J. theor. Biol. (1975) **50**, 87–105

The Control of Foot Formation in Transplantation Experiments with *Hydra viridis*

JOEL E. COHEN†

Harvard University, The Biological Laboratories,
16 *Divinity Avenue, Cambridge, Ma* 02138, *U.S.A.*

AND

HARRY K. MACWILLIAMS

Max-Planck-Institut für Virusforschung, 74 Tübingen,
Federal Republic of Germany

(*Received* 27 *March* 1974)

The results of three groups of experiments on the control of foot (basal disk) differentiation in *Hydra viridis* are well predicted by mathematical models of remarkable simplicity. In each of the experiments, a group of cells, called the *actor*, may or may not form a foot. The probability of foot formation is influenced by actor-specific properties and by properties of the rest of the animal, which may be called the *setting*.

A natural interpretation of the mathematical models is that the actor forms a foot or not according to the balance of an inhibition of foot formation determined by the setting, and a critical inhibition level or "threshold", an assumed actor characteristic. Threshold and inhibition are assumed to be normal random variables, the means of which depend on the parameters of the experiment. A foot is assumed to form in an individual case if the threshold is greater than the inhibition.

When transformed to a probit or logit scale, as described in the text, the mean intensity of inhibition appears to increase linearly with the quantity of foot tissue and to decrease linearly with increasing distance from the foot end of the hydra's body axis. The mean level of the threshold appears to increase linearly with time during which actor tissue is removed from inhibitory influences, but varies nonlinearly with position within the hydra.

The inhibition-threshold models are tested here through statistically more tractable but numerically indistinguishable models of different mathematical form. If interpreted directly these models suggested that the odds that a foot will form are the product of an actor-determined scalar and a setting-determined scalar. The logarithms of these scalars are linear functions of the experimental parameters.

† Present address: The Rockefeller University, New York, N.Y. 10021, U.S.A.

1. Introduction

A hydra holds fast to its substrate by means of its foot (basal disk). When the basal disk of *Hydra viridis* is cut off, the animal regenerates a new disk. The dependence of foot formation on several manipulable parameters has been investigated (MacWilliams & Kafatos, 1968; MacWilliams, Kafatos & Bossert, 1970; MacWilliams, 1972). The designs of the experiments vary. However, in each case one can identify an *actor*, a group of cells which may form a basal disk or may not, and a *setting*, a group of cells whose properties (varying independently, by experimental design, to those of the actor) may influence the probability of disk formation.

The data from these experiments have been shown (MacWilliams, Kafatos and Bossert, 1970; MacWilliams, 1972) to be consistent with a simple model in which foot formation is controlled by two factors, a continuously variable *inhibition* of disk formation determined by the setting, and a critical inhibition level, or *threshold* to inhibition, characteristic of the actor. In this model, the inhibition must exceed the threshold level in order to prevent the actor from forming a disk. For reasons which will become clearer later, we shall refer to this model as a probit model.

The computational procedure used to test probit models by MacWilliams *et al.* (1970) has the disadvantage that it does not necessarily accept all models which can account for the data. While this shortcoming does not weaken the conclusion that a particular probit model is acceptable as a description of the data, it does make it impossible to buttress the case for a particular model by rejecting some possible alternative formulations.

In this paper we solve this problem. We first introduce a general probit model of the control of foot formation in transplantation experiments, under which a variety of specific probit models may be naturally formulated. These models are well approximated numerically by testable models of a different mathematical form, which we shall call logit models. The logit models, which we shall describe more fully, are hierarchical log-linear models recently developed for the analysis of multidimensional, possibly incomplete, contingency tables. We test various logit models, including models that explicitly incorporate certain linear relationships between the values of the experimental parameters and the model variables, suggested by the probit models. The results make it possible to argue that certain of the logit models, and the empirically indistinguishable probit formulations, are the simplest acceptable interpretations of the experimental data.

The probit models used in this paper may be very briefly characterized as follows. Given a number of experimental variables (such as time, quantity, or location), assume there is some function Z of these variables such that a

hydra forms a foot if its $Z > 0$ and fails to form a foot otherwise. All random effects in the experiments are summarized by assuming that the probability of foot formation in repeated experiments with the same values of the experimental variables is $\Phi(Z)$, where $\Phi(.)$ is the cumulative distribution function of a standardized normal distribution (mean zero, variance one). The probability of foot formation increases sigmoidally with Z. This paper reports a search for the simplest acceptable functions Z. These functions turn out to be linear [equation (11)] or at worst cubic [equation (13)] functions of the experimental variables. Such simplicity makes possible a simple, still entirely speculative, physical interpretation of the models.

The following two sections describe the mathematics of the probit and logit models in greater technical detail. The section which describes "specific models and tests" gives a concrete account of the experiments and results of analysis.

2. The General Probit Model

We assume that a number of different things influence the actor's decision to form a foot or not. Each influence is modeled by a scalar variable Y_i. The scalars are assumed to be summed, and a foot formed if and only if the sum exceeds a certain critical value K, which is a characteristic of the experimental system. Thus in an individual experiment, given:

$$Z = -K + Y_1 + \ldots + Y_n, \tag{1}$$

a foot is formed if and only if Z is positive.

In systems with only two influences Y_1 and Y_2 a straightforward inhibition-threshold interpretation is possible. The criterion for foot formation may be rewritten:

$$Z = Y_1 - (K - Y_2) > 0.$$

A foot is now formed if and only if $Y_1 > (K - Y_2)$. Thus Y_1 may be regarded as the threshold and $(K - Y_2)$ as the inhibition. The choice of Y_1 or Y_2 as the threshold-modeling variable is of course arbitrary.

Each variable Y_i is associated with one or a combination of the experimental parameters. Each of the specific models we will propose and test is defined by the set of parameters and parameter combinations represented by its variables.

The stochastic nature of the experimental system is modeled by assuming that the Y_i vary from repetition to repetition of a single experiment. The variation of each Y_i is assumed to be normal. Hence Z is also a normally distributed random variable. Since a foot is assumed to form if Z exceeds zero, we have exactly the model of the dose–response relation adopted in

probit analysis (Finney, 1964). Hence we call models of the form of equation (1) probit models.

If the absolute value of the mean of Z is less than a few times its standard deviation, Z will be both positive and negative with non-negligible probability. For simplicity, we assume that within any data set of given experimental design, the standard deviation σ is constant. Since the frequency of foot formation depends only on the ratio of the mean to the standard deviation of Z, the model loses no generality if we set $\sigma = 1$. Then for any given mean of Z, the model predicts that the probability of foot formation

TABLE 1

Frequency of foot formation and predictions of a logit model in one sort of transplantation experiment

Condition of grafted basal disk (S)	Cut-regraft interval in minutes (T) 0	15	35	60	91	130	179
Intact							
Observed	2	2	4	7	12	9	9
Predicted	2·8	2·9	4·6	6·0	11·5	8·1	9·1
Out of	10	8	8	12	17	9	10
Excised							
Observed	9	7	11	11	12	12	10
Predicted	8·2	6·1	10·4	12·0	12·5	12·9	9·9
Out of	10	7	11	13	13	13	10

Model	d.f.	χ^2	G^2
1 (F) (T, S)	13	54·22	55·82
2 (F, T) (T, S)	7	31·70	37·76
3 (F, S) (T, S)	12	25·68	30·85
4† (F, T) (F, S) (T, S)	6	12·78	10·97

The bottom half of the body stalk (peduncle), including the foot, of a hydra was removed and replaced by a similar piece from another hydra after a varying interval (T). In the first series, the regrafted fragment included the basal disk; in the second, the graft was a bottom half-peduncle from which the entire basal disk had been excised. For each combination of the cut–regraft interval and the condition of the graft's basal disk several experiments were performed. The host formed a basal disk of its own in some but not all cases. Observed frequencies of host's foot formation, previously unpublished, are identical to those in MacWilliams & Kafatos (1968). Predicted frequencies are from model 4. "Out of" is the number of experiments performed; d.f. = degrees of freedom; χ^2 = Pearson's chi-square; G^2 = likelihood chi-square. For model 4, marked †, the probability (calculated from G^2) of a worse fit by chance between observed and expected frequencies is greater than 0·05, an acceptable fit; for the remaining models, the probability of a worse fit is less than 0·005, signifying a poor fit.

is the area to the right of zero under a normal probability density curve with the given mean and standard deviation of 1.

The effects of the experimental parameters on foot formation are modeled by allowing them to determine the mean values of the variables Y_i. No assumptions are made about the form of this dependence; instead, we assume an independent value for each variable Y_i for each possible value or value combination of the associated experimental parameter(s). For example, in model 4 for Table 2, a variable Y_S associated with basal disk

TABLE 2

Frequency of foot formation and predictions of logit models in a more elaborate transplantation experiment of the same sort described in Table 1

Grafted disk size (S)	Cut–regraft interval in minutes (T) 0	40	80	120	160
No disk					
Observed	8	9	13	15	15
Model 4	7·7	8·9	13·1	14·1	16·1
Model 4a	7·1	9·7	11·9	13·4	16·3
Out of	25	23	21	19	20
Quarter disk					
Observed	3	3	9	11	14
Model 4	3·6	4·4	8·5	11·4	12·1
Model 4a	3·7	5·5	7·9	11·1	13·0
Out of	25	23	22	22	20
Half disk					
Observed	2	4	4	5	7
Model 4	1·4	2·2	4·3	6·8	7·3
Model 4(a)	1·4	2·7	3·9	6·4	8·0
Out of	21	23	20	21	18
Whole disk					
Observed	0	0	1	3	2
Model 4	0·3	0·5	1·0	1·7	2·4
Model 4(a)	0·3	0·6	0·9	1·5	2·6
Out of	23	24	20	20	20

	Model	d.f.	χ^2	G^2
1	(F) (T, S)	19	120·30	131·25
2	(F, T) (T, S)	15	83·32	92·62
3	(F, S) (T, S)	16	51·70	55·06
4†	(F, T) (F, S) (T, S)	12	6·68	7·13
4(a)†	(S linear) (T linear) (T, S)	17	7·94	8·65

Two additional classes of regrafted fragment were introduced, bearing half and quarter basal disks. Observations are from MacWilliams *et al.* (1970); predictions are from models 4 and 4(a). For models 4 and 4(a) [equation (11)], marked †, the probability of a worse fit is greater than 0·1, signifying a good fit; for the remaining models, the probability is less than 0·001, signifying a poor fit.

TABLE 3

Foot formation frequencies and logit model predictions in a second sort of transplantation experiment

Final transplant position (P)		Initial transplant position (D) 7	6	5	4	3	2	1
		Host's basal disk intact ($H = 1$)						
	Observed	1	8	22	23	—	—	—
7	Model 4	0·8	8·8	20·7	23·4	23·7	24·4	24·9
	Model 4(h)	1·1	13·6	21·8	23·6	24·2	24·7	24·9
	Observed	—	5	16	21	24	—	—
6	Model 4	0·4	4·9	17·2	21·6	22·2	23·7	24·8
	Model 4(h)	0·3	5·7	15·7	20·1	22·0	23·7	24·8
	Observed	—	—	8	18	18	19	—
5	Model 4	0·1	1·6	9·3	15·8	17·2	20·7	24·4
	Model 4(h)	0·1	1·7	7·4	12·5	16·1	20·5	24·2
	Observed	—	—	—	5	4	15	24
4	Model 4	0·0	0·3	2·6	6·3	7·5	12·0	22·2
	Model 4(h)	0·0	0·4	2·3	4·9	7·7	13·2	21·9
	Observed	—	—	—	—	1	5	19
3	Model 4	0·0	0·1	0·8	2·1	2·7	5·1	17·3
	Model 4(h)	0·0	0·1	0·6	1·4	2·5	5·4	15·9
	Observed	—	—	—	—	—	0	4
2	Model 4	0·0	0·0	0·1	0·2	0·2	0·5	3·5
	Model 4(h)	0·0	0·0	0·2	0·4	0·7	1·6	7·5
	Observed	—	—	—	—	—	—	4
1	Model 4	0·0	0·0	0·1	0·2	0·3	0·5	4·0
	Model 4(h)	0·0	0·0	0·0	0·1	0·2	0·4	2·4
		Host's basal disk excised ($H = -1$)						
	Observed	2	16	23	—	—	—	—
7	Model 4	2·2	15·3	23·3	24·4	24·5	24·8	25·0
	Model 4(h)	1·4	15·0	22·4	23·9	24·4	24·7	25·0
	Observed	—	9	21	24	—	—	—
6	Model 4	0·9	9·4	21·1	23·5	23·8	24·4	24·9
	Model 4(h)	0·7	10·8	20·4	22·8	23·8	24·5	24·9
	Observed	—	—	19	21	22	—	—
5	Model 4	0·4	5·3	17·7	21·9	22·5	23·8	24·9
	Model 4(h)	0·4	7·0	17·3	21·1	22·7	24·0	24·8
	Observed	—	—	—	20	22	21	—
4	Model 4	0·2	3·1	14·0	19·7	20·6	22·8	24·7
	Model 4(h)	0·2	4·1	13·3	18·3	20·8	23·1	24·7
	Observed	—	—	—	—	16	19	22
3	Model 4	0·1	1·0	6·7	12·8	14·3	18·6	24·0
	Model 4(h)	0·1	2·3	9·1	14·5	17·8	21·5	24·4

TABLE 3 (*continued*)

Final transplant position (*P*)		Initial transplant position (*D*) 7	6	5	4	3	2	1
		Host's basal disk excised ($H = 1$)—*continued*						
	Observed	—	—	—	—	—	20	23
2	Model 4	0·1	1·0	7·0	13·2	14·7	18·9	24·1
	Model 4(h)	0·0	1·2	5·6	10·3	13·9	19·0	23·8
	Observed	—	—	—	—	—	—	24
1	Model 4	0·1	0·9	6·4	12·5	14·0	18·4	24·0
	Model 4(h)	0·0	0·6	3·2	6·5	9·8	15·4	22·7

		d.f.	χ^2	G^2
1	(F, D) (F, P) (F, H) (D, P, H)	26	62·87	68·06
2	(F, D, H) (F, P) (D, P, H)	20	50·50	48·46
3	(F, D, P) (F, H) (D, P, H)	17	53·37	56·75
4†	(F, P, H) (F, D) (D, P, H)	20	21·87	21·61
4(a)	(F, P, H) (D linear) (D, P, H)	25	88·95	85·32
4(b)†	(P lin.; P, H) (F, D) (D, P, H)	25	29·37	27·47
4(c)†	(P, H lin.; P) (F, D) (D, P, H)	25	33·98	32·40
4(d)	(P lin.; P, H) (D lin.) (D, P, H)	30	93·83	94·57
4(e)	(P, H lin.; P) (D lin.) (D, P, H)	30	93·07	87·11
4(f)†	(P lin.; P, H lin.) (F, D) (D, P, H)	30	40·02	37·24
4(g)	(P lin.; P, H lin.) (D lin.) (D, P, H)	35	99·65	98·39
4(h)†	(P lin.; P, H lin.) (D cubic) (D, P, H)	33	45·96	44·49
5	(F, D, H) (F, D, P) (D, P, H)	11	45·10	35·92
6†	(F, D, P) (F, P, H) (D, P, H)	11	16·48	15·53
7†	(F, D, H) (F, P, H) (D, P, H)	14	10·29	10·77
8†	(F, D, P) (F, D, H) (F, P, H) (D, P, H)	5	3·46	3·70

The eight apico-basal eighths of a hydra's body column were numbered from 0 at the basal end to 7 at the head. The seven planes between the pairs of body annuli were numbered from foot to head as 1–7. Annuli corresponding to the numbered eighths were removed from a donor hydra and grafted laterally into a wound made at one of the planes of section in a host, the basal disk of which was removed in some experiments. The annuli were scored after two days for foot formation. Observed frequencies of foot formation for various combinations of original transplant position (D), final position of transplant (P), and host condition (H) are from MacWilliams (1972). "—" means no experiments were performed; otherwise $n = 25$. Predictions are from model 4, equation (12), and model 4(h), equation (13). †, as in Table 2.

size S has four mutually unconstrained means, one for each of the disk sizes (0, 1/4, 1/2, 1) used in the experiment. In model 5 for Table 3, a variable which depends upon the combination of the transplant's original position (seven values) and the condition of the host basal disk (two values) has 14 means, which may be arbitrarily related to one another.

Our models do not specify *a priori* numerical values for the variables Y_i and K. Rather, the models may be considered hypotheses that satisfactory

numerical values exist. Given a specific model and any set of numerical values for the means of Y_i and K, predicted foot formation frequencies may be calculated. The ability of a particular model to account for the experimentally observed frequencies of foot formation is determined by seeking the set of numerical values for which the corresponding predicted frequencies best fit the actual frequencies of foot formation. If the difference between the best possible prediction and the actual results, as judged by statistical measures, can be reasonably attributed to sampling error, the model is considered acceptable.

3. The Log-linear Approximation: Logit Models

To test particular cases of the general probit model, we use the computationally well-understood hierarchical log-linear models recently developed for the analysis of multidimensional contingency tables (Bishop, 1967, 1969; Goodman, 1969, 1970, 1971*a*,*b*, 1972; Mantel, 1970; Fienberg, 1970, 1972). Our data may be viewed as contingency tables: in each data set the effects of several parameters on the frequency of foot formation are reported; the corresponding contingency tables have one dimension for each of the experimentally manipulated parameters and a further dimension for the outcome, this last with two possible values, foot or no foot. A log-linear model specifies that the logarithm of the number of experiments falling into each cell of this table is given by the sum of a set of scalar variables, each associated with one or a combination of the dimensions of the contingency table.

In the data of Table 2, for example, two experimental parameters are manipulated: the size of the basal disk of the setting, and the disk regeneration time in the actor. If F denotes the contingency table dimension corresponding to the result (foot formation, $F = 1$; or not, $F = 0$), S denotes the dimension corresponding to basal disk size, and T denotes the regeneration-time dimension, each cell in the contingency table may be specified by a subscript of the form ijk, where $i = 0, 1$ for the variable F; $j = 1, 2, \ldots, 5$ for the variable T; and $k = 1, 2, 3, 4$ for the variable S. Let X_{ijk} be the expected frequency for cell ijk. (All logarithms in this paper are to the base e.) The general log-linear model for this table is then

$$\ln X_{ijk} = \theta + \lambda_i^F + \lambda_j^T + \lambda_k^S + \lambda_{ij}^{FT} + \lambda_{ik}^{FS} + \lambda_{jk}^{TS} + \lambda_{ijk}^{FTS} \tag{2}$$

subject to the constraints

$$\begin{gathered} \sum_i \lambda_i^F = \sum_j \lambda_j^T = \sum_k \lambda_k^S = 0, \\ \sum_i \lambda_{ij}^{FT} = \sum_j \lambda_{ij}^{FT} = \sum_i \lambda_{ik}^{FS} = \sum_k \lambda_{ik}^{FS} = \sum_j \lambda_{jk}^{TS} = \sum_k \lambda_{jk}^{TS} = 0, \\ \sum_i \lambda_{ijk}^{FTS} = \sum_j \lambda_{ijk}^{FTS} = \sum_k \lambda_{ijk}^{FTS} = 0, \end{gathered} \tag{3}$$

where the superscripts indicate the experimental parameters, or combinations of experimental parameters, associated with each of the λ variables.

The odds of foot formation (defined as the ratio of expected numbers of cases of foot formation to expected numbers of cases of no foot formation) for any combination of the experimental parameters are given in the general log-linear model, as a consequence of equation (2), by

$$X_{1jk}/X_{0jk} = \exp\left[(\lambda_1^F - \lambda_0^F) + (\lambda_{1j}^{FT} - \lambda_{0j}^{FT}) + (\lambda_{1k}^{FS} - \lambda_{0k}^{FS}) + (\lambda_{1jk}^{FTS} - \lambda_{0jk}^{FTS})\right]. \quad (4)$$

The constraints in equation (3) on the λ parameters reduce equation (4) to

$$X_{1jk}/X_{0jk} = \exp\left(2\lambda_1^F + 2\lambda_{1k}^{FS} + 2\lambda_{1j}^{FT} + 2\lambda_{1jk}^{FTS}\right). \quad (5)$$

The frequency of foot formation f_{jk} = (odds/(odds + 1)) is then predicted by equation (5) to be

$$f_{jk} = \left[1 + \exp\left(-2\lambda_1^F - 2\lambda_{1j}^{FT} - 2\lambda_{1k}^{FS} - 2\lambda_{1jk}^{FTS}\right)\right]^{-1}. \quad (6)$$

This is an equation for the logistic curve

$$y = \left[1 + \exp\left(-2x\right)\right]^{-1} \quad (7)$$

and may be written in the simplified form

$$\text{logit}\,(f_{jk}) = \lambda_1^F + \lambda_{1j}^{FT} + \lambda_{1k}^{FS} + \lambda_{1jk}^{FTS} \quad (8)$$

where

$$\text{logit}\; y = (1/2)\ln\frac{y}{1-y}.$$

Because of the appearance of the logit function in equation (8) we shall refer to the particular cases of the general log-linear model which follow as logit models.

It is well known (Berkson, 1951) that when appropriate scales are chosen the logistic curve and the cumulative normal distribution are very similar, differing in value by a maximum of less than 1·5%. The dependence of the frequency of foot formation on the λ variables in logit models is thus indistinguishable from the dependence of the frequency of foot formation on the variables Y_i in probit models, equation (1), in experiments involving small sample sizes.

Specific logit models set various of the λ variables in the general model, equation (2), to zero. There is a logit model corresponding to each probit model. For instance, the model

$$\ln X_{ijk} = \theta + \lambda_i^F + \lambda_j^T + \lambda_k^S + \lambda_{ij}^{FT} + \lambda_{jk}^{TS} \quad (9)$$

which predicts foot formation frequencies f_{jk} which satisfy

$$\text{logit}\,(f_{jk}) = \lambda_1^F + \lambda_{1j}^{FT} \quad (10)$$

is isomorphic to a probit model which assumes that only the regeneration time (variable T, subscript j) influences foot formation. The variable λ_1^F

plays the role that K plays in equation (1). The λ variables without superscript F, whose values do not depend on F, are involved only in modeling the experimental design.

For each interesting probit model, we have tested the corresponding logit model. The logit models are referred to in the tables by their abbreviated list of parameters (Bishop, 1967; Goodman, 1970). Each parenthesized group of letters in an abbreviated list of parameters specifies a margin of the contingency table which the corresponding margin of the expected frequency table is required to fit exactly. A margin in a multidimensional contingency table is a sum of cell frequencies over one or more of the table's dimensions.

For each logit model, an iterative proportional fitting procedure finds the maximum likelihood estimate of the expected frequency corresponding to each observed frequency. One of us (J.E.C.), who carried out the numerical computations, used a slightly modified version of a computer program written by Bishop for the iterative fitting procedure. Additional computer programs were written to solve the estimated frequencies for the values of the λ variables, and to obtain the estimated frequencies when the λ variables were assumed subject to linear or cubic constraints.

The goodness of fit of the expected to observed frequencies is measured by Pearson's classical χ^2 and by the log-likelihood ratio G^2, both of which have the distribution of χ^2. The number of degrees of freedom is the difference between the number of measured frequencies and the number of independent scalars estimated from the data. It should be noted that in incomplete contingency tables mutual dependencies may exist among the non-zero λ variables beyond those implied by equation (3). When working with log-linear models, the degrees of freedom are customarily calculated by subtracting a fixed number for each fitted margin from the total number of contingency table cells. The number of degrees of freedom to be subtracted for each margin is a routine calculation except in Table 3, where the margin (D, P) has nine degrees of freedom, (F, D, P) has nine, and (D, P, H) has five.

4. Specific Models and Tests

(A) TABLES 1 AND 2

In the experiments reported in Tables 1 and 2 the foot and the basal half of the body stalk (peduncle) were removed from a hydra at time zero, initiating foot regeneration. The regenerating foot is considered the actor in this experiment. After a variable regeneration time a new basal half peduncle bearing a foot, a part of a foot, or no foot was grafted to the original animal. Foot formation by the actor was assayed. In these experi-

ments the actor is characterized by the regeneration time prior to grafting while the only other experimental variable, the size of the foot attached to the graft, determines the properties of the setting.

The models for Tables 1 and 2 are identical. Model 1 is the hypothesis that neither the regeneration time T nor the size of the transplanted foot S influences foot formation in the actor. The frequency of foot formation is therefore predicted to be the same for all values of the experimental parameters. The large values of χ^2 and G^2 relative to the degrees of freedom show that this hypothesis is unacceptable. Model 2 is the hypothesis that foot formation is influenced by the regeneration time T but not by the foot size S. Model 3 assumes an influence of S but not of T. Both models 2 and 3 are unacceptable as descriptions of the data.

Model 4 assumes two variables, one depending on the regeneration time T and one on the transplanted foot size S. No variable influencing foot formation which depends on the combination of T and S is assumed. Thus this model is given in logit form by equation (8) with $\lambda^{FST} \equiv 0$. The frequencies of foot formation expected according to this logit model are given under the observed frequencies in tables 1 and 2. The fit is acceptable in Table 1 and excellent in Table 2, in which the sample size is twice as large.

Model 4 in probit form is equivalent to an inhibition-threshold model proposed for the data of Table 2 by MacWilliams *et al.* (1970). The variable dependent on S (the setting property) corresponds to these authors' inhibition intensity. The T-dependent (actor) variable corresponds to the threshold. The probit model that MacWilliams *et al.* (1970) tested incorporated an additional assumption that on a probit scale the threshold increased linearly with regeneration time. Plotting the optimized inhibition values determined in the course of testing the model suggested to them a linear dependence of inhibition intensity (on a probit scale) on basal disk size. We now inquire whether the similar linearities are apparent with the present totally different and fundamentally sounder method of computation, and whether models incorporating these linearities describe the data of Table 2 acceptably.

In Fig. 1, the estimates of λ_{1k}^{FS}, the λ variable corresponding to inhibition, are plotted as a function of the transplanted basal disk size, and the estimates of λ_{1j}^{FT} (the threshold) as a function of the regeneration time. The deviations from linearity are strikingly small: the linear correlation coefficient is $-0{\cdot}999$ for S and $+0{\cdot}969$ for T.

Model 4(a) assumes both of these linearities. This model replaces the variable λ^{FS} by a linear expression of the form $a_S + b_S S$ and λ^{FT} by $a_T + b_T T$. The as and bs are estimated by least-squares fits of straight lines to the data plotted in Fig. 1. Here S is measured in units of basal disks and

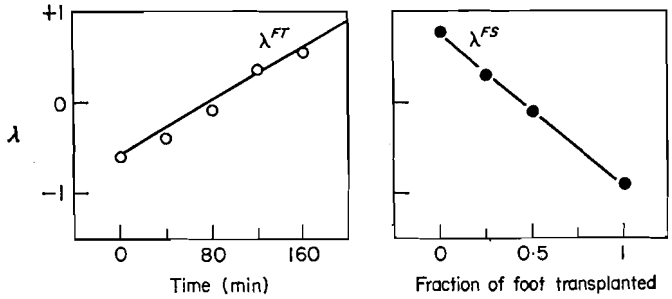

FIG. 1. Estimated parameters λ_{1j}^{FT} and λ_{1k}^{FS} in the logit form of model 4 for Table 2, as functions of the experimental variables time (T) and basal disk size (S). Straight lines are fitted by least squares.

T is in minutes. The result is a model with only three independent parameters:

$$\text{logit}\,(f_{jk}) = -0{\cdot}46 - 1{\cdot}69S + 0{\cdot}0075T. \qquad (11)$$

The numerical values of the constants in equation (11) are rounded to two significant figures, although the original calculations were done with five significant figures. The frequencies expected from equations (11) are given as the predictions of model 4(a) in Table 2.

The values of χ^2 and G^2 in Table 2 show that, for the 17 d.f. (= 20 measured frequencies of foot formation minus 3 parameters), the fit is excellent. Model 4 and model 4(a) may be compared by seeing if the difference in measure of fit for two models is significant for the corresponding difference in number of degrees of freedom. Table 2 shows that they differ by 5 d.f., but the difference in G^2 is only 1·52, not significant at the 0·1 level. The same is true of the difference in χ^2. Thus there is no significant improvement in fit from assuming that the λs are not linear functions of their corresponding physical variables.

A factor appropriate for converting from logit variables, such as S and T in equation (11), to probit variables such as Y_S and Y_T in equation (1), is $1{\cdot}19 = (0{\cdot}8413)^{-1}$, where 0·8413 is the integral of the standard normal density function from $-\infty$ to 1·0. The values of the logit model linear proportionality constants b_S and b_T given above, when multiplied by this factor, are the proportionality constants showing the dependence of the probit variables Y_S and Y_T on the experimental parameters S and T. Values of 0·36/40 min or −2·0/basal disk may thus be calculated. The values given in or apparent from MacWilliams *et al.* (1970) are 0·37/40 min and −2·1/basal disk.

From equation (11), it follows immediately that, for $0 < f < 1$,

$$\frac{\partial f}{\partial T} = +0{\cdot}0150 f(1-f), \qquad \frac{\partial f}{\partial S} = -3{\cdot}38 f(1-f).$$

Here the subscripts j and k, specifying respectively the values of T and S, have been omitted. Equations of this form describe self-limited autocatalysis, bimolecular reactions, and enzyme reactions (Berkson, 1951). If the frequency f is taken as directly proportional to the mass of some unknown chemical species which controls foot formation, these equations suggest a simple physical interpretation of equation (11). According to this speculative interpretation, the quantity of the hypothetical foot-forming substance increases autocatalytically in time; the quantity of the substance decreases with autocatalytic kinetics as far as permitted by the quantity of some supposed substance homogeneously distributed through the foot tissue; and the two reactions that increase and decrease the quantity of this hypothetical foot-forming substance proceed completely independently of each other.

In summary, the data of Tables 1 and 2 are well described by assuming independent effects of incubation time and transplanted foot size on the frequency of foot formation. In Table 2, these effects may be described as linear functions of the physical parameters of the experiment when frequency of foot formation is transformed to either the logit or the probit scale.

(B) TABLE 3

In each of the experiments reported in Table 3, an annulus (the "actor" in this experiment), consisting of a lengthwise eighth of the body column, was removed from a donor hydra and grafted laterally into a wound made in a host animal (the "setting") at one of the seven planes of section that divide the animal into eighths. The basal disk of the host was removed in some experiments. The transplanted annuli were numbered 0 to 7 according to position in the donor, starting basally. Only annuli 1–7 were used in the experiments. The final transplant positions were numbered from 1–7 starting basally. The actor annuli were scored after two days for foot formation.

Table 3 gives the observed number of cases of foot formation as a function of three parameters: the original position D of the actor, the final position P, and the condition of the host's basal disk H. D is the only actor-specific parameter; P and H characterize the setting.

Model 1 for these data is the hypothesis that the three experimental parameters influence foot formation independently; hence equation (1) is simply $Z = K + Y_D + Y_P + Y_H$. Since this model provides an unacceptable description of the data, it and simpler models assuming 0, 1, or 2 independent influences can be rejected. Consequently, any acceptable model must assume that at least one pair of the experimental parameters interact, that is, an acceptable model must include at least one variable which depends on two parameters simultaneously.

The three possible pairwise interactions are between D and P, D and H, and P and H. The first two of these entail the unattractive assumption that the influences of the actor and setting are not independent. Model 2 assumes a variable depending jointly on D and H and a second variable depending on P. Model 3 assumes a variable depending jointly on D and P and one depending on H. Model 5 assumes both of these interactions. All three models fail to describe the data acceptably.

The only remaining two-parameter interaction, that between P and H, the two parameters which describe the setting, is incorporated together with a variable associated with D in model 4. The model may be interpreted as the hypothesis that foot formation is controlled by two factors, one whose value is determined by the actor's original position and one whose value is determined by both setting-specific parameters. The model provides an excellent description of the data.

The predicted frequencies of foot formation, out of 25 trials, are given for each combination of experimental parameters as "model 4" underneath the observed frequencies in Table 3. The predictions of this model are also given for those regions of the table where experiments were not carried out.

The remaining models 6, 7 and 8 assume interactions in addition to those of model 4. Since they estimate more parameters from the data than does model 4, they naturally also fit well. But comparison of the degrees of freedom and of the measure of fit shows that in no case does the goodness of fit improve significantly. Hence the data provide no justification for the additional assumptions of these models. In particular, model 6 shows that nothing is gained by assuming an interaction between the transplant's original and final positions. Model 7 demonstrates that the assumption of an interaction of the host's condition H with the original position of the donor annulus D is also unnecessary. Model 8, which contains all possible pairwise parameter interactions, shows that the description of the data cannot be significantly improved by abandoning the assumption that the influences of host and donor (actor and setting) are independent.

Thus the assumptions of model 4 are minimal, in that any simpler model of the general form under discussion fails to describe the data acceptably, and sufficient, in that the data provide no evidence for additional assumptions.

Model 4 in probit form is equivalent to an inhibition-threshold model of MacWilliams (1972). In MacWilliams' model, the inhibition (the setting property) could be described satisfactorily as a linear function of position in both hosts with and hosts without basal disks. The actor-specific (threshold) property was found to be monotone with position, but changed more rapidly with position at the ends of the animal than in the middle, and could not be

acceptably described as linear. In logit form model 4 is

$$\text{logit}\,(f_{jkl}) = \lambda_1^F + \lambda_{1j}^{FD} + \lambda_{1k}^{FP} + \lambda_{1l}^{FH} + \lambda_{1kl}^{FPH}. \tag{12}$$

Here λ^F corresponds to K in the probit model, equation (1), while λ^{FD} represents the variable depending only on the original position of the transplanted piece. The last three terms represent the dependence of the frequency of foot formation on the setting parameters P and H. The parameters λ^{FH} do not depend on position. Figure 2 shows the estimated λ^{FD}, λ^{FP}, and λ^{FPH} as functions of position. It is not clear from the plots which of these functions can be acceptably approximated as linear.

Specific models were therefore tested. In each case, the λs which were not assumed to be linear were given the values obtained in model 4. For each set of parameters assumed to be linear functions of annulus number, slopes and intercepts were estimated by fitting lines by least squares in three ways:

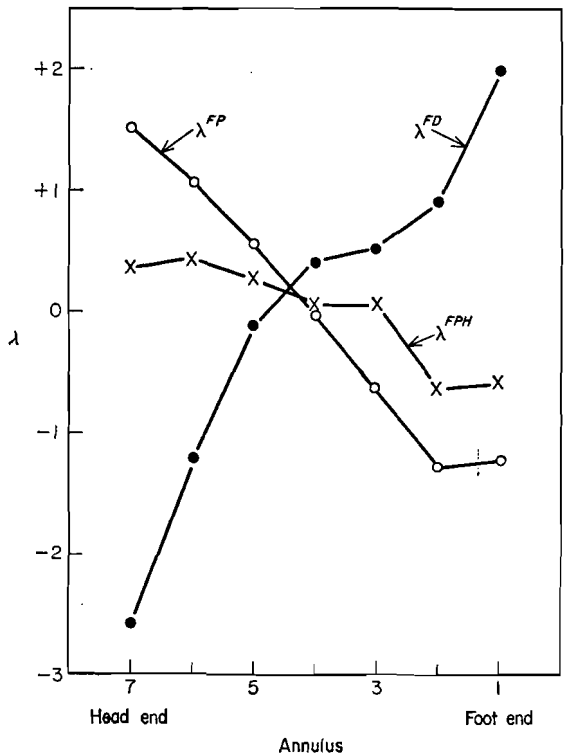

FIG. 2. Estimated parameters λ_{1j}^{FD}, λ_{1k}^{FP}, λ_{1kl}^{FPH} in the logit form, equation (12), of model 4 for Table 3, as functions of the annulus number in the donor and host hydra. The line segments between plotted points are intended only to guide the eye.

first to the parameters for all seven annuli, second to the parameters for all annuli except annulus 1 (closest to the basal disk), and third to the parameters for all annuli except numbers 1 and 7. Expected cell frequencies were calculated for each of these sets of slopes and intercepts. The parameters based on the slopes and intercepts (but not the λs not assumed to be linear) were then modified by a crude numerical search procedure to locate the fit which minimized G^2. In models which include most of the present models as special cases, Mantel & Brown (1973) use an iterative Newton–Raphson search procedure which is probably preferable. Because of the factorization of model 4, the use of the λs from model 4 within models 4(a)–(g) is appropriate. The minimization of χ^2, which has the same asymptotic distribution as G^2, gives asymptotically the maximum likelihood parameter estimates (Cramer, 1946). The resulting fits are optimal, subject to the limitations of the search procedure and finite sample sizes.

Models 4(a), 4(b) and 4(c) in Table 3 assume the linearity, respectively, of λ_{1j}^{FD}, λ_{1k}^{FP} and λ_{1kl}^{FPH} as functions of position. The measures of goodness of fit show that the actor parameters λ_{1j}^{FD} cannot be approximated linearly. Either of the remaining sets of variables, which characterize the setting, may be linearly approximated.

Models 4(d)–(e) assume the linearity of both λ_{1j}^{FD} and one of the remaining sets of variables, λ_{1k}^{FP} or λ_{1kl}^{FPH} respectively. Not surprisingly, these models fail to describe the data.

Model 4(f) assumes the linearity of both sets of setting variables λ_{1k}^{FP} and λ_{1kl}^{FPH} but not the linearity of the actor variables. This model does fit the data quite acceptably. Comparison of model 4(f) with models 4(b), 4(c) and 4 shows no significant improvement in fit from giving up the assumption that the setting variables are linear functions of position in the host.

Model 4(g) assumes that all parameters in model 4 which are functions of annulus number are linear functions. As expected, this model fails to describe the data.

Finally model 4(h) adds to model 4(f), which assumes that the setting variables are linear functions of position, the further assumption that the actor variable is a cubic function of position. A cubic equation was fitted by the method of orthogonal polynomials to the values of λ_{1j}^{FD} estimated in model 4. The resulting coefficients were then modified in a numerical search for the best fit. The final explicit form of model 4(h), with seven fitted coefficients, is (after rounding)

$$\text{logit}\,(f_{jkl}) = 1{\cdot}18 - 2{\cdot}17D + 0{\cdot}53D^2 - 0{\cdot}050D^3 + 0{\cdot}52P + (-1{\cdot}32 + 0{\cdot}18P)H. \quad (13)$$

Because of the asymmetry of the values of λ_{1j}^{FD} with respect to annulus 4,

as shown in Fig. 2, it is likely that a substantially larger number of observations might reveal that the present cubic approximation in equation (13) is inadequate. So far no acceptable model simpler that equation (13) has been found for Table 3.

The setting terms $0{\cdot}52P+(-1{\cdot}32+0{\cdot}18P)H$ predict a dependence of logit (f_{jkl}) on the final transplant position of 0·34 per eighth animal length when the disk is absent ($H=-1$) and 0·70 when the disk is intact ($H=1$), with a difference of 2·64 produced at position 0 when the disk is removed. Multiplying these by the conversion factor 1·19 gives 0·40, 0·84 and 3·14 as predicted coefficients for the probit model. The corresponding values determined for the probit model are 0·53, 0·99 and 3·46 (MacWilliams, 1972). The slightly poorer agreement in this case may reflect differences in assumptions and computational procedure or the deviations from linearity of the setting variables as functions of position (Fig. 2). The dependence of the actor variable on position (Fig. 2) is qualitatively the same as found by MacWilliams (1972): steep at both ends but nearly flat in the middle.

The number of cases of foot formation expected from model 4(h) for each possible combination of donor annulus, host position, and host condition, including but not restricted to those combinations for which measurements were actually made, appear underneath the data in Table 3 as "model 4(h)".

5. Discussion

This analysis confirms that previously proposed probit models, in which two normally varying scalars, "inhibition" and "threshold", control foot formation, provide a simple, acceptable interpretation of the experimental data. We are unable to propose a simpler general hypothesis consistent with the data. The acceptability of these simple models does not prove that more complex models are false, but suggests that they are uneconomical in the present state of knowledge.

The previously reported linear dependence of inhibition and threshold to inhibition on certain experimental parameters has also been confirmed. It is important to note, however, that the probability values assigned to the linearized model (4(a) in Table 2 and 4(a)–(h) in Table 3) do not have the interpretation customary in confirmatory statistics, where a model is chosen *before* the data are examined and the correspondence of the data to prediction is assayed. Here the linearity of the λs as functions of the experimental parameters appeared after examination of the data; the linearized models would probably not have been tested for fit had the parameters appeared grossly different. If models of the form of equation (11) and equation (13)

describe well future replications of the corresponding sets of experiments, the probability values associated with χ^2 and G^2 will have their classical interpretation.

This analysis in no way excludes non-linear dependencies in which the deviations from linearity are smaller than the resolution, which is limited by the sample size, of the statistical measures of fit which we used.

The models we have found satisfactory are not probit models directly but log-linear numerical approximations to them. Hence one must consider the possibility that the control of foot formation in transplantation experiments is fundamentally logistic. If it is, our findings suggest that the odds of foot formation are the product of two scalars, one specific to the actor and one to the setting. For example, model 4 for Tables 1 and 2 sets $\lambda^{FTS} \equiv 0$ in the right member of equation (5). Hence that exponential can be written as the product of one factor depending only on F and T, which is the actor-specific scalar, and another factor depending only on F and S, the setting-specific scalar. In the linearized forms of the logit models, such as model 4(a) for Tables 1 and 2, it is the logarithms of the scalars which depend linearly on the experimental parameter values.

If individual animals do perform some analog of multiplying the actor-specific scalar by the setting-specific scalar to obtain the odds of foot formation, the decision to form a foot or not could then be made by comparing the probability derived from these odds to an internally generated, uniformly distributed random variate.

Since their predictions are so similar, the choice between logit and probit models for a given data set is often made on the basis of simple personal preference. In this the two authors of this paper differ. J.E.C., following Berkson (1951), favors logit models, citing: (1) the ease of constructing mechanisms which behave logistically; (2) the fact that logit models do not require, as do probit models, an unexplained normal variation in the fundamental parameters of the system modeled; (3) the superior mathematical tractability of logit models. H.K.M. is attracted by the formal simplicity of the decision-making process postulated by the probit models, and finds the assumption of normal variation easy to accept, inasmuch as the system's fundamental parameter values are assumed here to be functions of experimental variables, which themselves cannot be perfectly controlled. The authors agree that in the absence of decisive evidence, both types of explanations must be regarded as viable, and it is necessary to keep an open mind.

All models presented here are phenomenological. Although suggestive, they do not suffice to specify a physical mechanism. Even if it were possible to associate these models with specific kinetic proposals, such as those of Gierer & Meinhardt (1972, 1974) direct evidence would still be required

to determine the biochemical basis of, and to select among, kinetic models. The function of our models is to highlight the regularities which physical mechanisms must explain.

We thank Yvonne M. M. Bishop and John Guckenheimer for very helpful discussions of previous drafts, and the U.S. National Science Foundation for partial support.

REFERENCES

BERKSON, J. (1951). *Biometrics* **7,** 327.
BISHOP, Y. M. M. (1967). Multidimensional contingency tables: cell estimates. Doctoral thesis: Harvard University, Department of Statistics.
BISHOP, Y. M. M. (1969). *Biometrics* **25,** 383.
CRAMER, H. (1946). *Mathematical Methods of Statistics.* Princeton, N.J.: Princeton University Press.
FIENBERG, S. E. (1970). *Ecology* **51,** 419.
FIENBERG, S. E. (1972). *Biometrics* **28,** 177.
FINNEY, D. J. (1964). *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve,* 3rd edn. London: Cambridge University Press.
GIERER, A. & MEINHARDT, H. (1972). *Kybernetik* **12,** 30.
GIERER, A. & MEINHARDT, H. (1974). *J. Cell Sci.* (in press).
GOODMAN, L. A. (1969). *Jl R. statist. Soc. Ser. B* **31,** 486.
GOODMAN, L. A. (1970). *J. Am. statist. Ass.* **65,** 226.
Goodman, L. A. (1971*a*). *Technometrics* **13,** 33.
GOODMAN, L. A. (1971*b*). *Am. statist. Ass.* **66,** 339.
GOODMAN, L. A. (1972). *Am. J. Sociol.* **77,** 1035.
MACWILLIAMS, H. K. (1972). Factors influencing foot formation in *Hydra viridis.* Doctoral thesis: Harvard University, Department of Biology.
MACWILLIAMS, H. K. & KAFATOS, F. C. (1968). *Science, N.Y.* **159,** 1246.
MACWILLIAMS, H. K., KAFATOS, F. C. & BOSSERT, W. H. (1970). *Devl Biol.* **23,** 380.
MANTEL, N. (1970). *Biometrics* **26,** 291.
MANTEL, N. & BROWN, C. (1973). *Biometrics* **29,** 649.