Aust. J. Biol. Sci., 1981, **34**, 81–95

Multiline Varieties and Disease Control. III* Combined Use of Overlapping and Disjoint Gene Sets

D. R. Marshall and J. J. Burdon

Division of Plant Industry, CSIRO, P.O. Box 1600, Canberra City, A.C.T. 2601.

Abstract

Alternate procedures, based on the use of combinations of overlapping (of the type AB, AC, BC, etc.) or disjoint (of the type AB, CD, EF, etc.) subsets of genes, for developing multiline components carrying two or more major genes for disease resistance are described. Simple mathematical models are used to examine the effects of multilines blended from such components on pathogen evolution and long-term disease control. It is found, in agreement with earlier studies, that the use of more than one resistance gene per component, where the genes are combined in either overlapping subsets or disjoint subsets, will be justified in practice only if it can be demonstrated that unnecessary genes for virulence in pathogens act additively to reduce the fitness of their carriers. If unnecessary virulence genes in the pathogen reduce fitness multiplicatively then multilines composed of components carrying multigenic resistances have only limited advantages (in the case of disjoint subsets of genes) or are at a positive disadvantage (in the case of overlapping subsets of genes) in long-term disease control compared with multilines in which each component carries a single unique gene for resistance.

Introduction

In previous papers (Marshall and Pryor 1978, 1979) we examined, using simple theoretical models, the effects of 'dirty crop' or 'partially resistant' (Marshall 1977; Marshall and Pryor 1978) multiline varieties on the long-term racial composition of pathogen populations. These studies indicated that the capacity of multilines to prevent the development of complex pathogen races which can attack more than a given small proportion (p) of multiline components, and hence, to provide effective disease control, depends on four factors:

(i) the level of selection against unnecessary virulence genes in the pathogen;
(ii) whether unnecessary genes for virulence act additively or multiplicatively in reducing the fitness of their carriers;
(iii) the number of resistance genes the plant breeder has at his disposal for use in the multiline;
(iv) the way, whether singly or in combinations, these resistance genes are incorporated in multiline components.

In particular, Marshall and Pryor (1979) showed that multiline varieties blended from components carrying multigenic resistances either as overlapping gene sets (of the type AB, AC, BC, etc.) or disjoint gene sets (of the type AB, CD, EF, etc.) would often be more effective in long-term disease control than multilines whose components carry only a single gene for resistance. Specifically, they showed that for overlapping gene

* Part II, *Euphytica*, 1979, **28**, 145–59.

sets, disease control would be better under the additive, but not the multiplicative, fitness-effects model if resistance genes are combined rather than used singly in multiline components. With disjoint gene sets, multilines composed of components with multigenic resistance would be as, or more, effective in disease control compared with a single resistance gene under both additive and multiplicative fitness models.

However, there are serious limitations to the practical use of both overlapping and disjoint gene sets *per se*. On one hand, the use of overlapping gene combinations is feasible only with small numbers of resistance genes, say six or fewer. Otherwise, the breeder must develop, maintain and mix impossibly large numbers of multiline components to exploit their enhanced potential for disease control. On the other hand, the use of disjoint sets is only feasible with large numbers of resistance genes, say 16 or more. Otherwise, the breeder would have too few component lines at his disposal to develop an effective multiline variety.

In this paper we describe alternative procedures, based on the use of combinations of overlapping and disjoint gene sets, of combining resistance genes in multiline components which would be of use with intermediate numbers of resistance genes. We also examine the effects of multilines blended from such components on pathogen evolution and long-term disease control.

Developing Multiresistant Multiline Components

We are concerned with the situation where a plant breeder has an intermediate number of resistance genes, say between 6 and 15, at his disposal and wishes to develop a multiline variety composed of 8–16 components, each carrying two or more genes for resistance. There are many ways a breeder could achieve this objective. We will consider two. The first, the use of combinations of overlapping subsets of genes, would appear to be of value where the breeder has 6–10 resistance genes at his disposal. The second, the use of combinations of disjoint subsets of genes, is likely to be applicable with larger numbers of resistance genes, say 10–16.

Combinations of Overlapping Subsets of Genes

Consider the case where the breeder has n ($6<n<10$) non-allelic resistance genes to incorporate two or more at a time in multiline components. Under this procedure, the n resistance genes are divided into h groups, the first group containing n_1 genes, the second n_2, and so forth. In general, h may take any value in the range $1 \leqslant h \leqslant n$ but, in practice, h will seldom be greater than 3. The genes within the ith group are combined g_i at a time as overlapping sets into subcomponents and the subcomponents from the h groups are intercrossed to develop the multiline components. This procedure would yield $l = \prod_{i=1}^{h} \binom{n_i}{g_i}$ component lines (where $\binom{n_i}{g_i}$ is the number of subcomponents developed from the ith group of genes), each line carrying $g = \sum_{i=1}^{h} g_i$ genes for resistance. To illustrate this procedure we will outline two numerical examples. In the first, we assume the plant breeder has $n = 7$ resistance genes on hand and that these are subdivided into $h = 2$ subsets with $n_1 = 4$ and $n_2 = 3$. We also assume all subcomponents carry only a single gene for resistance, that is, $g_1 = g_2 = 1$. In this example, the breeder would develop $l = 12$ lines each carrying $g = 2$ genes for resistance. To use all combinations of seven genes taken two at a

time would require the development of a $l = \binom{7}{2} = 21$-component multiline. In the second example, we assume the plant breeder has $n = 6$ resistance genes to use in a multiline and these are subdivided into $h = 2$ subsets with $n_1 = n_2 = 3$. We further assume that $g_1 = 2$ and $g_2 = 1$, that is, the resistance genes in the first subset are combined in each subcomponent in pairs, and the genes in the second subset are used singly. Intercrossing of these subcomponents would yield $l = 9$ lines each carrying $g = 3$ genes for resistance. If all possible three-gene sets were used in the multiline it would contain $l = \binom{6}{3} = 20$ component lines.

Combinations of Disjoint Subsets of Genes

Here we assume the breeder has 10–16 non-allelic resistance genes at his disposal which are divided into h groups containing $n_1, n_2, \ldots, n_h$ genes, respectively. The genes within each group are combined as disjoint sets into subcomponents which are then intercrossed to develop the multiresistant components of the multiline. In this case, the number of components would be $l = \prod_{i=1}^{h} (n_i/g_i) = \prod b_i$ (where $b_i = n_i/g_i$), each carrying $g = \sum_{i=1}^{h} g_i$ genes for resistance. As an example, assume a breeder has 16 genes at his disposal and these are divided into two equal groups ($h = 2$, $n_1 = n_2 = 8$) and combined two at a time as disjoint sets in subcomponents ($g_1 = g_2 = 2$), then the completed multiline would contain 16 lines each carrying four genes for resistance. The use of disjoint gene sets *per se* would result in a four-line multiline if each component carried four genes for resistance. As a second example, assume the breeder has 13 genes at his disposal which he divides into two groups with 3 and 10 genes respectively ($h = 2$; $n_1 = 3, n_2 = 10$). If the first group is incorporated into subcomponents singly while the second group is incorporated into subcomponents in pairs as disjoint sets ($g_1 = 1, g_2 = 2$), the completed multiline will contain 15 components each carrying three genes for resistance. Here again the use of disjoint sets *per se* would allow the development of only a four-component multiline if each line was required to carry at least three different resistance genes.

Effects on Pathogen Evolution

We will now examine the effects of the multilines, whose development is described above, on the long-term racial composition of an obligate pathogen population, using a simple theoretical model. The basic assumptions of the model are briefly that (Groth 1976; Groth and Person 1977; Marshall and Pryor 1978, 1979):

(i) the multiline is composed of equal proportions of l ($l > 2$) diploid host genotypes which are identical except that each has a unique complement of one or more homozygous dominant genes conferring resistance to the specified haploid pathogen;

(ii) the host–pathogen system conforms to the gene-for-gene relationship between host susceptibility/resistance and pathogen virulence/avirulence (Flor 1956);

(iii) the multiline mixture is reconstituted annually so its composition is stable over time, and further that it is grown over a large area so that it is the major factor influencing the evolution of virulence in the local pathogen population;

(iv) biotypes with all possible combinations of virulence genes exist initially, or will arise within a short time span, in the pathogen population;
(v) biotypes carrying unnecessary genes for virulence are less fit than biotypes carrying only effective genes for virulence, and that each unnecessary gene for virulence reduces their fitness by a constant amount (s);
(vi) in each generation, pathogen biotypes compete and reproduce at different rates (depending on the particular fitness model involved) on susceptible hosts and the resultant spores are distributed at random over the host population; this procedure is repeated until the pathogen population reaches equilibrium;
(vii) the pathogen populations are large so we are concerned only with deterministic equilibria.

In each case we consider two variations of this basic model. In particular, we assume that two or more unnecessary genes for virulence act either additively or multiplicatively in reducing pathogen fitness.

Combinations of Overlapping Subsets of Genes

Additive fitness effects

A multiline composed of $l = \prod_{i=1}^{h} \binom{n_i}{g_i}$ lines, each carrying g of n non-allelic genes for resistance, may be parasitized by $2(n-g+1)$ classes of pathogen biotypes, *viz*, those with $g, g+1, \ldots, k, \ldots, n$ effective genes for virulence with and without one or more unnecessary genes for virulence. We will restrict our attention to the $(n-g+1)$ classes of biotypes carrying only effective genes for virulence since, under our assumptions, biotypes with unnecessary genes for virulence invariably have reduced fitness and are eliminated from the pathogen population (Marshall and Pryor 1978).

For the sorts of mixtures considered here, not all members of a particular class of pathogen biotypes will necessarily be equally fit (virulent on the same number of components) as they were in the previous analysis of multilines containing all possible combinations of resistance genes (Marshall and Pryor 1979). The members of a class of biotypes which can grow on the most components will have the highest mean fitness and will eventually displace their less fit counterparts. As a result, we will further restrict our attention to those biotypes, within each of the $(n-g+1)$ classes of interest, which are virulent on the maximum number of multiline components.

Assuming additive fitness effects, the mean relative fitnesses of these biotypes on multilines whose components carry combinations of overlapping sets of genes are given in Table 1. The relative fitness of class-1 biotypes with g genes for virulence growing on hosts with g genes for resistance was taken to be $[1-(g-1)s]$ rather than 1 to ensure the present results will be comparable with earlier studies (Marshall and Pryor 1978, 1979). In effect we assume the fitnesses of all biotypes are measured relative to races with a single gene for virulence growing on hosts with a single gene for resistance. We also assume that a particular biotype has the same relative fitness on all susceptible host genotypes with the same number of resistance genes. Finally, since relative fitnesses cannot be negative by definition it is assumed that, if $(k-1)s \geqslant 1$ or $s \geqslant 1/(k-1)$, biotypes carrying k or more genes for virulence have zero fitness.

From the relative fitnesses in Table 1 it is clear that the simplest races with g genes for virulence have the highest mean fitness and will dominate the pathogen population if

$$[1-(g-1)s]/l > (g_1+1)(1-gs)/l, \tag{1}$$

that is, if

$$s > g_1/(1+g_1 g). \tag{2}$$

Since the relative fitness function $C_k[1-(k-1)s]/l$ has a single unique maximum value with varying k for a given s (Groth 1976) it follows that more complex races carrying k genes for virulence will dominate the pathogen population if

$$C_{k+1}(1-ks)/l < C_k[1-(k-1)s]/l > C_{k-1}[1-(k-2)s]/l, \tag{3}$$

where $C_k = \prod_{i=1}^{h} \binom{k'_i}{g_i}$, with k'_i chosen such that C_k is a maximum given the constraints $k = \sum_{i=1}^{h} k'_i$ and $n_i \geqslant k'_i \geqslant g_i$.

In simplifying inequality (3) we consider two distinct cases. Firstly, we consider the situation where

$$\begin{aligned} C_k &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i}{g_i}\cdots\binom{k'_h}{g_h}, \\ C_{k+1} &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i+1}{g_i}\cdots\binom{k'_h}{g_h}, \\ C_{k-1} &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i-1}{g_i}\cdots\binom{k'_h}{g_h}, \end{aligned} \tag{4}$$

that is, where the differences between C_k are due to differences in the ith group. In this case inequality (3) reduces to

$$g_i/[g_i(k-2)+k'_i] > s > g_i/[g_i(k-1)+k'+1]. \tag{5}$$

Secondly, we consider the case where differences in C_k involve two, say, the ith and jth, groups, so that,

$$\begin{aligned} C_k &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i}{g_i}\binom{k'_j}{g_j}\cdots\binom{k'_h}{g_h}, \\ C_{k+1} &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i}{g_i}\binom{k'_j+1}{g_j}\cdots\binom{k'_h}{g_h}, \\ C_{k-1} &= \binom{k'_1}{g_1}\binom{k'_2}{g_2}\cdots\binom{k'_i-1}{g_i}\binom{k'_j}{g_j}\cdots\binom{k'_h}{g_h}, \end{aligned} \tag{6}$$

and inequality (3) becomes on simplification

$$g_i/[g_i(k-2)+k'_i] > s > g_j/[g_j(k-1)+k'_j+1]. \tag{7}$$

Further, a super-race carrying n genes for virulence and capable of attacking all the component lines of the multiline will develop if

$$[1-(n-1)s] > C_{n-1}[1-(n-2)s]/l, \tag{8}$$

where $C_{n-1} = \binom{n_1}{g_1}\binom{n_2}{g_2}\cdots\binom{n_i-1}{g_i}\cdots\binom{n_h}{g_h}$. The ith group is chosen such that C_{n-1} is a maximum. Since

$$C_{n-1} = C_n\binom{n_i-1}{g_i}\Big/\binom{n_i}{g_i} = l(n_i-g_i)/n_i,$$

it follows that C_{n-1} has its maximum when g_i/n_i has its minimum; that is, when the ith group contains the smallest proportion of genes selected for use in the sub-components of the multiline. Thus,

$$C_{n-1} = l[1-(g_i/n_i)_{\min}], \qquad 1 \leqslant i \leqslant h.$$

Table 1. Relative fitnesses of pathogen biotypes on an l-line multiline carrying overlapping subsets of g genes for resistance

$l = \prod_{i=1}^{h}\binom{n_i}{g_i}$ where n_i is the number of resistance genes in the ith group and g_i is the number of these genes incorporated in each multiline component. For simplicity we assume $g_1 > g_2 > \ldots > g_i > \ldots > g_h$. s, Level of selection against unnecessary virulence genes in the pathogen

Pathogen biotype Class	No. of effective genes for virulence	Relative fitness on susceptible components: Additive model	Relative fitness on susceptible components: Multiplicative model	Max. No. of susceptible components	Mean fitness on multiline: Additive model	Mean fitness on multiline: Multiplicative model
1	g	$[1-(g-1)s]$	$(1-s)^{g-1}$	1	$[1-(g-1)s]/l$	$(1-s)^{g-1}/l$
2	$(g+1)$	$(1-gs)$	$(1-s)^{g}$	(g_1+1)	$(g_1+1)(1-gs)/l$	$(g_1+1)(1-s)^{g}/l$
$k-g+1$	k	$[1-(k-1)s]$	$(1-s)^{k-1}$	C_k[A]	$(C_k[1-(k-1)s]/l$	$C_k(1-s)^{k-1}/l$
$n-g$	$n-1$	$[1-(n-2)s]$	$(1-s)^{n-2}$	C_{n-1}	$C_{n-1}[1-(n-2)s]/l$	$C_{n-1}(1-s)^{n-2}/l$
$n-g+1$	n	$[1-(n-1)s]$	$(1-s)^{n-1}$	l	$[1-(n-1)s]$	$(1-s)^{n-1}$

[A] $C_k = \prod_{i=1}^{h}\binom{k'_i}{g_i}$ where k'_i are chosen such that C_k is a maximum given $k = \sum_{i=1}^{h} k'_i$ and $n_i > k'_i > g_i$ (additive model) $i = 1$ or $g_i \geqslant k'_i \geqslant n_i$ (multiplicative model).

We have, as a result,

$$\binom{n_i}{g_i}[1-(n-1)s] > \binom{n_i-1}{g_i}[1-(n-2)s], \tag{9}$$

which reduces to

$$g_i/[n_i+g_i(n-2)] > s > 0. \tag{10}$$

Finally, it should be noted that if $s = g_i/[g_i(k-1)+k'+1]$ in the case specified by inequality (5) or $s = g_j/[g_j(k-1)+k'_j+1]$ in the case specified by inequality (7), pathogen races carrying $(k-1)$ and k genes for virulence will co-exist in an equilibrium population. Since this equilibrium is neutral in practice, the relative proportions of each class of biotype will vary stochastically.

Multiplicative fitness effects

Here we take the same model as before and simply vary the fitnesses of the various classes of pathogen biotypes as shown in Table 1. Under this model the simplest races with g genes for virulence will dominate the pathogen population if

$$(1-s)^{(g-1)}/l > (g_i+1)(1-s)^g/l, \tag{11}$$

or

$$s > g_1/(g_1+1). \tag{12}$$

More complex races with k genes for virulence will win the struggle for dominance if

$$C_{k+1}(1-s)^k/l < C_k(1-s)^{k-1}/l > C_{k-1}(1-s)^{k-2}/l, \tag{13}$$

where, as before, $C_k = \prod_{i=1}^{l} \binom{k'_i}{g_i}$ with k'_i chosen such that C_k is a maximum given the constraints $k = \sum_{i=1}^{h} k'$ and $n_i \geqslant k'_i \geqslant g_i$.

Again we consider two distinct cases. In the first, the differences in C_k reside in the ith subgroup [see eqn (4)], in which case inequality (13) reduces to

$$g_i/k'_i > s > g_i/(k'_i+1). \tag{14}$$

In the second, the differences in C_k reside in two subgroups as in equation (6) and, in this case, inequality (13) becomes

$$g_i/k'_i > s > g_j/(k'_j+1). \tag{15}$$

Further, a super-race will develop if

$$(1-s)^{n-1} > C_{n-1}(1-s)^{n-2}/l, \tag{16}$$

where $C_{n-1} = \binom{n_1}{g_1}\binom{n_2}{g_2}\cdots\binom{n_i-1}{g_i}\binom{n_h}{g_h}$ and the ith subgroup is chosen such that C_{n-1} is a maximum. That is, when

$$C_{n-1} = l[1-(g_i/n_i)_{\min}], \qquad 1 \leqslant i \leqslant h.$$

In this case inequality (16) reduces to

$$g_i/n_i > s > 0. \tag{17}$$

Finally, if $s = g_i/(k'_i+1)$ in the case specified by inequality (14) or $s = g_j/(k'_j+1)$ in the case given by inequality (15), the pathogen population will be polymorphic for races carrying $(k-1)$ and k genes for virulence, with the relative proportions of the two classes of biotypes varying stochastically.

Numerical results

The above analysis shows that the use of multilines whose components carry combinations of overlapping subsets of genes may significantly alter the equilibrium racial composition of the pathogen population from that expected if the same resistance genes were used in components either singly or in all possible combinations. This point is illustrated in Table 2 which gives two examples of the levels of selection required to stabilize the composition of a pathogen population carrying k

($g \leqslant k \leqslant n$) genes for virulence, assuming the plant breeder uses the n resistance genes at his disposal in various ways. In the first example, it is assumed that $n = 7$ and the plant breeder incorporates these into multiline components either singly, or two at a time in all possible overlapping combinations, or two at a time as overlapping combinations of subsets of genes (with $h = 2$, $n_1 = 4$, $n_2 = 3$, $g_1 = g_2 = 1$). In the second example, $n = 6$ and we assume the plant breeder uses these genes singly, or three at a time in all possible combinations, or three at a time as combinations of gene subsets (with $h = 2$, $n_1 = n_2 = 3$, and $g_1 = 2$, $g_2 = 1$). The levels of selection required to ensure a particular level of virulence in the predominant pathogen biotype are higher in all cases for the multiplicative compared with additive models and where genes are used in combinations rather than singly (Table 2). The levels of selection required to ensure that the predominant pathogen biotype carries k genes for virulence were intermediate for combinations of gene subsets compared with their use singly or in all possible combinations.

Table 2. Levels of selection required to stabilize the racial composition of the pathogen populations

No. of virulence genes (k) in dominant pathogen biotype	Level of selection when available resistance genes are incorporated into multiline components: Singly		g at a time, overlapping sets of genes		g at a time, combinations of overlapping subsets of genes	
	Additive model	Multiplicative model	Additive model	Multiplicative model	Additive model	Multiplicative model
			Example 1 ($g = 2$)			
1	>0·50	>0·50	n.a.[A]	n.a.	n.a.	n.a.
⩽2	>0·25	>0·33	>0·40	>0·67	>0·33	>0·50
⩽3	>0·17	>0·25	>0·25	>0·50	>0·25	>0·50
⩽4	>0·13	>0·20	>0·18	>0·40	>0·17	>0·33
⩽5	>0·10	>0·17	>0·14	>0·33	>0·14	>0·33
⩽6	>0·08	>0·14	>0·12	>0·29	>0·11	>0·25
7	<0·08	<0·14	<0·12	<0·29	<0·11	<0·25
			Example 2 ($g = 3$)			
1	>0·50	>0·50	n.a.	n.a.	n.a.	n.a.
⩽2	>0·25	>0·33	n.a.	n.a.	n.a.	n.a.
⩽3	>0·17	>0·25	>0·33	>0·75	>0·29	>0·67
⩽4	>0·13	>0·20	>0·21	>0·60	>0·20	>0·50
⩽5	>0·10	>0·17	>0·17	>0·50	>0·14	>0·33
6	<0·10	<0·17	<0·17	<0·50	<0·14	<0·33

[A] n.a., Not applicable.

Combinations of Disjoint Subsets of Genes

Additive fitness effects

A multiline composed of components carrying g genes for resistance as combinations of disjoint subsets of genes will contain l [$= \prod_{i=1}^{h} (n_i/g_i) = \Pi_{i=1}^{h} b_i$] components. Such a multiline may again be parasitized by $2(n-g+1)$ classes of pathogen biotypes, *viz*, those with $g, g+1, \ldots, k, \ldots, n$ effective genes for virulence with and without unnecessary genes for virulence. Of these, only those with com-

binations of virulence genes which directly match the resistance genes in one or more of the component lines are of direct interest here. The number of such classes is a function of the number and composition of the subsets of resistance genes. The maximum number is $l = \prod_{i=1}^{h} b_i$.

The relative fitnesses of the biotypes of interest on each susceptible component as well as their mean fitness on the multiline as a whole are given in Table 3 under this model. The simplest race with g genes for virulence will dominate the pathogen population if

$$[1-(g-1)s]/l > 2[1-(g+g_h-1)s]/l, \tag{18}$$

or

$$s > 1/(g+2g_h-1). \tag{19}$$

Table 3. Relative fitnesses of pathogen biotypes on an l-line multiline mixture in which each component carries g genes for resistance as combinations of disjoint subsets of genes, assuming additive fitness effects

$l = \prod_{i=1}^{h}(n_i/g_i) = \prod_{i=1}^{h} b_i$ where n_i is the number of resistance genes in the ith subset and g_i is the number used in each multiline component. g_i are ranked such that $g_1 \geqslant g_2 \geqslant \ldots \geqslant g_i \ldots \geqslant g_h$

Pathogen biotype		Relative fitness	Max. No. of	Mean fitness
Class	No. of effective genes for virulence	on susceptible components	susceptible components	on multiline
1	g	$[1-(g-1)s]$	1	$[1-(g-1)s]/l$
2	$g+g_h$	$[1-(g+g_h-1)s]$	2	$2[1-(g+g_h-1)s]/l$
$m(=\prod_{i=1}^{h} m'_i)$	$k(=\sum_{i=1}^{h} m'_i g_i)$	$[1-(k-1)s]$	C_k[A]	$C_k[1-(k-1)s]/l$
$l-1$	$g-g_i$	$[1-(n-g_i)s]$	C_{n-g_i}	$C_{n-g_i}[1-(n-g_i)s]/l$
l	$n(=\sum_{i=1}^{l} b_i g_i)$	$[1-(n-1)s]$	l	$[1-(n-1)s]$

[A] $C_k = \prod_{i=1}^{h} m'_i$ where m'_i is chosen such that C_k is a maximum given $k = \sum_{i=1}^{h} m'_i g_i$ and $b_i \geqslant m'_i \geqslant 1$.

Alternatively, more complex races with k genes for virulence will dominate the pathogen population if

$$C_{k-g_i}[1-(k-g_i-1)s]/l < C_k[1-(k-1)s]/l > C_{k+g_i}[1-(k+g_i-1)s]/l, \tag{20}$$

where C_k represents the maximum number of lines susceptible to any race carrying k genes for virulence. Here $C_k = \prod_{i=1}^{h} m'_i$ where m'_i is chosen such that C_k is a maximum given $k = \sum_{i=1}^{h} m'_i g_i$ and $b_i \geqslant m'_i \geqslant 1$. The complexities involved in calculating C_k are illustrated in Table 4 which lists C_k, m'_1 and m'_2 for all important classes of pathogen biotypes virulent on a 15-line multiline developed using disjoint subsets of resistance genes where $n = 13$, $h = 2$, $n_1 = 3$, $n_2 = 10$ and $g_1 = 1$, $g_2 = 2$, $g = 3$. It should be emphasized that in this example pathogen races carrying 8, 10 and 12 genes for virulence will invariably be eliminated from the pathogen population since they are less fit than races with 7, 9 and 11 genes for virulence but can only grow on the same number of component lines.

In reducing inequality (20) to simpler terms we again consider two distinct cases. Firstly, we consider the case where

$$\begin{aligned} C_k &= (m'_1)(m'_2)\ldots(m'_i)\ldots(m'_h), \\ C_{k+g_i} &= (m'_1)(m'_2)\ldots(m'_i+1)\ldots(m'_h), \\ C_{k-g_i} &= (m'_1)(m'_2)\ldots(m'_i-1)\ldots(m'_h), \end{aligned} \tag{21}$$

that is where the differences among C_k are due solely to differences in the number of combinations taken from the ith subset of genes. Here inequality (20) reduces to

$$1/[(k-1)+g_i(m'_i-1)] > s > 1/[(k-1)+g_i(m'_i+1)]. \quad (22)$$

The second case we consider is where

$$\begin{aligned} C_k &= (m'_1)(m'_2)\dots(m'_i)(m'_j)\dots(m'_h), \\ C_{k+g_i} &= (m'_1)(m'_2)\dots(m'_i+1)(m'_j)\dots(m'_h), \\ C_{k-g_j} &= (m'_1)(m'_2)\dots(m'_i)(m'_j-1)\dots(m'_h), \end{aligned} \quad (23)$$

that is, where differences in C_k involve two, the ith and jth, groups. Here inequality (20) reduces to

$$1/[(k-1)+g_j(m'_j-1)] > s > 1/[(k-1)+g_i(m'_i+1)]. \quad (24)$$

Table 4. Maximum number of components (C_k) in a 15-line multiline susceptible to each class of pathogen biotypes

See text for additional details

No. of virulence genes in pathogen biotype (k)	m'_1	m'_2	C_k	No. of virulence genes in pathogen biotype (k)	m'_1	m'_2	C_k
3	1	1	1	9	3	3	9
4	2	1	2	10	3	3	9
5	3	1	3	11	3	4	12
6	2	2	4	12	3	4	12
7	3	2	6	13	3	5	15
8	3 (or 2)	2 (or 3)	6				

From Table 3 it is also clear that a super-race carrying n genes for virulence will dominate the pathogen population if

$$[1-(n-1)s] > C_{n-g_i}[1-(n-g_i-1)s]/l, \quad (25)$$

where $C_{n-g_i} = (b_1)(b_2)\dots(b_i-1)\dots b_h$, and b_i is chosen such that C_{n-g_i} is a maximum. In this case

$$C_{n-g_i} = \prod_{i=1}^{h} m'_j = \left(\prod_{\substack{j=1 \\ i\neq j}}^{h} b_j\right)(b_i-1) = l(b_i-1)/b_i,$$

and C_{n-g_i} is maximized when b_i is maximized. As in the overlapping subsets case, this occurs when $1/b_i = g_i/n_i$ is minimized, so that

$$C_{n-g_i} = l[1-(g_i/n_i)_{\min}], \qquad 1 \leqslant i \leqslant h.$$

This reduces to

$$b_i[1-(n-1)s] > (b_i-1)[1-(n-g_i-1)s], \quad (26)$$

or

$$s < 1/[(n-1)+g_i(b_i-1)]. \quad (27)$$

Multiplicative fitness effects

The relative fitness of the various classes of pathogen biotypes under this model are given in Table 5. Here the simplest races with g genes for virulence will dominate the pathogen population if

$$(1-s)^{(g-1)}/l > 2(1-s)^{(g+g_h-1)}/l, \tag{28}$$

or

$$s > 1-(1/2)^{1/g_h}. \tag{29}$$

Complex races with k genes for virulence will dominate the pathogen population if

$$C_{k-g_i}(1-s)^{(k-g_i-1)}/l < C_k(1-s)^{(k-1)}/l > C_{k+g_i}(1-s)^{(k+g_i-1)}/l, \tag{30}$$

where $C_k = \prod_{i=1}^{h} m'_i$ and m'_i is chosen such that C_k is a maximum given $k = \sum_{i=1}^{h} m'_i g_i$ and $b_i \geqslant m'_i \geqslant 1$.

Table 5. Relative fitnesses of pathogen biotypes on an l-line multiline where each component has g genes for resistance in disjoint subsets of genes, under the multiplicative fitness effects model

$l = \prod_{i=1}^{h}(n_i/g_i) = \prod_{i=1}^{h} b_i$, where n_i is the number of resistance genes in the ith subset and g_i is the number used in each multiline component. g_i are ranked so that $g_1 \geqslant g_2 \ldots \geqslant g_h$

Pathogen biotype: Class	Pathogen biotype: No. of effective genes for virulence	Relative fitness on susceptible components	Max. No. of susceptible components	Mean fitness on multiline
1	g	$(1-s)^{(g-1)}$	1	$(1-s)^{g-1}/l$
2	$g+g_h$	$(1-s)^{(g+g_h-1)}$	2	$2(1-s)^{(g+g_h-1)}/l$
$m(=\prod_{i=1}^{h} m'_i)$	$k(=\sum_{i=1}^{h} m_i g_i)$	$(1-s)^{(k-1)}$	C_k[A]	$C_k(1-s)^{k-1}/l$
$l-1$	$n-g_i$	$(1-s)^{(n-g_i)}$	C_{n-g_i}	$C_{n-g_i}(1-s)^{(n-g_i)}/l$
l	$n(=\sum_{i=1}^{h} b_i g_i)$	$(1-s)^{n-1}$	l	$(1-s)^{n-1}$

[A] $C_k = \prod_{i=1}^{h} m'_i$ where m'_i is chosen to maximize C_k given the restraints $k = \sum_{i=1}^{k} m'_i g_i$ and $b_i \geqslant m'_i \geqslant 1$.

As before we consider two separate cases. In the first C_k is given by equation (21) in which case inequality (30) reduces to

$$1-[(m'_i-1)/m'_i]^{1/g_i} > s > 1-[m'_i/(m'_i+1)]^{1/g_i}. \tag{31}$$

In the second, C_k is given by inequality (22), and inequality (30) becomes

$$1-[(m'_j-1)/m'_j]^{1/g_j} > s > 1-[m'_i/(m'_i+1)]^{1/g_i}. \tag{32}$$

Finally, a super-race will develop under this model if

$$(1-s)^{(n-1)} > C_{n-g_i}(1-s)^{(n-g_i-1)}/l, \tag{33}$$

where $C_{n-g_i} = (b_1)(b_2)\ldots(b_i-1)\ldots(b_h)$, and the ith group is chosen such that C_{n-g_i} is a maximum. As before, this occurs when

$$C_{n-g_i} = l[1-(g_i/n_i)_{\min}], \qquad l \leqslant i \leqslant h.$$

We have, as a result,

$$(b_i)(1-s)^{n-1} > (b_i-1)(1-s)^{(n-g_i-1)}, \tag{34}$$

or

$$s < 1-[(b_i-1)/b_i]^{1/g_i}. \tag{35}$$

As in the case of combinations of overlapping subsets of genes, if s is equal to any of the critical values given in inequalities (22), (24), (31) and (32), the pathogen population will be polymorphic, and will contain two classes of pathogen biotypes in a neutral equilibrium.

Table 6. Levels of selection (s) required to stabilize the racial composition of pathogen populations
See text for additional details

No. of virulence genes (k) in dominant pathogen biotype	Level of selection when available resistance genes are incorporated into multiline components: Singly		g at a time, overlapping sets of genes		g at a time, combinations of overlapping subsets of genes	
	Additive model	Multiplicative model	Additive model	Multiplicative model	Additive model	Multiplicative model
			Example 1 ($g = 4$)			
4	>0·13	>0·20	>0·09	>0·16	>0·14	>0·29
⩽6	>0·08	>0·14	n.a.[A]	n.a.	>0·11	>0·29
⩽8	>0·06	>0·11	>0·05	>0·10	>0·08	>0·18
⩽10	>0·05	>0·09	n.a.	n.a.	>0·06	>0·18
⩽12	>0·4	>0·08	>0·04	>0·07	>0·05	>0·13
⩽14	>0·04	>0·07	n.a.	n.a.	>0·05	>0·13
16	<0·03	<0·06	<0·04	<0·07	<0·05	<0·13
			Example 2 ($g = 3$)			
3	>0·17	>0·25	>0·13	>0·20	>0·25	>0·50
⩽4	>0·13	>0·20	n.a.	n.a.	>0·17	>0·33
⩽5	>0·10	>0·17	n.a.	n.a.	>0·13	>0·25
⩽7	>0·07	>0·13	n.a.	n.a.	>0·08	>0·18
⩽9	>0·06	>0·10	>0·05	>0·09	>0·06	>0·13
⩽11	>0·05	>0·08	n.a.	n.a.	n.a.	n.a.
⩽12	>0·05	>0·08	<0·05	<0·09	>0·05	>0·11
13	<0·04	<0·08	n.a.	n.a.	<0·05	<0·11

[A] n.a., Not applicable.

Numerical results

Two numerical examples are given in Table 6 to illustrate more clearly the effects of the use of combinations of disjoint subsets of genes on the equilibrium racial composition of a pathogen population. In the first example we assume the breeder has 16 resistance genes at his disposal which he incorporates into multiline components either singly, four at a time in disjoint sets, or four at a time as combinations of disjoint subsets of genes (we assume $k = 2$, $n_1 = n_2 = 8$ and $g_1 = g_2 = 2$). In the second example, we assume $n = 13$ and calculate the critical levels of stabilizing selection for cases where multiline components carry a single gene for resistance, three genes for resistance as disjoint sets (plus one unused gene), or three genes for resistance as combinations of disjoint subsets (in the latter case we assume $k = 2$, $n_1 = 3$, $n_2 = 10$ and $g_1 = 1$, $g_2 = 2$). Here, in contrast with the situation with overlapping gene sets, the levels of selection required to ensure the predominant pathogen race carries k genes for virulence are generally higher when the genes are

used as combinations of disjoint subsets of genes than when they are used either singly or in disjoint sets.

Effectiveness in Disease Control

As emphasized previously, the important question facing the practical plant breeder considering the use of different sorts of multilineal varieties is not which variety will maintain the simplest pathogen biotypes. Rather, it is whether such varieties differ in their effectiveness in disease control. The answer to this question certainly depends on the number of virulence genes carried by the predominant pathogen biotype. However, it also depends on the total number of lines in the multiline (l) and the acceptable proportion of susceptible components (p), that is, the proportion of lines which can be susceptible to the predominant pathogen biotypes while the crop still escapes significant disease damage.

Table 7. Levels of selection (s) required for adequate disease control with the breeding strategy of a single resistance gene per component (*1*), g genes per component in overlapping gene sets (*2*), or g genes per component as combinations of overlapping subsets of genes (*3*)

Examples 1 and 2 are those of Table 2

Breeding strategy	Level of selection for an acceptable proportion of susceptible host plants (p) of:					
	0·10–0·15		0·20–0·25		0·40–0·45	
	Example 1	Example 2	Example 1	Example 2	Example 1	Example 2
			Additive model			
1	>0·50	n.a.[A]	>0·50	>0·50	>0·17	>0·25
2	>0·25	>0·25	>0·25	>0·18	>0·18	>0·18
3	>0·30	>0·25	>0·25	>0·25	>0·17	>0·17
			Multiplicative model			
1	>0·50	n.a.	>0·50	>0·50	>0·25	>0·33
2	>0·50	>0·75	>0·50	>0·60	>0·40	>0·60
3	>0·67	>0·67	>0·67	>0·67	>0·67	>0·50

[A] n.a., Not applicable.

To evaluate the relative effectiveness of different types of multilines in disease control, we have calculated, for the examples given in Tables 2 and 6 for a range of values of p, the levels of selection against unnecessary genes for virulence required to stabilize the racial composition of the pathogen population and provide adequate disease control (Tables 7 and 8). Clearly, the lower the level of selection required to achieve a particular objective, the greater the likelihood that a breeder will have the necessary resistance genes at his disposal. The values of s required to achieve adequate disease control for the alternative breeding strategies and the two examples presented in Table 2 are given in Table 7. For the additive model, lower levels of selection are required to achieve adequate disease control if the resistance genes are combined in overlapping sets rather than used singly. Further, the use of overlapping subsets of genes appears to be as effective from this point of view as the use of all possible gene combinations. However, for the multiplicative model, greater levels of selection are required to ensure adequate disease control with more than one gene per component. Again, there is little difference between breeding strategies 2 and 3

in the critical values of s. These findings confirm a previous conclusion (Marshall and Pryor 1979) that the use of overlapping gene sets will only be of practical importance if it can be demonstrated unequivocally that unnecessary genes for virulence in pathogens act in an additive, or near additive, fashion. The levels of selection required to achieve adequate disease control where resistance genes are used singly or as disjoint gene sets, using the two examples given in Table 6, are compared in Table 8. Here the levels of selection necessary to achieve this objective are lower, under the additive fitness model, or very similar, under the multiplicative fitness model, for breeding strategy 3 compared with breeding strategies 1 and 2.

Table 8. Levels of selection (s) required for adequate disease control with the breeding strategy of a single resistance gene per component (*1*), g genes per component as disjoint gene sets (*2*) or g genes per component as combination of disjoint subsets of genes (*3*)

Examples 1 and 2 are those of Table 6

Breeding strategy	Level of selection for an acceptable proportion of susceptible host plants (p) of: 0·10–0·15		0·20–0·25		0·40–0·45	
	Example 1	Example 2	Example 1	Example 2	Example 1	Example 2
			Additive model			
1	>0·25	>0·25	>0·13	>0·17	>0·07	>0·10
2	n.a.[A]	n.a.	>0·09	>0·13	>0·09	>0·13
3	>0·11	>0·17	>0·08	>0·13	>0·06	>0·08
			Multiplicative model			
1	>0·33	>0·33	>0·20	>0·25	>0·13	>0·17
2	n.a.	n.a.	>0·16	>0·20	>0·10	>0·20
3	>0·29	>0·33	>0·18	>0·25	>0·18	>0·18

[A] n.a., Not applicable. These multilines contain only four components.

Discussion

The procedure described here, *viz*, the use of combinations of overlapping subsets of genes or combinations of disjoint subsets of genes, offers alternative means of developing multiline varieties in which the components carry two or more genes for resistance. Further, multilines developed in this way would, from the analysis given in Tables 7 and 8, appear to be as effective in long-term disease control as varieties developed using either overlapping or disjoint gene sets *per se*. Together with our earlier findings (Marshall and Pryor 1979), the present results indicate that the use of more than one resistance gene per multiline component is clearly advantageous only if unnecessary genes for virulence in pathogens act additively to reduce the fitness of their carriers. If unnecessary virulence genes reduce fitness multiplicatively, then multilines composed of multiresistant components have little or no advantage (in the case of disjoint sets of genes) or are disadvantageous (in the case of overlapping sets of genes) in long-term disease control compared with multilines in which the components each carry a unique single gene for resistance.

However, we must emphasize again that these conclusions are based on the use of simple models which embody many arbitrary and often unrealistic assumptions. In particular, we have assumed that unnecessary genes for virulence in the pathogen act either additively or multiplicatively to reduce the fitness of their carriers. Barrett

and Wolfe (1978) criticized the use of additive models in describing pathogen evolution on multiline varieties on the grounds that they represent a special case of the general multiplicative model for very small values of s. Although this is true, as shown both here and in a previous paper (Marshall and Pryor 1979) for large values of s, the way in which unnecessary genes for virulence combine to reduce the relative fitness of complex pathogen races substantially affects the potential of multiline varieties in terms of long-term disease control, particularly in the case of overlapping gene sets. Consequently, in this situation, the additive model cannot be regarded as a special case of the multiplicative model. Further, the exclusive use of multiplicative models may lead to erroneous conclusions concerning the value of multilines in disease control. The latter point stresses the urgent need for experimental data on the ways unnecessary virulence genes interact to reduce the fitness of complex pathogen races, as well as other aspects of stabilizing selection, before multiline varieties can be regarded as a viable alternative to other forms of disease control.

References

Barrett, J. A., and Wolfe, M. S. (1978). Multilines and super-races—a reply. *Phytopathology* **68**, 1535–7.

Flor, H. H. (1956). The complementary genic systems in flax and flax rust. *Adv. Genet.* **8**, 29–54.

Groth, J. V. (1976). Multilines and 'super-races': a simple model. *Phytopathology* **66**, 937–9.

Groth, J. V., and Person, C. O. (1977). Genetic inter-dependence of host and parasite in epidemics. *Ann. N.Y. Acad. Sci.* **287**, 97–106.

Marshall, D. R. (1977). The advantages and hazards of genetic homogeneity. *Ann. N.Y. Acad. Sci.* **287**, 1–20.

Marshall, D. R., and Pryor, A. J. (1978). Multiline varieties and disease control. I. The 'dirty crop' approach with each component carrying a unique single resistance gene. *Theor. Appl. Genet.* **51**, 177–84.

Marshall, D. R., and Pryor, A. J. (1979). Multiline varieties and disease control. II. The 'dirty crop' approach with components carrying two or more genes for resistance. *Euphytica* **28**, 145–59.

Van der Plank, J. E. (1963). 'Plant Diseases: Epidemics and Control.' (Academic Press: New York and London.)

Van der Plank, J. E. (1968). 'Disease Resistance in Plants.' (Academic Press: New York and London.)

Manuscript received 11 March 1980, accepted 16 October 1980