BU-1117-M

Application of Genetic Theory in Breeding for Multiple Viral Resistance[1]

Brian T. Scully[2]

Assistant Professor of Vegetable Crops
IFAS, University of Florida, Everglades Research Center,
Belle Glade, FL 33430

and

Walter T. Federer

Liberty Hyde Bailey Distinguished Professor of Biometrics
Biometrics Unit, Department of Plant Breeding and Biometry,
Cornell University, Ithaca, NY 14853-1902

---

1 Running Title: Breeding for multiple viral resistance

2 Formerly: Texas A&M University

Introduction

Among the broad array of disease agents that affect crop plants, viruses are potentially some of the most devastating. Because viruses are obligate parasites and are often insect-vectored, the methods used to control these diseases differ from the methods used to to control fungal and bacterial disorders. Viral diseases are commonly managed in three ways, including regulatory measures, pesticides targeted at vectors or alternative hosts, and/or genetic resistance. Regulatory controls such as indexing, crop free periods, and the elimination of alternate hosts have proven useful but are not guaranteed. Pesticides can augment the regulatory process by reducing the viral vectors and alternate hosts, although this is expensive, usually not effective, and raises environmental concerns. In nature, genetic resistance provides the best defense against plant disease, and this mechanism remains the safest and most economical approach for crop protection (Browning 1980). Genetic resistance is defined as any heritable trait that reduces the effect of the virus (Russell 1978). The highest level of resistance results from a lack of recognition between the pathogen and host and the subsequent inability of the virus to reproduce in the host (Gracen 1982). Tolerance provides a lower and less acceptable level of resistance, primarily because the virus replicates in the host organism. Tolerant plants show less severe disease symptoms, less damage to the economic organs, and/or suppression of yield by the pathogen (Russell 1978). Morphological traits that limit a vector's ability to transmit a virus mechanically also contribute to host plant resistance.

To develop genetically resistant varieties, resistant germplasm must first be identified. Natural populations provide the original

genetic resource from which to extract, identify, and characterize the genes that confer resistance. Locally adapted and foreign varieties are initially screened for resistant genes, followed by land races and wild accessions from the crop's center of diversity or origin. Because crop and disease often co-evolve, genetically diverse forms of resistance are often available from one of these germplasm sources. If resistance is not found within the species of interest, methods such as gene manipulation or transfer from related species are required.

Once resistant genotypes are identified, the inheritance of resistance must be determined. The mode of inheritance and the crop's reproductive biology establish the selection technique and breeding method most likely to maximize genetic gain from selection. Oligogenically inherited traits are more easily handled than polygenic traits, which require the use of more complex selection procedures (Mayo 1987). The choice of a selection procedure must also be guided by the genetic variability of the pathogen (Day 1974, Simmonds 1979). A basic understanding of viral epidemiology, as well as the mechanisms of resistance and pathogenicity, is useful in the development of resistant varieties. Ultimately, the virus and variety with the resistant genes must interact freely in a cropping system. Clearly, the final success of a resistant variety is judged by the productivity of a crop under field conditions.

Our purpose in this chapter is to compare selection techniques and breeding methods useful for the development of resistant varieties. We review successful breeding methods and present new permutations of these schemes. Breeding for resistance to one disease is relatively straightforward; the difficulty arises in breeding for multiple viral resistance (Khush 1980; Provvidenti 1985). Flexibility in the choice

and application of breeding methods is needed to "pyramid" resistant genes into breeding lines and varieties quickly and easily (Nelson 1973). We apply the theory that supports selection for multiple quantitative traits to selection procedures for multiple viral resistance. Multiple trait selection techniques are associated with a number of possible breeding methods. We develop these concepts for the transfer of mono/oligogenically inherited traits in autogamous diploid species. Additionally, we present an appropriate set of equations for the determination of population or sample size. An effective population size maximizes the probability of successfully finding a desired genotype, without demanding excessive resources.

Selection Techniques

Incorporation of resistance to a single virus is accomplished via selection over one or more generations. In contrast, breeding for multiple viral resistance presents three other possible approaches to selection, including tandem selection, independent culling levels, and the selection index (Turner and Young; 1967 Baker 1986). These techniques were developed primarily for quantitative traits and are used regularly in most plant and animal breeding programs; however, they also provide the theoretical foundation for the selection of qualitatively inherited traits.

In tandem selection, a single trait is selected discretely for one or more generations until the desired phenotype is obtained. Selection is then applied for a second, third, or more traits in the same way. For mono/oligogenically inherited traits, a single generation of selection should suffice to obtain the desired phenotype, provided the population

size is adequate. Homozygous genotypes are subsequently confirmed in a progeny test generation. With independent culling levels, two or more traits are selected concurrently over every generation of selection until the desired phenotype is obtained. Like tandem selection, independent culling levels can be applied to either oligogenic or polygenic traits. Variations in these two selection method include: selection for a set of multiple traits jointly in a single generation; and concurrent selection for subsets of two or more traits with each subset in a different generation. The selection index was developed specifically for quantitative traits (Smith 1936). It requires the use of genetic variances, covariances, and economic weights to rank and select individuals. Qualitatively inherited traits are occasionally incorporated into modified selection index models as "categorical" traits (Van Vleck 1979).

In breeding for multiple viral resistance, the biology of the virus and its interaction with the host species also affect the choice of a selection technique. Cross protection, synergism, variations in symptomology, and escapes influence selection for multiple viral resistance. Cross protection resulting from simultaneous inoculations of closely related viruses can confound resistant and susceptible phenotypes, thus affecting the accuracy of selection. In addition, the interaction of different viruses with the different genetic backgrounds of the host can cause symptomology to vary and become an unreliable criteria for selection. Synergism among viruses may require simultaneous inoculations with two or more viruses to identify resistance to a viral complex. When synergism exists, inoculation with the viral complex and the single viruses can clarify their pathogenicity

and interaction. Tandem selection removes the problem of cross protection but fails to address synergism, although resistance to a viral complex can be treated as a distinct trait. Inoculation of genetically different plants with a single virus does not remove possible differences in symptomology, but it does clarify the variation. Lastly, escapes affect all selection techniques, but can be detected with repeat inoculations within a generation or over successive generations.

No particular technique for multiple trait selection is best for the development of multiple viral resistant genotypes. The three approaches presented above are flexible; perhaps variations or different combinations of these techniques in different generations would yield the desired result. Tandem selection is the easiest, simplest, and the least complicated, under most circumstances.

Breeding Methods

In self-pollinated crops, the development of disease-resistant varieties is routinely accomplished with the backcross and pedigree methods (Allard 1960). Backcross methods rapidly introgress specific resistance genes from a donor parent into a desirable variety, concurrently reconstructing the original variety. A new variety is produced quickly, requires minimal testing and is nearly isogenic with the original. Backcross methods are particularly desirable when the donor parent is unadapted or genetically distant from the original variety. When disease problems arise unexpectedly, they are the methods of first choice.

Pedigree methods such as F_2 selection, single seed descent (Brim 1966; Empig and Fehr 1971), and the nested hierarchy (Cockerham 1954) are also useful methods for incorporating disease resistance. Pedigree methods are useful when both parents are adapted and carry genes for resistance to different diseases. Compared with the backcross, a more genetically diverse germplasm base is developed, but the time required to release a variety is much longer. In a pedigree method, recombination and segregation among the unselected genes permits the development of unique genotypes and phenotypes. The backcross and pedigree methods are forms of inbreeding that ultimately result in the development of pure lines. The level of homozygosity ($\mathcal{H}$) increases with each generation of selfing or backcrossing, and is defined as:

$$\mathcal{H} = \{(2^g-1)/(2^g)\}^{\mathcal{L}} \qquad [1]$$

where g is the number of backcrossed or selfed generations (for F_2, $g=1$; F_3, $g=2$, etc.); and $\mathcal{L}$ is the number of loci under selection (Allard 1960). The frequency of the desirable homozygous genotypes in any selfed generation is defined as $0.5\mathcal{H}$, in the absence of linkage.

Backcross Methods

There are numerous permutations of the backcross method; each addresses a given breeding objective. These permutations are divided into three categories according to the inheritance and number of traits under transfer. For resistance to a single virus conditioned by a dominant gene, the simple backcross is the standard method (Figure. 1). When resistance to a single virus is conferred by a recessive gene the

alternate backcross and self (Figure 2), continuous backcross (Figure 3), and the simultaneous backcross and self (Figure 4) methods are acceptable. For the transfer of resistance to multiple viruses, methods such as the sequential backcross (Fig 5), parallel backcross (Figure 6), and the multiple trait backcross (Figure 7) are effective.

The simple backcross procedure transfers a dominant gene from a single donor parent (D_1) to the original variety or recurrent parent (R) (Figure 1). The F_1 generation is composed strictly of resistant heterozygous individuals. The backcross generations are composed of susceptible homozygous recessive and heterozygous resistant individuals at expected frequencies of 1/2 each. The heterozygous resistant individuals are tested for, and selected in each backcross generation (BC-1 --> BC-v), then crossed to the recurrent parent. Heterozygous individuals are selected in the final backcross generation and selfed to produce homozygous and heterozygous resistant, and homozygous susceptible progeny at expected genotypic frequencies of 1/4, 1/2, and 1/4, respectively. Homozygous and heterozygous resistant individuals from this generation are confirmed and separated in the last generation (Figure 1). ☞ {Insert Figure 1}

The alternate backcross and self procedure transfers a recessive allele from the donor parent (D_1) to the recurrent parent (R) (Figure 2). In this case, each backcross generation is composed of homozygous and heterozygous susceptible individuals at expected genotypic frequencies of 1/2 each. A selfed generation is included after each backcross generation to reveal the recessive individuals. If more than 5 ($\underline{n} \geq 5$) backcross individuals are selfed to produce a backcross F_2 generation there is a ≥95% (P_α) chance that at least one individual is a

heterozygous carrier of the recessive allele (Table 1). Backcross F_2 families derived from each of the backcross individuals are tested for resistance, and a recessive individual selected for crossing to the recurrent parent. Homozygous recessive individuals are confirmed in the last generation (Figure 2). ☞ {Insert Figure 2}

The continuous backcross method transfers a resistant gene without the inclusion of a selfing generation (Figure 3). It applies for the transfer of a recessive allele masked in the heterozygous condition, but is not strictly limited to this case. In this method, the identification and selection of the desired genotype is delayed until the end of the breeding program. To compensate for the absence of selection, the number of crosses and population size is increased. If $\underline{n}$ individuals from the BC-1 generation are crossed to the recurrent parent, the probability that any one of these individuals is a heterozygous or homozygous equals 1/2. For a 95% chance of randomly selecting a heterozygote for a cross in BC-1, n equals 5 plants. To maintain this probability through repeated backcross generations, $\underline{n}$ more individuals are derived from each of the previous generation's $\underline{n}$ individuals. Thus, the number of plants increases exponentially as $\underline{n}$, $\underline{n}^2$, $\underline{n}^3$, $\underline{n}^4$, ... $\underline{n}^v$. By BC-5, $\underline{n}^5$ or 3125 individuals are required to maintain a probability of success (P_α) at 95%. As the number of individuals increases exponentially ($\underline{n}^v$), the resources required also increase compared to the alternate backcross and self procedure. More realistically, only a few backcross generations are done continuously before selfing to uncover the recessive gene. The process is again repeated until the trait is transferred. This approach is convenient when a few backcross generations are produced in the off season and the selfed generation planted in the field. ☞ {Insert Figure 3}

The simultaneous backcross and self method (Figure 4) is a permutation of the backcross that primarily transfers a recessive gene, but it requires less labor and space than the continuous backcross. Like the continuous backcross, the number of individuals (n) used for the crosses in each generation determine the probability that a heterozygous individual is chosen. In every backcross F_2 generation, the crosses with the heterozygous individuals are revealed and the homozygous crosses discarded. If partial, over- or co-dominant gene action controls a trait's expression, then the heterozygous individuals are more easily distinguished and these recessive gene methods may be unnecessary (see Munger, this volume). ☞ (Insert Figure 4)

The primary advantage of the continuous backcross and the simultaneous backcross and self method is that a breeding objective is attained much faster than is possible with the alternate backcross and self procedure (Figures 2, 3 and 4). Additionally, more crosses and the larger population sizes in these backcross schemes provide greater opportunity to break linkage, recover the original phenotype quickly, and find exceptional recombinants. There is a greater likelihood of these events in the continuous backcross because the number of crosses increases by the power of n compared to the simultaneous backcross and self scheme in which n remains constant.

The sequential backcross is a set of simple backcross cycles run in succession with each trait incorporated separately over time (Figure 5). From different donor parents (D_1, D_2, D_3, etc.), dominant or recessive genes can be transferred. Selection is practiced in the backcross generations (BC) for dominant genes, but recessive genes require a selfing generation. At the end of each cycle, a new donor

parent is incorporated into the breeding program and the line or variety derived from the previous backcross cycle becomes the recurrent parent. The process repeats with as many donor parents as needed. Progeny tests can be performed after each backcross cycle, and prior to the introduction of a new donor parent, or after all traits are incorporated. This method is the slowest of the multiple trait backcrosses but commonly used to deploy resistance genes for newly discovered diseases. ☞ {Insert Figure 5}

The parallel backcross technique is a set of simple backcrosses performed concurrently (Figure 6). Donor parents, each with a different dominant gene, are backcrossed to the same recurrent parent until the desired phenotype is obtained. At some point in the scheme, individuals with the desired trait from each donor are mated to produce the F_{1}' generation. The F_{1}' generation is used to produce the double cross F_{1}'' generation, which is selfed to produce the F_2 (Figure 6). Putative homozygous individuals are identified in the F_2, and F_3 families are produced. These F_2-derived F_3 families are partitioned into subfamilies, and homozygosity is confirmed for each trait (Figure 6). If viruses cross protect, multiple trait selection is delayed until the F_3. ☞ {Insert Figure 6}

In the parallel and sequential backcross methods (Figures 5 and 6), single traits are initially transferred independently and selected discretely in each backcross generation. In the parallel backcross, all traits are ultimately combined and selected together in the F_3 generation. Both methods are presented for the transfer of dominant genes, but can be modified for the transfer recessive genes. In the parallel backcross, selfing is integrated into the basic procedure.

Insertion of selfing generations between the last backcross and the F_1' generations and between the F_1' and the F_1'' generations permits the identification and selection of the recessive genotypes. Throughout the parallel backcross, the recurrent parent is consistent and the coancestry of the progenies converge on the recurrent parent. In the sequential backcross the recurrent parent changes with each cycle of backcrosses, and the progenies genetically diverge from the original recurrent parent, particularly as the number of cycles becomes large.

In the multiple trait backcross (Figure 7) genes are transferred jointly and selected in a single generation. The number of traits transferred is limited by the donor parent and the population size. As the number of genes under transfer increases, the population size must increase commensurately. This method is an extension of the simultaneous backcross and self method (Figure 4) except that the n backcrossed F_2 families are partitioned into subfamilies and evaluated for each trait. Partitioning is not needed for morphological traits, but is advised for identification of viral resistance genes. Unlike the sequential or parallel backcross, no modification of this method is required to handle recessive genes. ☞ (Insert Figure 7)

All of these backcross methods are generalized schemes that can be modified or combined to meet breeding objectives; all are subject to the usual assumptions assigned to the backcross (Simmonds 1979; Allard 1960). Every trait should have easily distinguishable classes (qualitative distribution), and the genes that condition a trait should be highly penetrant (Suzuki et al. 1985). The single trait methods provide a framework upon which the multiple trait backcross methods are constructed. With the multiple trait methods, greater resources and record keeping are required, but the breeding goal is accomplished more

quickly. The expression of a trait prior to flowering is desirable in certain applications of the backcross, but irrelevant in the recessive gene methods (Figures 2, 3, and 4) or the multiple trait backcross (Figure 6). In many viral resistance breeding programs, inoculation is done early in the plant's life cycle and susceptibility determined before flowering. Use of both backcross and simultaneous self methods (Figures 4 and 7) assumes multiple flowers, although maize (Zea mays L.) breeding techniques allow the same female inflorescence to be selfed and outcrossed (Sheridan and Clark 1987).

The continuous, parallel, and the simultaneous self and backcross methods were developed and refined by Henry M. Munger and identified as "Munger's permutations". In addition to these backcross methods, there is the double backcross (Walkof 1955, 1961); the inbred backcross (Wehrhan and Allard 1965; Dudley 1982; Cox 1984) and the congruity or alternating interspecific backcross (Haghighi and Ascher 1988; Barker et al. 1989; Superak and Scully, this volume). The double backcross is used to break linkage between two negatively correlated traits. The inbred backcross was initially designed to count genes, but it is now used to introgress needed genes from unadapted germplasm and improve adaptation in wide crosses. The congruity or interspecific backcross develops genetic bridges, increases fertility among interspecific crosses, or transfers desired genes from different species.

Pedigree Methods

Many of the varieties now in production have resistance to one or more viruses and could be intermated to produce breeding lines with a

broad spectrum of viral resistance. Selection for multiple viral resistance in the F_2 is the quickest way to develop a broad spectrum of resistance. Large populations are generally required with this method because the expected frequency of the desired homozygous genotype is only 1/4 for any single locus. As the number of required independent loci () increases, the frequency of desirable genotypes decreases by $(1/4)^{L}$ (Eq. 1). Cross protection and diverse symptoms also make selection in the F_2 inaccurate, but F_3 progeny tests should reveal any inaccuracies.

Single seed descent (Figure 8) is based on the principle of equal fecundity. Two parents are mated to produce an F_2 base population, but only one offspring is derived from each F_2 individual and carried forward to the F_3 generation. Likewise, a single progeny from each F_3 is contributed to the F_4, and so on. Thus, in all future generations of inbreeding, every individual traces back to a single and different F_2 progenitor (Brim 1966; Empig and Fehr 1971). This practice distinguishes three unique properties of single seed descent: constant allele frequency, constant population size, and changing genotype frequencies over generations. The proportion of homozygosity (H) increases as g (Eq. 1) increases, whereas the proportion of heterozygosity decreases as $1-H$. The proportion of homozygosity for all loci is 1 at F_∞. The original purpose of single seed descent was to develop inbred lines rapidly with a minimum of variation within lines, but maximum variation among lines. For a quantitative trait, the probability of finding a transgressive segregate increases compared to the pedigree method, in which selection is intense in the early generations. In single seed descent, selection is usually practiced in

F_7 or F_8 when "approximate homozygosity" is attained, although this is affected by the number of loci that condition a trait and their interaction (Snape and Riggs 1975; Hallauer and Miranda 1981). ☞ {Insert Figure 8}

In breeding for multiple viral resistance, F_2-derived F_g families are divided into subfamilies and each inoculated with a different virus (Figure 8). Resistance is then confirmed with a progeny test in F_{g+1} (Figure 8). Like the F_2 selection scheme, multiple virus resistance is identified in a single generation (F_g), with the work concentrated at the end of the program. Single seed descent is a slower process than F_2 selection, but requires a smaller population, and produces genetically pure lines.

The single seed descent method can be modified so that multiple traits are selected separately in each generation (F_2, F_3, F_4, etc.) (Figure 9). In breeding for multiple viral resistance, the F_2 generation is inoculated with one virus and the resistant individuals carried forward to the F_3. In the F_3 generation a second virus is inoculated and resistant individuals forwarded to the F_4. This process is repeated until all the desired traits are incorporated. At the end of the program, the number of lines to progeny test is much smaller than in either the F_2 or single seed descent schemes. The work is spread over all generations rather than concentrated in a single generation. This method is as slow as single seed descent, but there are fewer individuals to test with each succeeding generation. For the same probability of success, this method requires an initial F_2 population larger than single seed descent but smaller than F_2 selection.

☞ {Insert Figure 9}

The nested hierarchy (Cockerham 1954; Horner and Weber 1956;

Wricke and Weber 1986) is a mating design originally developed to partition genetic variances in self-pollinated crops, but may be a useful breeding method for incorporating multiple viral resistance (Figure 10). It is similar to single seed descent and Goulden's (1939) modified pedigree method, and is perhaps best described as double, triple, or quadruple (etc.) seed descent. The number of divisions chosen is flexible, as is the number of times these divisions are made. This method differs from single seed descent in that the population grows geometrically larger as a function of the number and size of the divisions. Inbreeding proceeds at the same rate. Given the same objectives as a single seed descent program, the nested hierarchy should produce the same results. The size of the original F_2 population can be smaller than in single seed descent; however, it is absolutely critical that all needed genotypes or alleles be represented in the F_2 generation. In the nested hierarchy, like single seed descent, multiple viral resistance is selected in the final generation, and progeny tested in the same way. ☞ (Insert Figure 10)

The nested hierarchy can be modified in a way analogous to the modified single seed descent. A single trait can be selected in each generation of inbreeding (Figure 11) until all resistant genotypes are identified. The size of the F_2 population will be larger than the nested hierarchy, but will not increase geometrically over generations because selection is applied in each generation. The number of lines tested in the final generation is smaller than in the single seed descent or nested hierarchy, given the same breeding objectives. ☞ (Insert Figure 11)

These pedigree and backcross methods are not presented at the exclusion of the bulk method, which is an easier and less expensive way to deal with large populations. In environments where viral diseases

are ubiquitous and predictable year after year, bulk methods provide an opportunity to select for both viral resistance and environmental adaptation. These pedigree methods are intended to be flexible and serve as a guide for more creative breeding schemes that include selection for adaptation and other desirable traits. Breeding methods that fail to incorporate selection for adaptation and horticulturally important traits are likely to result only in the release of germplasm and breeding lines, rather than finished varieties.

Population Size

Binomial Distributions

The success of a particular selection technique and breeding method is a function of the population size, which should maximize the probability of success without excessive demands on time and resources. In this section we present a set of working equations that allow breeders to determine population size ($\underline{n}$) and estimate the probability (P_α) that one, two or more plants have the desired genotype ($\underline{r}$) at an expected or constant genotypic frequency (f_d). These equations are drawn from binomial theory and are applicable to four selection situations. Selection for a single trait in a single generation provides the basic model and is used to construct three other multiple trait selection equations. Corollaries to the basic model include selection for different single traits separately in different generations; multiple traits jointly in a single generation; and subsets of two or more traits, each in different generations.

The use of binomial equations assumes that each trait is simply

inherited, mutually exclusive, and qualitatively distributed with only two categories. These categories are based on either phenotypic, genotypic, or allelic frequency. In breeding for viral resistance in autogamous species, selection is ultimately practiced for the genotype, so genotypic frequency is mostly used to determine a population's size. Equations based on genotypic frequency are often more conservative than those based on either phenotypic or allelic frequency, and require larger populations. Conversely, population sizes determined by phenotypic or allelic frequency are commonly smaller, but can also include unwanted genotypes within the desired category. Phenotypic and genotypic frequencies are equal when a trait is conditioned by additive gene action, but nonadditive gene action skews phenotypic frequencies away from genotypic frequencies. For allelic frequencies, the probability of finding an individual with at least one copy of each desired allele is much higher than finding a unique genotype, particularly for polygenic traits (Sneep 1977). As the number of loci that condition a trait increases, differences in population sizes between allelic and genotypic based models become large.

The assumption of only two genotypic categories is valid for a single gene in a backcross generation, but inappropriate for F_2 populations segregating for mono/oligogenically inherited traits. These populations have three or more genotypes and follow multinomial distributions. However, these distributions are easily collapsed into two categories that include a single desired genotype in one category and all other genotypes in the second category. By collapsing populations into two categories differences in gene action, inheritance, penetrance, and other genetic phenomena are more easily managed.

Application of binomial theory also requires that the inheritance of resistance be determined *a priori* to meet the assumption of known probability. This determination of expected genotype frequencies is essential for the assignment of genotypes to one category or another.

In cumulative binomial distributions, the probability of obtaining r plants of the desired genotype in a population of size n is defined by the expression (Larsen and Marx 1985, Mosteller et al. 1961):

$$P_\alpha = 1 - \sum_{x=n-r+1}^{n} \binom{n}{x} (f_d)^{n-x} (f_u)^{x} \qquad [2]$$

where: P_α is the selected probability of success; n is the number of plants in the population; r is the number of individuals with the desired genotype, with $x=n-r+1$ as index of undesirable genotypes; f_d is the expected frequency of the desirable genotype; and f_u is the expected frequency of the undesirable genotype.

Because the population is defined as a group of desirable and undesirable genotypes with constant frequency, then:

$$f_u + f_d = 1. \qquad [3]$$

The binomial coefficients $\binom{n}{x}$ in Eq. 2 are rewritten as:

$$P_\alpha = 1 - \sum_{x=n-r+1}^{n} [n!/\ (r-1)!\,(n-r+1)!]\ (f_d)^{r-1}\ (f_u)^{n-r+1} \qquad [4]$$

with components defined above. For a single trait in a single generation, the probability of having *at least* one ($r \geq 1$) desired individual in a population of size n is defined by Snedecor and Cochran (1981) as:

$$P_\alpha = 1-(f_u)^{n} \qquad [5]$$

or more commonly as:

$$P_\alpha = 1 - \{1 - (f_d)\}^{\underline{n}} \quad [6]$$

as adapted from Eq. 4. For the probability of *at least* two ($r \geq 2$) desired individuals:

$$P_\alpha = 1 - \{(f_u)^{\underline{n}} + [\underline{n}\,(f_d)\,(f_u)^{\underline{n}-1}]\}; \quad [7]$$

for *at least* three ($r \geq 3$) desired individuals:

$$P_\alpha = 1 - \{(f_u)^{\underline{n}} + [\underline{n}\,(f_d)\,(f_u)^{\underline{n}-1}] + [((\underline{n}^2-\underline{n})/2)\,(f_d)^2\,(f_u)^{\underline{n}-2}]\} \quad [8]$$

and so on, as Eq. 4 is expanded. Because most breeders require a minimum of one individual with the desired genotype, Eqs. 5 and 6 are acceptable and easily solved for $\underline{n}$. When $\underline{r} \geq 2$, no closed-end solution exists for $\underline{n}$ in Eqs. 7 or 8. Values for $\underline{n}$ can be extracted from a summed binomial distribution table, but simpler tables and computer programs that provide n given P_α and f_d are available (Table 1) (Mansur et al. 1990; Sedcole1977; Harrington 1952). ☞ [Insert Table 1]

Sedcole (1977) provides three computational methods to approximate population size when $\underline{r} \geq 2$. The simplest method multiplies $\underline{n}$ for the probability of obtaining one of the desired genotype, by $\underline{r}$. If 11 plants are needed in the F_2 for the probability of one individual, then 22 plants are required for $\underline{r} \geq 2$; 33 for 3 etc. This approximation clearly overestimate population size (see Table 1). A fourth method that approximates $\underline{n}$ is defined for any given r ($\underline{n}'_{\underline{r}}$), as:

$$\underline{n}'_r = \underline{n}_1 + [(r-1)\underline{n}_1/2] \quad [9]$$

where $\underline{n}_1$ is the population size for a minimum of one desired individual as computed from Eq. 5 or 6; and $\underline{n}'_{\underline{r}}$ is the approximate value of $\underline{n}$ given $\underline{r}$, with $\underline{r}$ defined previously. As an example, consider a recessive gene in the F_2, with a P_α = 95% and f_d = 0.25; the population size for $\underline{r} \geq 1$ is 11 plants (Table 1). For $\underline{r}$ = 2, 3, 4, 5, 6, 8 10 and 15, populations

are approximated as $\underline{n}$'=17, 22, 28, 33, 39, 49, 61, and 84, respectively. This technique more closely approximates n than Sedcole's simplest procedure. In general, this method overestimates n at P_α = 99%, but underestimates n for P_α = 95%, at r ≤ 6, and overestimates n as r becomes greater. These approximating methods are provided for frequencies (f_d) not covered in Table 1.

Single Trait: Single Generation

Selection for a monogenically inherited trait in a single generation is easily accomplished by a backcross method or selection in the F_2 generation. Consider an example with the alternate backcross and self procedure (Figure 2), in which a single recessive gene confers resistance to a particular virus. In the F_2 derived from each backcross generation, f_d = 0.25 and f_u = 0.75. Assuming the need for one ($r \geq 1$) homozygous recessive individual with a 95% (P_α) chance of success, the minimum population size (n) is computed with Eq. 6, such that

$$0.95 = 1 - \{1 - (0.25)\}^n$$

thus:

$$\underline{n} = \ln (1-P_\alpha)/\ln (1-f_d) = \ln (0.05)/\ln (0.75) = 10.41 \approx 11 \text{ plants.}$$

Eleven plants actually give a 95.78 % chance of success. Based on the genotypic frequency (f_d), two or three plants of the 11 (11*0.25) are expected to be homozygous recessive, and resistant to the virus. The chance (P_α) of obtaining these two or three resistant individuals in a population of $\underline{n}$=11 is 80% and 54.5%, respectively. If two or more ($\underline{r} \geq 2$) resistant individuals are required at a P_α = 95%, then $\underline{n}$ is available

from Table 1.

Multiple Traits: Single Generation

This basic probability equation can be extended to different selection problems. For two or more independent (i.e. unlinked) traits (j = 1 to t) selected in a single generation, the probability that one or more individuals possesses these traits is the product of the f_{dj} ($f_{d1} * f_{d2} * \ldots * f_{dt}$). For multiple viral resistance, the probability of obtaining a minimum of one individual with all desired traits is defined as

$$P_\alpha = 1 - \left\{1-(f_{d1} * f_{d2} * f_{d3} * \ldots * f_{dt})\right\}^n, \qquad [10]$$

where P_α and n are defined above; and $f_{d1}, f_{d2}, f_{d3}, \ldots f_{dt}$ are the expected genotypic frequencies for each trait $j = 1 \ldots t$ in a single generation. Thus:

$$P_\alpha = 1 - \left\{1-\left(\prod_{j=1}^{t} f_{dj}\right)\right\}^n \qquad [11]$$

This equation is appropriate for F_2 selection, single seed descent, the parallel and multiple trait backcross procedures (Figures. 6, 7 and 8). As an example consider four viral diseases with resistance to each conditioned by a single dominant gene, and selection practiced in the F_2. The expected frequency of the homozygous dominant individuals is $(1/4)^4$ or 1/64. The probability that *at least* one individual has the desired genotype is:

$$0.95 = 1 - \left\{1-(0.25 * 0.25 * 0.25 * 0.25)\right\}^n$$

thus

$$\underline{n} = \ln 0.05/\ln (1 - 1/64) = 190.2 \approx 191 \text{ plants.}$$

In a population of 191 plants an average of 3 individuals (191*1/64) are expected to have the desired genotype. With dominant gene action at all 4 loci the heterozygous and homozygous individuals are indistinguishable in the F_2, and 60 plants (0.75^4 *191) are expected to have resistance. Cross protection and confounding symptoms additionally complicate the selection of homozygous resistant genotypes. To avoid error, inoculation and selection are best practiced on four F_3 subfamilies derived from each of the 191 F_2 individuals.

For the nested hierarchy (Figure 10), Eq. 11 is adjusted to reflect the number and size of divisions made in the pedigree, such that

$$P_\alpha = 1 - \left\{1 - \left(\prod_{j=1}^{t} f_{dj}\right)\right\}^{\underline{n}(\underline{a}^{\underline{s}})} \qquad [12]$$

where P_α and f_{dj} are defined above; $\underline{n}$ is the number of plants required in the F_2 generation; $\underline{a}$ is the size of the division; and $\underline{s}$ is the number of generations over which these divisions are made.

The nested hierarchy is perhaps best applied when insufficient numbers of individuals are available from the F_2 generation. Rare genotypes or allelic combinations may not be represented if the F_2 population is too small; Equation 12 assumes that all possible genotypes are represented. The probability of having all the needed alleles represented in a large F_3 population derived from a few F_2 individuals is small, therefore the size of the F_2 generation is critical. Equation 12 is an approximating equation that becomes less accurate as the number ($\underline{s}$) and size ($\underline{a}$) of the divisions increase.

Single Traits: Multiple Generations

Another permutation of the binomial equation is used for selecting an array of single traits, each in a different generation (i =1 to g). This selection technique is practiced in the modified single seed descent program (Figure 9), and is computationally very similar to selection for multiple traits in a single generation, such that:

$$P_\alpha = 1 - \left\{1 - \left(\prod_{i=1}^{g} f_{di}\right)\right\}^{n} \qquad \text{[Eq. 13]}$$

where f_{di} is the frequency of the desired genotype in the ith generation, $i = 1 \ldots g$, where a given trait is selected; and all other components are defined above.

Consider a modified single seed descent program in which three recessive genes (aa, bb, and cc) each confer resistance to viruses A, B, and C, respectively. If they are selected separately in the F_2, F_3, and F_4 generations, the desired genotypes occur at frequencies of 1/4, 3/8, and 7/16, respectively, in each generation. The initial number of F_2 plants required for *at least* one individual with the aabbcc genotype is computed as

$$P_\alpha = 1 - \{1 - (1/4 * 3/8 * 7/16)\}^{n}.$$

When P_α is set at 0.99, then:

$$n = \ln(.01) / \ln(1 - 21/512) = 109.9 \approx 110 \text{ plants.}$$

With a population of 110 plants, the general expectation is that 4 to 5 (110*21/512) individuals will have the desired genotype in F_4. If resistance were conferred by three dominant genes, heterozygous

individuals would be carried forward into the next generation and a quarter of these would segregate into the homozygous resistant category. By the final generation, this would increase the number of resistant individuals in the population but also increases the amount of progeny testing required. For the modified nested hierarchy (Figure 11), the problems and the computational adjustments are the same as those used for the nested hierarchy (Eq. 12).

As an example, Scully et al. (1988) used the modified nested hierarchy to develop multiple viral resistance breeding lines of common beans (Phaseolus vulgaris L.) from an F_2 population of 1400 plants. A single division (s=1) of size 2 (a=2) was made in the F_3 generation. Selection was practiced for resistance to 6 viruses in generations F_2 through F_7. Single genes conditioned resistance to 5 of the viruses and 1 virus required two recessive genes, which was selected in the F_3. The F_2 population size was approximated as follows:

$$P_\alpha = 1-\{1-(1/4 * 9/64 * 7/16 * 15/32 * 31/64 * 63/128)\}^{n(2^1)}.$$

For an approximate P_α of 99.25%, 1422 plants were required in the F_2 generation.

Multiple Traits: Multiple Generations

The third permutation is best suited for a modified single seed descent program (Figure 9). It involves selection for subsets of two or more traits, with each sub-set in a different generation. In this case the size of the original F_2 population is defined as:

$$P_\alpha = 1 - \{1 - (f_{d1(1)} * f_{d2(1)} * f_{dt(1)} * \cdots * f_{dt(g)})\}^{\frac{n}{}}. \qquad [14]$$

Thus

$$P_{\alpha} = 1 - \left\{1 - \left(\prod_{i=1}^{g} \prod_{j(i)=1}^{t(g)} f_{dj(i)}\right)\right\}^{n} \qquad [15]$$

where P_{α} and n are defined above; and $f_{d(j)i}$ is the expected frequency of the jth trait nested within the ith generation.

If selection is practiced for multiple traits, it is best to combine traits that do not compound selection for each other.

Conclusion

The concepts presented here are intended to provide a useful guide for the development of autogamous genotypes with multiple viral resistance. Consideration must be given to the inheritance of resistance, the statistical properties that influence genetic gain from selection, and the biology of the virus and its interaction with the host plant. The use of simply inherited broad spectrum virus resistance genes can further expedite the process of developing multiple viral resistant genotypes (Kyle and Provvidenti, this volume). In addition to the standard breeding schemes, the multiple trait backcross, modified single seed descent, and both forms of the nested hierarchy were presented as supplemental breeding methods. They were developed as logical extensions of existing backcross and pedigree methods, and diversify the approaches for breeding for multiple virus resistance. A set of equations based on binomial theory were provided to determine population sizes for different breeding schemes. These equations can improve the efficiency of selection without overextending the time and resources allocated to a specific breeding objective.

Literature Cited

Allard, R.W. 1960. Principles of Plant Breeding. New York: Wiley and Sons Inc.

Baker, R.J. 1986. Selection Indices in Plant Breeding. Boca Raton, Fl: CRC Press Inc.

Barker, T., G. Varughese and R. Metzger. 1989. Alternative backcross methods for introgression of variability into triticale via interspecific hybrids. Crop Sci. 29:963-965.

Brim, C.A. 1966. A modified pedigree method of selection in soybeans. Crop Sci. 6:200.

Browning, J.A. 1980. Genetic protective mechanisms of plant-pathogen populations: Their coevolution and use in breeding for resistance. In: Biology and Breeding for Resistance to Anthropods and Pathogens in Agricultural Plants, Ed. M.K. Harris pp. 52-76, College Sta. TX: Texas A&M University Press.

Cockerham, C.C. 1954. An extension of the concept of partitioning heredity variances for analysis of covariance among relatives when epistasis is present. Genetics 39:859-878.

Cox, T.S. 1984. Expectations of means and genetic variances in backcross populations. Theor. Appl. Genet. 68:35-41.

Day, P.R. 1974. Genetics of Host Parasite Interactions. San Francisco: W.H. Freeman Inc.

Dudley, J.W. 1982. Theory for the transfer of alleles. Crop Sci. 22:631-637.

Empig, L.T., and W.R. Fehr. 1971. Evaluation of methods for generation advance in bulk hybrid soybean populations. Crop Sci. 11:51.

Goulden, C.H. 1939. Problems in plant selection. In: Proc. 7th Inter. Genet. Congr., Ed. R.C. Punnet, pp.132-133 Cambridge, UK: Cambridge University Press.

Gracen, V.E. 1982. Role of genetics in etiological phytopathology. Ann. Rev. Phytopathol. 20:219-233.

Haghighi, K.R. and P.D. Ascher. 1988. Fertile, intermediate hybrids between Phaseolus vulgaris and P. acutifolius for congruity backcrossing. Sexual Plant Reproduction 1:51-58.

Hallauer, A.R., and J.B. Miranda. 1981. Quantitative Genetics in Maize Breeding. Ames, IA: Iowa State University Press.

Harrington, J. B. 1952. Cereal breeding procedures. F.A.O. Development Paper No. 28. Rome: UN Food and Agricultural Organization.

Horner, T.W., and C.R. Weber. 1956. Theoretical and experimental study of self-fertilized populations. Biometrics 12:404-414.

Khush, G.S. 1980. Breeding for multiple disease and insect resistance in rice. In: Biology and Breeding for Resistance to Anthropods and Pathogens in Agricultural Plants, Ed. M.K. Harris pp. 341-345, College Sta. TX: Texas A&M University Press.

Kyle, M.M., and R. Provvidenti. 1991. Genetics of broad spectrum virus resistance in bean and pea. ~~See Kyle 1991, Ch. 6~~.

Kyle, M.M. , Ed. 1991. Resistance to Viral Diseases of Vegetables: Genetics and Breeding. Portland, OR: Timber Press.

Larsen, R.J., and M.L. Marx. 1985. An Introduction to Probability and its Applications. Englewoods Cliffs, NJ: Prentice-Hall, Inc.

Mansur, L.M., K.M. Hadder, and J.C. Suarez. 1990. A computer program for calculating the population size necessary to recover any number of individuals exhibiting a trait. J. hered. 81:407-408.

Mayo, O. 1987. Theory of Plant Breeding. Oxford, UK: Claredon Press,

Mosteller, F., R.E.K. Rourke, and G.B. Thomas. 1961. Probability with Statistical Applications. London: Addison-Wesley Publ. Co., Inc.

Munger, H.M. 1991. Breeding for viral disease resistance in cucurbits. ~~See Kyle 1991, Ch. 1.~~

Nelson, R.R. 1973. The use of resistance genes to curb population shifts in plant pathogens. In Breeding Plants for Disease Resistance: Concepts and Applications, Ed. R.R. Nelson, pp. 49-67, University Park, PA: Penn. State Univ. Press.

Provvidenti, R. 1985. Lectures on the Resistance to Viral Diseases in Vegetables. International seminar on virus diseases of horticultural crops in the tropics. Council on Agriculture of the Republic of China. Geneva, NY.

Russell, G.E. 1978. Plant Breeding for Pest and Disease Resistance. London: Butterworth & Co. Ltd. 485 pp.

Sedcole, J.R. 1977. Number of plants necessary to recover a trait. Crop Sci. 17:667-668.

Scully, B., D.H. Wallace, and R. Provvidenti. 1988. Breeding common beans for multiple virus resistance. HortScience 23:115

Sheridan, W.F., and J.K. Clark. 1987. Allelism testing by double pollination of lethal maize dek mutants. J. Heredity 78:49-50.

Simmonds, N.W. 1979. Principles of Crop Improvement. New York: Longman Group Ltd.

Smith, H.F. 1936. A discriminant function for plant selection. Ann. Eugen. 7:240-250.

Snape, J.W., and T.J. Riggs 1975. Genetical consequences of single seed descent in the breeding of self-pollinated crops. Heredity 35:211-219.

Snedecor, G.W., and W.G. Cochran. 1981. Statistical Methods. Ames, IA: Iowa State University Press. 7th ed.

Sneep, J. 1977. Selection for yield in early generations of self-fertilizing crops. Euphytica 26:27-30.

Superak, T.H., and B. T. Scully. 1991. Strategies for the deployment of new sources of plant virus resistance in Cucurbita. ~~See Kyle 1991, Ch. 3.~~

Suzuki, D.T., A.J.F. Griffiths, J.H. Miller, and R.C. Lewontin. 1985. An Introduction to Genetic Analysis. New York: W.H. Freeman and Co. 3rd ed.

Turner, H.N. and S.S.Y. Young. 1967. Quantitative Genetics in Sheep Breeding. Cornell University Press, Ithaca, NY. .

Van Vleck, L.D. 1979. Notes on the Theory and Application of Selection

Principles for Genetic Improvement of Animals. Ithaca, NY: Dept. of Animal Science, Cornell University.

Walkoff, C. 1955. An application of the double backcross method to tomato improvement. 14th Rept. Intern. Hort. Congr., Netherlands pp. 252-259.

Walkoff, C. 1961. Improvement of tomato fruit size and maturity by backcross breeding. Can. J. Plant Sci. 41:24-30.

Wehrhan, C. and R.W. Allard. 1965. The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics 51:109-119.

Wricke, G., and W.E. Weber. 1986. Quantitative Genetics and Selection in Plant Breeding. New York: Walter de Gruyter.

Table 1: The number of plants (n) needed in a population to recover a desired number of individuals (r) at a given genotypic frequency (f_d) and probability of success (P_α)§.

P_α	f_d	r= the number of desired individuals								
		1	2	3	4	5	6	8	10	15
95%	1/2	5	8	11	13	16	18	23	28	40
	1/4	11	17	23	29	34	40	50	60	84
	1/8	23	37	49	60	71	82	103	123	172
	1/16	47	75	99	122	144	166	208	248	347
	1/32	95	150	200	246	291	334	418	500	697
	1/64	191	302	401	494	584	671	839	1002	1397
99%	1/2	7	11	14	17	19	22	27	32	45
	1/4	17	24	31	37	43	49	60	70	96
	1/8	35	51	64	77	89	101	124	146	198
	1/16	72	104	132	158	182	206	252	296	402
	1/32	146	210	266	318	368	416	508	597	809
	1/64	293	423	535	640	739	835	1020	1198	1623

§ Adapted from Sedcole J.R. 1977. Crop Sci. 17: 667-668 with permission of the Crop Science Society of America.

Figure 1: The simple backcross procedure transfers a dominant gene from a donor parent (D_1) to the recurrent parent (R). Individuals are tested and selected in each backcross generation (BC-1 to BC-ν) and then progeny tested and confirmed in the selfed generations (BC-νF_2 and BC-νF_3).

Figure 2: The alternate backcross and self procedure incorporates a recessive gene from a donor parent (D_1) to a recurrent parent (R). Progenies are tested and selected in each backcross selfed generation (BC-$1F_2$ to BC-νF_2) and the gene is confirmed in the last generation (BC-νF_3).

Figure 3: The continuous backcross procedure primarily transfers a recessive gene. Initially n plants in the BC-1 generation are crossed to the recurrent parent (R) and crosses increased exponentially ($\underline{n}^2$, $\underline{n}^3$,...$\underline{n}^v$) with each backcross generation. Progeny testing, selection and confirmation of the trait is delayed until the last two generations (BC-vF_2 and BC-vF_3).

Figure 4: The simultaneous backcross and self method is primarily used to transfer a recessive gene from a donor parent (D_1) to a recurrent parent (R). A number of plants (n) are each crossed to the recurrent parent and concurrently selfed to produce an F_2, and uncover the recessive gene. A backcross parent with the recessive gene is selected for crossing in the next generation.

Figure 5: The sequential backcross procedure is used to incorporate dominant genes, each from a different donor parent (D_1, D_2, etc.)

Figure 6: The parallel backcross scheme primarily incorporates different dominant genes concurrently, each from a different donor parent (D_1 to D_4) into the same recurrent parent (R).

Figure 7: The multiple trait backcross method transfers two or more traits from the same donor parent (D_1). A number of plants (n) are crossed to the recurrent parent (R) and concurrently selfed and tested for each trait. The backcross parent with each trait is selected for crossing in the next generation.

Figure 8: A single seed descent program is used to incorporate a number of different traits. Selection, progeny testing and confirmation are performed in the last two generations of the program, usually F_7 or F_8.

Figure 9: In a modified single seed descent program, single traits are selected discretely in each generation. Progeny testing and confirmation of resistance are performed in the last two generations.

Figure 10: The nested hierarchy is similar to single seed descent, but in this case two (or more) offspring are carried forward in each generation. Selection, progeny testing, and confirmation of the trait are determined in the last two generations.

Figure 11: In the modified nested hierarchy, single traits are selected discretely in each generation and the offspring from resistant individuals are carried forward. Progeny testing and confirmation of the trait are determined in the last two generations.

Figure 1

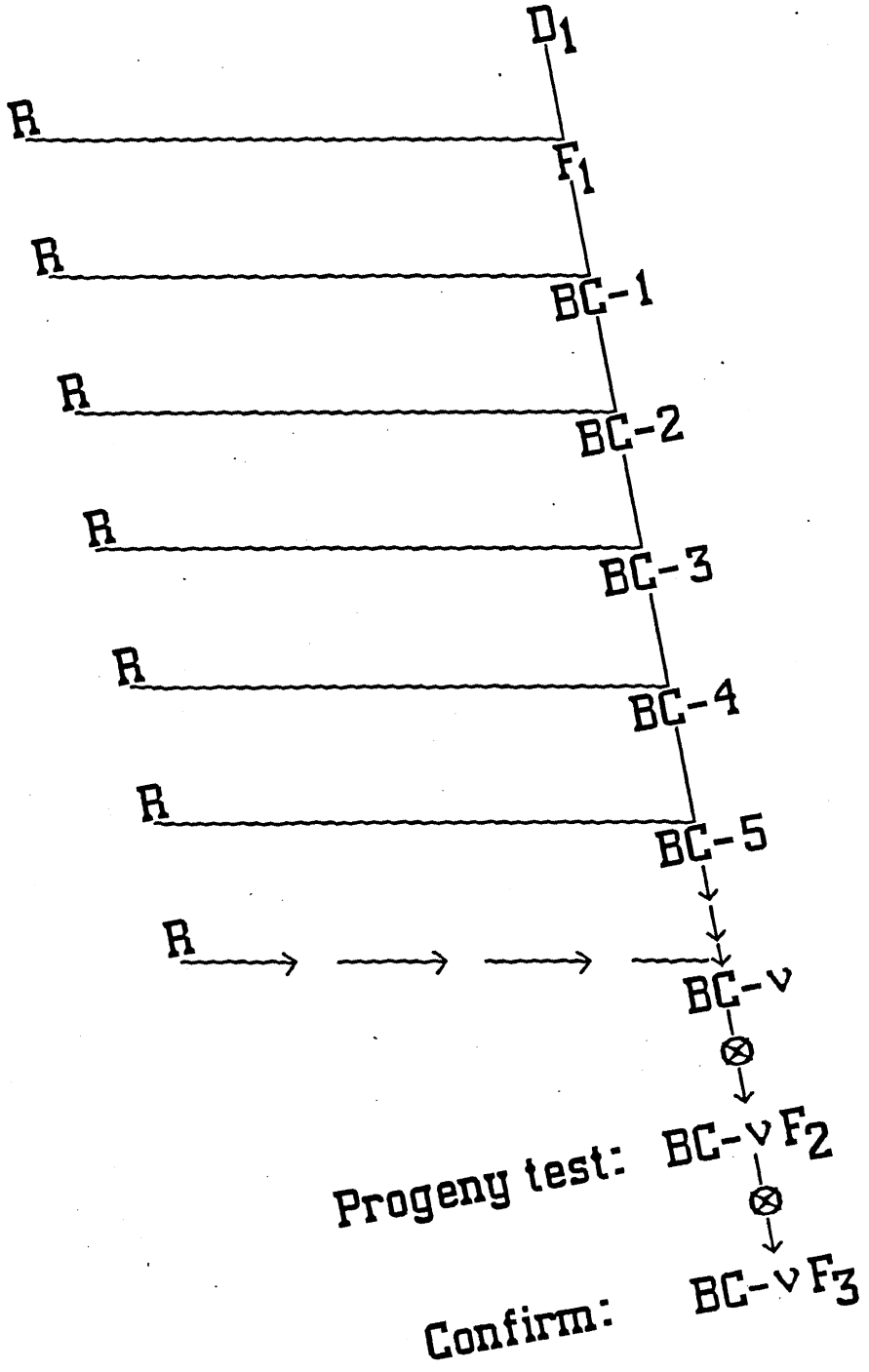

D1
R
F1
R
BC-1
R
BC-2
R
BC-3
R
BC-4
R
BC-5
R
BC-ν
⊗
Progeny test: BC-νF2
⊗
Confirm: BC-νF3

Figure 2

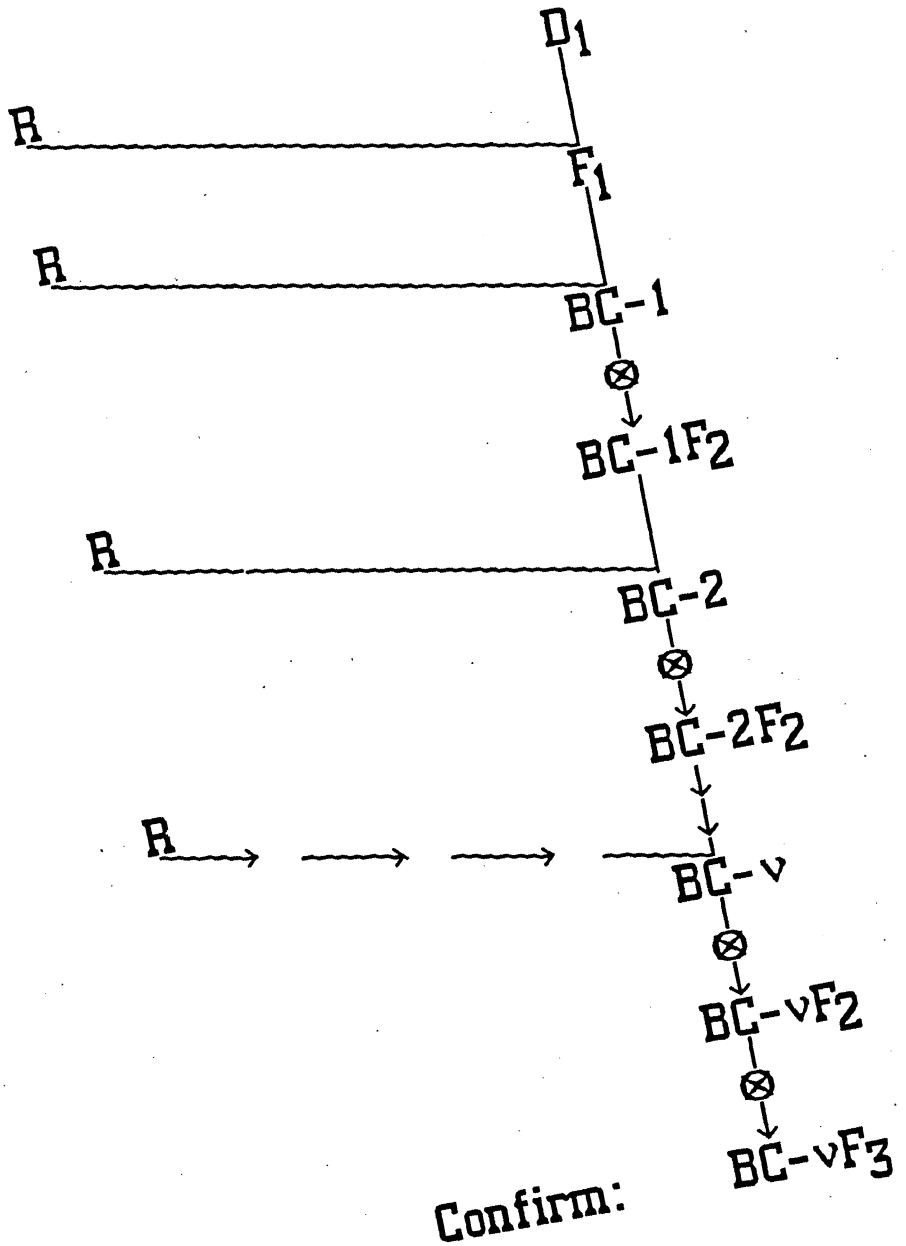


Figure 3

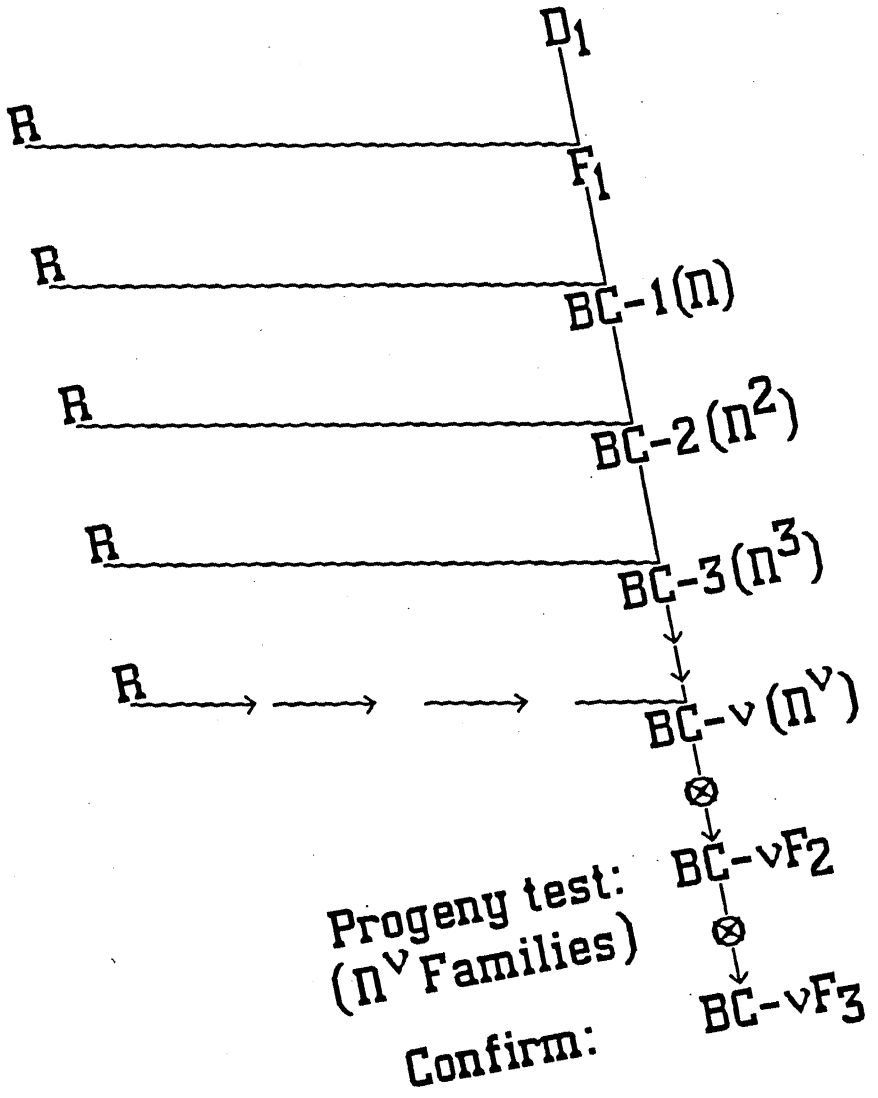

D1
R
F1
R
BC-1(n)
R
BC-2(n2)
R
BC-3(n3)
R
BC-ν(nν)
BC-νF2
BC-νF3
Progeny test:
(nν Families)
Confirm:

Figure 4

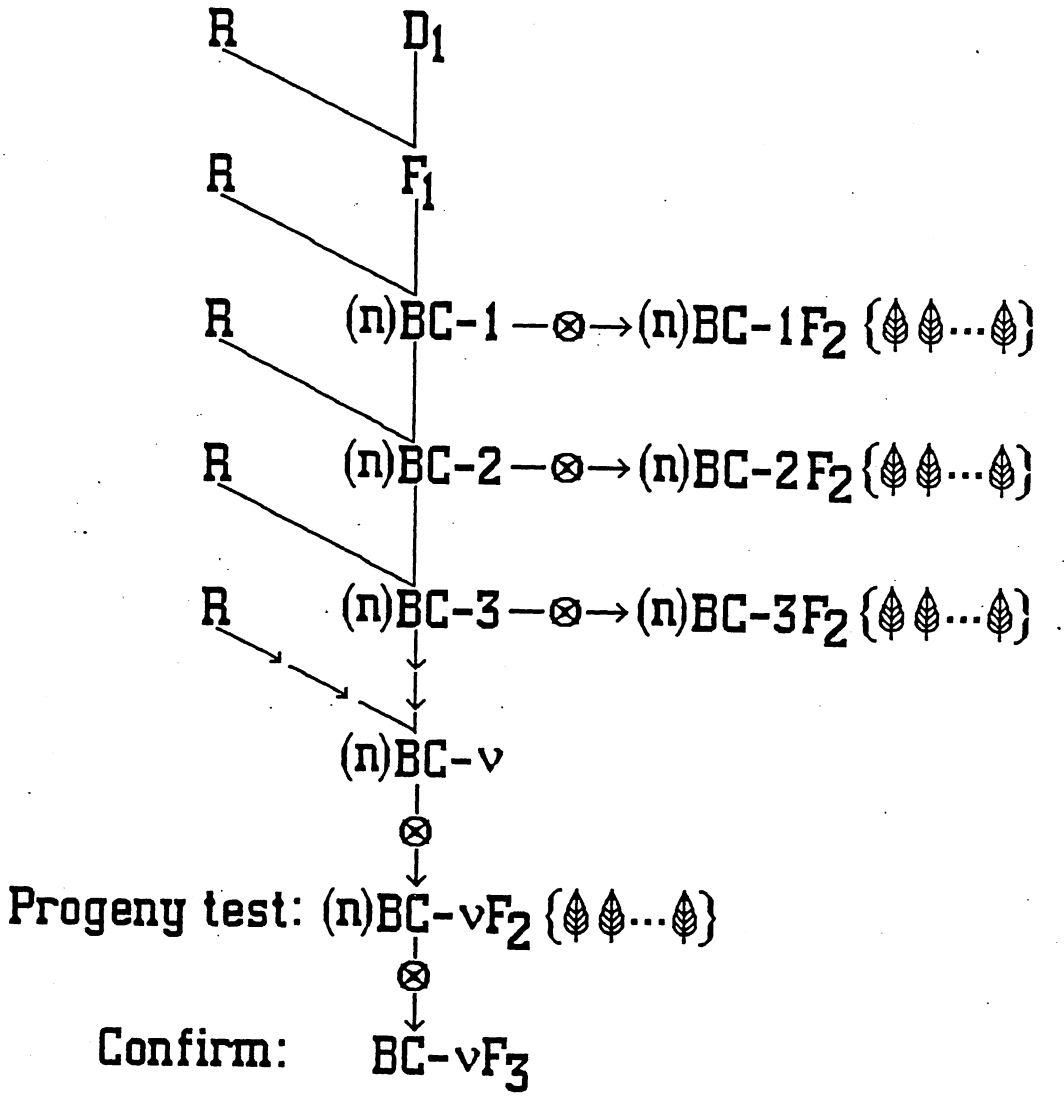

R
D1
R
F1
R
(n)BC-1 —⊗→ (n)BC-1F2 {...}
R
(n)BC-2 —⊗→ (n)BC-2F2 {...}
R
(n)BC-3 —⊗→ (n)BC-3F2 {...}
(n)BC-ν
⊗
Progeny test: (n)BC-νF2 {...}
⊗
Confirm: BC-νF3

Figure 5

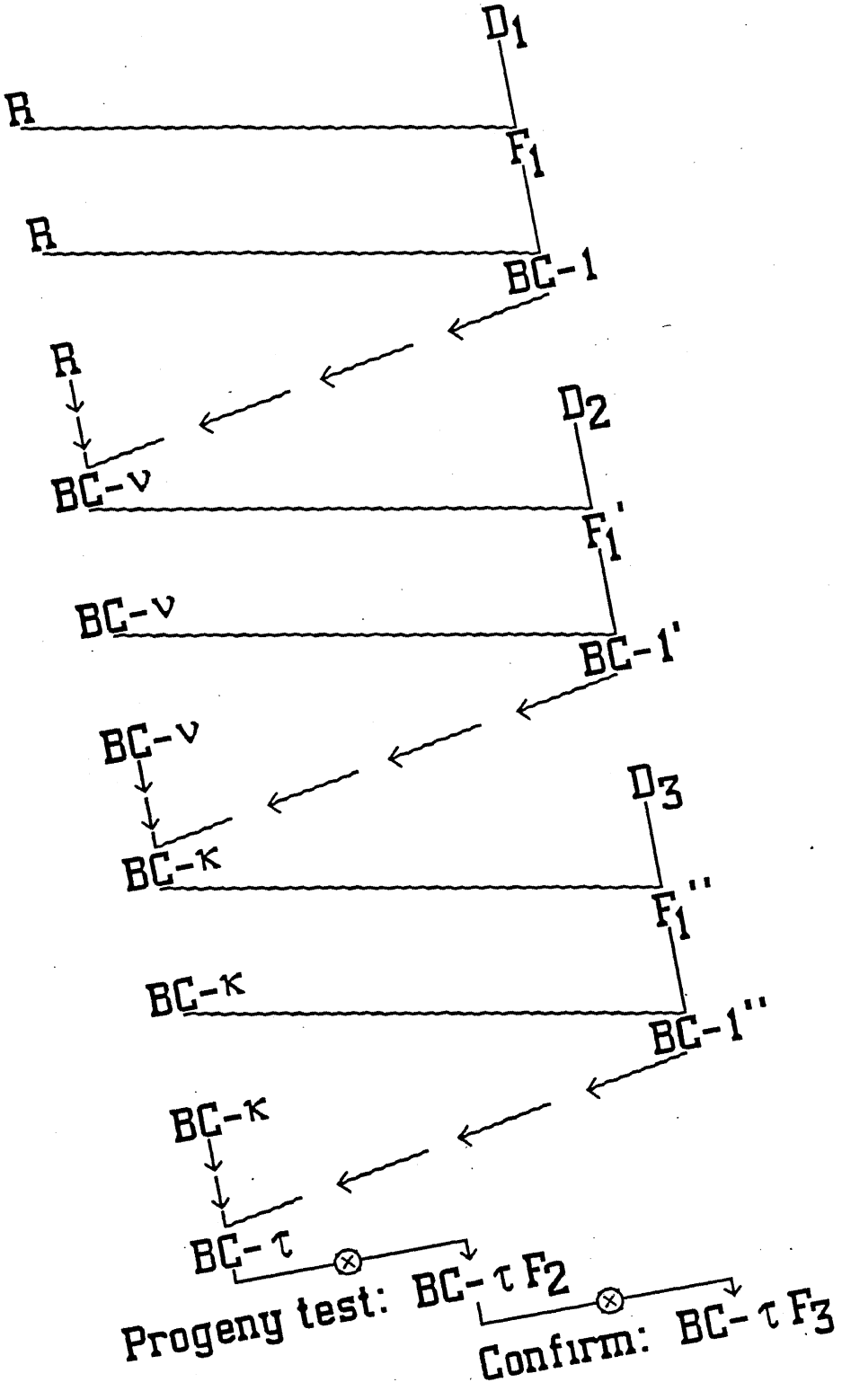
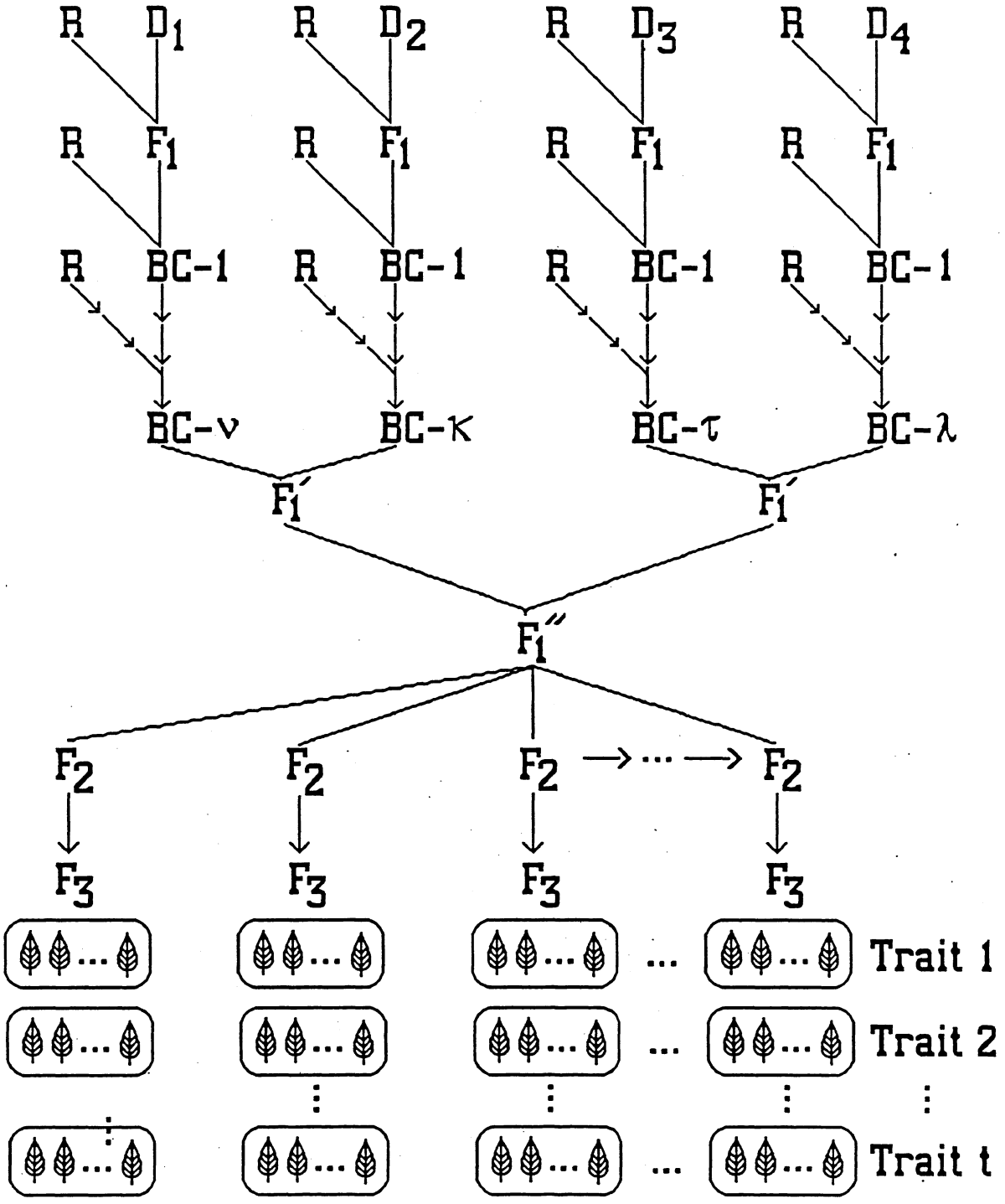

D1
R
F1
R
BC-1
R
D2
BC-ν
F1'
BC-ν
BC-1'
BC-ν
D3
BC-κ
F1''
BC-κ
BC-1''
BC-κ
BC-τ
Progeny test: BC-τF2
Confirm: BC-τF3

Figure 6



Progeny test each trait in F_3 sub-families

Figure 7

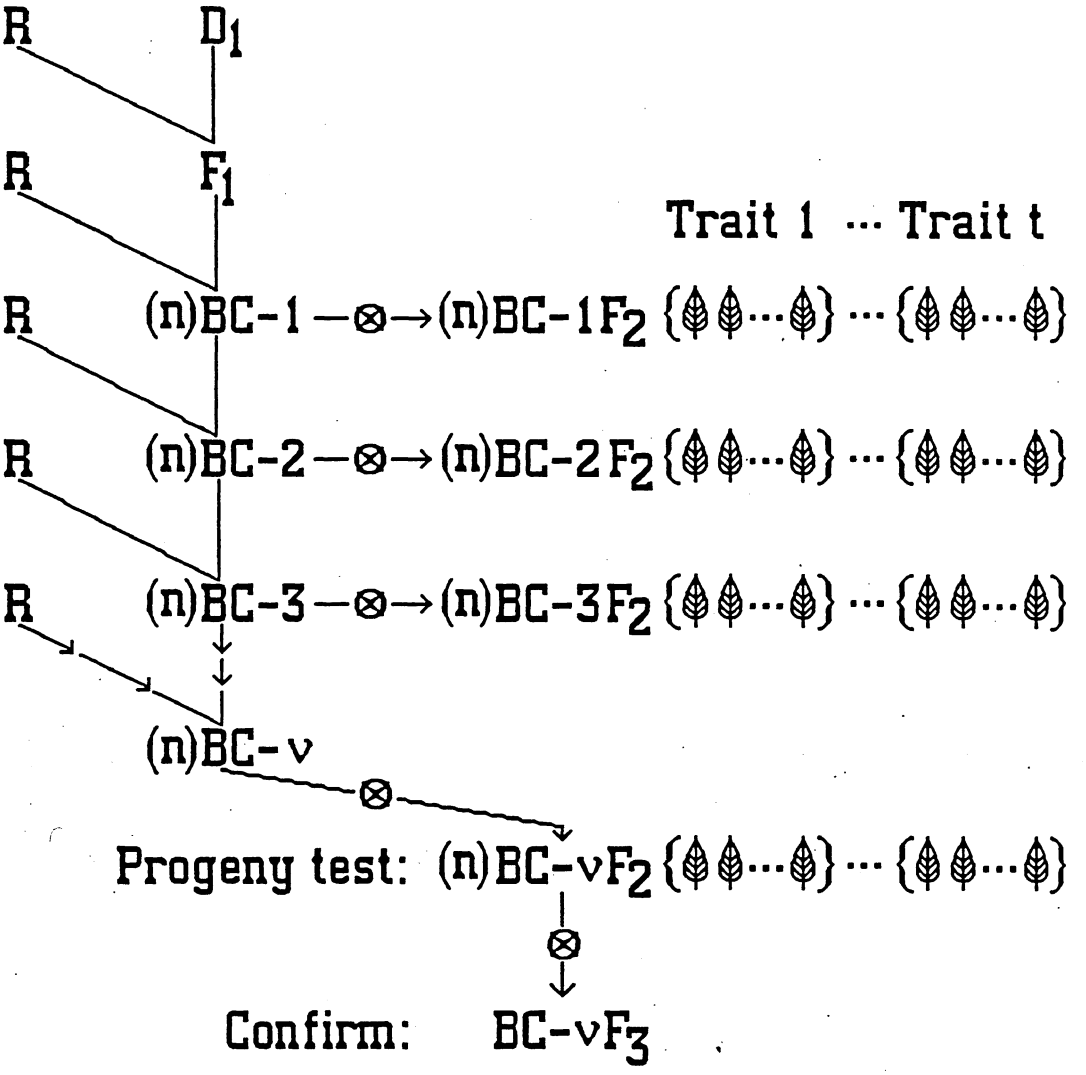

R
D1
R
F1
Trait 1 ··· Trait t
R
(n)BC-1 —⊗→ (n)BC-1F2 {···} ··· {···}
R
(n)BC-2 —⊗→ (n)BC-2F2 {···} ··· {···}
R
(n)BC-3 —⊗→ (n)BC-3F2 {···} ··· {···}
(n)BC-ν
⊗
Progeny test: (n)BC-νF2 {···} ··· {···}
⊗
Confirm: BC-νF3

Figure 8

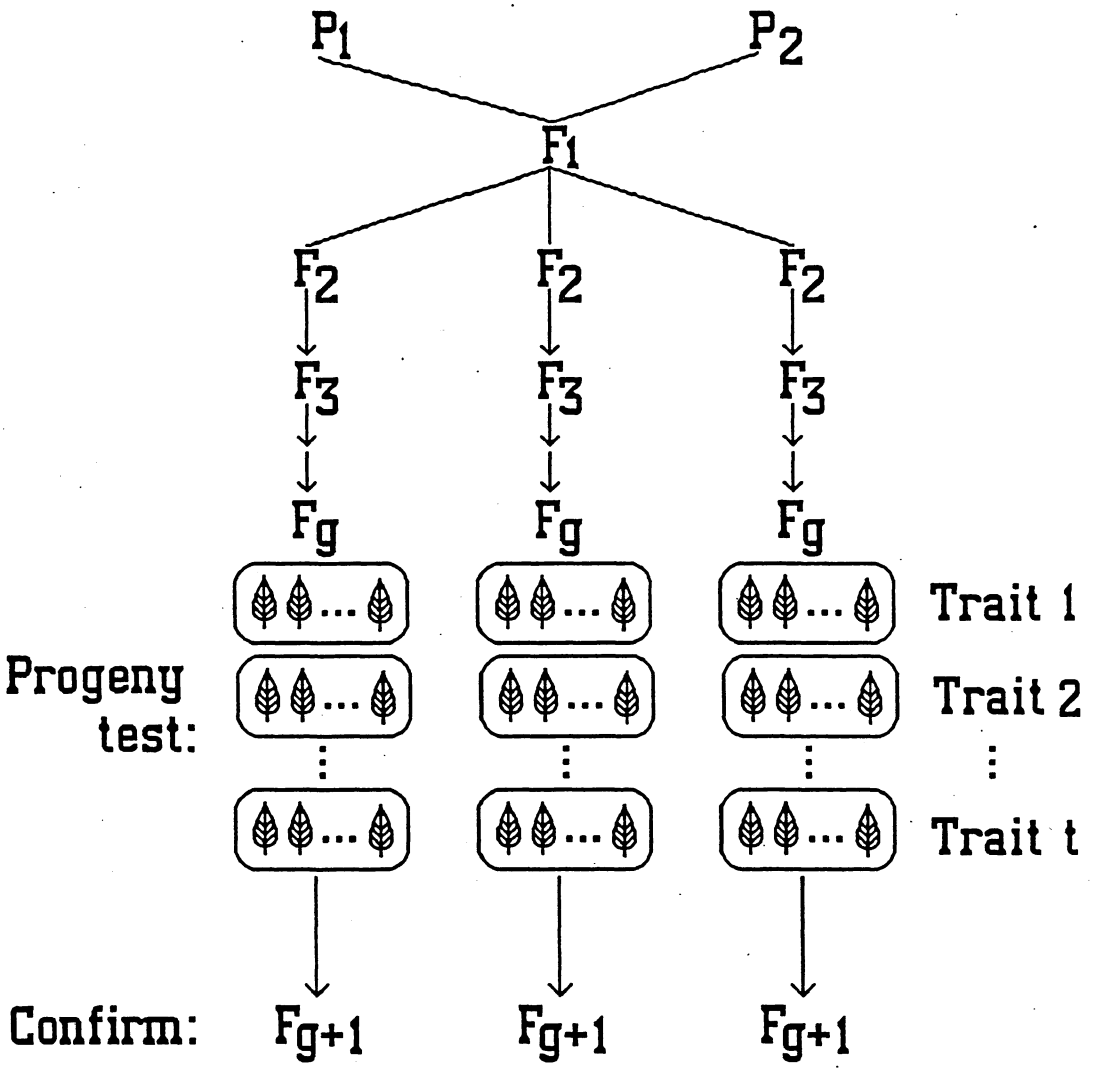

P1
P2
F1
F2
F2
F2
F3
F3
F3
Fg
Fg
Fg
Trait 1
Trait 2
Trait t
Progeny test:
Confirm:
Fg+1
Fg+1
Fg+1

Figure 9

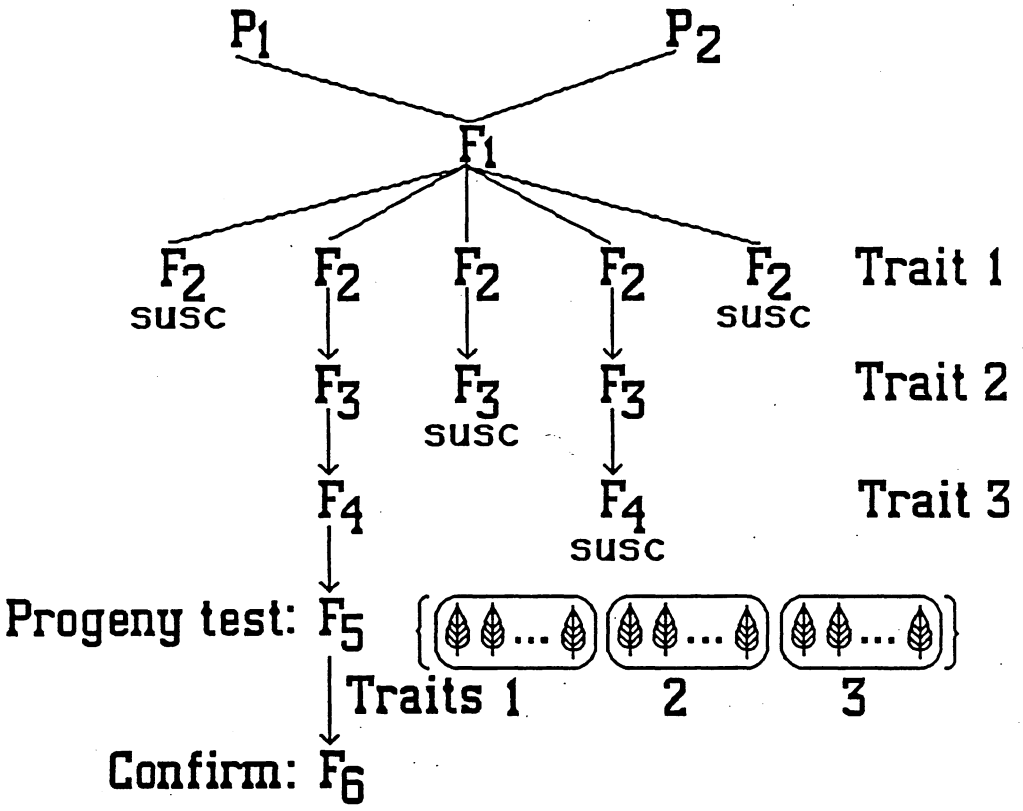
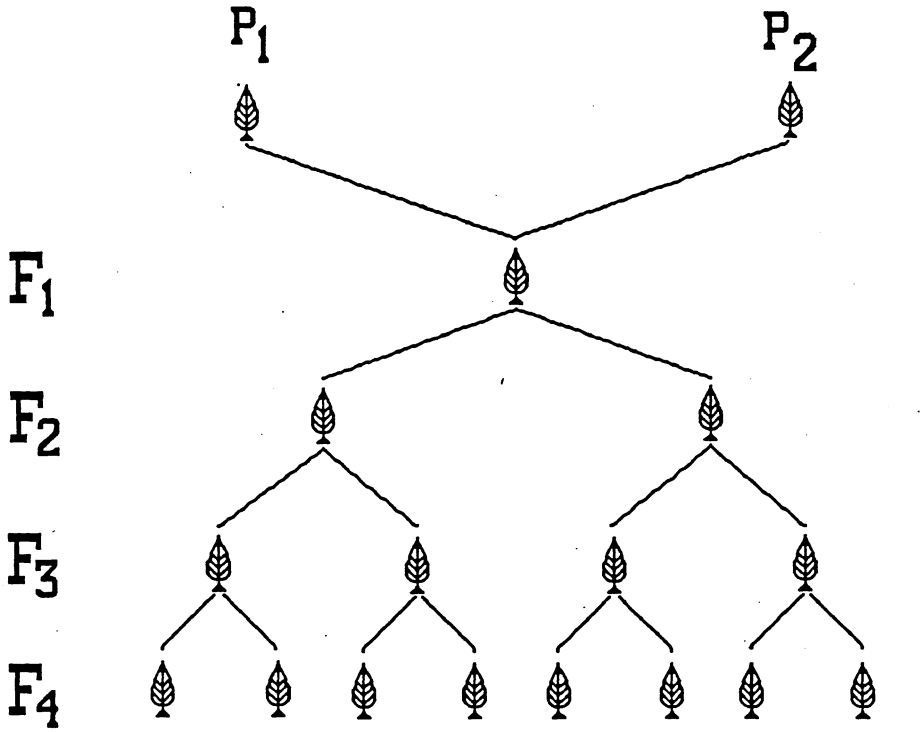
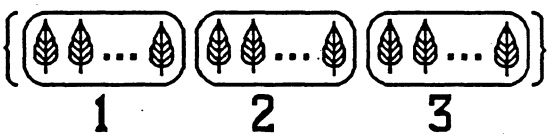

P1
P2
F1
F2 susc
F2
F2
F2
F2 susc
Trait 1
F3
F3 susc
F3
Trait 2
F4
F4 susc
Trait 3
Progeny test: F5
Traits 1
2
3
Confirm: F6

Figure 10

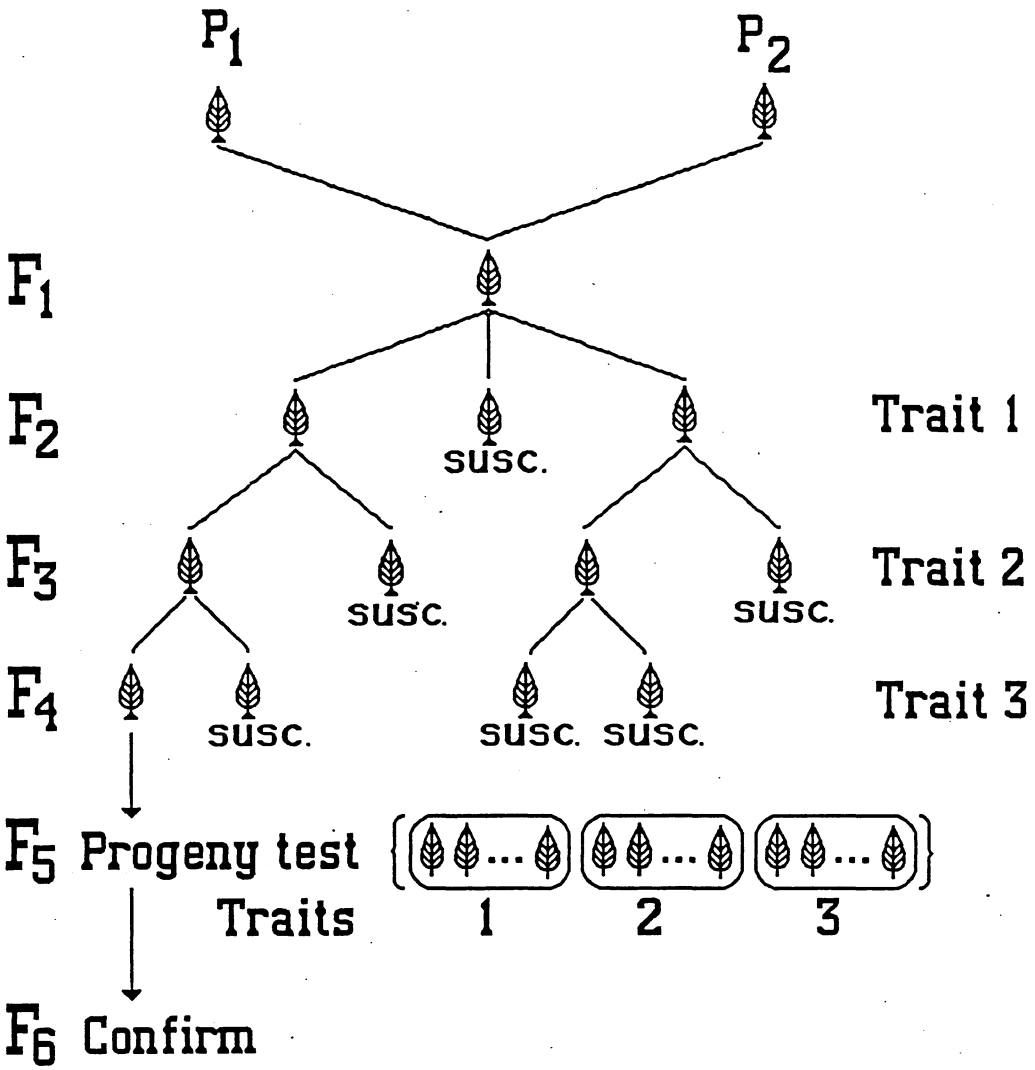


F5 Progeny test F_4 derived sub-families for each trait



F6 Confirm

Figure 11


P1
P2
F1
F2
susc.
Trait 1
F3
susc.
susc.
Trait 2
F4
susc.
susc.
susc.
Trait 3
F5 Progeny test
Traits
1
2
3
F6 Confirm