Large Numbers and Evolution

The Problem: Human intuition about numbers often fails when the numbers are extremely large. Consider the following statement: "Evolution tries every possible DNA molecule." It is not clear whether this is likely or unlikely because of the size of the numbers involved. How many DNA molecules will ever exist? How does this compare with the total number of possible DNA molecules? Suppose we restrict ourselves to the "viable DNA molecules", is it likely then?

My intuition[1] tells me that I should respect the power of exponentially growing numbers and say no, it is not likely that this happens. Consider numbers like 100 and 4^{100}, the latter is vastly larger than the former.

Assumptions: Let us be conservative[2] and consider only those possible molecules of a certain fixed length. Suppose the following: 1. The population of Escherichia coli is sufficient to cover the entire surface of the earth; 2.There is an entirely new generation every 20 minutes from the beginning of life on earth until the sun dies; 3. There are no duplicate trials, i.e. each E. coli organism counts toward trying all possibilities; 4. The E. coli genome is four million nucleotides long[3].

Population: E. coli is cylindrically shaped with length 2×10^{-6}m and diameter 10^{-6}m (approximate dimensions[4]). Thus the cross sectional area is 2×10^{-12} m^2. The surface area of the earth[5] (including the 70% that is water) is 5.096×10^8 $(km)^2$. Thus the population of E. coli that would cover the earth's surface is $[5.096 \times 10^8 \ (km)^2] / [2 \times 10^{-12} \ m^2] \approx 2.5\ 10^{26}$ individuals.

Number of generations: There is evidence of life on earth 3.5 billion years ago[6]. Being conservative, let us say that 4.5 billion years ago (when the earth was born)[7] the earth was already covered with E. coli. Life on earth is expected to end when the sun becomes a red giant in approximately five billion years[8]. Staying with our conservative approach let us round every thing up to 10 billion years. So the total number of generations is 10^{10} years / 20 minutes per generation $\approx 2.6 \times 10^{14}$ generations.

Total number of evolutionary trials: population times number of generations = 2.5×10^{26} individuals per generation times 2.6×10^{14} generations = 6.5×10^{40} trials. This is a wild over estimate of the number of trials likely.

---

1 At the time I am writing this paragraph I haven't done the estimates but am expressing only my intuition.
2 By "conservative" I mean to wildly over estimate the number of molecules tried and vastly underestimate the total number of possible molecules.
3 The genome is actually 4.64 million bases long: http://www.hgsc.bcm.tmc.edu/projects/microbial
4 http://www.nap.edu/html/ssb_html/NANO/nanopanel1moore.shtml
5 http://www.britannica.com/ebi/article-199816?tocId=199816
6 http://www.brembs.net/gould.html
7 http://www.brembs.net/gould.html
8 http://curious.astro.cornell.edu/question.php?number=48

Number of possibilities: Given that the genome is 4×10^6 bases long and each base can be any of Adenine, Guanine, Thymine, or Cytosine that means there are

$4^{4\times10^6}$ possibilities.

Fraction of possibilities that are tried: this would be the number of trials divided by the number of possibilities =

$$6.5\times10^{40} \big/ 4^{4\times10^6} = 10^{\log 6.5+40-4\times10^6 \log 4} \approx 10^{-2.4\times10^6}$$

To get an idea of what this means, suppose the possibility space is likened to the number of grains of sand on the earth[9] $\approx 10^{23}$. Then the number of grains that are tried would be:

$$10^{23} \times 10^{-2.4\times10^6} = 10^{-2.4\times10^6}$$

That's right—not even one grain is tried!

Even if there turned out to be as many E. coli organisms as there are electrons, protons, and neutrons in the universe[10], you still would not have tried one grain of sand from the possibility space.

P.S. If you find numerical or typographical errors in this article please contact me: Patrick@mr-ideahamster.com.

---

[9] http://www.madsci.org/posts/archives/jun99/927857197.Es.r.html

[10] $\sim 10^{130}$ http://en.wikipedia.org/wiki/Talk:Atom