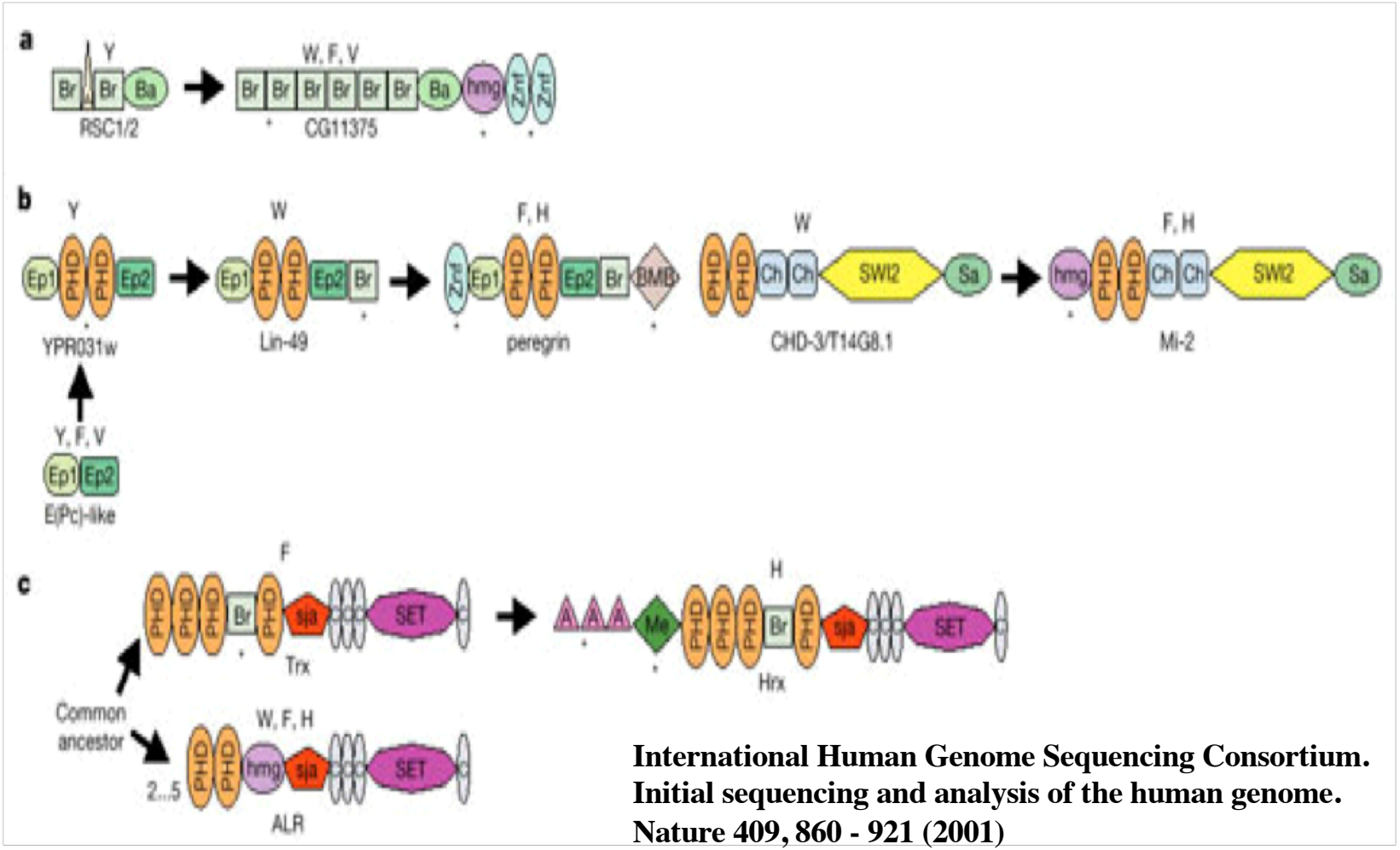
INNATE IMMUNITY

- **Lineage-specific ERVs dispersed IFN-inducible enhancers for inflammatory innate immunity response independently in diverse mammalian genomes (Chuong, et al. 2016).**

Ecological Disruption Stimulates Read-Write Genome Innovations:

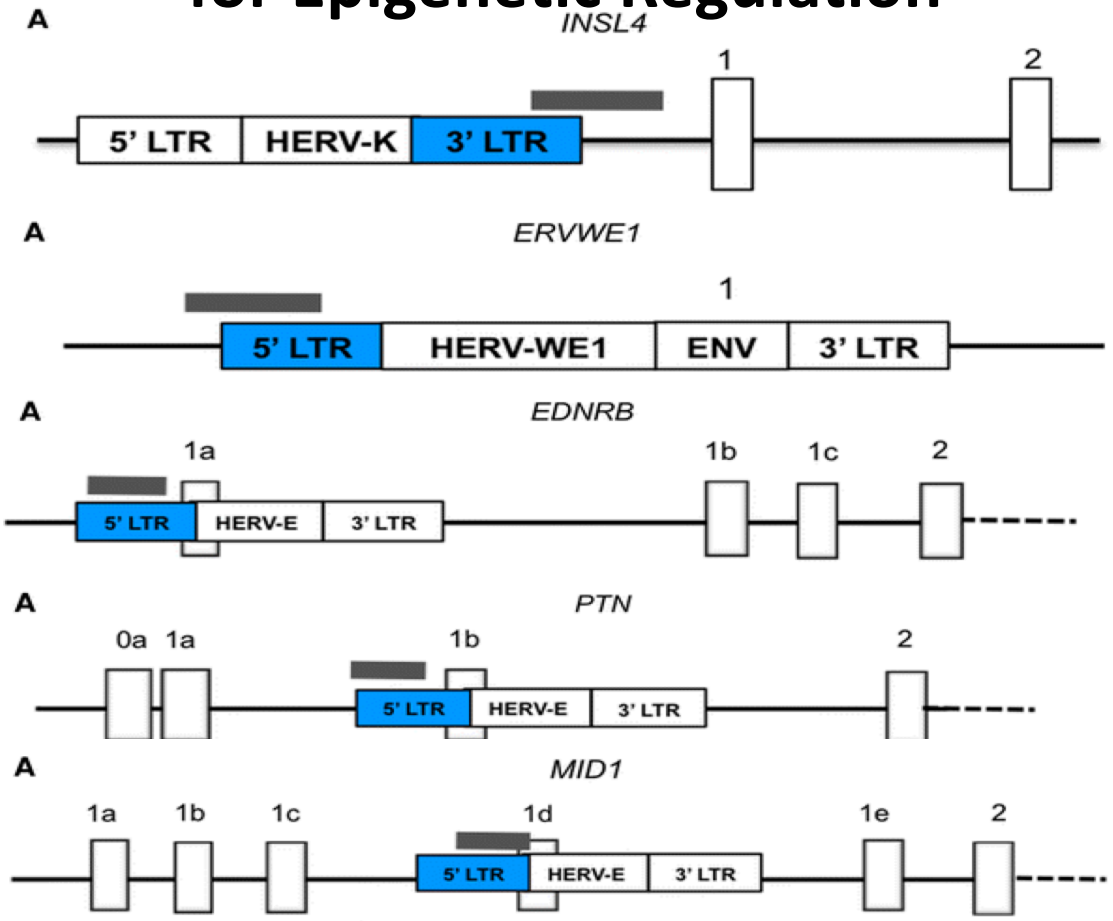
- Novel food sources, changes in endosymbiont/microbiome/virome populations (“holobiont” evolution) ==> altered symbiogenetic, horizontal DNA transfer potentials.
- Population depletion ==> increased mate selection outside normal species boundaries, genome destabilization and hybrid speciation events.
- Starvation and other stress factors are well documented to alter epigenetic control and trigger mobile DNA element activity and genome restructuring; http://shapiro.bsd.uchicago.edu/StimuliDocumentedActivateNGE.html

Genome writing by natural genetic engineering – Protein evolution by domain rearrangements



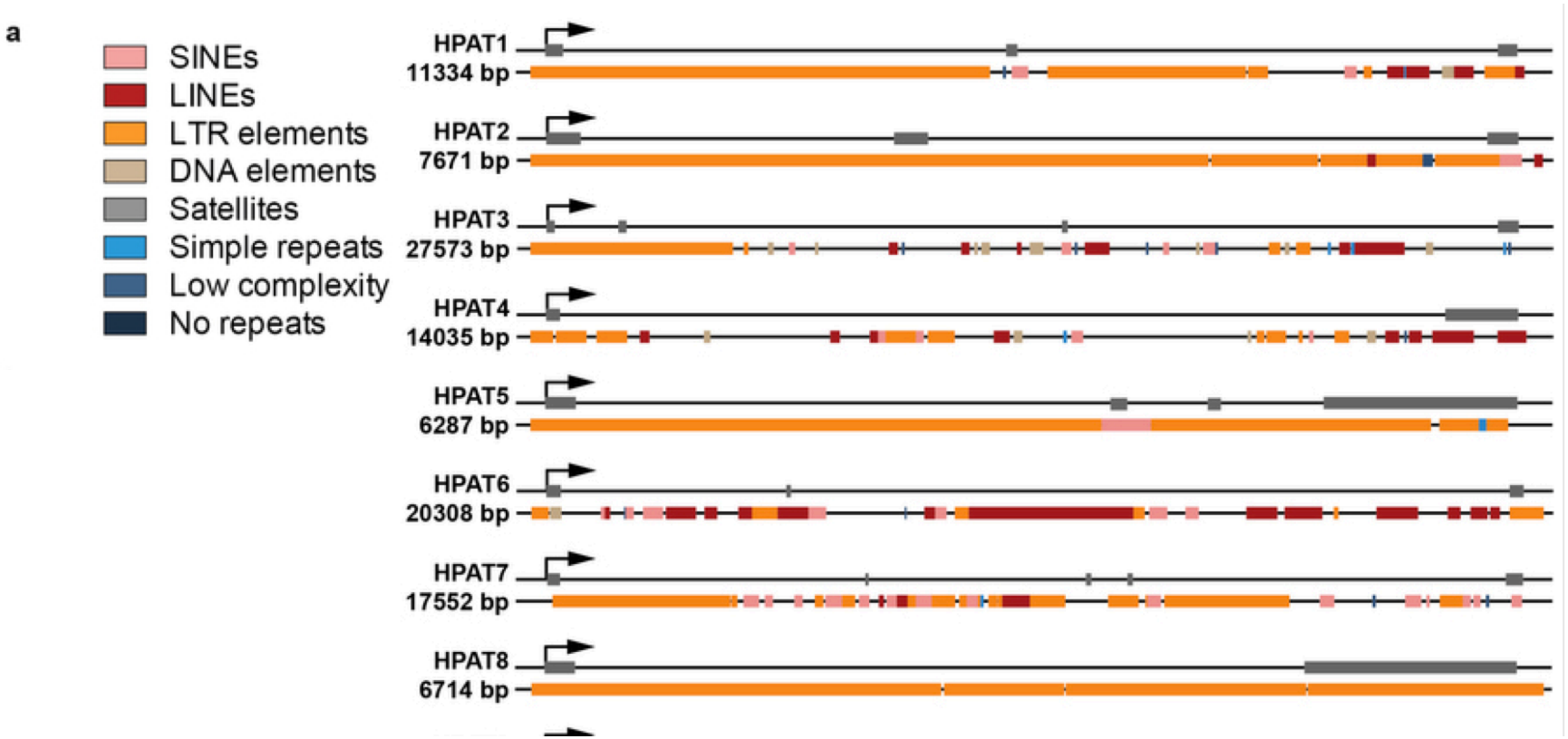
International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409, 860 - 921 (2001)

ERVs as Promoters for Human Placental-Specific Transcripts and Sites for Epigenetic Regulation



Macaulay EC, Weeks RJ, Andrews S, Morison IM. Hypomethylation of functional retrotransposon-derived genes in the human placenta. **Mamm Genome.** 2011 Dec;22(11-12):722-35. Chuong, E.B., 2013. Retroviruses facilitate the rapid evolution of the mammalian placenta. *Bioessays,* 35**,** 853-61.

Human Pluripotency-Associated Transcripts (HPATs) are significantly enriched for TEs



The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming

Jens Durruthy-Durruthy, et al. Nature Genetics 48, 44–52 (2016) doi:10.1038/ng.3449

Parsing The Fundamental Question in Evolution: How do heritable adaptive novelties and new groups of organisms arise?

- What is the difference between Variation and Selection as the source of evolutionary novelties?
- Does the process of variation require a gradual accumulation of “numerous, successive, slight modifications,” or do abrupt “saltational” changes with major biological effects occur repeatedly?
- Does variation arise in an unpredictable, stochastic fashion by random errors in hereditary transmission, or are there distinct cell processes that can repeatedly generate significant genomic novelty?
- Is evolutionary variation sensitive to ecological conditions?

What is the difference between Variation and Selection as the source of evolutionary novelties?

- “To put it in the terms chosen lately by Mr. Arthur Harris in a friendly criticism of my views: **‘Natural selection may explain the survival of the fittest, but it cannot explain the arrival of the fittest.’**” Hugo de Vries, *Species and Varieties: Their Origin by Mutation* (1904), The Open Court Publishing Company, Chicago, p. 826.
- In other words, Variation (Mutation) generates organisms with new traits, and Selection determines their survival and reproduction.

Is the process of variation always a gradual accumulation of "numerous, successive, slight modifications," or do abrupt "saltational" changes with major biological effects occur?

Symbiogenetic cell fusions creating

- Oxidative eukaryote ancestor (mitochondrial *alpha-proteobacteria* symbiosis)
- Photosynthetic eukaryotes (chloroplast *cyanobacteria* symbiosis)
- *Euglenids* and *Chlorarachniophytes* (*Cercozoa*) (green alga secondary symbiosis)
- *Chromalveolates*, including brown algae, dinoflagellates and diatoms (red alga secondary symbiosis)
- Photosynthetic animals, including corals, sea slugs and molluscs (algal symbiogenesis)
- Eye-like "ocelloid" in *Warnowiaceae* dinoflagellates (serial cyanobacterial symbiogeneses)
- Multiple reproductive isolations in organisms as diverse as *Paramecium* and *Drosophila* (bacterial symbiosis)
- Virus resistance in *Drosophila* (*Wolbachia* symbiosis)

Interspecific hybridizations coupled with **whole genome duplications** ensuring successful meiosis of the hybrid genome

- *Saccharomyces* yeasts
- Most domesticated crop species
- Hundreds + of wild plant species
- Dozens + of animal species, including mimetic butterflies, East African *cichlids*, and Darwin's finches

Is the process of variation always a gradual accumulation of "numerous, successive, slight modifications," or do abrupt "saltational" changes with major biological effects occur? (continued)

Horizontal DNA transfers

- Between prokaryotes, including antibiotic resistance, pathogenicity/virulence complexes and metabolic functionalities
- Between eukaryotes, including metabolic functions, mitochondrial and chloroplast genomes, mimicry patterns and mobile DNA elements
- From prokaryotic and eukaryotic microbes to multicellular eukaryotes, including digestive enzymes, phytopathogenicity determinants and energy metabolism
- From eukaryotes to bacteria, including regulatory "effector" domains essential to bacterial infectivity

Genome restructurings

- Whole genome duplications and other changes in chromosome number
- Chromosome rearrangements, including "chromothripsis" (chromosome shattering and reassembly)
- Amplifications and reductions of repetitive DNA arrays
- Integrations of infecting viral and bacterial genomes
- Bursts of mobile DNA transpositions throughout the genome

Does variation arise in astochastic fashion by random errors in hereditary transmission, or are there distinct cell processes that repeatedly generate significant genomic novelty?

- **Replication quality control**
  - Exonuclease proofreading
  - Mismatch detection and repair
- **Natural genetic engineering processes**
  - Mutagenic polymerases, reverse transcriptases and terminal transferases
  - Ectopic recombination between dispersed repeats (copy number variations, structural rearrangements)
  - Non-homologous end-joining (NHEJ), leading to chromosome rearrangements, deletions and insertions of exogenous DNA (chromosomal, organelle, cDNA, viral), mutagenic replication at the repair site, and "chromothripsis" (chromosome shattering and reassembly)
  - Nucleotide modification, excision and replacement (C --> U deamination and U excision), leading to hypermutation "kataegis" (mutational "thunderstorms") in vertebrates (specifically adapted for Immunoglobulin somatic hypermutation and class switch recombination in activated B cells)
- **Mobile DNA elements**
  - Integrons
  - Proviruses
  - DNA transposons
    - Cut-and-splice non-replicative
    - Replicative, including helitrons (rolling circle replication)
  - Retrotransposons
    - SINEs and LINEs
    - LTR elements
  - Retroviruses
  - Transduction of non-mobile DNA sequences

Is evolutionary variation sensitive to ecological conditions?

- **Ecological variables that modulate genome modification processes**
  - Nutrient depletion in bacteria [Maenhaut-Michel, 1994 #3450] yeast [Todeschini, 2005 #3951][Servant, 2008 #2090], and diatoms [Maumus, 2009 #3597]
  - UV irradiation in tobacco [Filkowski, 2004 #3408]
  - Peroxide in yeast [Sehgal, 2007 #2088]
  - Temperature in *Festuca arundinacea* [Ceccarelli, 2002 #6648], silk moth [Kimura, 1999 #3720][Kimura, 2001 #3721] and mice [Li, 1999 #6646]
  - Chlorine ions in *Arabidopsis* [Boyko, 2006 #7565][Boyko, 2010 #7568]
  - Heavy metal exposure in *Arabidopsis* [Rahavi, 2011 #15597]
  - Fungal metabolites in tobacco [Melayah, 2001 #6640]
  - Virus infection in tobacco, maize and *Arabidopsis* [Kovalchuk, 2003 #3400]
  - Air pollution in mice [Somers, 2002 #6645;Yauk, 2008 #3053]
  - Physical trauma (cutting) in tobacco [Sugimoto, 2000 #6643]

- **Ecological effects on genome variation observed at Evolution Canyon or Yehudiyya microsite in Israel**
  - Spontaneous mutation in soil fungi [Lamb, 2008 #16337]
  - Microsatellite expansion in wild emmer wheat [Li, 2003 #16400][Li, 2002 #16399]
  - Transposition in *Drosophila* [Hubner, 2013 #19427][Beiles, 2015 #24871]
  - Retrotransposition in wild barley [Kalendar, 2000 #6647]