DATA:

A DNA profile database with 10^6 entries.

Each entry is a 13-loci profile.

The probability of a random match at a single locus is 1/10.

CALCULATION:

The RMP for a 9 locus match is $1/10^9$.

There are $13!/[9! \times 4!] = [13 \times 12 \times 11 \times 10]/[4 \times 3 \times 2 \times 1] = 715$ possible ways to choose 9 loci from 13.

Hence the RMP for finding a match on ***any*** 9 of the 13 loci is $715/10^9$.

If you pick any profile in the database, the probability of a second profile not matching on 9 loci is roughly $1 – 715/10^9$.

Hence, the probability of all 10^6 database entries not matching on 9 loci is roughly

$$(1 – 715/10^9)^{1,000,000}.$$

Using the binomial theorem, and working for an eventual accuracy of 2 decimal places, this is approximately

$1 – 10^6 \times 715/10^9 + 10^{12}/2 \times 715^2/10^{18} – 10^{18}/6 \times 715^3/10^{27} + 10^{24}/24 \times 715^4/10^{36} – \ldots$

$= 1 – 715/10^3 + 715^2/2\text{x}10^6 – 715^3/6\text{x}10^9 + 715^4/24\text{x}10^{12} – \ldots$

$= 1 – 0.715 + 0.256 – 0.061 + 0.011 – \ldots$

$= 0.49$

To be safe, let’s check the next term in the expansion:

$+ 10^{30}/120 \times 715^5/10^{45} = + 715^5/120\text{x}10^{15} = + 0.0015.$

This does not alter the first two places. So, to 2 decimal places, the probability of there being a 9-locus match is the difference between 1 and 0.49, namely 0.51.

That’s a better than 50% chance of finding a random match. Oh my!