Evolutionary Anthropology 17:208–212 (2008)

A Formidable Shape. . .That Seem'd Woman

Mitochondrial Eve and her co-Eval stepsisters

KENNETH M. WEISS

In the beginning, when Satan left Hell to go wreak havoc on humankind, he had to barge his way through enclosing gates guarded by "a formidable shape" that "seem'd Woman to the waist and fair; but ended foul in many a scaly fold... about her middle round a cry of Hell Hounds." (Fig. 1). At least this is how John Milton told our origin myth in *Paradise Lost*.[1] When these Cerberean three-headed dog-like beasts were disturbed they would "creep...into her womb, and kennel there." These monsters were "hourly conceiv'd and hourly born" and, when in the womb, she proclaimed, they "howl and gnaw my Bowels, their repast; then bursting forth afresh with conscious terrors vex me round."

One would have thought that by then Satan would have had enough of woman but the Hell-gate guardian was not even his main target. That was Eve, whom Satan, disguised as a snake, tempted to eat the forbidden apple. However, here I'm presenting a tale not of religion, but of a different Eve, the so-called mitochondrial Eve, which is a tale of science and which, unlike Milton's, turns out not to be sexist.

Ken Weiss is Evan Pugh Professor of Anthropology and Genetics at Penn State University. E-mail: kenweiss@psu.edu

DOI 10.1002/evan.20184
Published online in Wiley InterScience (www.interscience.wiley.com).

A BATTERY OF GENERATORS

Our cells depend on small organelles called mitochondria, which play a vital role in energy-capture physiology and perform several other functions.

A typical cell is populated with many of these power generators (Fig. 2A). Some cells have only a few tens of mitochondria, but most have roughly a thousand. As opposed to the 50–70 mitochondria in a sperm cell, there are about 100,000, perhaps up to 200,000, in a human oocyte.[2]

Mitochondria contain their own genes. The human mitochondrial genome is a circular DNA molecule 16,569 nucleotide base-pairs long. These code for 37 different genes, 13 of which are parts of respiratory biochemical complexes. To translate these into protein, 22 mtDNA genes code for transfer RNA and 2 for ribosomal RNA (Fig. 2B). Within each mitochondrion are about 10 separate copies of this mtDNA. Individually, 16,569 bp is trivially small compared to the 3,100,000,000 bp of the nuclear chromosomes, of which there are two sets in each cell. But with the typical 1,000 mitochondria each having 10 copies, there are about 165,690,000 base pairs of mtDNA in a cell. That's not so trivial. With a thousand times more, an egg may contain 165 billion nucleotides, 25-fold *more* than the DNA in the nucleus.

A CELEBRATED ORIGIN MYTH ON THE EVE OF THE HUMAN DAWN

Mitochondria are mighty, not just as power cells, but also in their contribution to anthropological genetics. Variation in the mtDNA sequence has been important in the reconstruction of human evolutionary history. It has been used to address two main questions about modern human origins: Where did our species originate? And when? The answers, in an oversimplified nutshell, are that our species originated in Africa about 150,000 years ago. This result was presented in a now-classic paper in 1987 (Fig. 3).[3] In broad terms, the result has held up.

These inferences rest on two central facts. Most mtDNA sequences outside of Africa have closely related sequences found within Africa, but there are deep branches on the tree of sequence similarities; that is, there are relatively great sequence differences found only among Africans. It can be shown mathematically that all sequences present today in any gene are the diversified descendants of some single ancestral sequence, called the most recent common ancestor, or the coalescent. The rate at which sequence divergence accumulates in descendant copies of an ancestral coalescent sequence over time depends on how many copies are being transmitted at any given time. Given the mutation rate, the amount of mtDNA variation present today is estimated to have taken about 150,000 years to accumulate. That is why the report of mtDNA-based analysis of human origins referred to the bearer of the mtDNA coalescent sequence in a cute and newsworthy but badly misleading way that mixed religion and science, as "mitochondrial Eve."

The typical number of copies of a gene transmitted in each generation

Figure 1. Before the gates…a formidable shape. Detail from Gustav Doré's illustrations of *Paradise Lost,* II:648–649 (1866, Paris: Cassell & Co.).

is known as the effective population size and reflects, among other things, the amount of target DNA that is subject to mutation. As in any probabilistic process, divergent results are more likely if the sample is small: having Heads come up every time in a few coin flips is far more likely than a few hundred flips all coming up Heads. In any sexually reproducing population, in each pair of mates there are 4 copies of every chromosomal gene except those on the X chromosome of which there are 3 copies. But there is only a single copy of each mtDNA transmitted and a single copy of each Y chromosome (Fig. 4). mtDNA mutation rates are also higher than in the nucleus, because mitochondria do not contain the same nuclear mutation-repair mechanisms. Fewer copies and more variation means more signal to track per unit of historical time, making mtDNA particularly informative for reconstructing relatively recent evolutionary history.

In the case of mtDNA, the smaller number of copies transmitted per couple per generation is due to the fact that mitochondria are passed only through the maternal line, because none of the sperm's mitochondria make it into the egg. That's the theory, at least, but it relies on an additional assumption, that although the egg contains many mitochondria, each with many mtDNA copies, they all have the same sequence, a characteristic known as homoplasy. But is that theory accurate?

In fact, I've blithely slid past an obvious complication: While there may be a mitochondrial ancestress, it isn't a unitary Eve. Like the creatures hatched by the guardian of the gates of Hell, that primeval sequence had a lot of coeval stepsisters; they are in every cell, where their presence can be as evil as if they were eating her from within. Also, at least some copies do occasionally slip into the fertilized egg from sperm.

Each time any mtDNA replicates, new errors are possible. The mutation rate in nuclear DNA is about one per 10,000,000 nucleotides per cell division. At that rate, a cell with 165,690,000 mtDNA nucleotides should transmit around 16 new mutations when it divides; that's 0.1% of the entire mtDNA being altered by a new mutation in at least one copy in each new cell. Having more than one mtDNA genome in the cell at any given time is known as heteroplasmy (Fig. 4). An oocyte alone, which develops around 100,000 mitochondria during its follicular development stage,[2] may carry at least 1,600 new mutations; that is 10% of the nucleotides mutant in at least one copy of the mtDNA in each egg. But because mitochondria lack repair mechanisms, their mutation rate may be 1,000 to 100,000 times greater than that in the nucleus, with mind-boggling numbers of sequence variants per egg.[4]

When a mutation occurs in a cell, the cell's entire lineage in the body during the person's lifetime will inherit the variation. New mutations will arise with each new division in that tree of cellular descent. Counts are only approximate, but a human body contains some tens of trillions of cells. Thus, within and among the cells in a tissue, and among tissues in the body, a huge tree of mtDNA evolution should result. We are unpredictable genetic mosaics in a way that would make Gaudi proud. Mutations that occurred earlier in our embryological development will be found in more tissues than will mutations that arose later. The sequence divergence among our cells should grow markedly bushier with age. That suggests that our Ms. Eve who carried the ancestral sequence must also have carried and transmitted other sequences as well.

Could there really be this enormous amount of variation? Direct tests have generally suggested that the actual amount of heteroplasmy is small. But those tests usually are not sensitive to variation within individual cells, and at least some

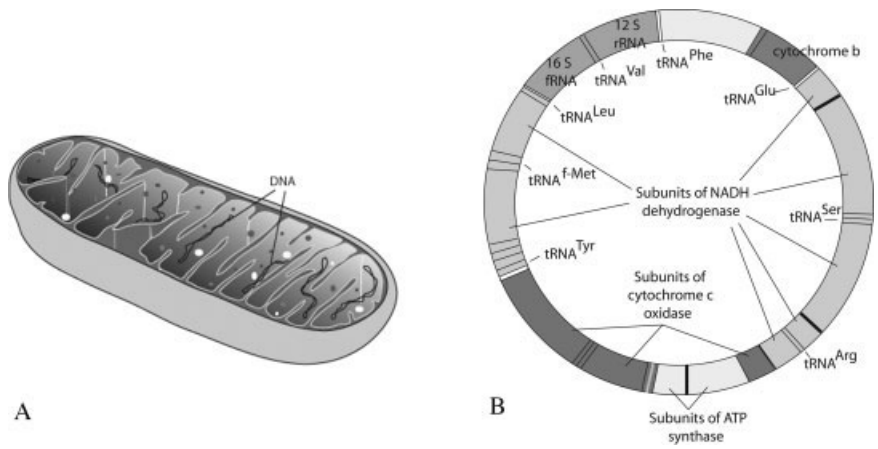

Figure 2. Mitochondria. A. A cutaway view showing the basic structural characteristics: internal and external membranes, ribosomes, and other functional units (white and darker dots), and the several circular mtDNAs. B. mtDNA genes. Drawings by Anne Buchanan.

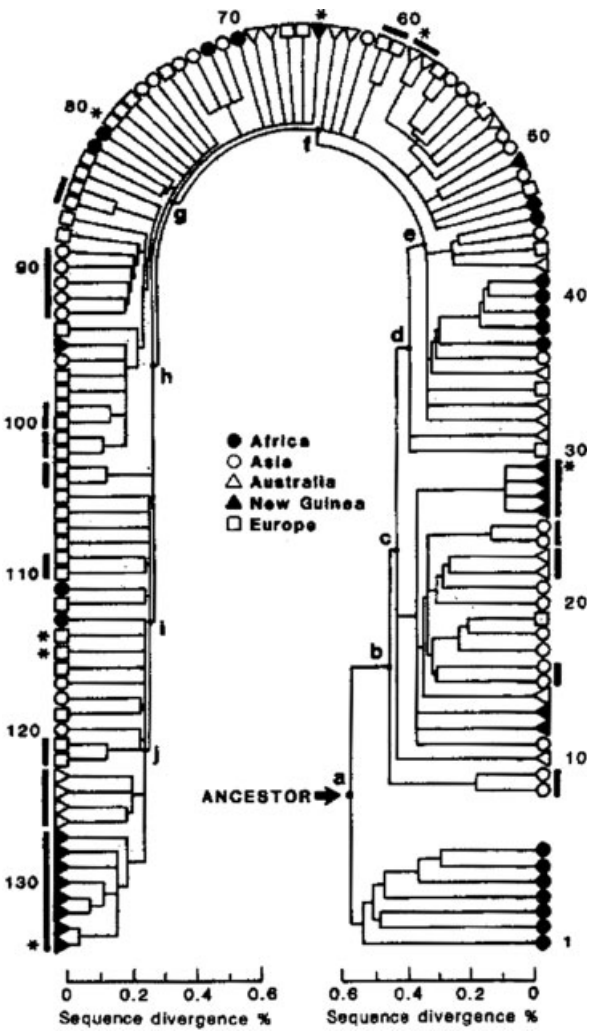

Figure 3. The famous human mtDNA tree. The inferred ancestral sequence is shown. Most of the tree shows close relationships between African and non-African sequences, but a deep Africa-only clade is on the bottom right. Reprinted from Cann, Stoneking, and Wilson[3] with permission and copyright by Nature.

variation is, in fact, known to be transmitted. Why isn't there a much greater amount or, if it arises, where does it go? Some recent papers have addressed this question.[5–7]

MIGHTY CHONDRIA STRUCK LOW

Like any part of our DNA mitochondrial genomes can contain mutations that are harmful. Roughly one in 4,000 children in the United States develop inherited mtDNA disease. These diverse disorders include muscle energy diseases (exercise intolerance), complications of diabetes, optic neuropathy leading to vision loss, and neural excitation disorders of the heart. One signpost to discovering these diseases is that they show strictly maternal-offspring inheritance. Inherited mitochondrial disease affects every cell and thus can, if severe or early enough, reduce reproductive fitness, so that natural selection removes the mutation from the population.

Mitochondrial DNA can cause harm in another way that is related to our understanding of evolution itself. Mitochondrial mutations are hourly conceived and hourly born in our huge number of cells, so each cell has thousands of copies that are not actually identical. In fact, some studies, for example, in mice where detailed data can be obtained, do find heteroplasmy. Variation has also been found among somatic cells, including those of humans.

Various mechanisms have been suggested to account for the apparent fact that there is a lot more variation within generations than is transmitted across generations. One idea is that every cell generation, in somatic as well as germline ancestral cells, there is a sampling, so that not all of its hundreds of mtDNAs are proportionately represented in its two daughter cells; each division produces two cells, each with only a fraction of the variation found in the parent cell. This is a kind of genetic drift that essentially reduces the mtDNA sample size and could lead to more rapid loss of variation, counteracting the input of new mutation.

Another idea is that the mtDNA in a cell are packaged into hypothesized structures called nucleoids, and it is these that undergo drift.[7] Rather than sample from all of a parent cell's 1,000 mtDNA, a daughter cell might sample a much smaller number of nucleoid packets, which could be a more severe form of variation-reducing drift.

A third idea is that in germline cells the mitochondrial count is not so large and may fluctuate, produc-

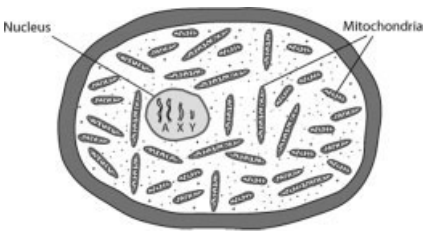

Figure 4. Typical cell. Relative proportional ''doses'' of different aspects of the genome: in the nucleus, 2 copies of each autosome (A), one or two sex chromosomes (XY in males or XX in females, not shown), and mitochondria. The latter are far more numerous, and are shown here in varying sizes to suggest their at-least-slightly heterogeneous mtDNA sequences. Drawing by Anne Buchanan.

ing several size bottlenecks, repeatedly reducing variation before a woman finally forms the mature oocytes, whose 100,000 mitochondria may thus be misleading.

These are all drift scenarios, based on statistical sampling of variation, in which small-sample bottlenecks reduce the total amount of variation. But there is another mechanism, too, which has to do with mitochondrial function. In all cells, including those that leave their mother, each sequence has to get along with its mutant stepsisters. If mutations in mtDNA accumulate, they may cause harm to the cell. If it is serious, the cell will die, taking with it all its mitochondrial variation.

Cells may accumulate changes that are harmful but not lethal to the cell. The person will not notice such changes unless, over time, they accumulate in too many cells. If that happens, the person's fitness may be reduced, as if he or she had inherited a single, harmful mtDNA sequence affecting all cells. Pathology can arise even from nondividing cells in muscle or neurons in the brain. As one ages, an increasing fraction of these cells may have increasing fractions of mutant mtDNA. Mutagenic compounds produced in normal energy metabolism in mitochondria may damage their DNA. They would accumulate over time, and signs of aging could result. Indeed, cells like these that do not normally divide during adult life may be especially vulnerable to such unrecoverable damage.

A recent example illustrates this point, along with the integrated nature of the mitochondrion in the cell. Mutations in an enzyme coded by a nuclear gene, which is required for the mtDNA to express its own genes, can cause mutations in mtDNA to accumulate. This, in turn, can lead to premature aging in different tissues such as muscle (heart) and brain.[8] The clonal expansion of copies of mutant mtDNA in a cell can lead to symptoms when they become a substantial fraction of all the mtDNA copies in the cell. Whether this actually produces an effective form of natural selection to remove this variation is debatable, because the effects occur largely later in life. In any case, these are mutations in somatic cells, not in the transmitted germ line.

However, the same processes can occur in the germline, where cells with damaged function may not yield very fertile eggs, if the cell even survives long enough to try.[2] Also, helpful mutations may have salubrious effects. In either case, this would constitute selective proliferation among a woman's cells that need not be related to her overall fitness since the loss of a few thousand defective germ-line cells would not be noticed. They would be only a trivial proportion of her millions of other germline cells.

This may seem to have little relevance beyond a technical fine point related to mitochondrial variation. Since it applies only to female lineages, it may seem sexist as well. However, that could be a mistake. Like the hounds spawned by the guardian of the gates of Hell, some may howl and gnaw at their host's bowels. To the extent that they cause harm to the cell, evil mitochondrial stepsisters in a cell are condemned. This may account for the consistency of evolutionary reconstructions of human ancestry that assume a mother transmits only a single mtDNA sequence to her children. In addition, somatic purging of mtDNA may also reveal the tip of a much deeper evolutionary iceberg, which applies to males as well as females, helping to answer a related question that was raised in the past but seems largely to have been forgotten.

BEASTS BATTLING SILENTLY WITHIN: A MAJOR EVOLUTIONARY FACTOR?

Almost every day, the media have stories about some newly discovered case of natural selection in humans, such as those associated with malaria, the evolution of skin color, and adult lactose tolerance. The evidence is in changes in genes that code for protein (classical "genes") or some other known function. These changes are significantly more rapid or slower than those in parts of the genome with no known function. Something has to be purging or favoring the variation that arises in these genes. While various explanations are possible, selection of some sort is clearly one of them.

In the 1960s, advances in genotyping methodology made us aware of an unexpected amount of genetic variation. If most mutations are harmful, what is maintaining this variation in the population? Two explanations for the amount of variation were vigorously contended. One held that, as in the sickle-cell story, balancing selection maintained two different variants at a gene at intermediate frequencies in the population because either allele by itself was relatively harmful: Aas were better than both AAs who had malaria and aas who had anemia. But balancing selection entails a lot of compensating reproductive excess to replace the AAs and aas who die young or fail to reproduce as well as the Aas do. If natural selection is keeping multiple alleles at thousands of genes in this way, there would have to be a lot of excess reproduction, a "genetic load" carried by the population to compensate for the selective loss of individuals who bore the less-favorable alleles. But there are so many genes. How could we bear that load? One idea was that each selective loss might simultaneously discard many different harmful genotypes.[9] The second explanation was the "neutral" theory, at the time called non-Darwinian evolution. According to this explanation, if genetic variants had no effect on reproductive success their frequency could change over time by genetic drift (chance) without causing any reproductive burden—no genetic load. As population genetics focused on neutral theory, the problems seemed to have gone away. But that may have been premature.

In recent years, extensive DNA sequencing has hatched the beast of genetic load in an inverse way. It is not that there is too much variation, but there is too little variation in too much of the genome to account for easily by natural selection. There are hundreds of millions

of nucleotides in our genomes that reflect restricted variation. For example, the amino acid coding regions for our 20,000 or so genes vary much less than do areas with little or no function. What is purging the mutations that must arise all over the genome in each individual in each generation?

The problem may even be worse, because nucleotides that have widely been regarded as evolving neutrally may be affected by selection after all. For example, genes typically have multiple protein-coding sections exons, which are separated by noncoding introns RNA is transcribed from the whole run of DNA, but then specific proteins actively cut out the introns and splice the exons together to form the mature messenger RNA that is translated into protein. But many genes, if not most, are spliced in different ways to use different subsets of their exons in different contexts. The pattern is often highly conserved, sometimes even among species. Mutations in an exon that do not change the amino acid coded for are called synonymous mutations. They were thought to evolve by drift, invisible to selective forces, while those forces can "see" the mutations that cause a code to specify a different amino acid than it did before the mutation. But selection may also screen the synonymous changes, because the mutated nucleotide may affect various functions, such as the ability of the splicing proteins to recognize the RNA splice sites.[10,11] The variation in nucleotides neighboring the supposedly neutral sites in amino acid codons is also constrained.[12] Thus, supposedly synonymous mutational changes may not be entirely neutral after all.

Discovery after discovery has been adding additional function to our genome, and hence corresponding conservative selective effects on sequence variation. Even classical "junk" DNA elements, such as copies of short sequences that litter our genomes and occasionally disperse to new sites, which originally came as viral or other intruding exogenous DNA, may have an important function in regulating our tissue-specific use of genes.[13] Why don't they cause a huge genetic load on our slow-reproducing population? Few people seem to be concerned with this problem, but our little mitochondrial stepsisters, gnawing away at our cells, may suggest to us a way out.

Prezygotic selection, which occurs before a sperm or egg cell is formed, may, based on cellular function, provide one relief from this hellish selective burden. mtDNA is not the only DNA that is subject to mutation. Each cell may have only two copies of the nuclear chromosomal genome, but each copy is billions of nucleotides long, and we have trillions of cells, millions of them in the male and female germline lineages. These cells have to pass the functional test: Cells that misbehave and try to devour their host from the inside, they will be purged. They will never survive to be turned into a fertilized egg that grows into an organism that has to compete for survival with other organisms in the classical Darwinian way. But the burden to these cells may pose no real genetic load to the individual as a whole.

Eve was the object of Satan's venom. (No one, to my knowledge, has asked where the mtDNA of Hell's gate-guardian would fall on the sequence tree). However, the fingers do not point only at women; similar kinds of discrimination go on somatically at all chromosomes, including the macho Y. Currently, we just know much less about that than we do about mtDNA. However, we do know that somatic mutation occurs. We can see it when a disease like cancer or age-related changes result in gonadal mosaicism, in which a fraction of sperm or egg from the same individual carries variation that the person did not inherit from his or her parents; or in the increase of genetic disease in the offspring of older mothers or fathers. We cannot say how important prezygotic selection may be, but there is clearly much to be learned. The struggle to control the beasts within before hatching them into the world is a story that, unlike Genesis, is still largely untold.

NOTES

I welcome comments on this column: kenweiss@psu.edu. I have a feedback and supplemental material page at http://www.anthro.psu.edu/weiss_lab/index.shtml. I thank Anne Buchanan, Heather Lawson, Sam Sholtis, and John Fleagle for critically reading this manuscript. This column is written with financial assistance from funds provided to Penn State Evan Pugh professors.

REFERENCES

1 Milton J. 1668. Paradise lost. London: Peter Parker.

2 Jansen RP, de Boer K. 1998. The bottleneck: mitochondrial imperatives in oogenesis and ovarian follicular fate. Mol Cell Endocrinol 145:81–88.

3 Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. Nature 325:31–36.

4 Sigurgardottir S, Helgason A, Gulcher JR, Stefansson K, Donnelly P. 2000. The mutation rate in the human mtDNA control region. Am J Hum Genet 66:1599–1609.

5 Cree LM, Samuels DC, de Sousa Lopes SC, Rajasimha HK, Wonnapinij P, Mann JR, Dahl HH, Chinnery PF. 2008. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. Nat Genet 40:249–254.

6 Khrapko K. 2008. Two ways to make an mtDNA bottleneck. Nat Genet 40:134–135.

7 Cao L, Shitara H, Horii T, Nagao Y, Imai H, Abe K, Hara T, Hayashi J, Yonekawa H. 2007. The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. Nat Genet 39:386–390.

8 Vermulst M, Wanagat J, Kujoth GC, Bielas JH, Rabinovitch PS, Prolla TA, Loeb LA. 2008. DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. Nat Genet 40:392–394.

9 Lewontin RC. 1974. The genetic basis of evolutionary change. New York: Columbia University Press.

10 Xing Y, Lee C. 2006. Alternative splicing and RNA selection pressure: evolutionary consequences for eukaryotic genomes. Nat Rev Genet 7:499–509.

11 Parmley JL, Hurst LD. 2007. How do synonymous mutations affect fitness? Bioessays 29:515–519.

12 Hardison RC, Roskin KM, Yang S, Diekhans M, Kent WJ, Weber R, Elnitski L, Li J, O'Connor M, Kolbe D, Schwartz S, Furey TS, Whelan S, Goldman N, Smit A, Miller W, Chiaromonte F, Haussler D. 2003. Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. Genome Res 13:13–26.

13 Tomilin NV. 2008. Regulation of mammalian gene expression by retroelements and non-coding tandem repeats. Bioessays 30:338–348.