Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

Edited by James F. Crow and William F. Dove

On Fisher's Criticism of Mendel's Results With the Garden Pea

E. Novitski[1]

Department of Biology, University of Oregon, Eugene, Oregon 97401

IT is generally agreed that the overall results from experiments with the garden pea reported by Gregor Mendel (1866) conform more closely with the ratios theoretically expected (such as 3:1, 1:2:1, etc.) than one might reasonably expect to obtain on a chance basis. Particularly troubling are those two groups of experiments in which Mendel's results are in close agreement with ratios that Mendel may have considered appropriate, but which were, according to Fisher (1936), incorrect. Wright (1966, p. 174) states that these two represent "the most serious evidence for fraud by Mendel, presented by Fisher."

Here I consider an alternative way of examining Mendel's procedures and results in those experiments (9 of 24 total) and suggest that Fisher, in his detailed and illuminating analysis of Mendel's results, may have erred in assigning specific expectations to those runs and that Mendel's expectations in fact may have been closer to the mark than Fisher's.

The seven pairs of characters described by Mendel can be assigned to two distinct categories. In the first group are five plant characters (length of stem, position of flowers, color of flowers, color of pods, and form of pods) and in the second group two seed characters (round *vs.* wrinkled and yellow *vs.* green). The seed characters reveal the properties of the next generation on the parental plant, making it unnecessary to grow that subsequent generation as individual plants. Since there are ~30 seeds per parental plant, each of those plants provides ~30 independent observations on the next (unplanted) generation. Capitalizing on the unusual properties of this second group of seed characters made it possible for Mendel to reach large numbers in his experiments on a relatively small plot of land. Further, since the results from runs with seeds are available sooner than those from plant characters, the seed experiments, involving larger numbers, would serve as the basis for the ratios that Mendel reports, as well as the rationale for those ratios, with the later, more limited results from runs with plant characters serving to determine whether those characters behave in their transmission like the seed characters. It is not surprising that more than three-fourths of Mendel's observations were made on the two seed characters and fewer than one-fourth were made on all five plant characters combined.

The experiments under consideration here involve the determination of the frequency of the genotypes in the phenotypically dominant class, *i.e.*, the ratio of the *AA* to *Aa* genotypes in the expected ratio of 1:2, as in the classic 1*AA*:2*Aa*:1*aa* distribution in the F_2 of a simple monohybrid cross. They will be considered in two parts: the 2:1 ratio and the trifactorial cross.

The 2:1 ratio: In crosses involving the seed characters of round *vs.* wrinkled and yellow *vs.* green, 1084 plants showed the dominant trait in the first generation (and so could be either *AA* or *Aa*). When these seeds were planted and the resulting plants allowed to self-fertilize, 359 produced the dominant character only, indicating that the parent was homozygous *AA*, and 725 plants produced both the dominant and recessive characters on each plant, indicating heterozygosity (*Aa*) for an overall ratio of 2.02:1.

Mendel tested the five different pairs of plant characters. For these, it was necessary to rely on progeny tests, and he specified the exact numbers to be reached in each case. He stated, "For each separate trial in the following experiments, 100 plants were selected which displayed the dominant character in the first generation, and in order to ascertain the significance of this, ten seeds of each were cultivated" (Mendel 1866, p. 12).

The set of five experiments (six, counting one repeat) with plant characters was unusual because, to make the distinction between *AA* and *Aa* plants, it was necessary to grow a number of progeny after selfing to see if a homozygous recessive plant appeared. This, however, would happen in only one-fourth of the offspring of a heterozygote when selfed, so that the chance that any

[1]*Address for correspondence:* 1690 E. 26th Ave., Eugene, OR 97403. E-mail: novitski@uoregon.edu

one of the progeny would not show the recessive is 3/4. A number of progeny would, therefore, have to be raised with the expectation that at least one of them would reveal the presence of the recessive when the parent was heterozygous. Apparently Mendel considered plants from 10 seeds adequate for the determination, because he noted, "ten seeds of each were cultivated."

Fisher pointed out that if the chance of not obtaining a homozygous recessive progeny is 3/4, then the chance that one will not appear in 10 progeny is simply 3/4 to the tenth power (or 0.0563). These heterozygotes, amounting to $0.0563 \times 2/3$ of the total, would then be misclassified as homozygous *AA* and would change the expectation from 2 to 1 to that of 1.8874 to 1.1126 (or ~1.7 to 1). It is the closeness of the ratio that Mendel reported as close to 2:1, rather than the 1.7:1 that Fisher calculated, that led Fisher to suspect that the actual results had been altered.

Surely Mendel would have known, at the outset, that there would inevitably be some (as well as the approximate proportion of) failures in his plantings, and if he wished to obtain 100 fruiting plants in every one of his six experiments, he would have to plant somewhat more than 100 seeds in each of those experiments. From each of the 100 (or more) fruiting plants in each experiment, 10 seeds were "cultivated." But there must have been some cases where, by failure to thrive or other mishap, only nine (or fewer) progeny survived in the test, and they would not, by Mendel's apparent criterion, be counted. However, it would be difficult to discount any case in which one or more homozygous recessives appeared, unambiguously identifying the parent as a heterozygote, regardless of the number in the test. In fact, most investigators, however conscientious, would, upon the appearance of a single homozygous recessive plant, immediately classify the parent as a heterozygote and not even bother to make further observations on that set or to count the total to see if they reached the specified number of 10. In short, they would not eliminate any test clearly identifying the parent as a heterozygote simply because the total in the test did not reach the number of 10.

If Mendel demanded 10 progeny for an adequate test only when the 9 or fewer existing progeny failed to reveal the heterozygosity of the parent, we can reasonably assume that when fewer than 10 progeny matured, he would be eliminating primarily *AA* individuals. [Consider, for instance, the case in which only 9 of the 10 plants survived. One-third of the time, the parent will have been *AA*, and all 9 progeny *AA*. Two-thirds of the time, the parent will have been *Aa*, and the probability of obtaining 9 *A*- progeny is $(3/4)^9$ or 0.075. Thus, the probability that the 9 *A*- progeny are from a homozygous parent is about seven times higher than the probability that they are from a heterozygous parent—a ratio of $(1/3)/[(2/3)\ (0.075)]$.] The selective elimination of *AA* individuals would shift the ratio calculated by Fisher from 1.8874 *Aa* to 1.1126 *AA* toward the ideal of 2 *Aa* to 1 *AA*. In fact, if the failure rate were high enough, the ratio might well exceed 2:1.

No information about the failure rate is available for these experiments, but Mendel did give it for a subsequent run, which involved three seed characters. Of 556 seeds, 11 did not yield plants; this is very close to 2% (0.0198). If he had had the same failure rate in his six experiments, each involving 100 sets of 10 seeds each, there would be a failure in 20 plants, affecting ~20 sets in each of the experiments (or 120 failures in all). It would have been quite reasonable for Mendel to accept parents as being heterozygous if any of the seedlings exhibited recessive characteristics, even if not all 10 plants had survived. Mendel would then be left, for example, with some sets of 9 plants all exhibiting the dominant trait, including all of those from really *AA* parents ($120 \times 1/3 = 40$ parents) and all of those really *Aa* but happening to have 9 successive *AA* offspring [$120 \times 2/3 \times (3/4)^9 = 6$ plants]. If Mendel then counted ~46 additional sets of 10 as replacement sets, then, applying Fisher's correction, there would be some identified as *AA* parents ($46 \times [(1/3) + 0.0375] = 17$ parents) and some as *Aa* parents ($46 \times [(2/3) - 0.0375] = 29$ parents). The 46 plants tossed out had 6 *Aa* plants, but the replacements had 29 *Aa* plants, a difference of 23. Thus, using additional plants to reach the arbitrarily selected value of 100 would cause Mendel to see 23 too many *Aa* plants, tending to counterbalance the 22 or 23 ($600 \times 0.0375 = 22.5$) too few *Aa* plants owing to Fisher's correction.

The similarity of the loss (22.5) and the gain (23) is purely fortuitous and not to be taken too literally. The essential point is that the two values are of similar magnitude and of opposite effect. They do show, however, how these two complicating factors, working in opposite directions, could give Mendel a final ratio closer to the original 2:1 theoretically expected, Fisher's analysis notwithstanding. The author is indebted to C. E. Novitski for the preceding calculations. However, a more precise mathematical formulation by him is found in the accompanying *Perspectives* article (Novitski 2004, this issue).

The trifactorial experiment: In the trifactorial experiment, Mendel used three pairs of factors. The first two pairs are the usual round-wrinkled (*A*, *a*) and yellow-green (*B*, *b*) seed characters. The third is, according to Fisher, the factor for colored-white flower color (*C*, *c*), which would necessitate progeny tests. Mendel obtained 687 seeds from the initial crosses, and of these 639 fruited. Every one of the 639 plants was entered into one or the other of his 27 classifications for the three conditions (two homozygotes and one heterozygote) for the three pairs of characters involved. Of the 639 plants, 473 were in the *CC*-*Cc* group.

Fisher used his 0.0563 correction to show that this experiment, insofar as the color of the flowers is concerned, also agreed with the ideal expectation of 2:1 more closely than Fisher's calculation of a corrected expectation. This is because Mendel's counts of the *CC* class are only 6 plants away from the ideal of 158 (out of 473), but are 23 plants away when Fisher's correction is applied. He suggested that here again Mendel obtained a reasonably good fit to the wrong expectation. It can be argued, however, that Mendel was almost certainly using the correct expectation, and it is Fisher who was using the incorrect one.

Mendel was apparently not using flower color, but rather seed coat color, which in his crosses had an identifiable characteristic in the hybrid (*i.e.*, heterozygote). Although from his description one gathers that the character in the individual pea is somewhat variable in the heterozygote, it is reasonable to assume that the maternal plant might usually be classified unambiguously as either a homozgote or a heterozygote, since all 30 or so peas on the plant would have the same maternal genotype, and even with considerable variability their overall similarity could give some assurance that the plants were classified correctly.

Mendel, in his description of the trifactorial cross, specifically stated that seed coat color was the third character involved. It seems that Mendel rather cleverly used three seed characters together so that the complete determination of the genotype of the parental plant could be made immediately upon inspection of the seeds borne by that plant. This would explain why the number 473 was both the number requiring classification and the number classified; that is, all plants that bore seeds were classifiable. In an addendum to *Perspectives* in this issue of Genetics, James Myers presents an ingenious and statistically acceptable proposal for another way in which Mendel could have used seed coat color for the third character in the trifactorial cross.

Discussion: We do not know, of course, which, if any, of the above considerations moderated Mendel's actions during the 2:1 experiment. We can speculate, however, that some of them may have. We shall consider two extremes: the simplest case and a more complicated sequence of events.

In the simplest hypothetical case, Mendel decided to set up enough tests in the 2:1 case for each of the six experiments to obtain 100 fruiting plants in each. He then classified the first 100 cases in which all 10 seeds produced classifiable plants, with the results reported in his article. He might have been completely unaware that in his six plant experiments there were two forces, acting in opposite directions, that would tend to give him a result close to the ideal 2:1 ratio that his [seed] experiments with seeds had indicated.

For the trifactorial experiment, it seems clear that Mendel was quite right in postulating a segregation of three seed characters, each classified in three different ways (two homozygous and one heterozygous genotypes), giving a total of 27 different categories. It seems likely that the chore of categorizing 473 plants, each in a specific one of the 27 categories, led him to describe this experiment in this way: "Among all the experiments it demanded the most time and trouble."

For the case for which we assume that Mendel was more knowledgeable, we might keep in mind that Mendel had spent 2 years at the University of Vienna studying scientific subjects, including mathematics and physics, and so we can assume that he was aware of the simple rules of probability. So, at this other extreme, let us imagine that Mendel became aware, as he made his counts on plant characters, that he had inadvertently blundered into a situation in which he could not, in theory, arrive at the result he had originally expected. He was certain, of course, that the 1:2 ratio of *AA* to *Aa* was generally true; the two fairly extensive experiments with seed characters counted earlier had already demonstrated that. But, as he counted, he must have realized that the number of 10 tested progeny for the six plant-character tests was inadequate. For one thing, in about one-fifth of all cases [the value of the second term of the expansion of $(3/4 + 1/4)^{10}$], he was classifying as a heterozygous parent instances when 1 and only 1 of the 10 progeny showed the homozygous recessive phenotype. While Mendel might not have known how many he was missing, he would have realized that he was simply approaching the theoretical value of 2:1 asymptotically. He might also have realized that the standard of 10 on which to base his determinations would predominantly eliminate homozygous dominants and that two disturbing factors were acting in opposite directions.

In a limited survey of relevant literature, several articles stand out. In his article on the rediscovery of Mendel's work, Correns (1900) took Mendel to task for overemphasizing the dominant-recessive relationship of the factors he used and pointed out that the phenotype of the heterozygote for the factor for seed coat color was distinct from either homozygote.

In a discussion of Fisher's analysis, Wright (1966) implicitly corrected Fisher by noting that Mendel used seed coat color (and not flower color) as the third character in the trifactorial test, and he noted that the character is identifiable in the heterozygote. But then he clearly accepted Fisher's conjecture that the determination of each plant was made on the basis of 10 progeny (totaling 4730 plants) because "the occurrence of segregation of *AA* and *Aa* would be obvious in a group of ten in the absence of recessives." This is a curious slip; clearly it would not be necessary to set up the "group of 10" progeny for each plant to distinguish between *AA* and *Aa* if the parental heterozygote *Aa* were distinguishable from the homozygote *AA* in the first place!

It seems ironic that the two foremost mathematical geneticists of the early part of the twentieth century, who had engaged in a vigorous debate on the nature and evolution of dominance, should both seem to err in their application of incomplete dominance in an actual experimental situation. But then, Sturtevant (1965), who surely was the most competent, genetically speaking, to examine Fisher's article, concludes that "Fisher's analysis of Mendel's data must stand essentially as he stated it," failing to note Fisher's likely error in suggesting that Mendel used flower color rather than seed coat color in the trifactorial experiment.

The analysis presented here does not alter Fisher's conclusions that, overall, Mendel's results are closer to theory than expected on a chance basis. I call attention to the work of A. W. F. Edwards (1986), who reviewed Mendel's data in their entirety, and that of C. E. Novitski (1995), who made a detailed analysis of some of Mendel's first experiments. The first of these studies provides convincing evidence of unusually close fits of data to theory. The second bypasses assumptions about statistical theory with computer simulations of some of Mendel's first crosses with seed characters and shows that simple repetition of experiments would probably not account for the closeness of his results to theory.

The defect in Mendel's overall data from the statistical point of view appears to be a deficiency of cases in which the results deviate markedly from expectation (a point on which both A. W. F. Edwards and C. E. Novitski agree). As Fisher has pointed out, great deviations occur more frequently than one might expect and therefore might be considered anomalous. Thus, Mendel might very well have repeated some runs in which the numbers did not appear to bear out his expectations, not with any intent to deceive, but to make certain for his own benefit that those runs were in fact not *bona fide* cases of exceptions to his rules. Having obtained additional data, he could have either replaced the earlier deviating data with the "better" numbers or combined the two sets of data, which would usually obscure the extent of the deviations in the first set. (As an illustration, if the repeat gave a similar deviation but of opposite sign, the addition of the two could give an almost perfect fit.) It might be observed that such procedures probably exist in preparing data for publication even today.

In addition to the possibilities mentioned above, the intriguing possibility that Mendel did actually select data, manipulate sets of results, or even alter a few of his numbers always remains. This would stem not from any desire to mislead, but as a concession to his ill-prepared audience. We can imagine that at the time of writing for oral presentation, Mendel changed, for didactic purposes, some specific results that might have distracted his audience from the main theme of the article because of their seemingly aberrant nature. Surely words like fraud or dishonesty should be used with caution. Perhaps his situation can be compared to that of the competent high school science teacher who, in explaining the structure of the atom to his students, falls back on the simple Bohr model, well aware that while it is not the correct picture, it is appropriate for the audience for which it is intended.

Finally, it should be kept in mind that Mendel not only anticipated but also would have welcomed repetitions of his experiments by others. He would not have benefited scientifically, financially, or ecclesiastically from any outright misrepresentation in his work.

In conclusion, Fisher's criticism of Mendel's data—that Mendel was obtaining data too close to false expectations in the two sets of experiments involving the determination of segregation ratios—is undoubtedly unfounded.

The author thanks James Myers for his invaluable contributions to my understanding of the interactions of the seed coat color genes; C. E. Novitski, who made many suggestions for changes in the original text, leading to a greater readability and precision of expression; and Raphael Falk, who has been a constant source of valuable criticism and encouragement throughout this work.

LITERATURE CITED

Correns, C., 1900 Mendel's Regel über das Verhalten der Nachkommenschaft der Rassenbastarde, reprinted in *The Origin of Genetics: A Mendel Source Book*, edited by C. Stern and E. R. Sherwood. W. H. Freeman, San Francisco, 1966 (English translation by L. K. Piternick).

Edwards, A. W. F., 1986 Are Mendel's results really too close? Biol. Rev. **61:** 295–312.

Fisher, R. A., 1936 Has Mendel's work been rediscovered? Ann. Sci. **1:** 115–137.

Mendel, G., 1866 Versuche über Pflanzen-Hybriden. Verh. Naturforsch. Ver. Brünn **4:** 3–47 (first English translation in 1901, J. R. Hortic. Soc. **26:** 1–32; reprinted in *Experiments in Plant Hybridization*. Harvard University Press, Cambridge, MA, 1967).

Novitski, C. E., 1995 A closer look at some of Mendel's results. J. Hered. **86:** 62–66.

Novitski, C. E., 2004 Revision of Fisher's analysis of Mendel's garden pea experiments. Genetics **166:** 1139–1140.

Sturtevant, A. H., 1965 *A History of Genetics*. Harper & Row, New York.

Wright, S., 1966 Mendel's ratios, pp. 173–175 in *The Origin of Genetics: A Mendel Source Book*, edited by C. Stern and E. R. Sherwood, W. H. Freeman, San Francisco.