

On the estimation of probabilities of segregation from a bacterium carrying two incompatible plasmids

By S. BAUMBERG

Department of Genetics, University of Leeds, Leeds LS2 9JT

(Received 2 July 1975)

SUMMARY

Equations are presented relating the number of generations elapsed to the fractions of bacteria carrying one or both plasmids in a clone derived from a cell containing two incompatible plasmids.

1. INTRODUCTION

Two bacterial plasmids are said to be incompatible if the inheritance of one is perturbed by the presence of the other in the same cell. A common finding is as follows. Strains A and B, carrying conjugative plasmids P and Q respectively, are mated and selection is made for transconjugants, with the chromosomal markers of A, that carry both plasmids P and Q. Cells from a colony derived initially from such an A.P⁺Q⁺ transconjugant are then allowed to grow in a non-selective medium. The segregation in this case of A.P⁺Q⁻ 'singles' at a frequency greater than that of A.Q⁻ cells in an A.Q⁺ culture, or of A.P⁻Q⁺ 'singles' at a frequency greater than that of A.P⁻ cells in an A.P⁺ culture, is taken as an indication that P and Q are incompatible.

The degree of incompatibility between plasmids can vary widely. For instance, R factors of the FII incompatibility group may segregate only slowly from 'doubles' (i.e. cells carrying both plasmids) (e.g. Uhlin & Nordström, 1975). However, those of the I α group show almost complete incompatibility; often, in an experiment such as that described above, no transconjugant 'doubles' can be obtained, even where entry exclusion is demonstrably less than 100% efficient as shown by the formation of transconjugants containing only one plasmid (Hedges & Datta, 1973). It is therefore perhaps surprising that quantitative estimates of frequencies of segregation from 'doubles', as measures of their degree of incompatibility, have rarely been made. One reason for this may be that the mathematical treatment of segregation from 'doubles' appears not to have been published. This paper attempts to provide such a treatment for the simplest possible case.

2. MODEL AND DERIVATION OF EQUATIONS

Consider a population consisting of a mixture of 'doubles' A.P⁺Q⁺ and 'singles' A.P⁺Q⁻ and A.P⁻Q⁺ derived as above from a single transconjugant 'double'

cell. Let us assume that there is a probability p per division that $A.P+Q^+$ will on dividing produce $A.P+Q^++A.P-Q^+$ (instead of $A.P+Q^++A.P+Q^+$), and a probability q per division that it will produce $A.P+Q^++A.P+Q^-$. The following assumptions will be made: (a) $A.P+Q^+$, $A.P+Q^-$ and $A.P-Q^+$ have identical generation time τ ; (b) segregation of completely plasmid-less cells $A.P-Q^-$ can be neglected; (c) at time zero, the three cell types are all sufficiently numerous that stochastic phenomena can be ignored; and (d) re-infection cannot occur – this may be achieved, for instance, by keeping cell density very low through periodic dilution or continuous culture, or by allowing growth in the presence of a low level of sodium dodecyl sulphate (Achtman & Helmuth, 1975). If the bacterial host is rec^+ , a further complication is the possibility of recombination between the incompatible plasmids in a ‘double’. Nevertheless, the segregation frequencies will still be useful parameters; and in any case this problem can be avoided by use of a $recA$ host.

Suppose that, after time t , the numbers of $A.P+Q^+$, $A.P-Q^+$ and $A.P+Q^-$ are X , Y and Z respectively, these having increased from X_0 , Y_0 and Z_0 at time zero; also that the corresponding total population sizes are N and N_0 , where

$$N = X + Y + Z, \quad N_0 = X_0 + Y_0 + Z_0; \quad \text{also} \quad N = N_0 e^{kt},$$

where $k = \ln 2/\tau$, and \ln stands for the natural logarithm.

Consider the infinitesimal time interval from t to $t + dt$. In this interval the Y and Z ‘singles’ of types $A.P-Q^+$ and $A.P+Q^-$ undergo $kY \cdot dt$ and $kZ \cdot dt$ divisions, respectively, giving rise to these numbers of new cells of the two types. Meanwhile the ‘doubles’ will undergo $kX \cdot dt$ divisions, of which fractions p and q will produce $A.P-Q^+$ and $A.P+Q^-$ as the only new cell. Hence the increases in the numbers of the three cell types in the interval dt will be

$$dX = kX(1 - p - q) dt, \tag{1}$$

$$dY = k(Y + pX) dt, \tag{2}$$

$$dZ = k(Z + qX) dt. \tag{3}$$

Equation (1) leads immediately to

$$X = X_0 e^{(1-p-q)kt}, \tag{4}$$

and substituting (4) in (2) and integrating, given the condition that $X = X_0$ when $t = 0$,

$$Y = Y_0 e^{kt} + \frac{pX_0}{p+q} \left\{ e^{kt} - e^{(1-p-q)kt} \right\}$$

which can be written as

$$Y = \left\{ Y_0 + \frac{pX_0}{p+q} \right\} e^{kt} - \frac{pX_0}{p+q}. \tag{5}$$

Replacing Y and Y_0 by Z and Z_0 and interchanging p and q in (5) gives a similar equation for Z .

If $p = q$, as might occur with two marked variants of the same plasmid, then (4) and (5) reduce to $X = X_0 e^{(1-2p)kt}$ and $Y = (Y_0 + \frac{1}{2}X_0) e^{kt} - \frac{1}{2}X_0$.

To estimate p and q experimentally, we need formulae for the observable quantities X/N and Y/N . Dividing (4) and (5) by $N_0 e^{kt}$ we obtain

$$\frac{X}{N} = \frac{X_0}{N_0} e^{-(p+q)kt}, \tag{6}$$

$$\frac{Y}{N} - \frac{Y_0}{N_0} = \frac{p}{p+q} \left(\frac{X_0}{N_0} - \frac{X}{N} \right). \tag{7}$$

Given a set of experimental determinations of N , X/N and Y/N for a series of values of t , we can estimate p , q and τ as follows:

Plot $\ln N$ against t : the slope of this line is k .

Plot $\ln (X/N)$ against t : the slope of this line is $-(p+q)k$.

Plot Y/N against X/N : the slope of this line is $p/(p+q)$.

From these estimates p and q can be calculated, and τ can be estimated as $\ln 2/k$.

The three plots referred to above should give straight lines if the assumptions of our model are correct, but assumption (a) in particular would be wrong if doubles divide more slowly than singles. This might occur for reasons unconnected with the compatibility relationships of the plasmids, and segregation would then follow, mimicking incompatibility. In this case, let τ_1 and τ_2 be the generation times for doubles and singles, and write k_1 and k_2 for $\ln 2/\tau_1$, and $\ln 2/\tau_2$.

Let $(k_2 - k_1)/k_2 = \alpha$. Then equations (1) and (2) become

$$dX = k_1(1-p-q)X \cdot dt, \quad dY = k_2Y - pk_1X \cdot dt$$

with solutions

$$X = X_0 e^{(1-p-q)k_1t}, \tag{8}$$

$$Y - \frac{pX}{p+q+\alpha} = \left(Y_0 + \frac{pX_0}{p+q+\alpha} \right) e^{k_2t}, \tag{9}$$

$$Z - \frac{qX}{p+q+\alpha} = \left(Z_0 + \frac{qX_0}{p+q+\alpha} \right) e^{k_2t}. \tag{10}$$

Dividing by $N = X + Y + Z$ no longer gives simple equations from which p , q , etc., can be estimated by linear regressions, but it would be possible to obtain estimates of all the parameters from a set of experimental determinations of X , Y and Z for a series of values of t in a growing culture. These data could then be subjected to a computer numerical optimization procedure to solve equations (8)–(10).

Alternatively one could plot $\ln X$ against t to estimate $(1-p-q)k_1$ as the slope of the regression line obtained, and try out a series of values of the fractions $p/(p+q+\alpha)$ and $q/(p+q+\alpha)$ to find those which gave the closest approximations to straight lines when the logarithms of the left hand sides of (9) and (10) are plotted against t . This would provide estimates of the two fractions and of k_2 , and all the parameters could then be evaluated.

I wish to thank Dr E. C. R. Reeve for helpful suggestions.

REFERENCES

- ACHTMAN, M. & HELMUTH, R. (1975). The F factor carries an operon of more than 15×10^6 daltons coding for deoxyribonucleic acid transfer and surface exclusion. In *Microbiology - 1974* (ed. D. Schlessinger). Washington, D.C.: American Society for Microbiology.
- HEDGES, R. W. & DATTA, N. (1973). Plasmids determining I pili constitute a compatibility complex. *Journal of General Microbiology* **77**, 19-25.
- UHLIN, B. E. & NORDSTRÖM, K. (1975). Plasmid incompatibility and control of replication: Copy mutants of the R-factor R1 in *Escherichia coli* K12. *Journal of Bacteriology* **124**, 641-649.