

Probabilities in Forensic Genetics

Background: Given the genotype of the parents, the chance that a certain allelic combination will occur in the offspring can be predicted using the rules of probability.

The properties of probability can be demonstrated by flipping coins or rolling dice. For example, when flipping a coin, the probability that coin will land on any one side out of a possible two sides is one out of two, or 0.5. When rolling a standard gaming dice, the probability that the dice will land on any one side out of a possible six sides is one out of six, or 1/6.

To calculate the probability that two dice throws will match, multiply the two probabilities together, that is $1/6 \times 1/6 = 1/36$. In other words, there are 36 different combinations of numbers that can be obtained with two six-sided dice (for example, 1,1; or 1,6, etc.). This prediction is based on the **rule of multiplication**, which can determine the chance that two or more independent events will occur in some specific combination. Since there are two parents, there are two ways in which alleles can combine to produce a mixed or heterozygous result. To illustrate, when two coins are tossed, they have the chance of landing on heads/heads, heads/tails, tails/heads, or tails/tails. Each of these events is equally likely to happen, so each represents one of all the possible outcomes. The probability of flipping heads with tails or tails with heads is $1/4 + 1/4 = 1/2$. Thus, the probability of an event occurring in two or more ways is the sum of the separate probabilities for each, and is known as the **rule of addition**. In forensics, if our only information is a blood spot, 1,2 and 2,1 are indistinguishable.

One way to think about this problem is to consider a matrix to make all possible combinations of alleles taken two at a time. In a one gene, two allele system, with the alleles designated as 1 and 2, the following matrix can be generated:

	1	2
1	1,1	1,2
2	--	2,2

The total number of combinations is 2^2 , which is 4, but since two of these are repeats, the number of combinations is reduced to 3. If we consider a locus with 5 alleles, then the following matrix applies:

	1	2	3	4	5
1	1,1	1,2	1,3	1,4	1,5
2	--	2,2	2,3	2,4	2,5
3	--	--	3,3	3,4	3,5
4	--	--	--	4,4	4,5
5	--	--	--	--	5,5

Although 25 (5^2) possibilities exist, there are only 15 that are unique. If n is the number of alleles for a single locus, then the number of possible combinations of pairs of alleles is n^2 . However, if order is ignored, the number of possible combinations of n alleles taken two at a time is:

$$\frac{n(n+1)}{2}$$

Why is this important? The number of combinations provides an estimate of the probability of matching a genotype by chance alone. For example, in the 2-allele system, with 3 possibilities, there is a 1 in 3 chance of matching by chance alone. In a court of law, this is not even close to being beyond the shadow of a doubt. In other words, only 2/3 of the population can be eliminated. The population of the world is over 6 billion, this leaves 2 billion suspects!

As mentioned above, with a loci that has 5 alleles, the number of possibilities is 15, which means 1 in 15 chance of matching by chance, eliminating 14/15 of the population. Of 6 billion people, this still leaves 400 million suspects.

The rule of multiplication applies here also, if both the 2 allele and 5 allele genes are considered at the same time. The probability of a match by chance is $1/3 \times 1/15 = 1/45$. One-forty-fifth of six billion is 130 million suspects. That is less, but still not beyond the shadow of a doubt. As the number of alleles increase and loci increase, then the number of suspects can be reduced.

Questions:

1. The following table shows the number of alleles that each loci in the AmpFLSTR® kit can detect. Using the formula: $n(n+1)/2$, calculate the number of combinations that are possible with the number of alleles and the probability of each combination. Fill your results in on the table:

Loci	# Alleles (n)	# Combinations	Probability of each combination
D3S1358	8		
vWA	11		
FGA			
Amelogenin*	2	2	0.5 or 1/2
D8S117	12		
D21S11	22		
D18S51	21		
D5S818	10		
D13S317	8		
D7S820	10		

2. *Amelogenin is a special case since it detects the X and Y chromosome. Although one would expect 3 combinations, there are only 2. Why?

3. If a match is made between a suspect and a blood stain on the wall, what are the odds of this event happening by chance alone? (Hint: This uses the multiplicative law) (Note: This assume that each allele is equally frequent in the population)