

Re: MRCA question

Source: <http://sci.tech-archive.net/Archive/sci.bio.evolution/2005-02/0803.html>

From: Joe Felsenstein (joe_at_removethispart.gs.washington.edu)

Date: 02/26/05

Date: Fri, 25 Feb 2005 23:57:07 -0500 (EST)

In article <cvmjrl\$1jam\$1@darwin.ediacara.org>, Perplexed in Peoria <jimmenegay@sbcglobal.net> wrote:

>
> "Joe Felsenstein" <joe@removethispart.gs.washington.edu> wrote in
> message news:cvm8eu\$1f80\$1@darwin.ediacara.org...
>> *However the chance is very small that you both have any genes that came from
>> that ancestor. So it's a genealogical ancestor but probably not a genetic
>> ancestor.*
>>
>> *Going back 20 generations (500 years or so), you may have as many as a
>> million ancestors.*
>
> Assuming no inbreeding.

I did say "as many as", intending to cope with the finiteness of the population size.

>> *But your genome will have come from maybe about 300 of
>> them, and the others contributed no genes to you.*
>
> *This estimate (300) seems too low to me, particularly if there is little
> inbreeding. 1/300 of a genome works out to about 1/6 of a chromosome,
> and I would have thought that recombination would slice and dice things
> much more finely than this over 20 generations.*
>
> *I don't have online access to either the article or the News and Views.
> Is the estimate of 300 justified in one of them? If so, then I need to
> make a trip to the library. If not, could you sketch a justification for
> this number or provide a reference?*

I can sketch it. Consider a place in one copy of the genome. It comes from some ancestor at a remove of T generations (say). Another place far enough away from it to have recombination fraction r will come from the same ancestor, up the same lineage, if there is no recombination in any of those T generations. So the probability of being in the same chunk is $(1-r)^T$. If the recombination is uniformly spread with c recombination fraction per base, then $1-r = (1-c)^b$ when one is b bases away. Putting this together, one can see that the length of the chunk on this side of the original

sci.bio.evolution: Re: MRCA question

place then has exponential distribution with mean length $1/(cT)$ bases.
On the other side of the marker the length of the chunk is also that.

So if, as in humans, c approximately equal to $10^{(-8)}$, at a remove of 20 generations the average chunk length is about $2 \times 10^8 / 20 = 10^7$ bases. As we have 3.2×10^9 bases per haploid set and 23 pairs of chromosomes, a chromosome length averages 139 Megabases. This chunk length is 10 Megabases, so about 13.9 chunks per chromosome. I think you got 1/6 by taking a diploid individual, while I took a haploid. So you were closer as we would want to count how many chunks in a whole genome, not just a haploid genome.

Of course, the chunks are not of uniform size, and also they coalesce with each other, and recombination has "hot spots". Still, this shows the basis for this surprisingly small number.

--
Joe Felsenstein joe@removethispart.gs.washington.edu
Department of Genome Sciences and Department of Biology,
University of Washington, Box 357730, Seattle, WA 98195-7730 USA