Feature Article: **JAF4351**

ADAM AND EVE REDUX

by Ann Gauger

This article first appeared in CHRISTIAN RESEARCH JOURNAL, volume 35, number 01 (2012). For further information or to subscribe to the CHRISTIAN RESEARCH JOURNAL go to: http://www.equip.org/christian-research-journal/

This last year a flurry of new articles and blogs have been published that challenge the idea that we are all descended from just two first parents, traditionally known as Adam and Eve. Evangelical scientists have argued that human genome analysis has made a historical Adam impossible. This apparent abandonment of traditional teaching has made headlines, with reports appearing on NPR[1] and in *Christianity Today*,[2] to the distress of many. The confusion has provided fodder for atheists, who now declare that this controversy reveals the complete victory of science over religion.[3]

These blogs and articles all assume that the scientific argument against two historical first parents is solid. There are many problems with this story, starting with the assumption that our common descent from ancient ape-like ancestors is true. But when making a case, it is always good to begin by challenging the strongest of your opponent's arguments *on his or her own terms*. So in this article, I will challenge the assumption that current DNA polymorphism (differences in DNA sequence between individuals) provides enough evidence to *prove* Adam and Eve never existed, *by strictly critiquing their own arguments*. Specifically, I will address the claims that allelic diversity in modern humans is too high to have passed through a bottleneck of two, and that linkage disequilibrium also indicates ancient population sizes of a thousand or more.

POPULATION GENETICS—METHODS AND ASSUMPTIONS

Genetic diversity is the basis for all calculations of past genetic history. This diversity can be single base differences, called *single nucleotide polymorphisms (SNPs)*, or differences in sequence elements like mobile genetic elements, or variation in specific genes.

Models and equations that describe the distribution of genetic variants (alleles) in populations must make a number of assumptions to simplify the mathematics. Some of these assumptions are actually presuppositions. For example, the use of tree-drawing algorithms begs the question of whether or not a tree exists. This is an assumption drawn from Darwinism. Similarly, the assumption that random processes are the *only* causes of genetic change over time is an assumption drawn from philosophical naturalism. These assumptions may or may not be justified.

In addition, these equations describe an *ideal* population, one with random breeding and all other variables held constant. But because populations are not ideal, Sewall Wright defined a term called the *effective population size, N_e*, to indicate the size of an ideal breeding population whose distribution of allele frequencies and mating behavior is similar to an actual population of size *N*. N_e is substituted in the equations in place of *N to make the equations work*. Estimating N_e is a tricky business, because it requires significant information about a population's breeding behavior, and such information is impossible to obtain directly for ancient populations. Nonetheless, values for N_e are commonly used to challenge the existence of Adam and Eve.

TWO METHODS TO ESTIMATE PAST EVENTS

Let's consider a familiar example of human diversity. Many humans lack the ability to digest milk as adults. This is known as lactose intolerance, and is wide-spread in non-European populations. It appears to have been the ancestral state for all humans. Then at some point in our history, a mutation occurred in the lactase gene of some individual, probably in Scandinavia, so that the gene was active in adulthood as well as infancy. The appearance of this lactase persistence allele, as it is called, gave that individual the ability to digest milk as an adult. This would have been a beneficial trait in a community that keeps dairy animals.

How long did it take that new beneficial allele to spread in the population? To estimate that from first principles, we would need to know the population size where the mutation began, the mutation rate, the average generation time and reproductive success per individual, the amount of inbreeding in the population, and how much advantage the new allele gave its bearer in terms of ability to reproduce, plus the frequency of migrations in or out of the population.

Using Allelic Diversity

Much of that data is obviously unavailable to us here in the present. So geneticists have developed ways of looking at sequence data to make these estimates. They use SNP data to draw evolutionary trees for the DNA sequences in and around the gene in question, and from those trees (and assumed mutation rates and generation times), they can estimate at what point in time the trees *coalesce to a single allele*. Once that has been determined, they can also estimate what Ne must have been when the allele first arose, to explain the observed allele frequency of the lactase persistence gene today.

This estimate for N_e depends on an awful lot of assumptions that may not be justified, and on estimates that may not be accurate. For example, a common assumption is that there has been no variation in the rate of mutation among DNA sequences over time. This assumption has been called into question by recent studies. Another assumption is that the DNA being analyzed is either not under selection (has no effect on reproductive success), or that there has been no *change* in selection over time. Yet in the case of the lactase persistence allele, the strength of selection likely varied over time, depending on what was available to eat. Other assumptions, such as

constant population size, or no migration in or out of the population, are also questionable.

Using Linkage Disequilibrium

Linkage disequilibrium can also be used to calculate when the lactase persistence allele first appeared. Genes under selection tend to travel together with other closely linked genes. The linkage between the gene under selection and its hitchhiking neighbors disappears over time, depending on the amount of recombination between them. This rate of *linkage disequilibrium*, together with other information, can be used to estimate the age of the allele, *subject to the same qualifications as above*.

Estimates using linkage disequilibrium are finer scale, and can detect more recent population changes. Nonetheless, these estimates also depend on assumptions about rates of recombination, mutation, population size, selection, and so on, and should be viewed with caution.

Comparing the Results

The analysis based on linkage disequilibrium puts the first appearance of the lactase persistence allele range anywhere from two thousand to twenty thousand years ago,[4] while analysis based on associated sequence polymorphism places it at seven thousand to twelve thousand years.[5] This is a fairly large range of values for an event perhaps no more than twenty thousand years ago.

DEEP TIME AND ADAM AND EVE

What does all this have to do to estimates of population size at the time of the emergence of true humans? If it is hard to sort things out for events that occurred less than twenty thousand years ago, things are even hazier for populations in deep time. There are global problems associated with using measures of allelic diversity and/or linkage disequilibrium for calculating effective population sizes. First, any region of DNA chosen for analysis may be under selection of one sort or another. For example, regions of non-coding DNA previously assumed to have no function are increasingly being found to have important regulatory functions. So we don't know which regions of DNA are under selection and which are not.

Second, the rate of selection has increased in the last ten thousand years of human microevolution, presumably due to changing environmental factors.[6] This can throw off calculations that assume a constant rate of selection.

Third, models that assume a certain rate of loss due to "genetic drift" (random accidents that eliminate individuals before reproduction) may overestimate the effective population size needed to preserve polymorphism after a bottleneck has occurred. During rapid population expansion after a bottleneck, rare variants can be preserved more than usual in what's called "allele surfing."

Fourth, most models assume a constant population size with no migration into the population. Yet some population geneticists are arguing that modern humans have

undergone episodes of rapid migratory expansion, and possible cross-breeding with ancient humans. This is an active area of dispute.[7]

As a result of these problems, different models give different results. A study published in 2011 used multiple genetic loci to estimate an ancestral effective population size of about fourteen thousand (but ranging from two thousand to twenty thousand, depending on the model used).[8] In contrast, linkage disequilibrium estimates have reported effective population sizes for humans of between one thousand and two thousand, with bottlenecks as small as a hundred.[9] So the question arises, with what seriousness should we regard these estimates?

AN EXAMPLE OF WHAT CAN HAPPEN

In 1987 A. C. Wilson and colleagues published a paper claiming that all women could trace their ancestry to a single African woman one hundred sixty thousand years ago, based on comparisons of mitochondrial DNA.[10] (Mitochondrial DNA is passed from generation to generation by women, mother to child.) They dubbed her mitochondrial Eve. Not long after, Y chromosome sequencing data was used to construct a patrilineal lineage that appeared to lead back to a single African male.

Controversy quickly sprang up surrounding these results, as many rushed to explain that this didn't really mean that we are all descended from a single pair, Adam and Eve. In 1995 an eminent population biologist named Francisco Ayala sought to challenge their existence directly.[11] He examined alleles of a gene called HLA-DRB1, for which many variants exist in both primate and human populations. This gene encodes a blood cell protein involved in immune defense. He calculated that there were thirty-two separate versions of HLA-DRB1 present in the ancestral population that gave rise to chimps and humans, and that the size of that ancestral population was no fewer than a hundred thousand.[12]

It turns out that the particular DNA sequence from HLA- DRB1 that Ayala used in his analysis was guaranteed to give an overestimate of population size. HLA-DRB1 is one of the most highly variable genes in the human genome. It is under strong selection for heterozygosity, meaning having two different versions of the gene gives you a better chance of dealing with parasites and disease. Not only that, the particular region of the gene Ayala studied is a hotspot for gene conversion (a kind of mutation particularly likely to confound assumptions in tree-drawing).

A later, more careful study[13] examined the same HLA-DRB1 gene, but used a portion of the gene where mutations were more likely to occur at a baseline rate. This study concluded that only seven versions of the gene existed in the ancestral population at the time of the divergence between the hominid lineage and great apes, around five million years ago. About thirteen versions existed by the time true humans came along. And now, there are more than a thousand variants of the HLA-DRB1 gene.[14] Obviously this gene is under strong selection for increasing diversity, so is a poor choice for making estimates about ancient genetic events.

But the story gets better. Additional research indicates that there may be *as few as three or four ancestral versions* of the region surrounding the HLA-DRB1 gene, when all

sequence comparisons are taken into account.[15] This is tremendously exciting, since Ayala's claim of thirty-two ancestral alleles has frequently been used to argue that Adam and Eve cannot have existed. But with additional study, those claims have been shown to be false, and the genetic argument now no longer contradicts the possibility of two first parents. Three or four alleles can be carried by just two individuals, especially if after those two first parents came a period of rapid population expansion.

THE LIMITS OF SCIENCE

Truth be told, we have no idea when the first true humans appeared. The fossil records don't record things like the emergence of intellect and will, so inferences about fossils and dates are indirect at best. And the further back in time we go, the more confounding factors can intervene. As we have seen, population bottlenecks, changing selection, nonrandom breeding, mutational hotspots, changing mutation rate, or migratory behavior—all these can affect population genetics calculations. Parallel mutations or back-mutations can obscure evolutionary trees. Because of complications like these, it is an open question whether present genetic diversity is sufficient information from which to draw solid conclusions about ancient populations. A recent paper has acknowledged that significant migration in or out of a population can make it *impossible* to estimate effective population sizes in the distant past, based solely on current genetic diversity.16 Another researcher has admitted that drawing conclusions about events in deep time, before true humans arose, may be impossible.[17]

My point here is that there are limits to what can be proven about past events, based on population genetics models and current human genetic diversity. A little caution and humility is called for. Given the potential number of unknown variables, and the limitations of the models involved, no one should be making claims of certitude about what happened in humanity's distant past.

Ann Gauger, Ph.D. is senior research scientist at the Biologic Institute. Her current work challenges the credibility of Darwinian explanations of life, ranging from protein evolution (www.bio-complexity.org), to butterflies (Illustra Media's movie Metamorphosis), to human origins.

NOTES

1 http://www.npr.org/2011/08/09/138957812/evangelicals-question-the-existence-of-adam-and-eve.
2 http://www.christianitytoday.com/ct/2011/june/historicaladam.html.
3 http://whyevolutionistrue.wordpress.com/2011/09/08/adam-and-eve-theologians-squirm-and-sputter/; http://scienceblogs.com/evolutionblog/2011/09/what_does_original_sin_mean_in.php.
4 T Bersaglieri et al., "Genetic Signatures of Strong Recent Positive Selection at the Lactase Gene," *American Journal of Human Genetics* (2004) 74:1111–20.
5 M. Coelho et al., *American Journal of Human Genetics* 117 (2005): 329–339.
6 J. Hawks et al., www.pnas.org_cgi_doi_10.1073_pnas.0707650104 (2007).
7 http://johnhawks.net/weblog.

8 MGB Blum, M Jakobsson, "Deep Divergences of Human Gene Trees and Models of Human Origins," *Molecular Biology and Evolution* (2011) 28:889–98.
9 Albert Tenesa et al., "Recent Human Effective Population Size Estimated from Linkage Disequilibrium," *Genome Research* (2007) 17:520–26.
10 Rebecca Cann et al., "Mitochondrial DNA and Human Evolution," *Nature* 325 (1987): 31–36.
11 Francisco Ayala, "The Myth of Eve: Molecular Biology and Human Origins," *Science* 270 (1995): 1930–36.
12 Ibid.
13 Thomas Bergström et al., "Recent Origin of *HLA-DRB1* Alleles and Implications for Human Evolution," *Nature Genetics* 18 (1998): 237–42.
14 http://www.ebi.ac.uk/imgt/hla/stats.html, accessed October 19, 2011.
15 Gaby Doxiadis et al., "Reshuffling of Ancient Peptide Binding Motifs between HLA-DRB multigene Family Members: Old Wine Served in New Skins," Molecular Immunology (2008) 45:2743–2751.
16 Paula Sjödin et al., "On the Meaning and Existence of an Effective Population Size," *Genetics* (2005) 169:1061–70.
17 http://johnhawks.net/weblog.