Humans and their closest living relatives ...

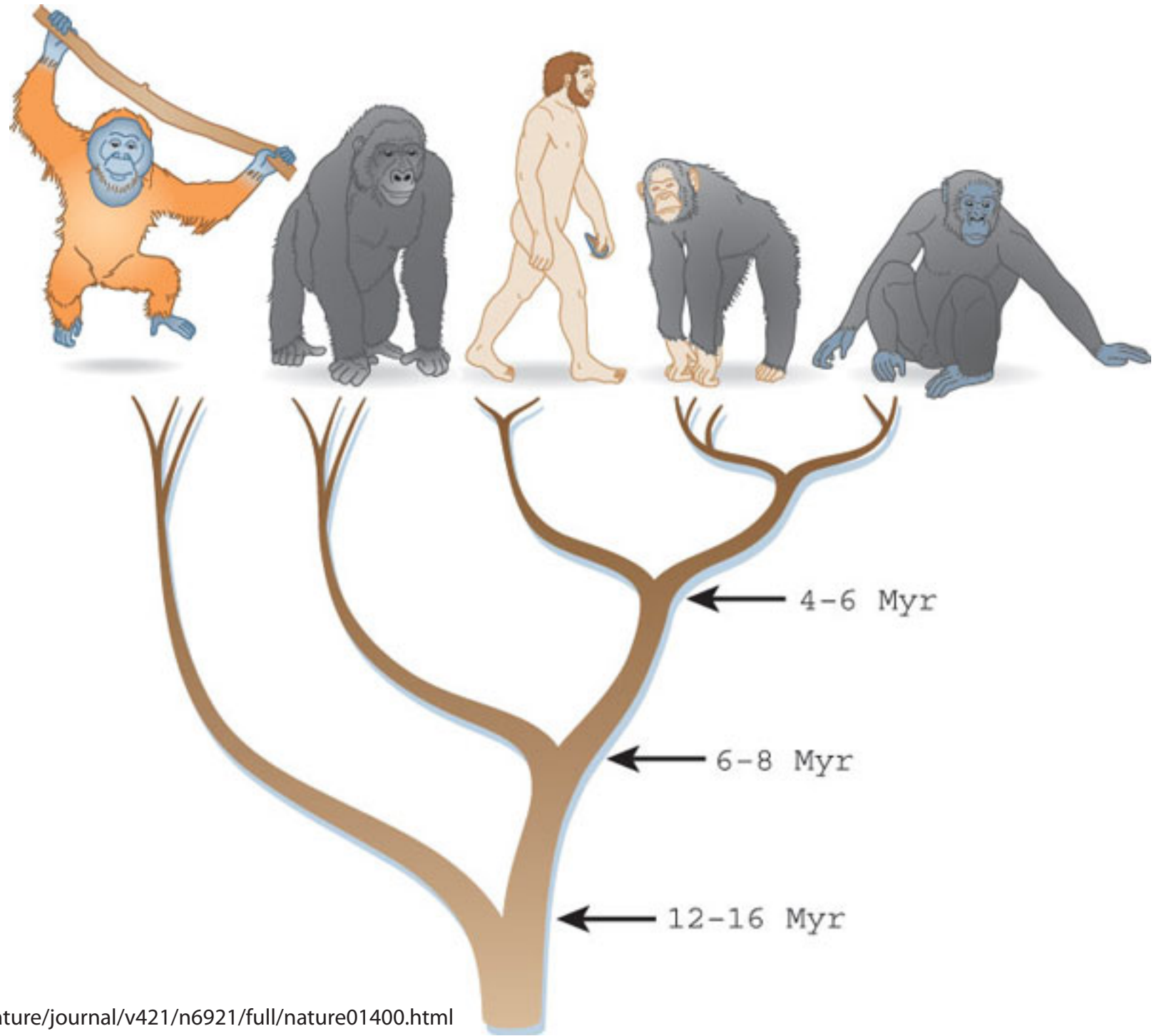
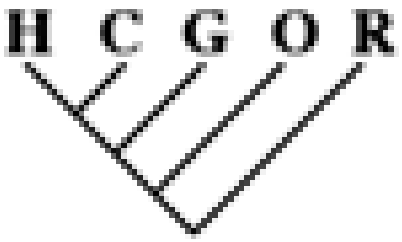
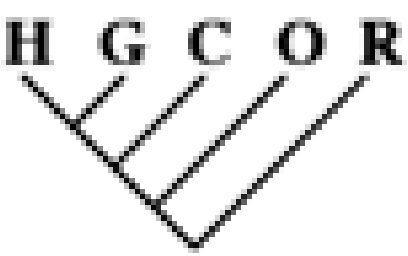


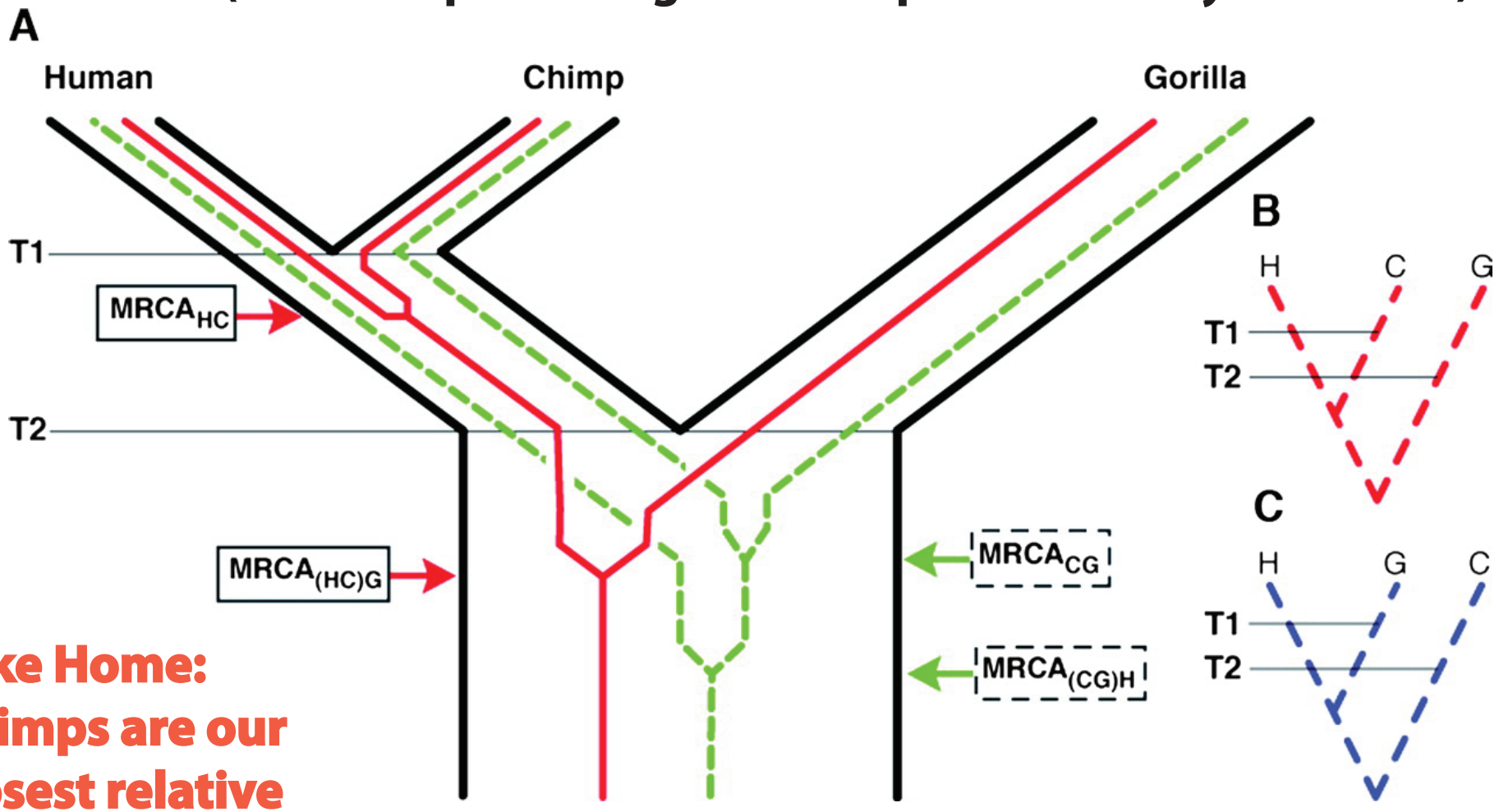
Image from Figure 1 of
http://www.nature.com/nature/journal/v421/n6921/full/nature01400.html





About 80% of genes **About 10%** **About 10%**

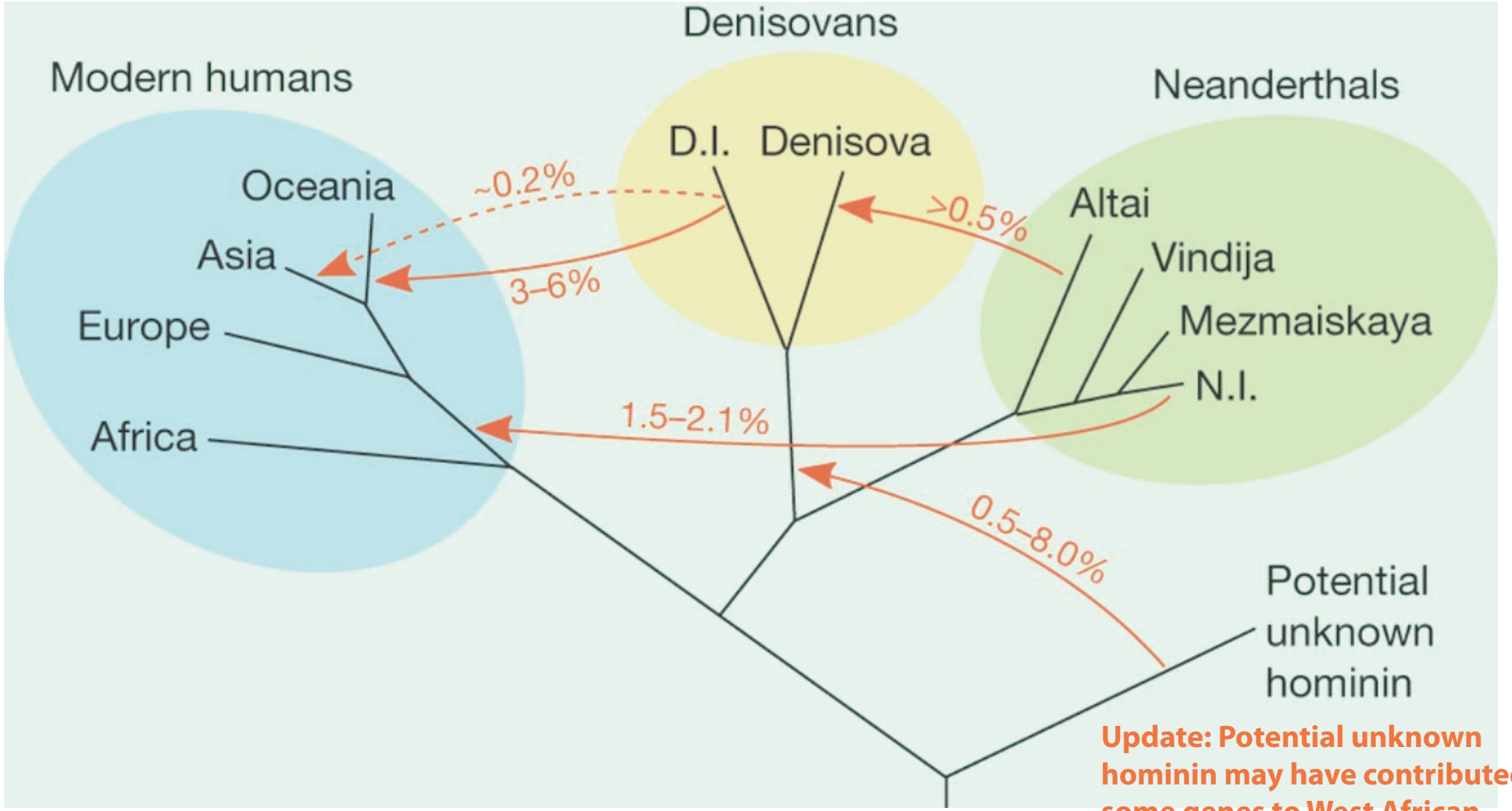
(detailed percentages will depend on analysis details)



Take Home:
Chimps are our closest relative but not for some genes!

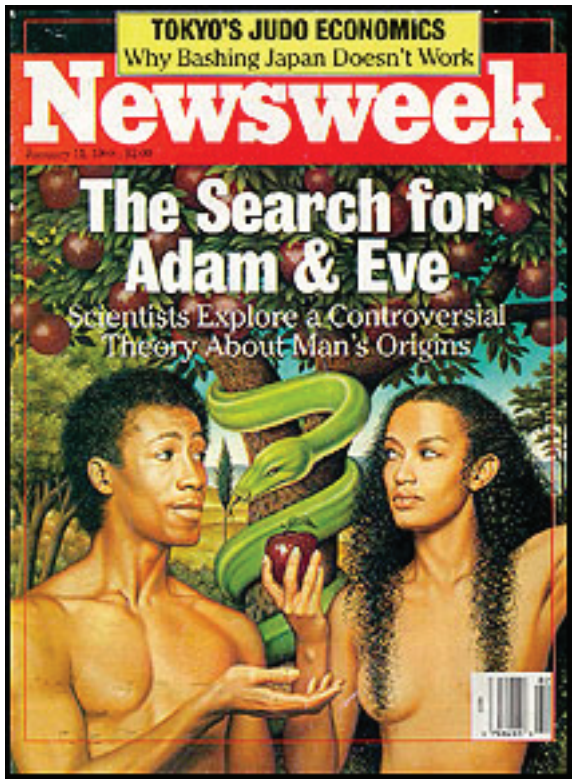
Figures From: Ingo Ebersberger et al. Mol Biol Evol 2007;24:2266-2276

Modern humans appear to have some genes that originated in extinct human species...

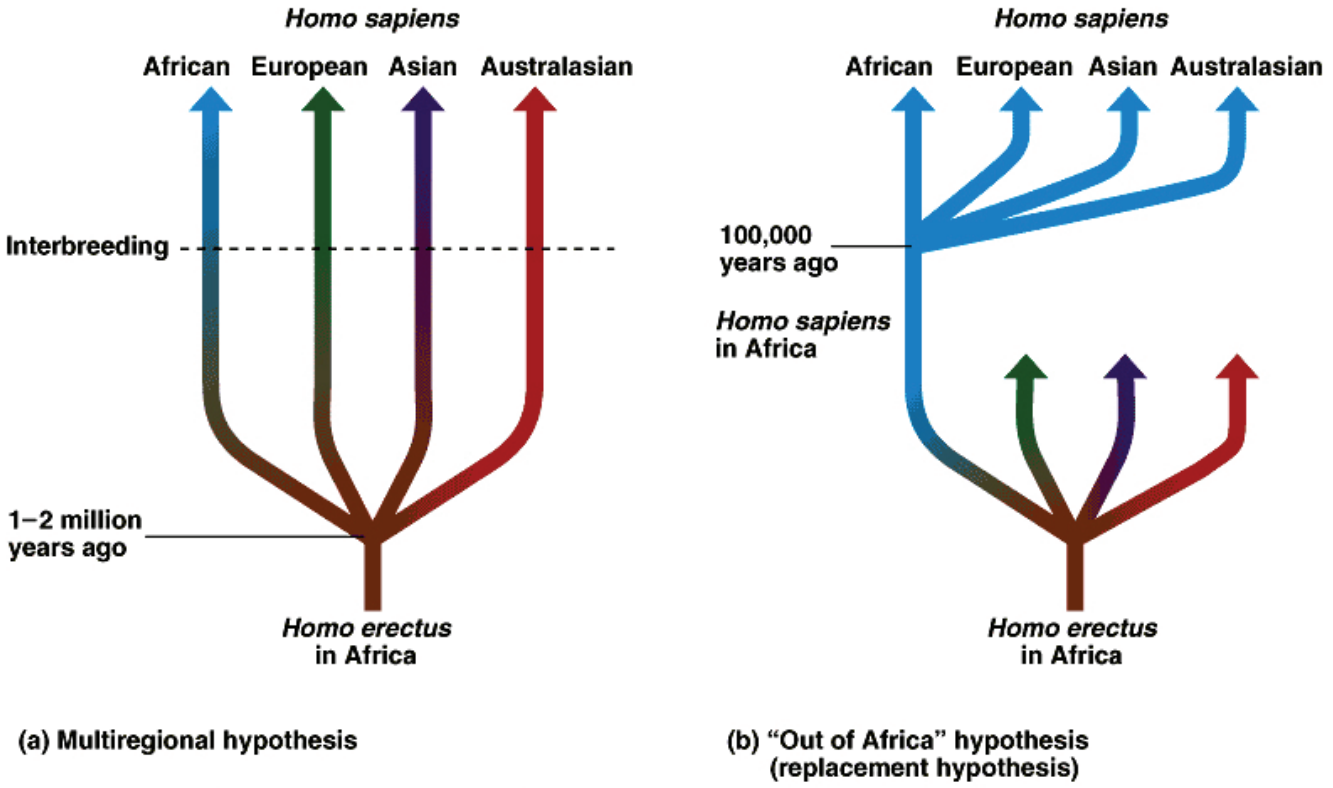


above image is from http://cdn4.sci-news.com/images/enlarge/image_1624_2e-Neanderthals.jpg

Update: Potential unknown hominin may have contributed some genes to West African populations of modern humans (12 Feb. 2020. Science Advances. Vol. 6, No. 7, https://doi.org/10.1126/sciadv.aax5097)

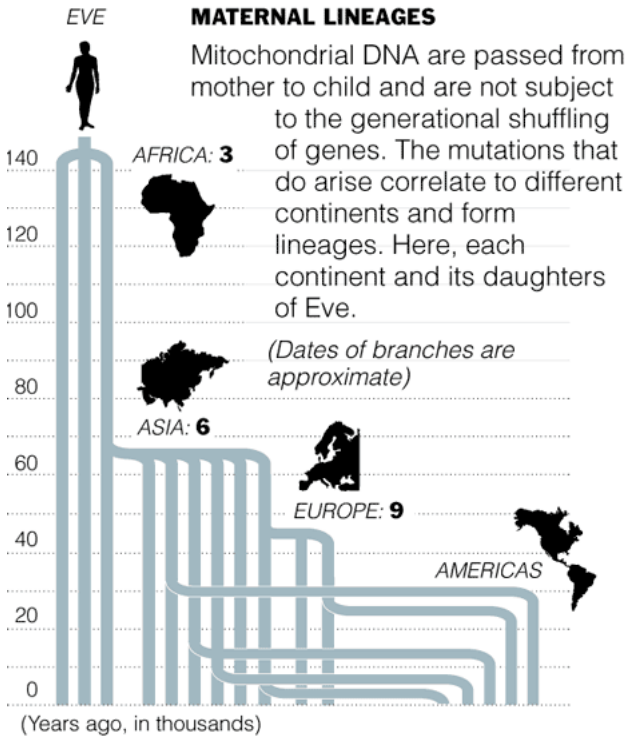


"Mitochondrial Eve" and the "Out of Africa" Hypothesis



Allan Wilson

"Out of Africa" for anatomically modern humans seems to be largely correct with major modification to include Neandertal / Denisovan introgression.



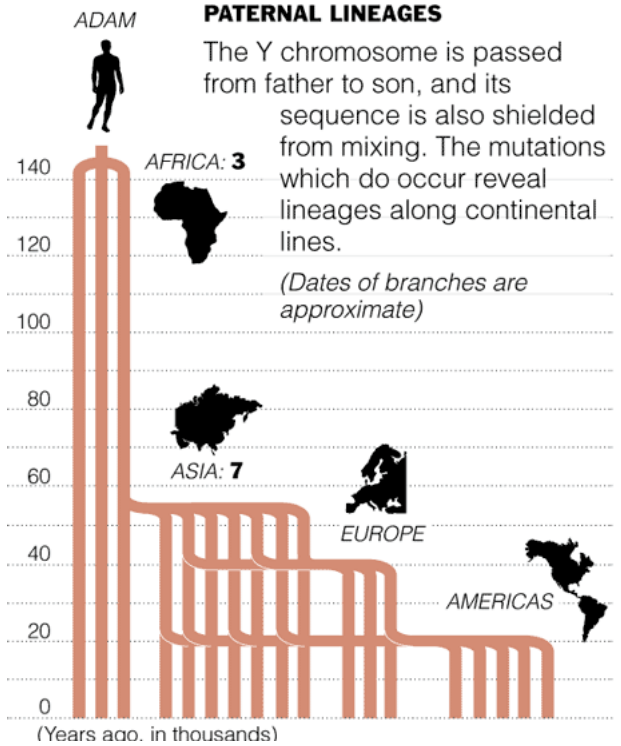


figure from http://www.ramsdale.org/dna7.htm

"... any two humans differ, on average, at about 1 in 1,000 DNA base pairs (0.1%). Human genetic diversity is substantially lower than that of many other species, including our nearest evolutionary relative, the chimpanzee. Genetic diversity is a function of a population's "age" ..."

"... To put the 0.1% genetic diversity estimate into perspective, it is useful to remember that humans have approximately 3 billion base pairs in a haploid cell. Thus, any pair of humans differs by approximately 3 million base pairs. ..."

Quoted from essay by Dr. Lynn Jorde on "Genetic variation and human evolution"
(see https://www.ashg.org/wp-content/uploads/2019/09/genetic-variation-essay.pdf)

"... Using a common definition that groups populations into major continents (Africa, Asia, Europe, and North and South America), many studies have shown that approximately 90% of genetic variation can be found within these populations, and only about 10% of genetic variation separates the populations.

Thus, the great majority of genetic differences can be found between individuals from any one of the major continents, and, on average, only a small proportion of additional differences will be found between individuals from two different continents.

Furthermore, because human history is a history of population movement, and because humans are extraordinarily adept at sharing their DNA, the genetic boundaries between populations are typically indistinct. For any given DNA sequence or gene, two individuals from different populations are sometimes more similar to one another than are two individuals from the same population. ..."

Quoted from essay by Dr. Lynn Jorde on "Genetic variation and human evolution"
(see https://www.ashg.org/wp-content/uploads/2019/09/genetic-variation-essay.pdf)

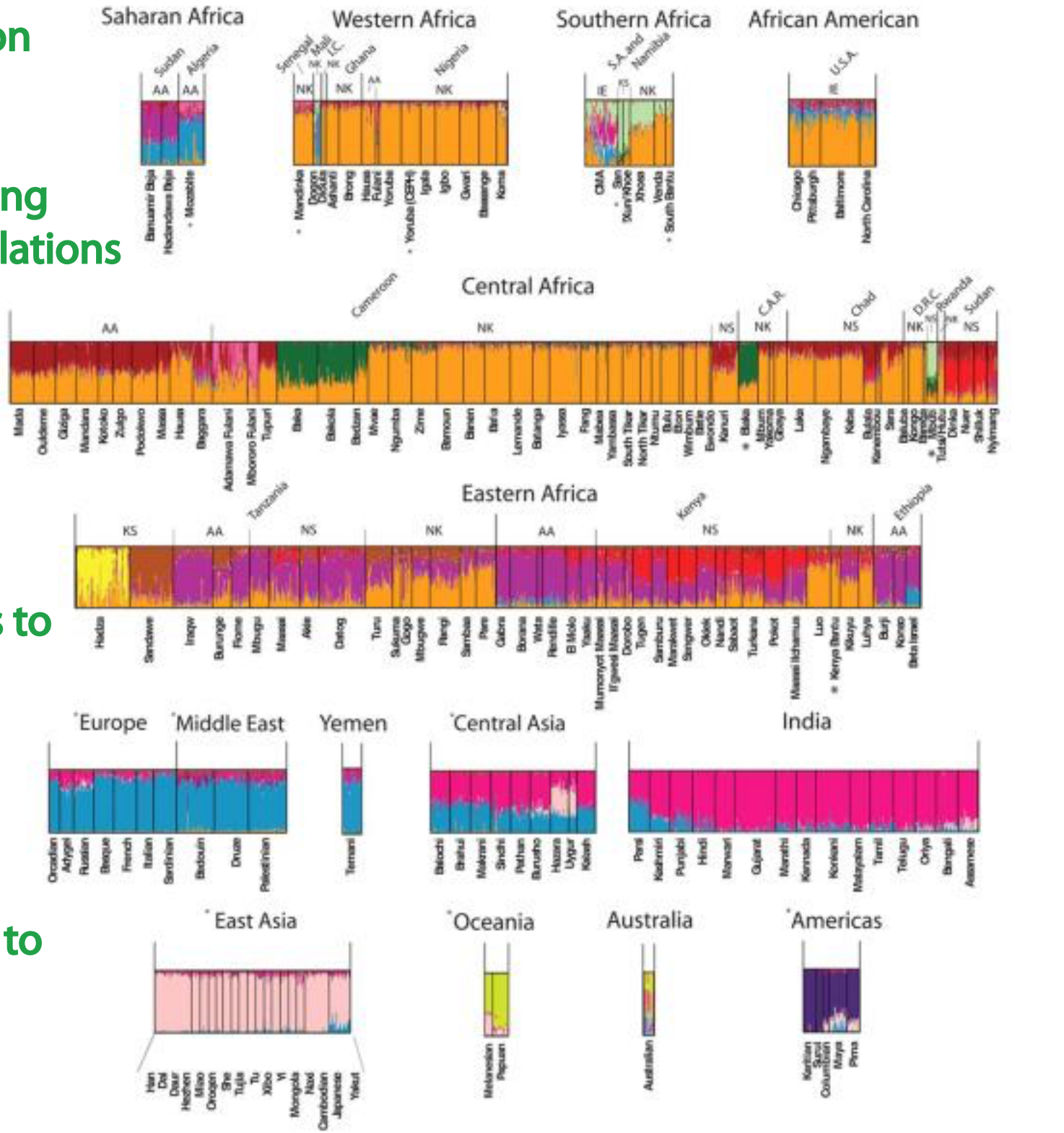
Human Genetic Variation Generalizations:

(1) Mixed ancestry among genes in existing populations

(2) Much genetic variation among African populations

(3) Within-population genetic variation tends to be higher in African populations than in others.

(4) Within-population genetic variation tends to decline with increasing distance of population from Africa



Above is Figure 4 from Tishkoff et al. 2009. Science 324:1035-1044. Above from "Structure Analysis with 14 populations. Summary also from same paper.

"Lessons from the genomic analysis of natural selection"
(from Nielsen et al. 2017. Nature 541: 302-310)

1. ***"...As humans migrated out of Africa and encountered new environments, introgression with other hominins that had already adapted to these environments seems to have been an important factor in facilitating rapid acceleration. This might be particularly true for genes related to immunity and defence against infection, as anatomically modern humans probably encountered pathogenic agents that could jump from other hominins, and to which humans did not yet have immunity. ..."***

“Lessons from the genomic analysis of natural selection”
(from Nielsen et al. 2017. Nature 541: 302-310)

2. “...much of the selection that has affected the human genome has been in response to changes in the environment that were induced by people. These include changes in diet that were driven by cultural innovations and an increase in the pathogen load of the population owing to changes in social structure and the emergence of cities. As we modify our environment, the resulting changes in conditions induce new selective pressures. Biological evolution and cultural evolution are therefore intimately linked. ...”

"Lessons from the genomic analysis of natural selection"
(from Nielsen et al. 2017. Nature 541: 302-310)

3. "...a close relationship often exists between genetic variants that have been under selection and those that have a strong influence on human health. Studies of human evolution are therefore of increasing relevance for medical genetics. For example, variants found to be selected for by adaptation to high altitudes provide a model for studying hypertension ..."

"Lessons from the genomic analysis of natural selection"
(from Nielsen et al. 2017. Nature 541: 302-310)

4. "...the genes that show the greatest difference in allele frequency between continental groups (indigenous Africans, Europeans, Americans and Australians) are enriched for associations with visible traits such as skin, hair and eye pigmentation. An interesting consequence is that the geographic groups are more different from each other in terms of pigmentation than they are, on average, at the level of the genome. Humans from various parts of the world are therefore more genetically similar than might be predicted on the basis of observed hair colour, skin colour or other visible traits. ..."

"Lessons from the genomic analysis of natural selection"
(from Nielsen et al. 2017. Nature 541: 302-310)

5. "...genetic variants with large effects, such as those that influence eye colour, hair colour or lactase persistence, are unusual. Instead, ... [Most traits subject to natural selection] "... seem to be highly complex and may be influenced by many loci across the genome. ... It is probable that a considerable component of selection in humans is polygenic and is yet to be discovered by studies that scan for genomic regions that are under selection."

SNP - single nucleotide polymorphism

Genetic variation among humans has a spatial organization along chromosomes. Specifically, there is often linkage disequilibrium between SNPs (i.e., if a chromosome has one SNP variant at one position then it tends to have another specific SNP variant at a nearby polymorphic position)

This means we can get a good guess as to what SNP variants are at positions in chromosome without actually experimentally determining the variants at each position.

"Haplotype Blocks" / "Linkage Disequilibrium Bins"

Human Genetic Variation has a "Block Structure"

Haplotype blocks in 500 kb region of human chromosome 5

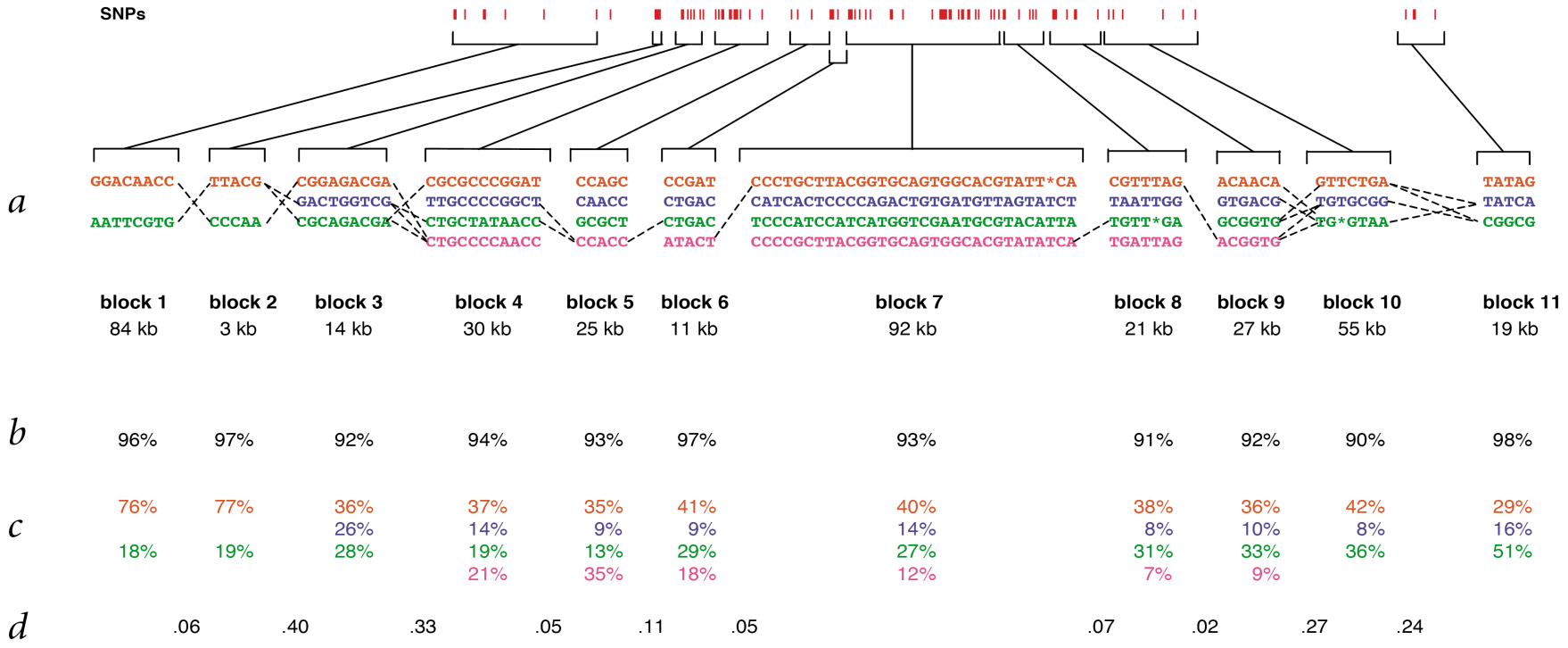
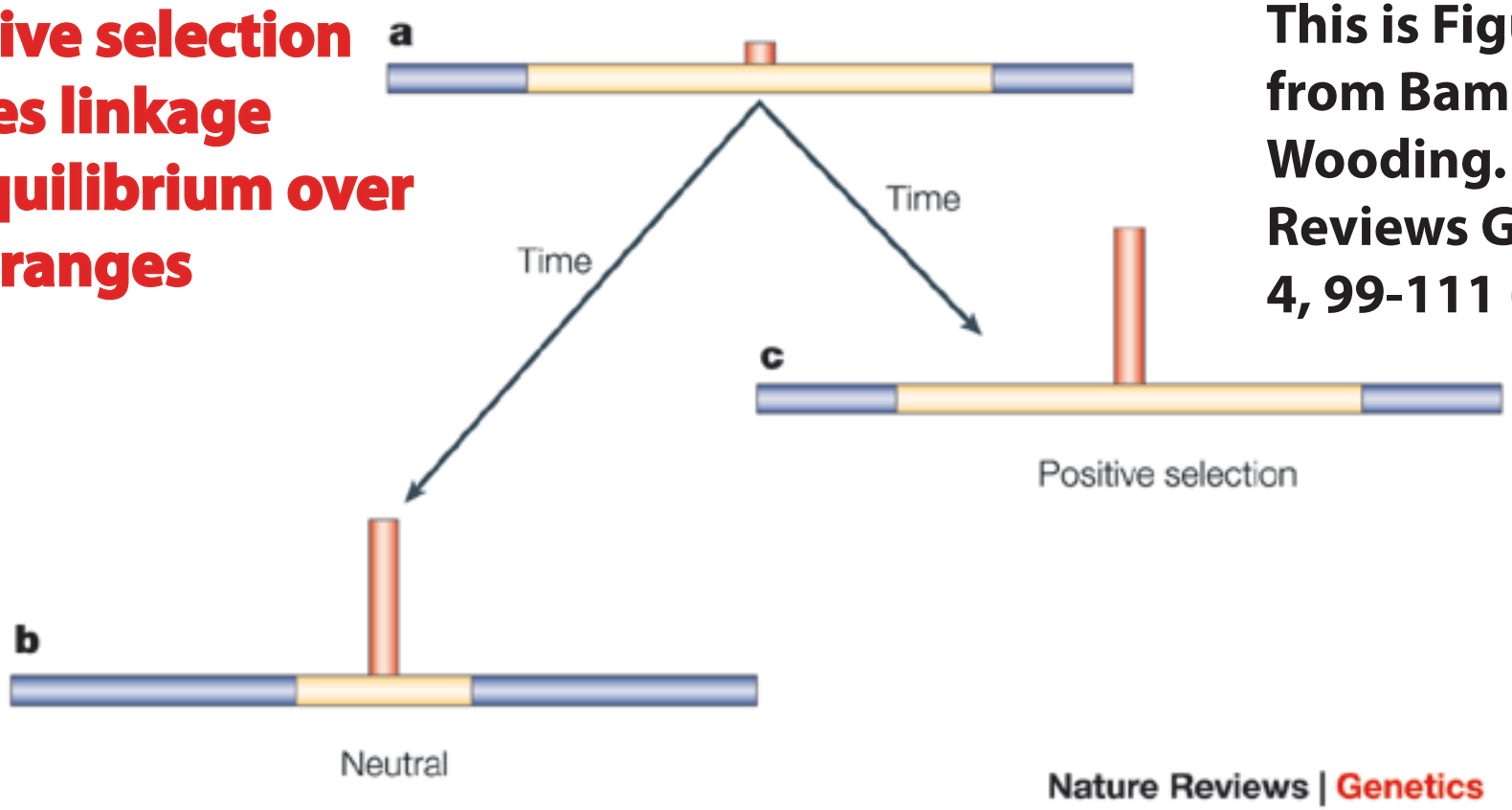


Fig. 2 Block-like haplotype diversity at 5q31. ***a***, Common haplotype patterns in each block of low diversity. Dashed lines indicate locations where more than 2% of all chromosomes are observed to transition from one common haplotype to a different one. ***b***, Percentage of observed chromosomes that match one of the common patterns exactly. ***c***, Percentage of each of the common patterns among untransmitted chromosomes. ***d***, Rate of haplotype exchange between the blocks as estimated by the HMM. We excluded several markers at each end of the map as they provided evidence that the blocks did not continue but were not adequate to build a first or last block. In addition, four markers fell between blocks, which suggests that the recombinational clustering may not take place at a specific base-pair position, but rather in small regions.

From Daly et al. 2001 (Nature Genetics 29:229-232)

Positive selection causes linkage disequilibrium over long ranges



This is Figure 4 from Bamshad & Wooding. Nature Reviews Genetics 4, 99-111 (2003)

"... a) A new allele (red) exists at a relatively low frequency (indicated by the height of the red bar) on a background haplotype (blue) that is characterized by long-range linkage disequilibrium (LD) (yellow) between the allele and the linked markers.
b) Over time, the frequency of the allele increases as a result of genetic drift, and local recombination reduces the range of the LD between the allele and the linked markers (that is, it creates short-range LD).
c) An allele influenced by recent positive selection might increase in frequency faster than local recombination can reduce the range of LD between the allele and the linked markers. **..."**

V_T - variance of quantitative trait (phenotypic variance)
V_g - genetic variance
V_e - environmental variance

$$V_T = V_g + V_e$$

Broad-sense heritability H^2 is

$$H^2 = V_g/V_T$$

$$V_g = V_a + V_d + V_i$$

V_a - additive variance
V_d - dominance variance
V_i - epistatic (interaction) variance

Narrow-sense heritability h^2 is

$$h^2 = V_a/V_T$$

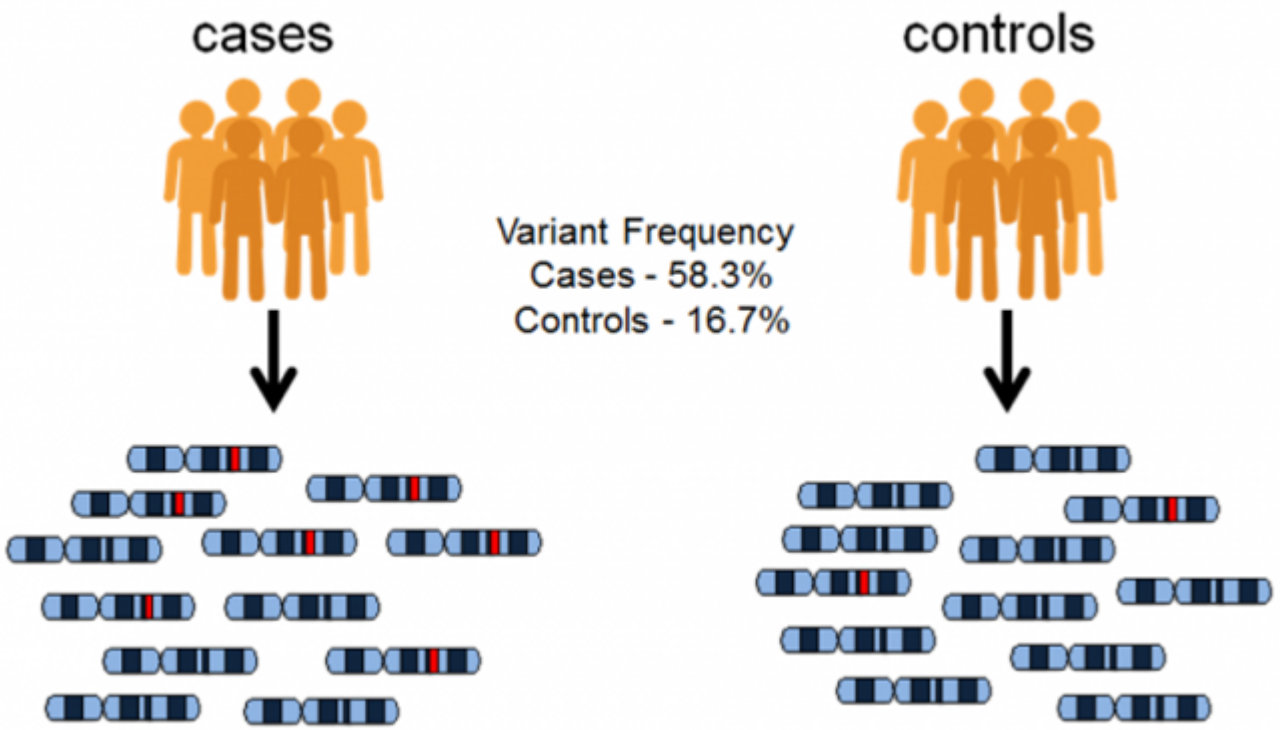
Estimates of Narrow-Sense Heritability
(ratio of additive genetic variance to phenotypic variance)

eyeball approximation by Jeff Thorne from Figure 2 in Boomsma et al. 2002, Nature Rev. Genetics 3:872-882

Trait	Females	Males
Smoking (#cig/day)	0.85	0.85
HDL-Cholesterol	0.72	0.72
Heart Rate	0.45	0.45
Thrill Seeking	0.65	0.65
Neuroticism	0.60	0.46
Testosterone Level	0.42	0.67
Boredom Suscept.	0.55	0.50
Religion	0.00	0.00
Intelligence*	.3,.5,.58,.62,.84,.88	

*heritability seems to increase with age of child in years 5-18

(estimates from monozygotic vs. dizygotic twin studies)



Genome-wide association studies (GWAS):

"Cases" may be more likely to have a certain nucleotide type at a specific genome position than "Controls"

image is from:
https://www.ebi.ac.uk/training/online/courses/gwas-catalogue-exploring-snp-trait-associations/what-is-gwas-catalog/what-are-genome-wide-association-studies-gwas/

"Manhattan Plots" are often employed to depict GWAS results. They help visualize which SNPs seem most associated with the phenotype of interest.

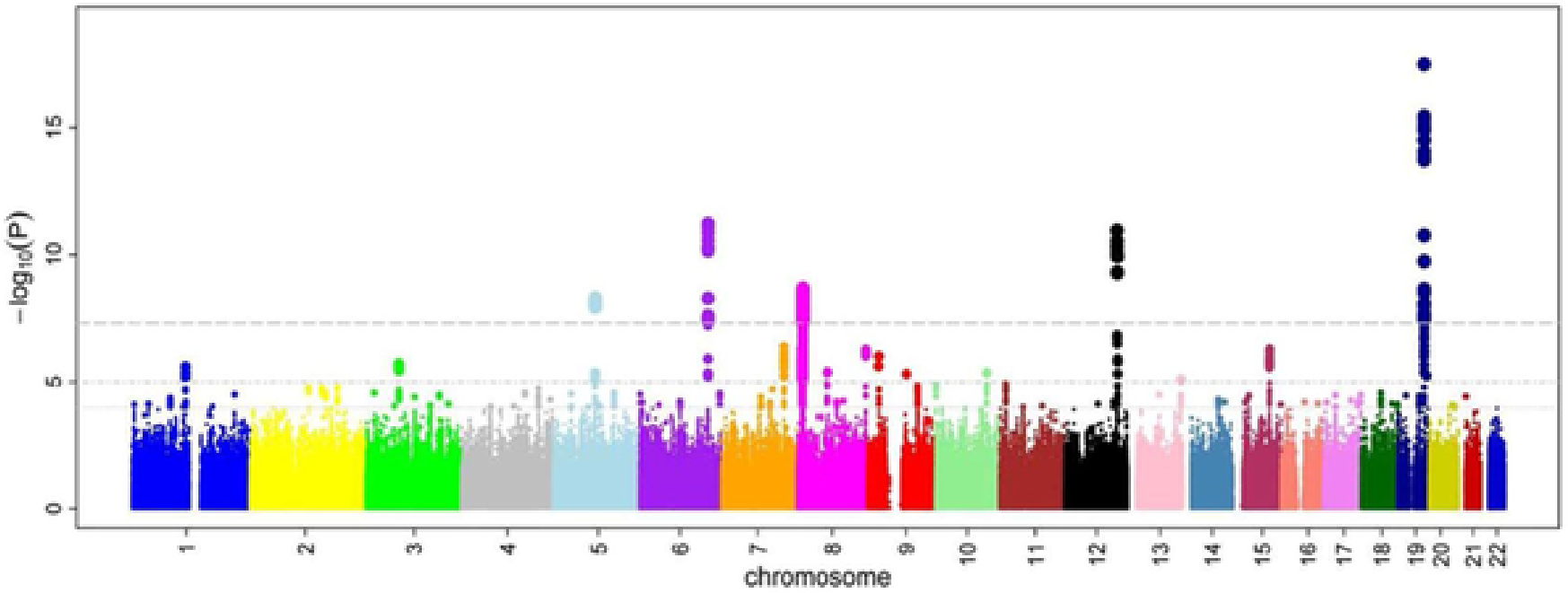


Image is from: https://en.wikipedia.org/wiki/Manhattan_plot

Missing Heritability: (formerly) The big mystery in gene mapping

By looking at correlations in trait values (e.g., height, disease status, etc.) between relatives and by knowing the proportions of genes shared among relatives, good estimates of the genetic contribution to trait values can be made.

Despite an abundance of available DNA markers, GWAS (Genome-Wide Association) Studies have been unable to find the genes responsible for much of the genetic contribution to traits.

Why? Some possibilities that are all at least partially correct ...

Do rare genetic variants make contributions? Low penetrance of variants (i.e., variants only have effect in some people)? Epistasis (interactions between genes) ? Epigenetics? Additive genetic variation of tiny effects?

Some Difficulties in gene mapping ...

Linkage Analysis in Pedigrees can lead to general area in genome of gene but hard to find precise location (due to insufficient recombination)

Probands / Ascertainment Bias in pedigrees

Incomplete Penetrance

Complex Genetic Traits

Many Gene Markers and Multiple Testing

Population Structure (a difficulty for association mapping)