The FUTURE Of Our History

Molecular Evolution, Phylogenetic Studies, and Admixture Mapping

By Nguyen Nguyen

Within each of us, microscopically packed among histones, are the answers to humanity's deepest questions. Where do we come from? How long have we been on earth? What makes us different from other animals? The answer is in our very own DNA, simple strands of genetic code that have been passed down from generation to generation. And it is this DNA, analyzed by an evolutionist with a historical perspective, which has sparked the field of molecular evolution. Molecular analysis of evolution combines recent genetic analyzing technologies to find evidence of change generated within the human genome. Natural selection and adaptation are reconsidered, and by mapping and studying on a large scale human variation in DNA, scientists are able to conclude when and how new genes evolved. Studies in molecular evolution have been used to infer that chimpanzees are humans' closest relatives. Even viral evolution and the origin of SIV and HIV are illuminated through studies of molecular evolution. Taken a step further, scientists have been able to link susceptibility to diseases, such as prostate cancer, to specific sites in the human genome. Thus, molecular evolution is being used to explore our history as well as the origins of diseases and viruses, through the use of modern technology.

Molecular evolutionists focus on two interrelated areas, one concerning the evolution of the genetic material, such as DNA and RNA sequences, and the second involving the reconstruction of the evolutionary history of genes and organisms (1). Molecular analysis allows us to explore the overlap of these two areas to create a new, innovative field.

Before the advent of DNA analyzing techniques, scientists studied evolutionary history through phylogenetic trees. These phylogenies related the ancestry of animals in a linear, branching fashion. However, these trees were only based on observable evidence; scientists would use morphology – measuring the sizes and shapes of body parts – and developmental traits. Microscopes and computer simulations also enabled scientists to precisely measure features and categorize organisms. Furthermore, evidence from the fossil record gave environmental context about where organisms lived. These trees were limited in scope because morphological and developmental evidence is based on qualitative categorization (2).

However, advancements that have occurred in the last forty years allow us to use molecular data to construct a more accurate phylogenetic tree. In particular, the introduction of protein sequencing methodologies, which involves finding the order of amino acids, has been useful in piecing together evolutionary connections. DNA has also been useful because it provides a strictly heritable basis of evidence. The sequences that are revealed are those passed down from ancestor to progeny. The study of DNA is helpful because it has not been influenced greatly by environmental factors and it mutates in a predictable manner that can be modeled with mathematics. Thus, the benefits of using DNA sequences in molecular evolution have been applied to numerous historical and practical fields. We will explore this process through three examples: how viral evolution can be tracked in a case of the HIV virus; the case of human's closest relative; and, finally, clinical applications of disease-related molecular mapping.

Molecular Modeling of HIV-1

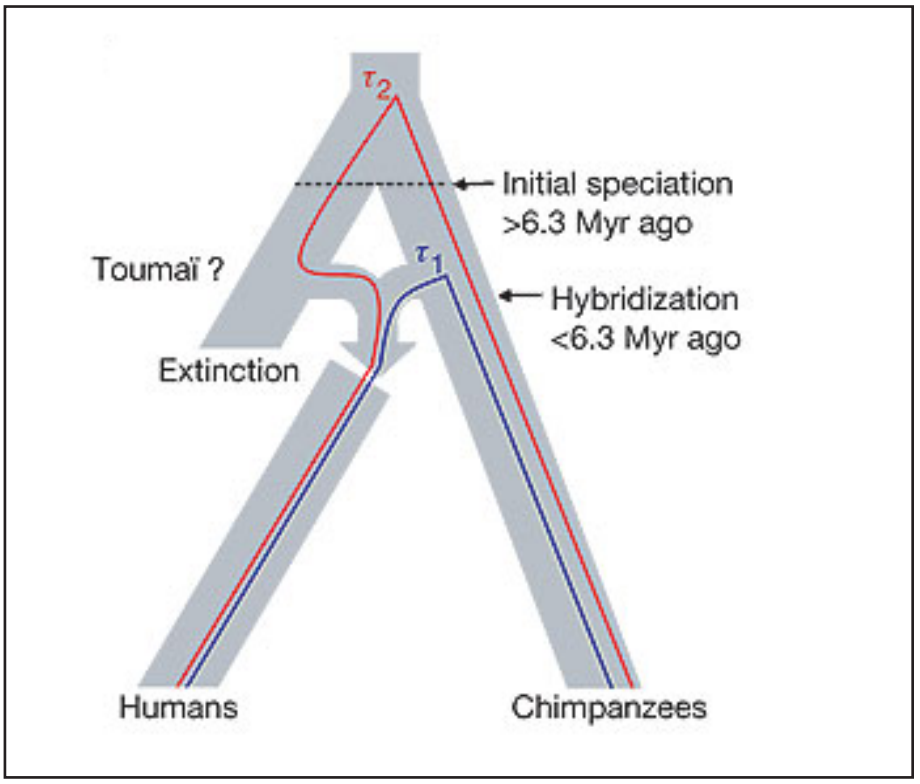
How can we use raw DNA information to create branching phylogenetic trees? Once the molecular data has been gathered, the process begins with aligning the specific DNA bases from different samples for comparison. Then, specific bases that differ between species tell scientists where change has occurred. These sites are marked and analyzed with several mathematical models that map out which sequence was the common ancestor (1).

This phylogenetic tree building is particularly useful in tracking down the quickly evolving HIV-1 strains. Recently, scientists have been able to characterize the strains of the immigrant population living in New York City by obtaining DNA samples from immigrants and sequencing the viral DNA. It was found that 43.4% of the cases were newly recombined forms of the virus, atypical to NYC; and the other 56.6% were similar to common North American subtypes. The recombined forms of the virus were actually found to be typical to the country of origin of immigrants, suggesting that many immigrants brought the virus from their home country (3).

Being able to genetically analyze origins of the HIV-1 virus, in a specific case such as New York, has allowed for public health officials and doctors to take note of the complex strains, possibly helping with their assessment of the disease and the nature of treatment. By making use of current molecular evolutionary technology, scientists can map the genetic diversity and evolution of viruses, while gaining a deeper understanding of the origins of HIV.

Chimps Tell the Story

Humans have been known to be related to chimpanzees and gorillas for hundreds of years, but until recently it was difficult to determine precisely which species was our closest relative. Since Darwin, many have contested the dominant thought that African apes were the closest related to humans.



credit: adapted from Ref. 6

Figure 1. Harvard Medical School'ss Reich was able to estimate that it was less than 6.3 million years ago that speciation, the split between human and chimpanzee to create two species, occurred.

It was not until the last two decades that geneticists and anthropologists could gather the evidence to prove that humans were most related to chimpanzees. In 1989, using DNA-DNA hybridization methods, Adalgisa Caconne and Jeffrey Powell mixed together single strands of DNA to conclude that chimpanzee and humans were most closely related genetically (4). By 1997, Maryellen Ruvolo, a professor in Harvard's Biological Anthropology department, had extensively studied multiple independent genes to show that humans were most closely related to chimpanzees (5).

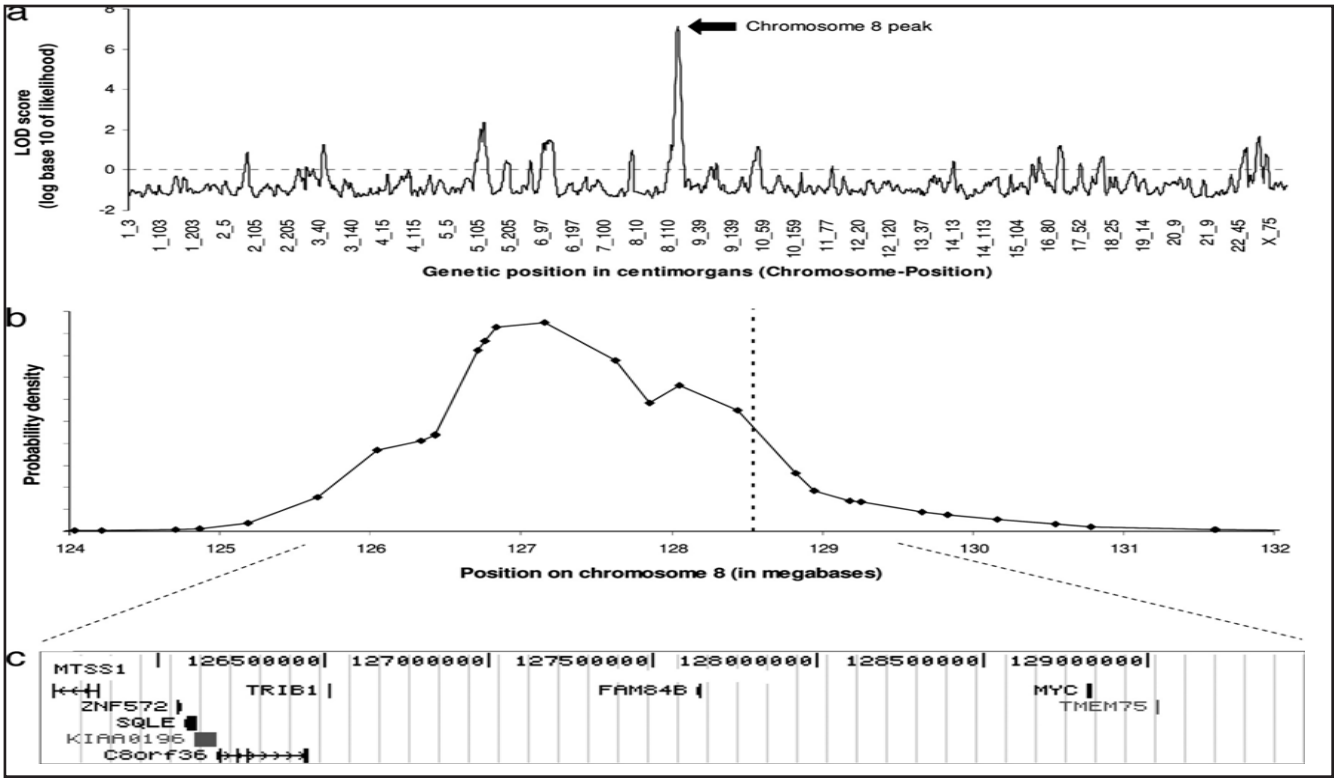
Most recently, Harvard Medical School's David Reich was able to estimate that it was less than 6.3 million years ago that speciation, the split between human and chimpanzee to create two species, occurred. Interestingly, the estimated time of divergence is shorter than what scientists believed by several million years, conflicting with some interpretations that had been based on ancient fossils (6).

Reich's work involved aligning 20 million base pairs, taking the sequences of DNA (found in nucleotide base pairs) from humans and chimpanzees and comparing the homologous chromosomes, proteins and individual genes. He concludes from his analysis that one uniquely human sex chromosome, the X chromosome, had actually evolved only recently and had an "extremely young divergence time" (6). From his evidence he hypothesizes that chimpanzees and humans may have diverged earlier, then exchanged particular genes over time through hybridization. Ultimately, Reich's work represents a unique application of molecular evolution because it enables scientists to rethink interpretations of human history.

What the Future Holds

Molecular evolution is a unique field that is quickly expanding, and potential applications in other fields are among the most distinguishing features of this technology.

One growing field is admixture



credit: adapted from Ref. 7

Figure 2. Admixture mapping tracks down the exchange of genetic code from population mixing and pinpoints genetic diversity that is produced through reproduction.

mapping of human genes. Admixture mapping involves comparisons of chromosomes from historically distinct ancestry, and looking for differences due to "mixing" of these distinct alleles. Specifically, admixture mapping tracks down the exchange of genetic code from population mixing and pinpoints genetic diversity that is produced through reproduction (7). Scientists do this by taking DNA from a group of organisms that share a particular trait, and using markers to track down similar sequences of DNA. This technique allows them to infer that the trait is linked to the DNA at a particular location, or locus.

The process of admixture mapping involves sequencing the entire genome so that all of a person's nucleotide bases are known. Then, scientists "stick" markers, which are complementary to particular target sequences, to the DNA. When this process is carried out over a large sample group, associations can be found between particular loci across chromosomes for a particular trait.

It is easy to imagine that admixture mapping might be very relevant to disease mapping (7). In fact, scientists have recently found links between region 24 on the long arm of chromosome 8 and prostate cancer in African American males. Prostate cancer was studied because African American males have the highest risk of prostate cancer in the U.S. among the ethnic groups, and mapping was done in order to assess the genetic factors. Admixture mapping is useful because of the relatively recent mixing of European and African populations, so that associations between known sequences of DNA could be found using only several thousand markers (7).

With the advance of admixture mapping, the wide world of genetics and molecular analyzing techniques can be applied to disease detection. Twenty years ago this technology would not have been possible; the first admixture mapping was done in 2005 based on recent advances related to DNA markers and mathematical analysis. By fusing molecular technology and an evolutionary perspective, scientists are approaching science from a new vantage point. Molecular evolution has shed light on the genetic basis of common diseases, and this knowledge will lead to improved detection and treatment of these diseases, benefiting a large number of those affected. H

—Nguyen Nguyen '09 is a Human and Evolutionary Biology Concentrator in Leverett House.

References

1) Graur, Dan and Li, Wen-Hsiung. *Fundamentals of Molecular Evolution*. (2002): 190-202.
2) Purves, Sadava, Orians, Heller, *Life*. (2004): 500-503
3) Lin H, Gaschen BK, Collie M, El-Fshaway M, Chen Z, Korber BT, Beatrice ST, Zhang L. "Genetic characterization of diverse HIV-1 strains in an immigrant population living in New York City." *J Acquir Immune Defic Syndr*. (2006): 399-404.
4) Caccone A, Powell J. "DNA Divergence Among Hominoids." *Evolution: International Journal of Organic Evolution*. (1989): 925-934.
5) Ruvolo, M. "Molecular Phylogeny of the Hominoids: Inferences from Multiple Independent DNA Sequence Data Sets." *Mol. Biol. Evo*. (1997): 248-265
6) Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D. "Genetic Evidence for Complex Speciation of Humans and Chimpanzees." *Nature*. (2006):1103-8.
7) Freedman ML, Haiman CA, Patterson N, McDonald GJ, Tandon A, Waliszewska A, Penney K, Steen RG, Ardlie K, John EM, Oakley-Girvan I, Whittemore AS, Cooney KA, Ingles SA, Altshuler D, Henderson BE, Reich D. "Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men." *PNAS*. (2006): 14068-73.