Edward East on the Mendelian Basis of Quantitative Trait Variation

Michael Turelli
Department of Evolution and Ecology, University of California, Davis, California 95616
ORCID ID: 0000-0003-1188-9856 (M.T.)

Studies of Size Inheritance in *Nicotiana*
Edward M. East
*GENETICS* March 1, 1916 1: 164–176

East’s (1916) analysis of corolla length in *Nicotiana longiflora* was a capstone study integrating Mendelian genetics with a continuously varying phenotype. Gregor Mendel and Francis Galton produced profound insights into heredity by employing mathematics novel to biology (in Galton’s case, novel to science). Both used *Pisum sativum* to illustrate different patterns of inheritance. Mendel established his ratios by studying qualitative variants [what Darwin (1859) called “sports”]. Those discrete differences were determined by individual genes, and generally displayed dominance and independent assortment, features atypical of the natural variation that attracted students of hybridization, breeding, and evolution. Mendel’s fundamental discoveries were ignored for decades, at least in part because of their questionable relevance to common phenotypic differences. In contrast, Galton studied the distribution and inheritance of continuous variation using characters such as seed-pod length in *P. sativum*, and human height. Galton showed that phenotypic variation could often be approximated by Gaussian distributions. Comparing the average phenotypes of offspring with their parents (adjusting, as needed, for sexual dimorphism), Galton discovered “regression toward mediocrity,” namely that offspring are generally less “exceptional” than their parents (i.e., they tend to deviate less from their population means). Mendel struggled to convince himself, and his contemporaries, that he had uncovered general principles, failing miserably with his second, and final, publication (Mendel 1870) on interspecific *Hieracium* “hybrids” (in fact, produced asexually, Olby 1966, Chap. 4). In contrast, Galton (1889) observed normal distributions and “regression towards mediocrity” everywhere. (Unfortunately, Galton’s misunderstanding of the implications of “regression” led to decades of muddled controversy concerning the efficacy of selection on continuous variation, Provine 1971.)

When Mendel’s principles were rediscovered in 1900, two central problems needed to be resolved: determining the generality of Mendelian inheritance, and understanding its connection to continuous variation. Famously, Morgan began his *Drosophila* experiments skeptical of Mendelism’s generality (Shine and Wrobel 1976, Chap. 4). One of the great triumphs of early 20th century genetics was reconciling Mendelian genetics with Galton’s “biometrical” observations.

Both key elements of the reconciliation, environmental effects and polygenic inheritance, appear in Mendel (1866). In discussing flowering time in *P. sativum* hybrids, Mendel recognized that hybrids can be intermediate, and that subtle environmental differences, such as planting depth and temperature, can critically affect phenotypes. Considering the inheritance of flower and seed-color differences between *Phaseolus nanus* and *Ph. multiflorus*, Mendel also clearly understood that, when multiple loci contribute additively to a trait, a wide range of phenotypes can appear in the *F₂*. Shortly after Mendel’s rediscovery in 1900, Bateson and Saunders (1902) conjectured that polygenic inheritance might explain continuous variation. However, it took about 15 years to accumulate convincing data, and it took Fisher’s (1918) brilliance to develop...
the relevant mathematics [following important contributions by Pearson and Yule (Provine 1971)].

A trio of plant geneticists, Johannsen, Nilsson-Ehle, and East, were largely responsible for demonstrating the complementary roles of multiple Mendelian factors plus nongenetic, “environmental,” effects. Wright (1968, Chap. 15) gives a concise and authoritative overview of their experiments. Johannsen (1903, summarized in Yule 1903) demonstrated that average bean weight in highly inbred “pure lines” of Phaseolus vulgaris differed because of hereditary factors, whereas weight differences within pure lines were nonheritable, and did not respond to selection. Working with varieties of hexaploid bread wheat, Nilsson-Ehle (1909) found that the range of kernel colors from dark-red to white was determined by roughly equal and additive effects of alleles at three (presumably homologous) loci. East’s (1916) work exquisitely unified environment effects and polygenic inheritance by examining crosses between inbred lines of Nicotiana longiflora with markedly different corolla lengths. Ignoring his small 1913 samples, East found that the parentals had mean corolla lengths of 40.54 and 93.30 mm, with corresponding variances 3.53 and 5.11 mm². As expected with predominantly additive allele effects, the means of the $F_1$ (63.53 mm) and $F_2$ (68.65 mm) were both close to the parental average (66.92 mm). The $F_1$ showed substantially larger phenotypic variance (8.53 mm²) than the parentals, perhaps suggesting that the parentals were not fully homozygous (or that their hybrids were developmentally unstable). However, the $244 F_2$ showed greatly increased variance (40.52 mm²), but none produced phenotypes recovering the parental means—as expected with segregation and assortment of several relevant loci. With five loci contributing to the parental difference, the frequency of each parental genotype in the $F_2$ would be about $10^{-3}$.

As a graduate student, Sewall Wright realized that data like East’s could be used to estimate $n$, the number of loci contributing to the parental difference. Assuming no linkage, equal and additive allelic effects, with all trait-increasing “$+$” alleles in one parental line, and all “$−$” alleles in the other, the “Wright-Castle estimator” (Castle 1921 first published the result without acknowledging his student) is

$$n = \frac{(\bar{P}_1 - \bar{P}_2)^2}{8(V_{P1} - V_E)}, \quad (1)$$

where the $\bar{P}_i$ indicate parental means, and $V_E$ is an estimate of the nongenetic variance experienced by the $F_2$. Under some (but not all) more realistic conditions, including linkage and unequal allelic effects, Equation (1) underestimates the number of loci responsible for the parental difference (Wright 1968, Chap. 15). If we conservatively estimate $V_E$ as the average of the parental variances, East’s data produce $n = 9.6$. Hence, we expect that at least 10 loci contribute to the corolla length difference between East’s varieties.

For those of us who use East’s data to illustrate the reconciliation of Mendelian and Galtonian approaches to inheritance, it is jarring to recognize that East saw his study as capping a triumph of Mendel over Galton (“...Galtonian regression in the original sense is now entirely discredited...,” p. 174). The data that supposedly contradict Galton result both from genetic differences among $F_2$ families, and the high variance of regression estimates based on relatively few observations, a fact East acknowledges later on the same page. Unfortunately, East’s standing as an eminent Harvard geneticist, and his willingness to make unsubstantiated pronouncements, such as rejecting Galtonian regression, also contributed to his influential role in promoting racial anxiety in the United States eugenics movement (East 1920; Provine 1973, 1986).

**Literature Cited**


**Further reading in GENETICS**


Other GENETICS articles by E. M. East