



## Stages in the evolution of the Hardy-Weinberg law

Alan E. Stark

*Balgowlah NSW, Australia*

### Abstract

The Hardy-Weinberg law has been used widely for about one hundred years with little question as to the foundations laid down by its originators. The basic assumption of random mating, that is choice of mates by a process akin to that of a lottery, was shown to produce genotypic proportions following the “binomial-square” rule, the so-called Hardy-Weinberg proportions (HWP). It has been assumed by many that random mating was the only way of pairing genes capable of producing HWP. However it has been shown that HWP can be obtained and maintained by non-random mating. The steps along the way to this revelation and some implications are reviewed.

*Key words:* Hardy-Weinberg law, non-random mating.

Received: May 24, 2006; Accepted: August 16, 2006.

It is a strange fact that the most basic law of population genetics, which is attributed to Hardy (1908) and Weinberg (1908), is poorly understood by many scientists who use it routinely. Support for this apparently extravagant statement can be found in numerous textbooks and web sites. It is unfair to take Hancock (2004), which appears to be an excellent monograph on its topic - plant evolution - to illustrate the point. But, because it is such an admirable book, it does provide a good example. In the section entitled “Random Mating and Hardy-Weinberg Equilibrium” on page 24 is the following statement: “What these two men [Hardy and Weinberg] discovered mathematically is that genotype frequencies will reach an equilibrium in one generation of random mating in the absence of any other evolutionary force. The frequencies of different genotypes will then depend only upon the allele frequencies of the previous generation. If gene frequencies do not accurately predict genotype frequencies, then plants are crossing in a non-random fashion or some other evolutionary force is operating”. While some of the quoted text is valid, the overall impression that random mating and Hardy-Weinberg proportions (HWP) are inextricably linked is not correct. Stark (2006, in press) shows that, provided the population has discrete and non-overlapping generations, HWP can be attained in one round of non-random mating and that random mating is a single point in a continuum of such possibilities. Furthermore Li (1988) and Stark (2005) give mating schemes which show that HWP can be maintained by non-random mating.

To illustrate the problem further, Hancock (p. 35) writes “While the Hardy-Weinberg formula is very useful in describing the completely randomized situation, it is a rare natural population that is in fact in Hardy-Weinberg equilibrium. More commonly, mating is not random and the populations are subjected to other evolutionary forces, such as migration, genetic drift and natural selection”. As noted above, HWP are consistent with (some) non-random mating.

The object of this paper is to trace some of the history of basic population genetics to see how this misunderstanding has arisen and to point out some of the misconceptions. In the proceedings of the Mendel Centennial Symposium sponsored by the Genetics Society of America, which was held in 1965, Wright (1967) wrote: “It is not surprising that the first attempt to formulate the statistical properties of Mendelian populations and to reconcile the views of Galton with those of Mendel was made by a biometrician, Yule (1902). He showed that the unselected, randomly bred descendants of a cross would maintain indefinitely the 1:2:1 ratio of  $F_2$ . Castle (1903), in the course of a criticism based on a misunderstanding of Yule’s postulates, worked out for the first time the effect of selection in a Mendelian population. He also showed that if selection ceases, the composition of the randomly bred descendants remains constant thereafter, with genotypic frequencies according to the now familiar binomial-square rule. Unfortunately, he did not stress this as a basic principle of population genetics, and it did not attract attention. It is now known as the Hardy-Weinberg law because of independent restatements by Hardy and by Weinberg in 1908”. We see here not only mention of the fore-runners of Hardy and Weinberg but

also the emphasis on random mating as recently as the centenary of Mendel's work.

Cotterman (1940) wrote: "the fact first noted by Hardy (1908), that random mating produces an equilibrium in the distribution of the three genotypes,  $AA$ ,  $Aa$ ,  $aa$ , the frequencies of which are then proportional to the terms of the squared binomial of the gametic proportions of the two genes. The derivation also reveals the additional fact, later noted by Wentworth and Remick (1916), that such an equilibrium is reached in a single generation of random mating". This quotation appears on page 198 of the reproduction of Cotterman's thesis in Ballonoff (1974). That Hardy-Weinberg proportions were reached in one generation of random mating was first made explicit in 1916, according to Cotterman. As noted above, this is possible with non-random mating also.

Dobzhansky (1998) contributed to the workshops on the "Evolutionary Synthesis" held in 1974. His submission, along with those of other attendants was published in 1980 and again in 1998. In his paper, which dealt with scientists from the Soviet Union, he wrote: "The personal backgrounds and the scientific traditions of this Anglo-American trinity [Fisher, Wright, and Haldane] were quite different from Chetverikov's, as well as from each other. It is well known that Mendelism and genetics in general were, by a singular miscomprehension, regarded as contradictory to Darwin's theory of natural selection in the early decades of the twentieth century. The short but important theoretical paper by Hardy in 1908 should have dispelled this miscomprehension, and in fact it served as a point of departure for Chetverikov as well as for the other founding fathers (the parallel work of Weinberg was not then considered)". This quotation illustrates the significance of the H-W law and also the fact that Weinberg's independent contribution was not recognized until Stern (1943) drew attention to it. Even then it was not universally known for Schmalhausen (1986, p. 98) cites Hardy (1908) as the originator of the formula. Schmalhausen's (1986) text was translated with minor changes from the original Russian manuscript available around 1947. Also Malécot (1948) refers to "loi de Hardy".

Hardy's (1908) paper contains some questionable statements. For example he wrote "suppose that the numbers are fairly large, so that the mating may be regarded as random". There is no necessary relation between population size and randomness of mating. Presumably Hardy wished to exclude genetic drift as a mechanism capable of changing the status of the population. The genesis of the fallacy that random mating was a necessary requirement for the production and maintenance of Hardy-Weinberg proportions may be due to Hardy himself, for he wrote "I have, of course, considered only the very simplest hypotheses possible. Hypotheses other than that of purely random mating will give different results". As will be seen, this statement is not generally true.

Weinberg (1908) invoked random mating in his study of the inheritance of twinning in humans. The translated version of his paper includes the following: "I have therefore tried to construct a formula for the frequency of dominant and recessive traits among the mothers, daughters, and sibs of persons affected with such traits of the same character, under the assumption that absolute panmixis is present". As will be shown later, the kind of inference that can be made validly, using the data which Weinberg collated, depends critically on what assumption is made as to the mating system which generates HWP.

Li (1988) made a valuable contribution when he showed that Hardy-Weinberg (H-W) frequencies can be maintained in large populations with non-random mating. Crow (1988) made the following comment on Li's article: "we are still learning things about the Hardy-Weinberg relationship. C.C. Li shows that random mating is a sufficient, not a necessary condition for H-W ratios". This fact was noted by Stark (1977a) and is implicit in Stark (1977b). However, to add weight to his model, Li attributed additive genotypic values to male and female parents from which he calculated that the coefficient of correlation between mating pairs is zero. In this paper it is shown that this arbitrary assignment of values is questionable.

The origin of the confounding of random mating and HWP can be traced to the (correct) principle enunciated by Mendel that, in the formation of gametes, the genetic material is halved and then restored to its full complement in the production of offspring. It is not immediately obvious that, in this sequence of events, the pair of uniting gametes may be independent in a probabilistic sense while the genotypes of its parents not be independent. There are numerous examples in the genetics literature where random union of gametes and random pairing of parents are taken to be equivalent. As will be seen, this is not necessarily so.

The next section contains a description of Li's model. That is followed by an analysis of dependence and independence, as mentioned in the previous paragraph. Then Li's model is expressed in a canonical form which allows an analysis of the assignment of genotypic values. The final section reviews Weinberg's (1908) study of twinning.

### Li's parametrization of non-random mating

Consider a population with respect to a single locus having alleles  $A$  and  $B$  with respective frequencies  $q$  and  $p$ , the same in males and females. Denote frequencies of genotypes  $AA$ ,  $AB$  and  $BB$  by  $f_0$ ,  $f_1$  and  $f_2$ . Table 1 gives Li's (1988) symmetrical mating model which he introduces with the remark: "When reciprocal crosses have the same frequency, the general pattern will be symmetrical". Thus the roles of males and females can be reversed without changing the model. This case is simpler than Li's more general model but is suitable for the present purposes. The

**Table 1** - Li's symmetric non-random mating model.

Male vs. female	AA	AB	BB
AA	$q^4 + a$	$2pq^3 - a - b$	$p^2q^2 + b$
AB	$2pq^3 - a - b$	$4p^2q^2 + 4b$	$2p^3q + a - 3b$
BB	$p^2q^2 + b$	$2p^3q + a - 3b$	$p^4 - a + 2b$

3 x 3 matrix of cell frequencies will be denoted by  $[f_{ij}]$ ,  $i = 0, 1, 2; j = 0, 1, 2$ .

Both row totals and column totals give Hardy-Weinberg proportions:  $f_0 = q^2, f_1 = 2pq, f_2 = p^2$ . Thus the parental population is in H-W form and it is simple to show that the distribution of genotypes among offspring is the same. Note that  $f_{11} = 4f_{02}$ .

Parameters  $a$  and  $b$  are constrained by the requirement that the elements of  $[f_{ij}]$  be non-negative. Thus, for example,  $a$  cannot be less than  $-q^4$ . It is convenient to replace the symbol  $b$  by  $k$ , defined as  $k = b/a$ , in the section in which Li's model is analysed. The introduction of  $k$  entails  $a > 0$ . However, this is not a practical constraint since  $|a|$  can be made arbitrarily small.

### Randomness and independence

There exist many examples where the derivation of the H-W distribution appeals to a random union of egg and sperm. In fact, this is correct but, as Li has shown, not necessarily because of random pairing of mates. This may be the basic cause of the invalid claim that random pairing of mates is a necessary condition for the existence of H-W frequencies. Suppose that a given zygote was derived from an egg of type  $A$ , a meaningful question is: what is the probability that this egg was fertilised by a sperm of type  $A$ ? Using the mating matrix  $[f_{ij}]$  of Table 1, by application of the probability calculus, the required probability is found to be equal to  $q$ . The probability that an egg is of type  $A$  is  $q$ , so the joint probability is  $q^2$ , thus deriving the H-W frequency of offspring of type  $AA$ . Corresponding calculations produce the remaining H-W zygotic frequencies.

The preceding paragraph may be summarized in another way. But first it is necessary to recall the definitions of conditional probability and independence. Two events  $X$  and  $Y$  are said to be independent if the joint probability of  $X$  and  $Y$ , denoted  $\Pr(XY)$ , is equal to the products of the probabilities of the separate events, that is  $\Pr(XY) = \Pr(X) \cdot \Pr(Y)$ . This property of independence can be described in another way through the conditional probability of  $X$ , given  $Y$ , denoted  $\Pr(X|Y)$ . By definition,  $\Pr(XY) = \Pr(Y) \cdot \Pr(X|Y)$ , so that independence of  $X$  and  $Y$  requires the identity  $\Pr(X|Y) = \Pr(X)$ . Relating this to the derivation of the probability of genotype  $AA$  in the previous paragraph, it can be said that the union of egg and sperm is the joint occurrence of independent events. In other terms, egg and sperm are said to pair 'randomly'.

However, pairing of mates in the scheme given by Table 1 is not random. As can be seen by looking along the rows of the matrix, the conditional probabilities relating to male type, given the type of female, are not identical to the unconditional probabilities relating to male types, that is the types of female and male mates are not independent.

### The canonical form of Li's model

It is instructive to examine  $[f_{ij}]$  through its canonical form

$$f_{ij} = f_i f_j (1 + \rho x_i x_j + \sigma y_i y_j), \quad (i = 0, 1, 2; j = 0, 1, 2) \quad (1)$$

Formula (1) is a particular example of the representation of a discrete bivariate probability distribution which Lancaster (1969, p. 90) refers to as "Fisher's Identity". Some important properties of this representation are given below. An example of its use in classifying mating systems is Stark (1980).

Without loss of generality take  $q$  in the interval  $0 < q < 1/2$ . Putting  $k = b/a$ , some simplification is achieved by introducing the following symbols:

$$n = 1 + p$$

$$d = 1 + p(p - q) - 2q^2k$$

$$r = \sqrt{(p^2 + q^2 - 4q^2k + 2q(1 + q)k^2)}$$

Then the terms in (1) are obtained from:

$$= a(p - q + 2qk - r)/(2p^2q^2)$$

$$= a(p - q + 2qk + r)/(2p^2q^2)$$

$$x_2 = \sqrt{(1/2q(n + d/r))/p}$$

$$y_2 = \sqrt{(1/2q(n - d/r))/p}$$

$$x_0 = -p \sqrt{(1/2q(n + d/r))/(qn)} - I_k \sqrt{(p(n - d/r))/(qn)}$$

$$x_1 = -p \sqrt{(1/2q(n + d/r))/(qn)} + I_k \sqrt{(p(n - d/r))/(2pn)}$$

$$y_0 = -p \sqrt{(1/2q(n - d/r))/(qn)} + I_k \sqrt{(p(n + d/r))/(qn)}$$

$$y_1 = -p \sqrt{(1/2q(n - d/r))/(qn)} - I_k \sqrt{(p(n + d/r))/(2pn)}$$

$I_k$  is an indicator function defined as:

$$I_k = 1, k > 0; I_k = -1, k < 0.$$

The vector of values  $\{x_0, x_1, x_2\}$  will be denoted by  $x$  and  $\{y_0, y_1, y_2\}$  by  $y$ .

The connection between Li's formulation and the canonical form is complicated. Therefore there is some value in considering some special cases in order to illustrate in a simpler way the connection and to reveal key properties of the canonical form. If  $k = -p/(1 + q)$ , then

$$\rho = -2a/(p^2q(1 + q)), \quad \sigma = a/(pq^2(1 + q)),$$

$$x_0 = 0, x_1 = -1/\sqrt{(2q(1 + q))}, x_2 = p^{-1}\sqrt{(2q/(1 + q))}.$$

$$y_0 = -q^{-1}\sqrt{p(1+q)}, \quad y_1 = y_2 = -y_0^{-1}.$$

For this case,  $\rho/\sigma = -2q/p$ . Calculation shows that, with respect to the H-W distribution, the mean of  $\mathbf{x}$  and the mean of  $\mathbf{y}$  are each zero and the standard deviation of each is one. Furthermore, with respect to the mating matrix  $[f_{ij}]$ , the correlation of  $\mathbf{x}$  in females and  $\mathbf{x}$  in males is  $\rho$  and the correlation of  $\mathbf{y}$  in females and  $\mathbf{y}$  in males is  $\sigma$ . The correlation of  $\mathbf{x}$  in females and  $\mathbf{y}$  in males is zero. These are properties of the general form of Fisher's Identity.

Narrowing this example further, when  $q = 1/3$ ,  $a = 2pq^3$  and  $k = -p/(1+q)$ ,

$$\rho = -1/2, \quad \sigma = 1/2,$$

$$x_0 = 0, \quad x_1 = -3\sqrt{2/4}, \quad x_2 = 3\sqrt{2/4},$$

$$y_0 = -2\sqrt{2}, \quad y_1 = \sqrt{2/4}, \quad y_2 = \sqrt{2/4}.$$

This case illustrates readily another property of Fisher's Identity, namely that the regression of  $\mathbf{x}$  values in females on that of males is linear with coefficient equal to  $\rho$  and the regression of  $\mathbf{y}$  values in females on that of males is linear with coefficient  $\sigma$ . This property was the condition taken by Hirschfeld (1935) as the starting point for investigating a two-way contingency table of frequencies. He reached essentially the same representation as Fisher's Identity.

As has been noted already, the case of Fisher's Identity used here is special in that it is being applied to a symmetric matrix, the roles of males and females being exchangeable without altering the properties. In an example used by Fisher (1940) and later by Maung (1942), where Fisher's formula is given, the object was to study the connection between eye colour and hair colour in Scottish children. They calculated sets of numerical values to stand for the grades of hair colour and eye colour, one set for each. That is they were dealing with two different but supposedly related characters. Here, given a symmetric mating table, two sets of values,  $\mathbf{x}$  and  $\mathbf{y}$ , are calculated for each sex, but the same for each sex. Thus there is a difference from the analysis of Li who assumed one set of *additive* values for each sex.

A natural question to ask is "What is gained by introducing Fisher's Identity?". On the face of it, Li's model is very simple and shows very easily and generally how H-W proportions can be maintained with non-random mating. But Li went further and stated that the correlation between mates is zero. He did this by assigning *additive* values to genotypes, that is essentially 0 to type *AA*, 1 to type *AB*, and 2 to type *BB*. This is intuitively plausible and was used frequently by Sewall Wright. However, there is no compelling reason to assign additive values. The important point is that, using Fisher's Identity, the pair of conditions  $\rho = 0$  and  $\mathbf{x}$  be a vector of *additive* values is true if and only if  $b = a$  ( $k = 1$ ). Then (1) reduces to:

$$f_{ij} = f_j f_i (1 + \sigma y_i y_j), \quad (i = 0, 1, 2; j = 0, 1, 2), \quad (2)$$

where  $\sigma = a/(p^2 q^2)$ , and  $\mathbf{y} = \{p/q, -1, q/p\}$ . Taking  $q = 1/2$ ,  $\sigma$  can take values in the interval  $(-q^2/p^2, q/p)$ , the corresponding interval for  $a$  being  $(-q^4, pq^3)$ , as noted by Stark (2005).

Li's assignment of additive values is purely arbitrary, except in the sense that they are the count of genes of type *B* in the respective genotypes, whereas  $\mathbf{x}$  and  $\mathbf{y}$  are intrinsic to the structure of (1).

## Weinberg's analysis of twin data

Crow (1999) discussed the reasons for the relative neglect of Weinberg's work, noting *inter alia* that Weinberg (1908) concluded that dizygotic twinning was inherited. I show here that Weinberg's analysis, though ingenious, can be questioned on the ground that HWP can be maintained with non-random mating. As will be seen, the relative proportions of a particular phenotype among different categories of relatives may be confounded with non-random mating parameters.

Weinberg selected a single locus model with either dominant or recessive propensity to produce twins at a higher than normal rate, which for convenience is referred to here as a high propensity phenotype, denoted by  $\underline{A}$ . He calculated the proportions of high propensity versus normal groups ( $\underline{B}$ ) among three types of relatives of high propensity women: parents, children and sibs. Following Weinberg's reasoning, but using Li's mating model rather than random mating frequencies as the basic mating system, the probabilities relating to high propensity trait being dominant are as follows:

among parents and children (of  $\underline{A}$  mothers):

$$(q(1+pq) - 1/2(a-b))/(q(1+p)); \quad (3)$$

among sibs:

$$(4q(1+pq) + p^2 q^2 - (2a-3b))/(4q(1+p)). \quad (4)$$

The probabilities relating to high propensity trait being recessive are as follows:

among parents and children:

$$(q^3 + 1/2(a-b))/q^2; \quad (5)$$

among sibs:

$$(q^2(1+q)^2 + 2a-b)/(4q^2). \quad (6)$$

Each of formulae (3) - (6) gives the conditional probability that the particular relative has higher propensity to have twins, given that the index mother has higher propensity ( $\underline{A}$ ).

Weinberg used the differences between observed and hypothetical rates in sibs and the other two categories of relatives to support his case. However, the hypothetical differences may depend on the values of  $a$  and  $b$ . To illustrate



this with admittedly extreme choices of  $a$  and  $b$ , the difference in hypothetical rates can even be zero.

Bulmer (1970) reviewed data relating to the inheritance of twinning in humans. It is impossible, in a few sentences, to do justice to the thorough treatment of this question given by Bulmer. He compiled tables of twinning rates per thousand maternities among mothers of twins and first degree relatives of such mothers and in the general population. He found, for example, that the repeat dizygotic twinning rate of high propensity mothers is nearly four times the dizygotic twinning rate in the general population.

Bulmer used data from various sources including his own collection and tables published by Weinberg in 1901 and 1909 from family registers in Stuttgart relating to Württemberg. However he did not refer to Weinberg (1908) which gives twinning rates in various categories. Some of these are given in Table 2 where they have been converted to rates per thousand maternities. The hypothetical rates are for the model of recessive inheritance preferred by Weinberg and can be obtained from formulae (5) and (6) by setting  $a = 0$  and  $b = 0$  and inserting estimated values of  $q$  and twinning rates from observed twinning rates. Weinberg inferred  $q = 1/2$ . Twinning fractions of  $1/32.5$  for high propensity mothers and  $1/130$  for the remaining mothers produce hypothetical rates close to Weinberg's values which are given in the third column of Table 2.

The final paragraphs of the translation of Weinberg's (1908) paper are:

"The situation found in the inheritance of twinning best finds its explanation in the assumption that the trait for twinning is inherited according to the Mendelian rule and is recessive.

This investigation (a more thorough presentation based on a new collection of material that I have in the meantime obtained from another source will follow) may show that one can gain an insight into the nature of human inheritance by suitable changes in the investigative methods".

Weinberg's analysis of twinning rates depended on the assumption of random mating used to construct the mating matrix. Li's model shows that HWP can be found in the general population with non-random mating. Formulae (3) - (6) show that the values of parameters  $a$  and  $b$  are required to calculate hypothetical twinning rates, for comparison with observed rates. Weinberg's analysis exploited the fact

that sibs may have zero, one or two genes at a locus identical by descent whereas parent and child have one. However the mating system has bearing on the probabilities of the respective events. This needs to be taken into account when making inferences about the mode of inheritance of a trait.

## References

- Bulmer MG (1970) *The Biology of Twinning in Man*. Clarendon Press, Oxford, 205 pp.
- Castle WE (1903) The laws of Galton and Mendel and some laws governing race improvement by selection. *Proc Amer Acad Arts Sci* 39:233-242.
- Cotterman CW (1940) A calculus for statistico-genetics. PhD Thesis, Graduate School of The Ohio State University. In: Ballonoff PA (ed) *Genetics and Social Structure: Mathematical Structuralism in Population Genetics and Social Theory*. Dowden, Hutchinson & Ross, Inc., Stroudsburg, pp 157-271.
- Crow JF (1988) Eighty years ago: The beginnings of population genetics. *Genetics* 119:473-476.
- Crow JF (1999) Hardy, Weinberg and language impediments. *Genetics* 152:821-825.
- Dobzhansky T (1998) The birth of the genetic theory of evolution in the Soviet Union in the 1920's. In: Mayr E and Provine WB (eds) *The Evolutionary Synthesis: Perspectives on the Unification of Biology*. Harvard University Press, Cambridge, Massachusetts and London pp 229-242.
- Fisher RA (1940) The precision of discriminant functions. *Ann Eugen Lond* 10:422-429.
- Hancock JF (2004) *Plant Evolution and the Origin of Crop Species*. 2<sup>nd</sup> edition. CABI Publishing, Wallingford, Oxfordshire and Cambridge, MA, 313 pp.
- Hardy GH (1908) Mendelian proportions in a mixed population. *Science* 28:49-50.
- Hirschfeld HO (1935) A connection between correlation and contingency. *Proc Camb Phil Soc* 31:520-524.
- Lancaster HO (1969) *The Chi-Squared Distribution*. John Wiley & Sons, Inc., New York, 356 pp.
- Li CC (1988) Pseudo-random mating populations. In celebration of the 80<sup>th</sup> anniversary of the Hardy-Weinberg law. *Genetics* 119:731-737.
- Malécot G (1948) *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris, 60 pp.
- Maung K (1942) Measurement of association in a contingency table with special reference to the pigmentation of hair and eye colours of Scottish school children. *Ann Eugen Lond* 11:189-223.
- Schmalhausen II (1986) *Factors of Evolution: The Theory of Stabilizing Selection*. The University of Chicago Press, Chicago and London, 327 pp.
- Stark AE (1977a) Models of correlation between mates and relatives and some applications. Unpublished Ph.D. Thesis, School of Community Medicine, The University of New South Wales.
- Stark AE (1977b) Dwa uogólnienia prawa Hardy'ego-Weinberga w genetyce populacji (The Hardy-Weinberg Law of Population Genetics and two Generalizations). *Matematyka Stosowana* IX:123-137.
- Stark AE (1980) Inbreeding systems: Classification by a canonical form. *J Math Biol* 10:305.

**Table 2** - Weinberg's observed and hypothetical twinning rates.

Relative of mother (of twins)	Observed twinning rate (per 1000 maternities)	Hypothetical twinning rate (per 1000 maternities)
Mother	19.8	19.2
Daughter	16.4	19.2
Sister	26.4	20.4

- Stark AE (2005) The Hardy-Weinberg principle. *Genet Mol Biol* 28:485-485.
- Stark AE (2006) A clarification of the Hardy-Weinberg law. *Genetics*, in press.
- Stern C (1943) The Hardy-Weinberg law. *Science* 97:137-138.
- Weinberg W (1908) Über den Nachweis der Vererbung beim Menschen. *Jahresh Verein f vaterl Naturk Württem* 64:368-382. English version: On the demonstration of heredity in Man. In: Boyer SH (ed) *Papers on Human Genetics*, Prentice-Hall, Englewood Cliffs, pp 4-15.
- Wentworth EN and Remick BL (1916) Some breeding properties of the generalized Mendelian population. *Genetics* 1:608-616.
- Wright S (1967) The foundations of population genetics. In: Brink RA and Styles ED (eds) *Heritage from Mendel*. The University of Wisconsin Press, Madison and London, pp 245-263.
- Yule GU (1902) Mendel's laws and their probable relation to intraracial heredity. *New Phytol* 1:193-207, 222-238.

*Associate Editor: Paulo A. Otto*