GENE CONVERSION: A HITHERTO OVERLOOKED PARAMETER IN POPULATION GENETICS[1]

HERBERT GUTZ[2] AND JOHN F. LESLIE[3]

Institute for Molecular Biology, The University of Texas at Dallas, Richardson, Texas 75080

Manuscript received July 28, 1975
Revised copy received December 1, 1975

ABSTRACT

Gene conversion causes deviations from the 2:2 segregation of allele pairs in meiosis. Thus, gene conversion is a potential cause for changes of allele frequencies in populations. Equations are derived for the effects of conversion in a large random-mating population. The influence of gene conversion on allele frequencies is compared with that of spontaneous mutation and meiotic drive.

GENE conversion is a mechanism which causes deviations from the 2 : 2 segregation of allele pairs (**A, a**) in meiosis. This phenomenon has been primarily studied in Ascomycetes, since tetrad analyses can be easily performed with these fungi (FINCHAM and DAY 1971; FOGEL, HURST and MORTIMER 1971; GUTZ 1971; LEBLON 1972). For example, 3**A** : 1**a**, 1**A** : 3**a**, 5**A** : 3**a**, or 3**A** : 5**a** segregations have been observed. (Note: in the higher Ascomycetes, 8 spores are formed per ascus.) The total frequency of "aberrant" segregations can be as high as 10%.

The occurrence of gene conversion has also been demonstrated in *Drosophila melanogaster* (BALLANTYNE and CHOVNICK 1971; CHOVNICK, BALLANTYNE and HOLM 1971; CHOVNICK 1973) and, in all probability, in *Zea mays* (NELSON 1975). It is reasonable to suppose that gene conversion takes place in all eukaryotes.

Since individual conversion events cause a deviation from the 2**A** : 2**a** segregation, gene conversion is a potential cause for changes in allele frequencies in populations. Its implication for population genetics has been briefly mentioned by CHOVNICK (1973). To our knowledge, however, the role of gene conversion has not yet been studied in detail by population geneticists. In the present paper we derive equations for the effects of gene conversion in a random-mating population. Furthermore, we discuss whether conversions can have an effect on allele frequencies comparable to that of spontaneous mutations. We will use the texts of LI (1958) and of CROW and KIMURA (1970) as our general references on population genetics.

[1] This project was supported by Grant GM-19849 awarded by National Institute of General Medical Sciences, DHEW. Part of the work was performed when the junior author (J.F.L.) was assisted by a Clark Foundation Fellowship.

[2] Present address: Lehrstuhl für Genetik, Technische Universität Braunschweig, Postfach 3329, D-3300 Braunschweig, Germany. To whom requests for reprints should be sent.

[3] Present address: Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706.

THEORY

We assume a large random-mating population of a diploid species with non-overlapping sexual generations. To simplify the treatment, we will consider only 3**A** : 1**a** and 1**A** : 3**a** conversion ratios.

For an evaluation of the effect of gene conversion in the population, the total conversion frequency per meiosis (f) is of importance, as well as the relative frequencies of 3**A** : 1**a** segregations (i) and 1**A** : 3**a** segregations (j) ($i + j = 1$). Meiotic gene conversion would have no effect on the allele frequencies if $i = j$. The parameter k will be called the net gene conversion constant; k relates f, i, and j:

$$k = f(i - j) = f(2i - 1). \tag{1}$$

Gene conversion can take place only in heterozygous individuals (**Aa**). According to the Hardy-Weinberg law, the heterozygotes are present with a frequency $2pq$ in a random-mating population (p: frequency of **A**, q: frequency of **a**). The gametic output from heterozygotes is:

$$p_H = f(\tfrac{3}{4}\, i + \tfrac{1}{4}\, j) + \tfrac{1}{2}(1 - f)$$
$$p_H = \tfrac{1}{2} + \tfrac{1}{4}\, f(2i - 1) = \tfrac{1}{2} + \tfrac{1}{4}\, k.$$

Therefore, the change in the allele frequency among the gametes from the heterozygotes is

$$\Delta p_H = p_H - \tfrac{1}{2} = \tfrac{1}{4}\, k.$$

Since the heterozygotes represent $2pq$ of the population, Δp due to gene conversion is:

$$\Delta p = 2\, pq \cdot \tfrac{1}{4}\, k = \tfrac{1}{2}\, kpq = \tfrac{1}{2}\, kp(1 - p). \tag{2}$$

If Δp is small compared to p, equation (2) can be written as the differential equation

$$\frac{dp}{dt} = \tfrac{1}{2}\, kp(1 - \mathrm{p})\ , \tag{3}$$

where t is the time measured in generations, n is the number of generations, p_0 and q_0 are the frequencies of **A** and **a** in the initial population, and p_n is the frequency of **A** in the nth generation. In each generation, $p + q = 1$. Integration of (3) with respect to t gives

$$\frac{p_n}{(1 - p_n)} = \frac{p_0}{(1 - p_0)} \exp\left(\frac{kn}{2}\right)\ ,$$

which can be transformed to

$$p_n = \frac{(p_0/q_0)\exp(kn/2)}{1 + (p_0/q_0)\exp(kn/2)}\ . \tag{4}$$

Equation (4) is the general equation for changes in allele frequencies caused by gene conversion at a single site in a large random-mating population. For a given $k \neq 0$, the effect of gene conversion is maximized with $p = q$; the effect of

conversion decreases as the number of heterozygotes decreases. If no counteracting forces (e.g., selection and/or mutation) are at work, it is apparent that $\lim_{n\to\infty} p_n = 1$ if $k > 0$, and that $\lim_{n\to\infty} p_n = 0$ if $k < 0$. If $k = 0$, the allele frequencies do not change.

A prerequisite for the operation of gene conversion is that at least two different alleles are present in the population. Before conversion can occur, mutations must take place.

When the occurrence of spontaneous mutations in the population is also taken into consideration, then the effect of gene conversion is to counteract the mutation pressure from **A** to **a** and to act in the same direction as the mutation pressure from **a** to **A**, if $k > 0$. The opposite situation prevails if $k < 0$. An instructive way of looking at the importance of gene conversion is to compare the equilibrium of a random-mating population in which mutation but no conversion takes place, with that of a population in which both events take place. We neglect the fact that mutation creates diverse alleles which may show different conversion characteristics (this point is explained in more detail in the DISCUSSION).

If in the population only mutations from **A** to **a** and from **a** to **A** are considered (mutations rates per generation: μ and ν, respectively), then the net change in the frequency of **A** per generation is:

$$\Delta p = \nu q - \mu p. \tag{5}$$

At equilibrium the following relationships hold true:

$$\Delta p = \nu q - \mu p = 0, \tag{6a}$$

$$\hat{p} = \frac{\nu}{\mu + \nu}, \quad \hat{q} = \frac{\mu}{\mu + \nu} \tag{6b}$$

(the "hat" denotes the equilibrium values of p and q).

By adding equations (2) and (5) we get, for the combined change of p caused by mutation and gene conversion,

$$\Delta p = \nu q - \mu p + \tfrac{1}{2} kpq. \tag{7}$$

At equilibrium ($\Delta p = 0$), equation (7) yields

$$k = \frac{2(\mu\hat{p} - \nu\hat{q})}{\hat{p}\hat{q}} = \frac{2[(\mu + \nu)\hat{p} - \nu]}{\hat{p} - \hat{p}^2}, \tag{8a}$$

$$\hat{p} = \frac{k - 2\mu - 2\nu \pm \sqrt{(2\mu + 2\nu - k)^2 + 8k\nu}}{2k}. \tag{8b}$$

An interesting relationship between gene conversion and the mutation pressure becomes apparent when only one of the two directions of mutation is considered ($\nu = 0$ or $\mu = 0$). If we assume $k \neq 0$, $\mu > 0$ and $\nu = 0$, equation (8b) simplies to

$$\hat{p} = \frac{(k - 2\mu) \pm (2\mu - k)}{2k} \tag{9}$$

For $k = 2\mu$, equation (9) has one solution: $\hat{p} = 0$. For $k \neq 2\mu$, equation (9) has two solutions, one of which is $\hat{p} = 0$. The second solution is either $0 < \hat{p} < 1$, if $k > 2\mu$, or $\hat{p} < 0$, if $0 < k < 2\mu$, or $\hat{p} > 1$, if $k < 0$. The solution $\hat{p} = 0$ is trivial (it corresponds to a population with only **aa** individuals); the solutions $\hat{p} < 0$ and $\hat{p} > 1$ are biologically meaningless. Gene conversion is sufficient to prevent the extinction of **A** only when $k > 2\mu$. An analogous situation exists, if we assume $k \neq 0$, $\nu > 0$ and $\mu = 0$. An equilibrium $0 < \hat{p} < 1$ results only for $k < (-2\nu)$.

The situation is more complex if $\mu > 0$ and $\nu > 0$. An informative way to evaluate the effect of gene conversion under these conditions is to assume specific values for $\hat{p}$ and $\hat{q}$. For example, if $\hat{p} = \hat{q} = 0.5$, equation (8a) yields $k = 4(\mu - \nu)$. Thus, if the net gene conversion constant is four times as large as the difference of the spontaneous mutation rates, the equilibrium $\hat{p} = \hat{q}$ will result.

To illustrate the importance of k in comparison to μ and ν, we present an additional example. If $\mu = 10^{-5}$ and $\nu = 10^{-6}$, then mutation alone (equation 6b) would result in $\hat{p} = 0.0909$ and $\hat{q} = 0.9091$. What value of k is needed to shift the equilibrium to $\hat{p} = 0.2$ and $\hat{q} = 0.8$? Using equation (8a), we find $k = 1.5 \times 10^{-5}$, a value which is of the same order of magnitude as μ.

DR. J. F. CROW drew our attention to the fact that the algebra derived for gene conversion shows a close similarity to that of meiotic drive. Our equation (2) is analogous to the special case of equal meiotic drive in both sexes without selection (cf. the first equation on page 436 of HIRAIZUMI, SANDLER and CROW 1960). If, as conventionally symbolized in the literature on meiotic drive, K represents the proportion of **A** alleles among the successful gametes of **Aa** heterozygotes, the following analogy exists between K and the net gene conversion constant: $4K - 2 = k$. By using this relationship, the equations derived for the interaction of meiotic drive and selection (HIRAIZUMI, SANDLER and CROW 1960) could also be applied to gene conversion and selection.

DISCUSSION

Our theoretical considerations show that gene conversion has an important effect on the allele frequencies in a large random-mating population if $|k| \geqslant 4|\mu - \nu|$. Even if k is of the same order of magnitude as the mutation rates, the influence of conversion is still significant. The question arises whether such k values exist in natural populations. For the discussion of this question, we will make the reasonable assumptions that $\mu \leqslant 10^{-5}$ and $\mu > \nu$.

The net gene conversion constant k is defined by f, i, and j. The values of these parameters vary with the organism, the locus, and the specific mutation studied. In Ascomycetes, the frequency of gene conversion per meiosis (f) ranges from 10^{-1} to 10^{-3} (see the references cited at the beginning of this paper). For rosy mutants of *Drosophila melanogaster*, $f \approx 10^{-5}$ has been reported by CHOVNICK, BALLANTYNE and HOLM (1971). As to the relative frequencies of 3**A** : 1**a** segregations (i) and 1**A** : 3**a** segregations (j), $i \approx j$ seems to be the general rule in *Saccharomyces cerevisiae* (FOGEL, HURST and MORTIMER 1971). In other fungi, for example in *Schizosaccharomyces pombe* (GUTZ 1971) and in *Ascobolus*

immersus (LEBLON 1972), mutants are known for which i and j differ greatly ($i > j$ or $i < j$).

An extreme, but illustrative, example is the *ade6* mutant *M26* of *Schiz. pombe*. From different crosses performed with *M26* (GUTZ 1971), the following average values for f, i, and j result: $f = 0.04$, $i = 0.923$, and $j = 0.077$. For these values, equation (1) gives $k = 3.38 \times 10^{-2}$. If a mutant allele such as *M26* were present in a population, if $p_0 = q_0 = 0.5$, and assuming that there were no mutations, then the frequency of the wild-type allele (**A**) would be 0.5042 in the first generation, 0.5542 in the tenth generation, and 0.8445 in the hundreth generation (equation 4).

From the published data on gene conversion, it appears that $f \approx 10^{-3}$ (or even $f > 10^{-3}$) is frequently observed. Less information is available with regard to the question whether the relationship $i \neq j$ is also frequently true. Those mutants for which large differences between i and j have been reported may be exceptional. We ask therefore: if $f = 10^{-3}$ is assumed, how great a difference must exist between i and j to result in $k = 3 \times 10^{-5}$, a value which would cause a significant change in the population? For the above values, equation (1) gives $i = 0.515$ and $j = 0.485$, or $i = 1.062\, j$.

Thus, if $f \approx 10^{-3}$, only a relatively small difference between i and j would be sufficient to cause a significant effect in the population. It should be noted that in order to verify $i \neq j$ (if $i = 1.062\, j$) in experiments with an ascomycete, at least 4267 *conversion tetrads* would be necessary for a 95% level of confidence, and 7367 *conversion tetrads* for a 99% level of confidence.

From the above arguments we conclude that, in addition to mutation, gene conversion also has a significant influence on the allele frequencies in natural populations.

In our theoretical equations, it was not possible to consider that the gene conversion parameters may vary for nonidentical alleles. The k values are only valid for one specific pair of alleles (e.g., for the wild type and one particular mutant). Due to recurrent mutation pressure, alleles with different conversion characteristics may be present in a natural population.

We thank DRS. DIMITRIJ LANG and JAMES H. MEADE for critical readings of the manuscript.

LITERATURE CITED

BALLANTYNE, G. H. and A. CHOVNICK, 1971 Gene conversion in higher organisms: Non-reciprocal recombination events at the rosy cistron in *Drosophila melanogaster*. Genet. Res. **17:** 139–149.

CHOVNICK, A., 1973 Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. Genetics **75:** 123–131.

CHOVNICK, A., G. H. BALLANTYNE and D. G. HOLM, 1971 Studies on gene conversion and its relationship to linked exchange in *Drosophila melanogaster*. Genetics **69:** 179–209.

CROW, J. F. and M. KIMURA, 1970 *An Introduction to Population Genetics Theory*. Harper and Row, New York, Evanston and London.

FINCHAM, J. R. S. and P. R. DAY, 1971 *Fungal Genetics*. Third edition. Blackwell Scientific Publications, Oxford and Edinburgh.

FOGEL, S., D. D. HURST and R. K. MORTIMER, 1971 Gene conversion in unselected tetrads from multipoint crosses. pp. 89–110. In: *Stadler Genetics Symposia,* Vol. 1 and 2. Edited by G. KIMBER and G. P. RÉDEI, Columbia, Missouri.

GUTZ, H., 1971 Site specific induction of gene conversion in *Schizosaccharomyces pombe.* Genetics **69:** 317–337.

HIRAIZUMI, Y., L. SANDLER and J. F. CROW, 1960 Meiotic drive in natural populations of *Drosophila melanogaster.* III. Populational implications of the segregation-distorter locus. Evolution **14:** 433–444.

LEBLON, G., 1972 Mechanism of gene conversion in *Ascobolus immersus* I. Existence of a correlation between the origin of mutants induced by different mutagens and their conversion spectrum. Molec. Gen. Genetics **115:** 36–48.

LI, C. C., 1958 *Population Genetics.* Second impression. University of Chicago Press, Chicago.

NELSON, O. E., 1975 The waxy locus in maize. III. Effect of structural heterozygosity on intragenic recombination and flanking marker assortment. Genetics **79:** 31–44.

Corresponding editor: J. F. CROW