

Report #4: Probability in biology - the basics of inheritance

Back in the 19th century Mendel disproved the then-popular (and wrong) blending theory of inheritance, which stated that parental traits were mixed, or “blended”, into the offspring. Mendel’s experiments were conducted on peas and relied on diligently noting the traits of the different generations and recognising the inheritance patterns. The data he gathered over years of breeding plants allowed him to establish laws through which the probability of inheritance of a certain trait could be calculated. By the end of his experiments, he could reliably predict the percentage of the offspring that would have a certain trait if the parental traits were known. His work is a good example of using empirical probability to find a certain pattern.

Without getting too much into the details of plant physiology, the traits Mendel tracked were all easy to observe with the naked eye, which made them easy to track. Furthermore, for each characteristic, each plant exhibited only two different traits, also known as phenotypes in biology. The following table presents Mendel’s data:

Characteristic	Contrasting P ₀ Traits	F ₁ Offspring Traits	F ₂ Offspring Traits	F ₂ Trait Ratios
Flower color	Violet vs. white	100 percent violet	<ul style="list-style-type: none"> • 705 violet • 224 white 	3.15:1
Flower position	Axial vs. terminal	100 percent axial	<ul style="list-style-type: none"> • 651 axial • 207 terminal 	3.14:1
Plant height	Tall vs. dwarf	100 percent tall	<ul style="list-style-type: none"> • 787 tall • 277 dwarf 	2.84:1
Seed texture	Round vs. wrinkled	100 percent round	<ul style="list-style-type: none"> • 5,474 round • 1,850 wrinkled 	2.96:1
Seed color	Yellow vs. green	100 percent yellow	<ul style="list-style-type: none"> • 6,022 yellow • 2,001 green 	3.01:1
Pea pod texture	Inflated vs. constricted	100 percent inflated	<ul style="list-style-type: none"> • 882 inflated • 299 constricted 	2.95:1
Pea pod color	Green vs. yellow	100 percent green	<ul style="list-style-type: none"> • 428 green • 152 yellow 	2.82:1

(Mendel, 1865. Table adapted from Openstax Textbooks)

P0 stands for the parental generation, while F1 and F2 are the first and second offspring generations, respectively. P0 were bred through hybridisation, meaning that a true-breeding individual with one phenotype was mated with the true-breeding individual with the other phenotype. Self-fertilization was not allowed for the P0 generation. As it can be seen on the table, despite having two parents of different phenotypes, the F1 generation only exhibited one of them. This phenomenon later became known as the Law of Dominance - the appearance of only one of the phenotypes despite the presence of the genes for both. Thus, the phenotypes that the F1 generation exhibited are the dominant phenotypes and the ones that did not show are the recessive ones.

When the F1 generation was left to self-fertilise, the resulting F2 generation exhibited a different pattern of inheritance. It can be seen that for all characteristics, the observed ratio of the two phenotypes was approximately 3:1 in favour of the dominant one. Similar results were obtained when the experiment was repeated.

The role of probability in this study came during the analysis of the data. Working backwards from the results and applying the rules of probability allowed Mendel to establish the remaining two Laws of Inheritance and to accurately predict the ratios of progenic phenotypes.

The first rule to consider is the product rule of probability. The rule states that the probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event:

$$P(A \cap B) = P(B | A) \cdot P(A)$$

It implies the Law of Segregation, according to which inherited traits depend on the parental genes, which each have two versions, known as alleles (usually, uppercase letters are used for dominant alleles, while lowercase ones are used for recessive alleles). The progeny randomly inherits one allele from each parent. The number of plants with the recessive phenotype was $\frac{1}{4}$ of the total, which in other words means that a fourth of the plants inherited the recessive allele from each parent, leaving them with two recessive alleles and thus the recessive phenotype. Those results are in accordance with the (correct) assumption that the allele from each parental cell has a 50% chance to be inherited independently of the other and thus

$$P(a \cap a) = (\frac{1}{2} * \frac{1}{2}) = \frac{1}{4}$$

The Law of Dominance carries the notion of heterozygosity - the inheritance of a dominant and a recessive allele, which results in the dominant phenotype (or more generally - just the presence of alleles that produce different phenotypes). When the progeny inherits two of the same-phenotype-carrying alleles, this is called homozygosity. Those two notions allow us to go beyond phenotypic inheritance and consider genotypic inheritance.

Punnett squares are often used when teaching Mendel's laws to visualise both phenotypic and genotypic inheritance patterns.

Dominant and Recessive
(T = Tall & t = short
Cross: Tt x Tt)

	T	t
T	TT	Tt
t	Tt	tt

Genotypic ratio: 1 : 2 : 1 (TT=25% Tt=50% tt=25%)
Phenotypic ratio: 3 : 1 (Tall=75% Short=25%)

Source: Greeleyschools

When the product rule is considered in terms of multiple genes, (PAa × PBb) gives the probability of the dominant phenotypes and heterozygous genotypes of characteristics A and B occurring together. Since the data matched the predicted results if the alleles for each gene are inherited independently of one another, the Law of Segregation was established, which states that the inheritance of one trait is independent of the inheritance of other traits. In other words, whether the plant is tall or short will not affect whether it's green or yellow. Using the product rule for multiple genes allows us to calculate the frequency of multiple phenotypes occurring together. Punnett squares for multiple genes can also be constructed.

Source:
Greeleyschools

Dihybrid Cross
Dominant and Recessive
T=Tall, t=short
B=Black, b=white
Cross: TtBb x TtBb

	TB	Tb	tB	tb
TB	TTBB	TTBb	TtBB	TtBb
Tb	TTBb	TTbb	TtBb	Ttbb
tB	TtBB	TtBb	ttBB	ttBb
tb	TtBb	Ttbb	ttBb	ttbb

Genotypic ratio: 1:2:2:1:4:1:2:2:1
Phenotypic ratio: 9:3:3:1

However, when working with more than two genes, punnett squares become tedious to create. Using the product rule for calculations is a much more efficient way to find the genotypic and phenotypic distribution.

The sum rule, on the other hand, states that the probability of at least one of two independent events occurring is the sum of their individual probabilities:

$$P(A \cup B) = P(A) + P(B)$$

The sum rule is useful when it comes to calculating the chance of one phenotype occurring within the progeny of parents with known genotypes or phenotypes.

Finally, it should be noted that in order to be able to apply the product and sum rules, one needs to be working with a large sample size in order to avoid deviation in results due to variance. This general principle becomes even more important when we consider inheritance patterns that go beyond the scope of Mendelian genetics. As it turns out, there are many factors that determine how inheritance works, especially when a single phenotype depends on multiple genes. Still, probability can be applied in those cases, some of which include predictions of disease inheritance or plant traits that are important for agriculture, to give a general idea of the expected phenotypes. Those calculations, however, are usually based on a combination of theoretical postulates like Mendel's Laws and large-scale studies that have accumulated enough data to give relatively accurate predictions.

Sources:

<https://opentextbc.ca/biology2openstax/chapter/mendels-experiments-and-the-laws-of-probability/>

<https://www.greeleyschools.org/cms/lib2/CO01001723/Centricity/Domain/5219/punnett%20sq%20cheat%20sheet.pdf>