ANALYSIS OF GENE NUMBER AND DEVELOPMENT IN POLYGENIC SYSTEMS

(polygenes, variation, selection, pattern)

JAMES N. THOMPSON, JR.

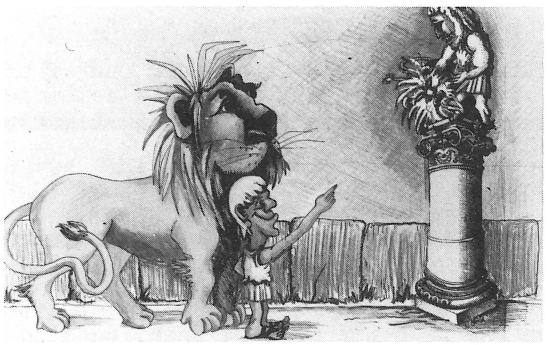
Department of Zoology
University of Oklahoma
Norman, Oklahoma 73019

SUMMARY

The assumption that quantitative variation is produced by a large number of genes is re-examined. In fact, one finds that often only a small number of loci are involved. This, therefore, opens to careful study the developmental effects associated with polygenic variation.

INTRODUCTION

Aesop, the storyteller, writes that once upon a time a man and a lion were walking through the forest arguing about which



of them was the more powerful. Each bragged about impressive

feats of strength. At a bend in the road they came to a small open area in which stood a statue of a man and a lion fighting. The man was obviously winning. "See!" cried the man, "That proves that I am clearly much stronger than you are!" "The only thing that *is* clear," replied his companion, "is that the statue was not made by a lion."

We often, quite inadvertently, build biases into the things we construct. Even geneticists are not entirely free from this error. It is probably seen most frequently in the names we assign new phenomena and processes and in the way we first define new terms. We choose names that we believe are descriptive, but gradually the name or definition comes to be thought of as an explanation -- fact rather than hypothesis. I believe this is precisely what has happened in the development of quantitative genetics.

The term "polygene" was coined by MATHER (1941), in the context of biometrical genetics, to describe the genes contributing to quantitative or continuous phenotypic variation, in which each gene has an effect that is small relative to other sources of variation. Polygenes can be distinguished from "major genes" by the relative magnitude of their phenotypic effects. Major gene mutants have phenotypes that enable them to be classified unambiguously into discontinuous groups, thus allowing their transmission to be followed easily and leading to a solid understanding of the Mendelian rules of inheritance.

One of the primary differences between genetic studies of quantitative variation and of classical Mendelian differences is in the way the data are handled statistically. The ability to classify discontinuous characters into discrete groups leads naturally to statistical analyses using the chi-square test of goodness of fit. The continuous distributions found for quantitative traits, on the other hand, lead to comparison of variances between and within classes and, thus, to a biometrical approach.

The development of biometrical genetics, like that of any other analytical construct, involved a number of assumptions about the genetic systems being studied. One of the early, and quite necessary, assumptions which biometrical models made about polygenic systems was that quantitative variation was produced by many genes (often assumed to number hundreds or even thousands affecting any particular character) having small, similar effects (MATHER 1943, 1949; MATHER and JINKS 1971). Unfortunately, some have now come to think of this as an established fact rather than as a simplifying assumption (see, for example, VETTA 1976, and response by THOMPSON 1976). The confusion is, of course, compounded by the biases inherent in the term "polygene" to the point where a survey of dictionary definitions leaves one with the feeling that, outside of a limited area of biometrical genetics, the term is considered to be meaningless (see LERNER 1972, RIEGER *et al.* 1968).

This situation is quite unfortunate, since many of the most important medical, behavioral and agricultural characters are

polygenic, and a clear understanding of the nature and function of the genes controlling quantitative traits could make a significant contribution to these fields and complement the elegant contributions that biometrical genetics has already made (MATHER and JINKS 1971). But again, one is often blocked by confusion and biases. Questions about polygene number are at times considered irrelevant, and attempts to isolate and analyze individual polygenes are looked upon with suspicion. Indeed, more than once it has been said that if one extracts a single polygene so that it can be studied in relative isolation from other sources of variation, it is no longer a polygene. From the definition of the term and from the assumptions built into biometrical models, then, it appears impossible to ask how many genes contribute to quantitative variation and impossible to isolate individual factors and ask how they produce their effects. Obviously, this is an unworthy paradox, and a number of people, most notably J.M. THODAY and his colleagues, have begun to re-examine basic assumptions in an attempt to understand the underlying basis of quantitative genetic variation.

In order to summarize the results of that re-examination, I would like to begin by asking whether one is justified in assuming that quantitative variation necessarily involves a very large number of genes. The answer will be that the assumption is not justified, either theoretically or experimentally. I shall then go on to discuss ways in which one can dissect the development of a quantitative trait and learn how individual polygenes influence it. Finally, we shall attempt to see briefly how the insights to be gained from this type of developmental analysis affect our understanding of the phenotype as a whole.

THE NUMBER OF GENES IMPLIED BY CONTINUOUS VARIATION

There seems to be a general tendency to conclude that if a trait is complex or if the phenotypic distribution is continuous, the segregation of a large number of polygenes must be involved. This conclusion cannot, however, be justified, because the degree to which a phenotypic distribution will reflect the underlying genetic distribution depends upon a number of factors, including heritability. When heritability is limited, a continuous distribution can even be produced by the segregation of a single pair of alleles (ALLARD 1960, THODAY and THOMPSON 1976). In Figure 1, for example, samples of F_2 individuals have been taken from a theoretical distribution in which only one locus is segregating and in which the heritability is about 50 per cent.

Indeed, heritability is not the only factor that can smooth out the distribution of phenotypes. Sampling can also do this. Observed distributions are random samples of limited size from some underlying theoretical distribution. Most experimental samples are actually much smaller than those necessary to distinguish small gene from large gene number segregations, as is illustrated in Figure 2. This sample of 150 individuals has been taken from a theoretical distribution in which three loci

are segregating and the genes have effects of 1x, 2x, and 3x; heritability is 100 per cent. Over 90 per cent of the variance is accounted for by only two loci. In Figure 3, t (kurtosis) is plotted against sample size of random samples from the theoretical distribution generated by these three loci. Rather

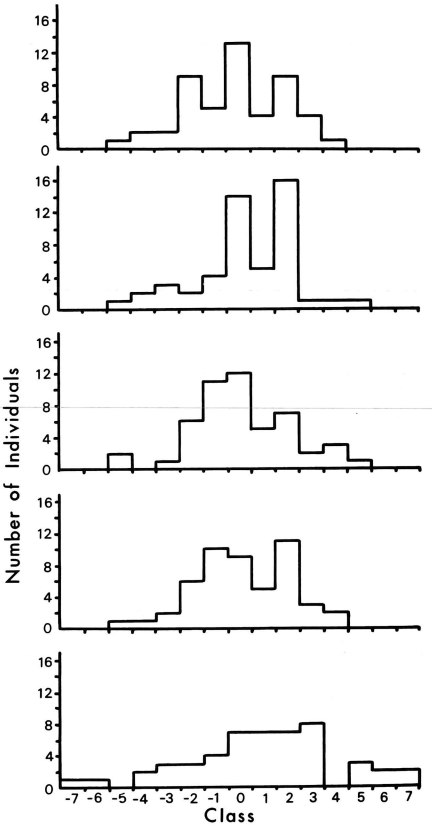


Figure 1. Random samples of 50 individuals taken from a theoretical distribution in which only one locus is segregating and the heritability is 50 per cent.

large sample sizes would clearly be needed for one consistently to detect departures from the normal distribution expected if polygene number were large (THODAY and THOMPSON 1976). Superimposing environmental variation upon this distribution, as is typical of quantitative traits, would blur phenotypic classes even more. Clearly, there is no justification for automatically assuming that a continuous distribution reflects the segregation of a large number of polygenes.

One further difficulty arises in discussions about polygene

number. It quickly becomes apparent that different people mean different things by the question, "How many genes are involved?" The various forms the question can take have been examined by THODAY and THOMPSON (1976). Briefly, we can distinguish four

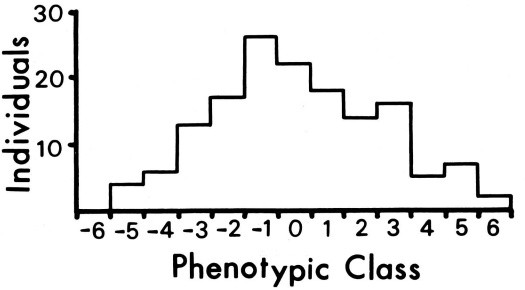


Figure 2. Random sample of 150 individuals taken from a theoretical distribution in which 3 genes are segregating and over 90% of the variance is accounted for by only two loci.

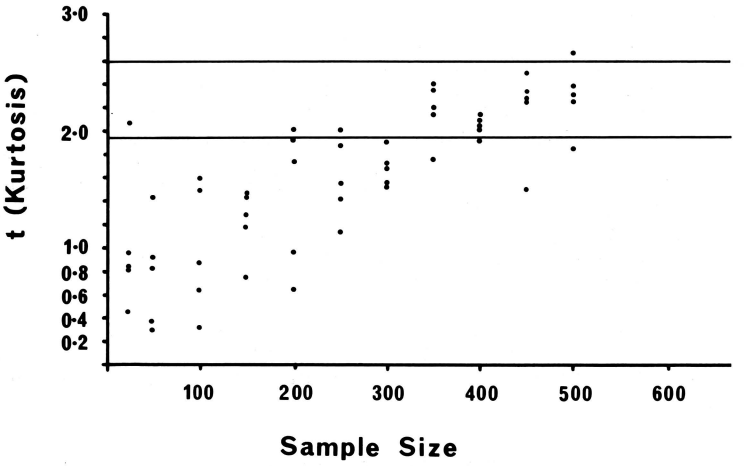


Figure 3. Plot of t (kurtosis) against sample size of random samples taken from the theoretical distribution as in Figure 2.

versions, each leading to a smaller expected number of genes. The first version is, "How many genes are involved in influencing variable x in species y?" This question implicitly includes all genes, even homozygous loci, involved in the development of the trait. Though interesting, it is essentially unanswerable with present genetic techniques.

Second, "At how many loci in species y is there allelic segregation affecting variable x?" Though a large number of genes are still included, this clearly involves fewer loci than the first question since homozygous loci are eliminated. A similar, but perhaps more realistic, question refers, not to the species, but to a given population. Here the number of genes is even smaller, because one would expect the particular heterozygous loci to vary from population to population. A number of people have begun to study polygene polymorphism in nature in order to answer this version of the question (*e.g.*, HOSGOOD and PARSONS 1967, 1968; MILKMAN 1965; PARSONS and HOSGOOD 1967; PARSONS *et al.* 1967), and we have recently attempted to identify wild polygene alleles affecting a wing character (THOMPSON and THODAY 1976) and eye pigmentation (THOMPSON *et al.* unpublished). The general conclusion is that the number of polygenes per character in a population is not unworkably large, and there is a real possibility of identifying individual alleles affecting certain characters, allowing allele frequencies to be estimated for the first time.

The third, and from my point of view the most appropriate, form of the question is,"How many loci are responsible for the major part of the genetic variance of variable x in population y_1?" This version brings us nearer the development of the character and focuses upon the limited number of polygenes having comparatively large effects and intermediate allele frequencies. These are the loci whose developmental effects are potentially open to investigation.

The final version is,"At how many loci is there segregation *directly* affecting the development of the character x in population y_1?" This may look identical to the previous question, but it is not. It is, in fact, much more sophisticated, for it draws attention to the fact that there may be loci that influence viability or which interact with each other or through linkage and thereby contribute to the genetic variance without affecting the development of the character directly. These loci would be recognized by biometrical analyses, but would be only indirectly relevant to developmental analyses.

Our studies of the nature and developmental effects of polygenes focuses upon the set of loci and alleles delimited by the third and fourth versions of the question,"How many genes are involved?"

ISOLATION OF POLYGENES

Although studies of continuous and discontinuous variation are separated historically and methodologically, the distinction between polygenes and major genes is usually considered to be empirical rather than qualitative. Polygenes, like major genes, show both segregation and linkage (MATHER 1943) and otherwise appear to be inherited in the classical Mendelian fashion. Indeed, since most characters are influenced to some extent by polygenic variation, it is not surprising that there is an important and intimate relationship between polygenes and development.

This is particularly evident if one views the individual as a highly organized complex of biochemical processes that can be influenced at many stages by minor genetic differences and by environmental factors.

A number of geneticists share the rather Baconian belief that the basic questions of population genetics can be answered by counting alleles and allele combinations. But others believe that it is at least as important to understand the precise nature of the influences that variation, including polygenic variation, can exert upon the development of an individual (*e.g.*, ROBERTSON 1966, THODAY 1967, THOMPSON 1975b). Such questions can be answered properly, however, only when it is possible to isolate and manipulate the individual members of a polygenic set.

Early attempts to locate the genes influencing the expression of quantitative traits relied primarily upon an analysis of their linkage with major genes. For example, BRIDGES (1919) demonstrated that there were eight genes on the second and third chromosomes of *Drosophila melanogaster* that affected the expression of the mutant *eosin*; SAX (1923) demonstrated the linkage of seed size factors to a pigmentation locus in *Phaseolus*; and RASMUSSON (1935) demonstrated that factors affecting flowering time were linked to a flower color locus in the garden pea.

An analysis of the effects of whole chromosomes, such as that used by MATHER and HARRISON (1949), allowed the gross location of factors accumulated in stocks selected for changes in certain modifiable characters. Linkage relationships of the modifier factors could, then, be assayed by the loss or maintenance of presumably linked marker genes during selection (SISMANIDIS 1942), by assays of recombinants of the selected lines and standard marker stocks (WIGAN 1949), by the construction of synthetic chromosomes in which marker genes were used during the intermediate steps to follow relevant segments during the breeding program (BREESE and MATHER 1957, 1960), or by redefining the trait until Mendelian segregation could be discerned in what was originally a quantitative distribution (SPICKETT *et al.* 1967).

Most of the precise polygene locations have been carried out in *Drosophila* because of the wealth of genetic markers and balancer chromosomes available. Techniques have, however, been designed for more complex situations. These include the location of factors by the study of variation, produced by recombination, in the derivatives of single chromosome heterozygotes of chromosome substitution lines in the hexaploid wheat *Triticum aestivum* (LAW 1966, 1967; LAW and WORLAND 1973), and by an inbreeding and backcross program used by WEHRHAHN and ALLARD (1965) to study the genetic control of heading time in wheat.

THODAY (1961) greatly refined the process of polygene location by integrating the use of chromosome markers, as practiced by SAX, WIGAN, and others, with the method of progeny testing first applied to the analysis of quantitative characters by JOHANNSEN (1909). Assaying the phenotypes of marked recombinant segments allows one to determine which segments are important to the expression of the character, while progeny testing serves the

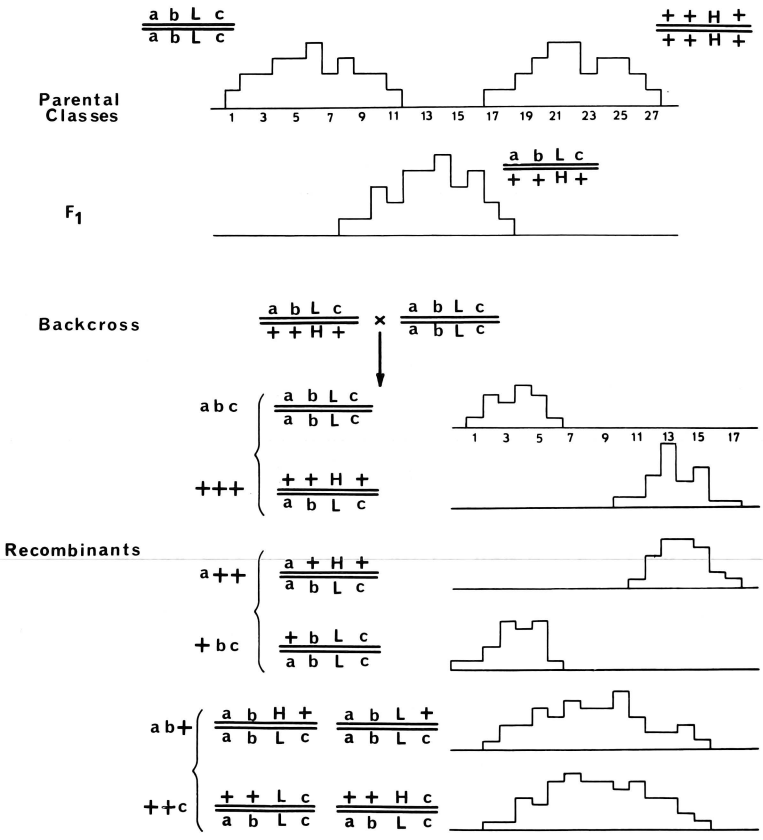


Figure 4. Hypothetical location of a polygenic locus within the chromosome region marked by the three recessive mutants *a*, *b*, and *c*. See text for full discussion.

dual role of verifying the phenotypic classification of each recombinant chromosome and, by inbreeding, of further increasing the homogeneity of the genetic background. One can also determine the number of linked loci between the marker genes, since the number of loci will be one fewer than the number of classes of recombinant chromosomes within a region.

A hypothetical example of polygene location analysis is illustrated in Figure 4. An inbred assay stock, marked with the recessive mutants *a*, *b* and *c*, differs from a selected strain (+ + +) in the expression of some quantitative character. As shown in this figure, the *a b c* strain is homozygous for a factor (*L*) that decreases the expression of the character, relative to the *H* factor in the selected line. For example, the *H*, or High, line may have longer wing veins than the inbred standard *L*, Low.

Although I have shown the lines differing by a single factor, it is obvious that in most circumstances a number (though not necessarily a *large* number) of loci will contribute to the phenotypic variation. Not all loci are equally effective, however. The locus postulated here has a significant effect upon the phenotype and is located between known markers (*i.e.*, between the recessive mutants *a* and *c*). Other effective loci would have been eliminated from the tester strain by chromosome substitution

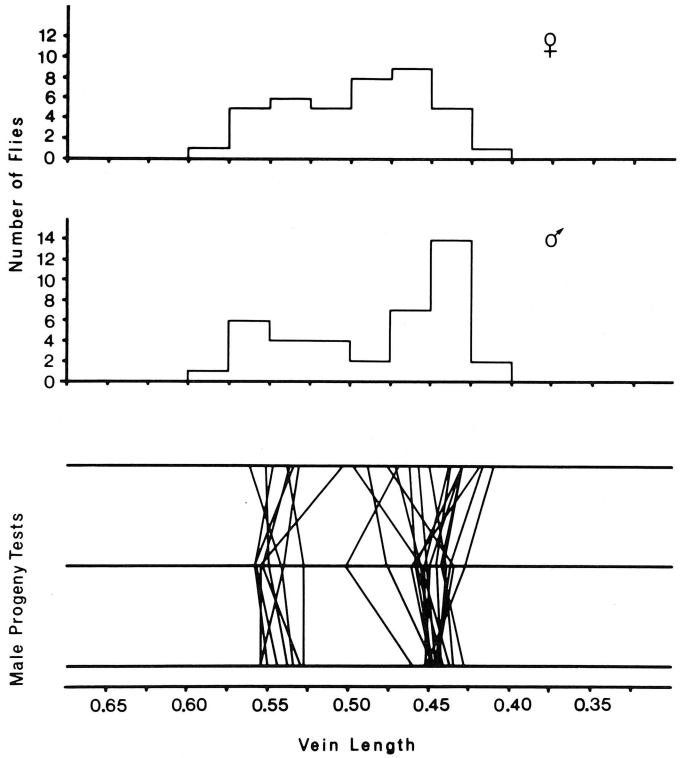


Figure 5. Phenotypic distributions and progeny testing for the isolation of a polygenic locus affecting vein length in *Drosophila* wings carrying the mutant *veinlet*. (from THOMPSON 1975a).

and recombination prior to this stage in order to focus attention upon a single representative polygene. The small variance in the two parental strains is uncontrolled environmental variation.

The F_1 progeny of a cross between the two parental lines are intermediate in expression of the trait. They are phenotypically wild type for the three recessive markers and are heterozygous for the polygenic locus (only postulated as a single

gene locus at this point in the analysis). A simple testcross of heterozygous females to homozygous standard assay strain males will result in a series of recombinant chromosomes that can be recognized phenotypically: *a b c*, *+ + +*, *a + +*, *+ b c*, and so forth. Only single recombinants are illustrated in Figure 4.

Individuals in each recombinant class can then be measured for the expression of the quantitative trait; only heterozygous polygene effects are assayed here, though homozygous effects can easily be measured by modifying the program. The resulting distributions are shown at the right of Figure 4. Let us look first at the *a + +* recombinant chromosome class. If the effective region is between *b* and *c*, all *a + +* recombinant chromosomes will carry the *H* allele (*i.e.*, they will be *a + H +*). From the backcross they will also carry the *a b L c* chromosome and will have intermediate expression of the character (a mean of about 13 units in this example). The *+ b c* chromosomes will carry the *L* allele resulting in a phenotypic distribution like that of the standard (parental) strain. The recombinants between *b* and *c*, however, can be of two types: those with the break occurring between *b* and the polygenic locus and those with the break between the locus and *c*. Thus, the *a b +* chromosome class is actually composed of two chromosome types, resulting in a wider distribution of phenotypes.

The question that remains is,"How many of the chromosomes in the heterogeneous distribution represent recombination to the left of the polygene and how many represent recombination to the right?" This can be established by progeny testing, that is, by testcrossing and reassaying each recombinant chromosome for several generations. An example is shown in Figure 5.

As discussed earlier, terminology is one possible source of confusion in the distinction between major genes and polygenes. To speak of "a" polygene or "located" polygene (SPICKETT and THODAY 1966) is really self-contradictory, since"polygene"is a term that describes the *system* of genes affecting a particular character. A solution to this difficulty was offered by THOMPSON and THODAY (1974), who introduced the term "polygenic locus" to refer to genes isolated, by the above techniques, from the system of genes responsible for the genetic component of variation in some quantitative character. To be precise, a *polygenic locus* is defined as a genetic locus composed of one or more closely linked genes at which allelic substitutions contribute to the variance in a specified quantitative character. A standard nomenclature has been proposed, and some of the problems and limitations associated with locations and their interpretation have been discussed by various authors (THOMPSON and THODAY 1974, ROBERTSON 1967, McMILLAN and ROBERTSON 1974).

Though this approach obviously involves more work than does localization of a Mendelian factor, it is potentially as accurate and has allowed workers to isolate and study a number of polygenic loci affecting a variety of traits (summarized in THOMPSON and THODAY 1974). In addition to those isolated using the progeny testing series, located polygenes have been described by MAC BEAN *et al.*(1971), BOYER *et al.* (1973), and SCHARLOO

(1970), who has identified several loci in an elegant series of studies of wing vein modifiers.

THE NATURE AND FUNCTION OF POLYGENES

With the use of theoretical models, we have seen that a normal distribution, characteristic of quantitative variation, does not necessarily mean that large numbers of genes are involved in producing the observed variation. Then we discussed techniques that would allow us to isolate, count and study individual polygenic loci. The obvious next step is to ask what experimental analysis of polygenic variation has shown us about polygene number and about the nature of the genes that produce the effects we lump together as polygenic variation.

The first thing that is evident from polygene analysis is that, as predicted by theoretical models, one often finds that few loci are responsible for the majority of the variance. THODAY and his colleagues (see references in THOMPSON and THODAY 1974) found that only five loci accounted for more than 85 per cent of the difference between a high line of *Drosophila melanogaster* having a mean of about 40 sternopleural bristles per fly and a control line having about 20 bristles per fly. Similarly, WEHRHAHN and ALLARD (1965) reported that the majority of the variance in heading time in wheat could be accounted for by four loci. Quantitative variation in eye pigmentation in a *Drosophila* stock showing variegated position effect can be attributed to a single locus (THOMPSON unpublished). These results should be contrasted with biometrical estimates which suggest, for example, that 100 or more polygenes influence a similar trait, abdominal bristle number in *Drosophila* (reviewed in FALCONER 1955).

Thus from a growing variety of experimental sources there appears to be no compelling evidence to suggest that the number of polygenes influencing many quantitative characters is exceptionally large. Indeed, this conclusion should not surprise any geneticist, because we are frequently faced with indirect evidence that this is true.

One source of evidence is found in accelerated responses to artificial selection. After a number of generations of selection, one often finds that the response reaches a plateau in which no further progress occurs for perhaps many generations. Then over a small number of generations the phenotype makes a step-wise jump to a new plateau. This is generally interpreted to be the result of rare recombination between closely-linked polygenes in a $\frac{+\ -}{-\ +}$ heterozygote to produce $\frac{+\ +}{}$ and $\frac{-\ -}{}$ chromosomes that allow selection to continue until one is fixed and the genotype of the population is again stabilized (see THODAY *et al.* 1964). The implication is that the observed phenotypic change is attributable to the cumulative effects of these two linked loci and any interactions they may have.

The advance of biochemical genetics is even providing new types of polygenic variation to analyze. The answers from

studies of enzyme activity modifiers (WARD and HEBERT 1972, WARD 1975), for example, support the hypothesis that polygene number is not *necessarily* exceptionally large.

A survey of the literature, however, shows that some complex characters *are* apparently influenced by a very large number of genes (FALCONER 1955). Is this fact really irreconcilable with the evidence presented so far? Are there different kinds of quantitative traits? These questions can best be answered by considering what polygenes are and what influences they may be exerting during development.

But before discussing what quantitative genetics can tell us about polygenes and development, it would be wise to think about what it can *not* tell us. Population genetics and studies of quantitative inheritance are primarily concerned with phenotypes and the changes in gene frequency which come about by selecting different phenotypes. But phenotype frequencies can tell us nothing about primary gene action in the sense of, "What do genes do?" On the other hand, the study of quantitative traits can tell us a lot about the way in which the phenotype is organized. It can show relationships between general measures of gene expression, such as dominance, penetrance, and expressivity, and it can clarify the role of pleiotropy among quantitative traits. Perhaps most important, it can help us understand the processes involved in pattern formation and the interaction of developmental sequences that results in the characters we see.

A number of investigators have approached the problem of the development of quantitative characters without tracing the observed effects to specific polygenic loci. One series of studies, for example, concentrates upon the genetic interrelationship among different characters, such as sternopleural bristle variation versus abdominal bristle variation in *Drosophila*. REEVE (1961), ROBERTSON (1962), FRANKHAM (1970) and many others have used statistical correlations, sometimes accompanied by chromosome substitutions, to measure the genetic component of variation shared by two or more traits. A more direct way to examine pleiotropy and avoid the confounding effects of linkage is to select for modification of two traits in opposite directions (*e.g.*, DAVIES and WORKMAN 1971). In general, these studies show that phenotypically similar traits often differ significantly in the polygenes that influence them, implying that the quantitative traits are, to a large extent, a result of variation in quite independent developmental processes.

Similar studies of the wing vein system of *Drosophila melanogaster* (THOMPSON 1975a, 1975b) have confirmed this conclusion and provided a useful insight into the genetic control of a simple pattern. Chromosome substitutions between selected lines of different vein mutations (THOMPSON 1973, 1975a; Table 1) have shown that some polygenic modifiers of the expression of vein mutants produce their effects independently of the mutant. Selected chromosomes carried polygenes that increased the length of a specific vein, such as the L4, and consequently influenced the phenotype of any mutant in which that vein was variable. These polygenic loci apparently have minor influences upon the

same developmental processes and only indirectly modify the mutant phenotype. Vein-specific polygenes suggested that the individual veins are largely under separate genetic control, and this is confirmed by an analysis of categories of major gene mutations (THOMPSON 1974, 1975a).

Table 1. A sample of the observed effects upon vein length when single whole chromosomes from a selected line are substituted into the unselected background of a non-homologous mutant.

Source*	Vein	Second Mutant	Effect in First Mutant	Effect in Second Mutant
ve Long-*II*	*L2*	*ri*	+	+
ve Long-*III*	*L2*	*ri*	+	+
ri Long-*I*	*L3*	*ve*	none	none
ri Short-*IV*	*L3*	*ve*	none	none
ve Long-*II*	*L4*	*ci*	+	+
ve Short-*II*	*L4*	*ci*	-	-
px High	*fragments*	*shv*	+	+
px Low	*fragments*	*shv*	-	-
net High	*fragments*	*shv*	+	+
net Low	*fragments*	*shv*	-	-

**For description of mutants and selection lines and for levels of significance, see* THOMPSON (1974, 1975a).

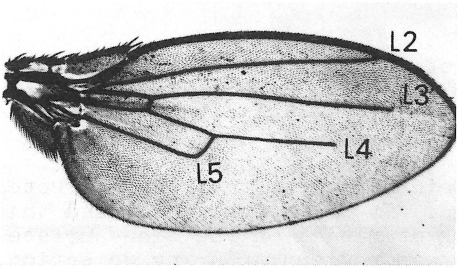


Figure 6. The recessive mutant *veinlet* and vein nomenclature.

The genetic correlation between veins was further tested by artificial selection using the mutant *veinlet*, which has terminal gaps in all longitudinal veins (Figure 6). In one line, the L2 and L4 veins were selected to be shortened, while the L3 was selected to be longer. In the other, the L2 and L3 were selected to be longer, while the L4 was shortened. Both lines responded readily to selection (THOMPSON 1975a), confirming that

at least some polygenes affecting vein length variation act upon individual veins rather than upon vein length in general. Indeed, even different regions of a single vein appear to be under separate genetic control (THOMPSON and THODAY 1975, SCHARLOO 1962).

Selection in a different mutant extended our understanding of polygenic variation and its effects upon vein pattern (THOMPSON 1974). The recessive mutant *plexus* is characterized by vein fragments in many parts of the wing (Figure 7). Two selection lines were maintained: one selected for an increased

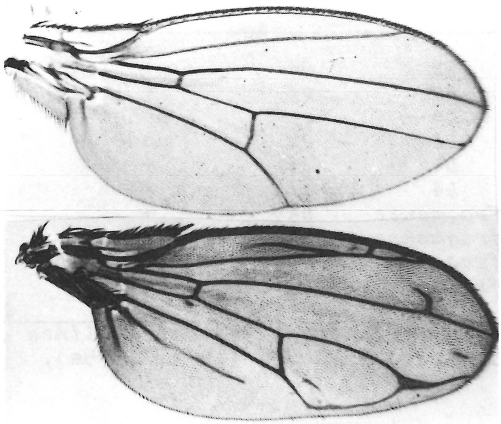


Figure 7. Wings of *Drosophila melanogaster*: normal, above; *plexus*, below. The position of branches and fragments is quite regular and predictable.

amount of extra vein material, the other selected to approach a normal phenotype. It was quickly realized that the fragments were not appearing at random on the wing surface, but were found only in certain regions. During selection, they increased or decreased in size in a predictable manner. By measuring the position and frequency of vein fragments, the regions of preferential vein appearance can be summarized as a series of subthreshold developmental potentials (Figure 8), and all vein mutant phenotypes can be explained in terms of changes in these developmental potentials (probably equivalent to a developmental prepattern) or in terms of changes in an expression threshold which determines the amount of vein material formed in all parts of the wing.

Since the subthreshold vein pattern appears to be a general characteristic of *Drosophila* wing development, one possibility is that it marks regions in which ancestral veins once appeared.

If this is true, one might expect to find correlations between subthreshold veins and the positions of normal veins in related families of Diptera, similar to the bristle correlations reported by MAYNARD SMITH and SONDHI (1961) in a *Drosophila subobscura* selection line. In fact, a large number of correlations are found. The longitudinal vein between the L5 and the posterior

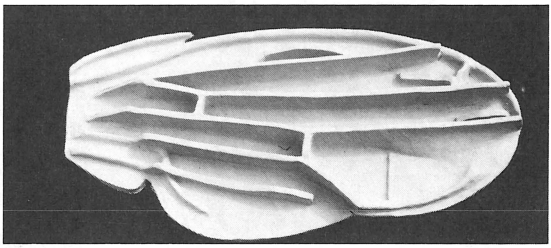


Figure 8. The subthreshold vein pattern represented in this clay model as a series of peaks and valleys with heights corresponding to the frequency with which extra vein fragments are found in each area.

margin of the wing, for example, is found in the wings of even distantly related Diptera. The conclusion implicit in this observation is that one way the *Drosophila* vein pattern might have evolved was by the selection of polygenes or mutant alleles which decreased the amount of necessary enzymes in specific regions or altered the presumptive prepattern in localized ways. In this way some ancestral veins were essentially erased, leaving only the subthreshold vein potentials, which are either closely associated with normal development or are not eliminated simply because no phenotype normally exists for selection to act upon.

WADDINGTON (1973) has stressed that in order to understand the causal structure of a pattern one must be able to make controlled changes in it. This brief summary should prove that artificial selection is a simple, useful tool in this type of developmental analysis. Selection and the isolation of polygenes can make a further contribution to our understanding of the phenotype: it can indicate the types of contributing processes that interact to make up the phenotype we study. Such an approach is elegantly illustrated by the work of SPICKETT (1963). Sternopleural chaetae are the small bristles on the side of *Drosophila* adults (Figure 9). SPICKETT isolated and analyzed three polygenic loci accounting for the majority of the difference between a selected high chaetae number line and an inbred control line. His results are summarized in Figure 10, taken from a review of polygene analysis by THODAY (1967). One polygenic locus caused a general increase in cell number, thus indirectly increasing the number of bristle-forming cells. Another locus

balanced the effect of the first by decreasing cell size, so there was no net increase in fly size. This latter might be thought of as an example of a locus contributing to variation in a character without directly influencing the development of that character. A second locus with direct effect increased the distribution of small ventral bristles, perhaps by influencing

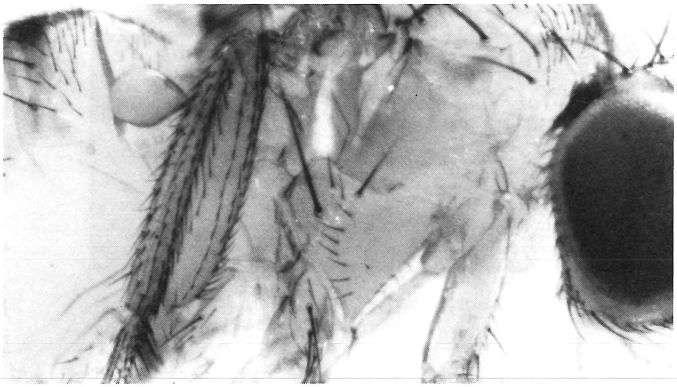


Figure 9. Photograph of the side of an adult *Drosophila melanogaster* female showing the sternopleural bristles.

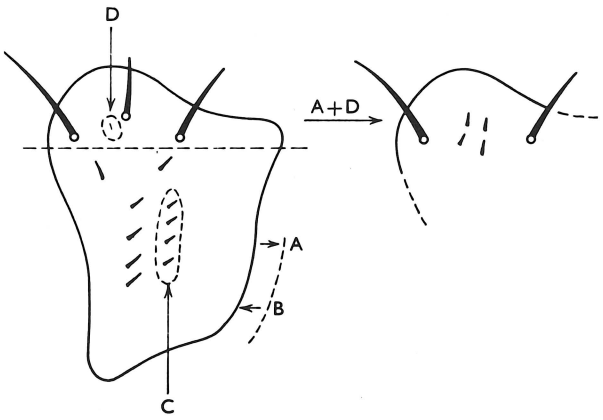


Figure 10. Summary of SPICKETT's analysis of polygenic loci affecting sternopleural bristle number. *A*, increase in cell number; *B*, decrease in cell size; *C*, increase in bristle distribution; *D*, change in timing of cell division. (from THODAY 1967).

the distribution of some bristle precursor. A third locus modified the timing of cell division in one region, so that the

central dorsal bristle-forming cell divided one or more extra times before bristles were formed. Each locus affected a different developmental process, illustrating the complexity inherent in quantitative variation. One of the polygenic loci influencing vein development has also been isolated and described in detail (THOMPSON and THODAY 1975, THOMPSON 1975), confirming the hypothesis that quantitative traits are the result of complex interactions involving a number of contributory processes. Other examples could also be cited (JACOBSON 1967 Quoted in FRASER 1967, HOOPER 1976, and others).

This perspective of quantitative phenotypes suggests that the more "complex" a trait, *i.e.*, the more contributory processes involved, the more polygenes one would expect to influence it. Thus, although a trait may appear highly complex and affected by a large number of polygenes, it is still likely that informative locations and developmental analysis could be done (*c.f.* SPICKETT *et al.* 1967). Indeed, there is evidence that polygenes are nothing more than minor mutants, or isoalleles, of loci also known through major mutations (see SCHARLOO 1970), though the situation is not necessarily simple (THOMPSON 1975c).

Although we still have much to learn about the nature of polygenic variation and the number of genes involved, there can be no doubt that we now have appropriate techniques to answer the questions. Hopefully, our better understanding will lead us to a meaningful appreciation of both the complexity and the simplicity of quantitative characters.

ACKNOWLEDGMENTS

I would like to thank Dr. Ron Woodruff, Dr. Jenna Hellack, Brent N. Hisey, G. Bradley Schaefer and Timothy N. Kaiser for helpful comments on the manuscript and Wendy Oberlin for the illustration of Aesop's fable. I am also happy to acknowledge that many of the ideas summarized here were generated in collaboration with Prof. J.M. Thoday, who also kindly gave permission to reproduce Figure 10. This paper was written during the tenure of DHEW-NIH grant 1-RO1-ES01439-01.

LITERATURE CITED

ALLARD, R.W. 1960 Principles of Plant Breeding. Wiley. London.

BOYER, B.J., D.L. PARRIS, and R. MILKMAN 1973 The crossveinless polygenes in an Iowa population. Genetics 75: 169-179.

BREESE, E.L. and K. MATHER 1957 The organisation of polygenic activity within a chromosome in *Drosophila*: I. Hair characters. Heredity 11: 373-395.

BREESE, E.L. and K. MATHER 1960 The organisation of polygenic activity within a chromosome in *Drosophila*: II. Viability. Heredity 14: 375-399.

BRIDGES, C.B. 1919 Specific modifiers of eosin eye color in *Drosophila melanogaster*. J. Exp. Zool. 28: 337-384.

DAVIES, R.W. and P.L. WORKMAN 1971 The genetic relationship of two quantitative characters in *Drosophila melanogaster*.

I. Responses to selection and whole chromosome analysis. Genetics 69: 353-361.
FALCONER, D.S. 1960 Introduction to Quantitative Genetics. Oliver and Boyd. London.
FRANKHAM, R. 1970 The effects of different chromosomes on four bristle number characters in *Drosophila melanogaster*. Australian J. Biol. Sci. 23: 503-505.
FRASER, A. 1967 Variation of scutellar bristles in *Drosophila*. XV. Systems of modifiers. Genetics 57: 919-934.
HOOPER, A.C.B. 1976 Longitudinal growth of skeletal muscle fibres in lines of mice selected for high and low body weight. Growth 40: 33-39.
HOSGOOD, S.M.W. and P.A. PARSONS 1967 The exploitation of genetic heterogeneity among the founders of laboratory populations of *Drosophila* prior to directional selection. Experientia 23: 1066.
HOSGOOD, S.M.W. and P.A. PARSONS 1968 Polymorphism in natural populations for the ability to withstand temperature shocks in *Drosophila*. Experientia 24: 727-728.
JOHANNSEN, W. 1909 Elemente der exakten Erblichkeitslehre. Fischer. Jena.
LAW, C.N. 1966 The location of genetic factors affecting a quantitative character in wheat. Genetics 53: 487-498.
LAW, C.N. 1967 The location of genetic factors controlling a number of quantitative characters in wheat. Genetics 56: 445-461.
LAW, C.N. and A.J. WORLAND 1973 Chromosome substitutions and their use in the analysis and prediction of wheat varietal performance. Proc. IVth Int. Wheat Genet. Symp. 41-49.
LERNER, I.M. 1972 Polygenic inheritance and human intelligence. Evolutionary Biology 6: 399-414.
MATHER, K. 1941 Variation and selection of polygenic characters. J. Genet. 41: 159-193.
MATHER, K. 1943 Polygenic inheritance and natural selection. Biol. Rev. 18: 32-64.
MATHER, K. 1949 Biometrical Genetics. Methuen. London.
MATHER, K. and B.J. HARRISON 1949 The manifold effects of selection. Heredity 3: 1-52, 131-162.
MATHER, K. and J.L. JINKS 1971 Biometrical Genetics. Chapman and Hall. London.
MAYNARD SMITH, J. and K.C. SONDHI 1961 The arrangement of bristles in *Drosophila*. J. Embryol. Exp. Morph. 9: 661-672.
MAC BEAN, I.T., J.A. McKENZIE, and P.A. PARSONS 1971 A pair of closely linked genes controlling high scutellar chaeta number in *Drosophila*. Theoret. Appl. Genet. 41: 227-235.
McMILLAN, I. and A. ROBERTSON 1974 The power of methods for the detection of major genes affecting quantitative characters. Heredity 32: 349-356.
MILKMAN, R.D. 1965 The genetic basis of natural variation. VII. The individuality of polygenic combinations in *Drosophila*. Genetics 52: 789-799.
PARSONS, P.A. and S.M.W. HOSGOOD 1967 Genetic heterogeneity among the founders of laboratory populations of *Drosophila*. I. Scutellar chaetae. Genetica 38: 328-339.
PARSONS, P.A., S.M.W. HOSGOOD, and B.T.O. LEE 1967 Polygenes and polymorphism. Molec. Gen. Genetics 99: 165-170.
RASMUSSON, J.M. 1935 Studies on the inheritance of quantitative

characters in *Pisum*: I. Preliminary note on the genetics of time of flowering. Hereditas 20: 161-180.

REEVE, E.C.R. 1961 Modifying the sternopleural hair pattern in *Drosophila* by selection. Genet. Res. 2: 158-160.

RIEGER, R., A. MICHAELIS, and M.M. GREEN 1968 A Glossary of Genetics and Cytogenetics. Springer-Verlag. Berlin.

ROBERTSON, A. 1966 Artificial selection in plants and animals. Proc. Roy. Soc. Lond., B, 164: 341-349.

ROBERTSON, A. 1967 The nature of quantitative genetic variation. pp. 265-280. Heritage from Mendel. (Brink, R.A., Ed.) Univ. Wisconsin Press. Madison.

ROBERTSON, F.W. 1962 Changing the relative size of the body parts of *Drosophila* by selection. Genet. Res. 3: 169-180.

SAX, K. 1923 The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8: 552-560.

SCHARLOO, W. 1962 The influence of selection and temperature on a mutant character (ci^D) in *Drosophila melanogaster*. Arch. Neerl. Zool. 14: 431-512.

SCHARLOO, W. 1970 Stabilizing and disruptive selection on a mutant character in *Drosophila*. II. Polymorphism caused by a genetical switch mechanism. Genetics 65: 681-691.

SISMANIDIS, A. 1942 Selection for an almost invariable character in *Drosophila*. J. Genet. 44: 204-215.

SPICKETT, S.G. 1963 Genetic and developmental studies of a quantitative character. Nature 199: 870-873.

SPICKETT, S.G. and J.M. THODAY 1966 Regular responses to selection. 3. Interaction between located polygenes. Genet. Res. 7: 96-121.

SPICKETT, S.G., J.G.M. SHIRE, and J. STEWART 1967 Genetic variation in adrenal and renal structure and function. pp.271-291. Endocrine Genetics (Spickett, S.G. and J.G.M. Shire, Eds.) Mem. Soc. Endocrin. 15.

THODAY, J.M. 1961 Location of polygenes. Nature 191: 368-370.

THODAY, J.M. 1967 Summing up: Uses of genetics in physiological studies. pp. 297-308. Endocrine Genetics (Spickett, S.G. and J.G.M. Shire, Eds.) Mem. Soc. Endocrin. 15.

THODAY, J.M. and J.N. THOMPSON, JR. 1976 The number of segregating genes implied by continuous variation. Genetica 46: 335-344.

THODAY, J.M., J.B. GIBSON, and S.G. SPICKETT 1964 Regular responses to selection. 2. Recombination and accelerated response. Genet. Res. 5: 1-19.

THOMPSON, J.N., JR. 1973 General and specific effects of modifiers of mutant expression. Genet. Res. 22: 211-215.

THOMPSON, J.N., JR. 1974 Studies on the nature and function of polygenic loci in *Drosophila*. II. The subthreshold wing vein pattern revealed in selection experiments. Heredity 33: 389-401.

THOMPSON, J.N., JR. 1975a Studies on the nature and function of polygenic loci in *Drosophila*. III. veinlet modifiers having region-specific effects upon the vein pattern. Genetics 81: 387-402.

THOMPSON, J.N., JR. 1975b Quantitative variation and gene number. Nature 258: 665-668.

THOMPSON, J.N., JR. 1975c A test of the influence of isoallelic variation upon a quantitative character. Heredity 35:

401-406.
THOMPSON, J.N., JR. 1976 Reply to "Evidence for polygenes". Nature 261: 525-526.
THOMPSON, J.N., JR. and J.M. THODAY 1974 A definition and standard nomenclature for "polygenic loci". Heredity 33: 430-437.
THOMPSON, J.N., JR. and J.M. THODAY 1976 Estimating polygene allele frequencies in a natural population of *Drosophila melanogaster*. Genetics 83: s76.
VETTA, A. 1976 Evidence for polygenes. Nature 261: 525-526.
WADDINGTON, C.H. 1973 The morphogenesis of patterns in *Drosophila*. pp. 499-535. Developmental Systems: Insects, Vol. 2. (Counce, S.J. and C.H. Waddington, Eds.) Academic Press. New York.
WARD, R.D. 1975 Alcohol dehydrogenase activity in *Drosophila melanogaster*: a quantitative character. Genet. Res. 26: 81-93.
WARD, R.D. and P.D.N. HEBERT 1972 Variability of alcohol dehydrogenase activity in a natural population of *Drosophila melanogaster*. Nature New Biology 236: 243-244.
WEHRHAHN, 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.
WIGAN, L.G. 1949 The distribution of polygenic activity on the X chromosome of *Drosophila melanogaster*. Heredity 3: 53-66.

Front row: Drs. Feldman, Riddle, Thompson, Janet Stein and Gary Stein

POLYPEPTIDE COMPOSITION OF FRACTION I PROTEIN AS AN AID IN THE STUDY OF PLANT EVOLUTION

(chloroplast DNA, age of genera and species, origin of genomes, composition of proteins as affected by amphiploidy)

HIROFUMI UCHIMIYA, KEVIN CHEN AND S. G. WILDMAN

Department of Biology
Molecular Biology Institute
University of California
Los Angeles, California 90024

SUMMARY

Fraction 1 protein (F-1-P), found in all green plants, consists of eight large subunits (LS) coded by chloroplast DNA and eight small subunits (SS) coded by nuclear DNA. By electrofocusing in 8M urea many different species of F-1-P from all parts of the Plant Kingdom, the LS is invariably found to resolve into three polypeptides of different isoelectric points and the SS into from one to four polypeptides. During evolution of new plant species by amphiploidy the composition of the LS is always determined by the maternal parent, but both parents make equal contributions to the composition of the SS. In the case of N. tomentosiformis (n=12; 1SS) ♂ x N. sylvestris (n=12; 1SS) ♀ → N. tabacum (n=24; 2SS) ♂ x N. glutinosa (n=12; 2SS) ♀ → N. digluta (n=36; 4SS) a F-1-P evolved containing 4SS polypeptides. None of these arose by point mutation during speciation, although the SS of F-1-P of N. digluta could have eight differences in amino acid sequence compared to N. sylvestris or N. tomentosiformis, the differences being the consequence of amphiploidy. Using this example, together with F-1-P composition in parasexual hybrids, it is hypothesized that the genetic information for more than one SS polypeptide is non-allelic and most likely located on heterologous chromosomes. The study of F-1-P in 62 species of Nicotiana provides a model system whereby the LS is an indicator of the evolutionary age of one genus relative to another and the SS an indicator of the age of one species relative to another. The SS can also serve to designate a new species of plant, while the LS has been used to trace the origin of genomes in amphiploids.

MACROMOLECULAR NATURE OF FRACTION I PROTEIN

Fraction I protein is the major soluble protein of plant leaves which assimilate carbon dioxide by the Calvin-Benson path-

way. During photosynthesis the protein catalyzes the combination of carbon dioxide from the atmosphere with ribulose 1,5 diphosphate within the leaf cells to produce two molecules of 3-phosphoglyceric acid. Some of the phosphoglyceric acid is used to resupply the ribulose 1,5 diphosphate, whereas the remainder is mainly converted to starch. Fraction I protein is located in the mobile phase of chloroplasts. In living cells of leaves the visible manifestation of its catalytic activity is the rapid accumulation and growth in size of starch grains, also located in the mobile phase of chloroplasts (WILDMAN *et al.* 1974). In succulent leaves approaching major expansion in area, like those of tobacco, spinach, pea, alfalfa, soybean and sugar beet, for example, Fraction I protein will constitute up to one per cent of the fresh weight of the leaves. In tobacco, tomato, petunia and egg plant, Fraction I protein can be isolated in high yield as a pure, crystalline protein using a remarkably simple method (CHAN *et al.* 1972).

In all preparations from eucaryotic organisms so far analyzed, Fraction I protein is composed of eight large subunits combined with eight small subunits. In the case of the crystalline protein, electron microscopy and X-ray crystallography have indicated how the large and small subunits are put together (BAKER *et al.* 1977). As shown by the diagram in Figure 1, the eight large subunits are organized as a cube, with pairs of small subunits probably occupying four out of six faces of the cube. Each monomeric large subunit contains about 450 amino acids and has a molecular weight of ca. 55,000 daltons, so that the cube of eight large subunits produces a mass of ca. 440,000 daltons. The small subunits are composed of ca. 100 amino acids (ca. 12,500 daltons), so that these eight add another 100,000 daltons to the macromolecule, which in its native state has a molecular weight in excess of a half million daltons.

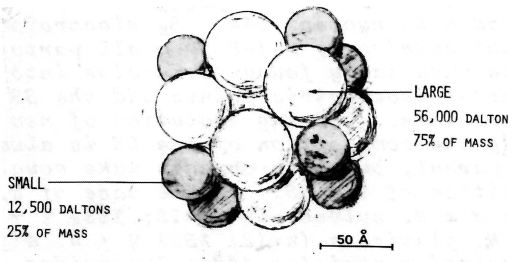


Figure 1. Structure of Fraction I protein.

HETEROGENEITY IN THE POLYPEPTIDE COMPOSITION OF THE LARGE AND SMALL SUBUNITS

Whereas the large and small subunits of Fraction I protein are homogeneous in regard to molecular weight, the large subunit is invariably, and the small subunit sometimes, heterogeneous in regard to the isoelectric points of their polypeptides (CHEN *et al.* 1976a). One type of heterogeneity is illustrated in Figure 2. Here pure crystalline Fraction I protein from *Nicotiana*

tabacum leaves was dissociated into its large and small subunits by 8M urea, carboxymethylated and then subjected to isoelectric focusing. This procedure caused the protein of the large subunit to resolve into three polypeptides whose isoelectric points are slightly different from one another. The electrofocusing also caused the small subunit protein to separate into two polypeptides of different isoelectric points. The heterogeneity displayed by the polypeptides has proved to be remarkably useful in studies on the evolution of Fraction I protein and also on the origin of plant species. This comes about because the number and isoelectric points of the polypeptides serve as phenotypic markers for genes located in chloroplast DNA as well as genes contained in nuclei. The origin of the isoelectric points of the three large subunit and two small subunit polypeptides of the *N. tabacum* Fraction I protein illustrates the mode of evolution of Fraction I protein.

EVOLUTION OF *N. tabacum* FRACTION I PROTEIN

CLAUSEN (1932), GOODSPEED (1954) and KOSTOFF (1938) had concluded on cytological, morphological and geographical grounds that *N. sylvestris* (n=12) and *N. tomentosiformis* (n=12) were the diploid parents of the *N. tabacum* (n=24) amphiploid. Several recent studies (GRAY *et al.* 1974; KAWASHIMA *et al.* 1976; STRØBAEK *et al.* 1976) on the polypeptide composition of the Fraction I protein contained in the three plant species in question have converged to authenticate the previous conclusion. The chemical analyses have carried the knowledge an additional step forward by demonstrating that *N. sylvestris* had to be the female parent and that the genetic information for the small subunit was contained on heterologous chromosomes, one donated by each parent in the original hybridization. The coding information for the Fraction I protein large subunit polypeptides is contained in chloroplast DNA (CHAN and WILDMAN 1972) and is inherited only via the maternal line. The maternal mode of inheritance has been demonstrated for species of Nicotiana (SAKANO *et al.* 1974), Avena (STEER 1975), Triticum (CHEN *et al.* 1975) and, in our unpublished experiments, for species belonging to the

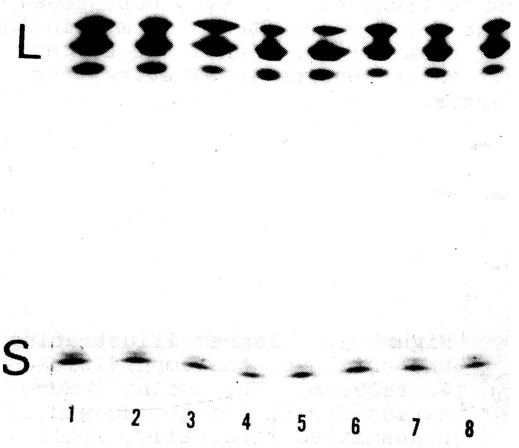


Figure 2. Composition of Fraction I protein from cultivars of *N. tabacum* shown by electrofocusing. 1,2,3,4: n, 2n, 3n and 4n cultivars from U.S.; 5,6,7,8: cultivars from Europe and Japan.

Cruciferae and Malvaceae. Thus, on the basis of present knowledge, only one of the two parents in an interspecific hybrid can contribute the genetic information controlling the primary structure of the large subunit polypeptides during evolution of a new species of plant. In contrast, the genetic information for the small subunit polypeptides is contained in nuclear DNA, permitting both parents to make an equal contribution to the new species. The diagram reproduced in Figure 3 illustrates the specific case of how the Fraction I protein contained in *N. tabacum* originated.

The large subunits of the Fraction I protein in the parents of *N. tabacum* contain a difference in chymotryptic peptides (KAWASHIMA *et al.* 1976). Only *N. sylvestris* large subunit protein has a peptide composition identical to *N. tabacum* large subunit protein. Similarly, *N. sylvestris* has three polypeptides whose isoelectric points correspond to those in *N. tabacum*, whereas those in the other parent do not. Therefore the differences in composition of the large subunit obtained by electrofocusing have their counterpart in differences contained in the primary structure of the amino acid sequence of the large subunit protein.

The small subunit of *N. tabacum* Fraction I protein consists of two polypeptides of different isoelectric points. These have been separated and partially sequenced and then compared to the sequences for the single component polypeptides found in the two parents of *N. tabacum* (STRØBAEK *et al.* 1976). The comparison is shown in Figure 4. There is some dispute as to whether the N-terminal amino acid is methionine (IWAI *et al.* 1976) but agreement in regard to order of amino acids thereafter. If methionine is the beginning, then, reading from left to right, the seventh amino acid in the small subunit is isoleucine in *N. sylvestris* and tyrosine in *N. tomentosiformis*.

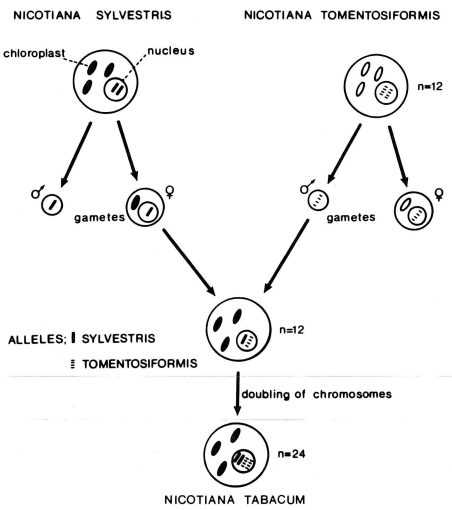


Figure 3. Diagram illustrating the origin of the amphidiploid *N. tabacum*. The coding information for the large subunit is transmitted maternally, while the coding information for the small subunit has biparental transmission. From STRØBAEK *et al.* (1976).

Both of these amino acids occupy the seventh position in the unseparated polypeptides of the *N. tabacum* protein. The important point to be made out of these data is that the change from an isoleucine to a tyrosine cannot be accounted for by a point mutation in the genetic code for the small subunit, that is, the coding triplet for tyrosine (UAU or C) cannot be transformed to the triplet for isoleucine (AUC or A) by changing just one nucleotide. Rather, it is evident that small subunit polypeptides distantly related in regard to primary structure can be combined into the structure of the oligomeric Fraction I protein by the single step of hybridization between different species of plants. Furthermore, the data permit the prediction that information for each of the two polypeptides in *N. tabacum* Fraction I protein will be located on heterologous chromosomes, one chromosome having come from *N. sylvestris* and the other from *N. tomentosiformis*.

SPECIES	NUMBER OF POLYPEPTIDES	AMINO ACID SEQUENCE
N. tomentosiformis	1	H_2N-MET.....TYR-GLY.....
N. tabacum	2	H_2N-MET.....TYR-ASN..... / ILE GLY
N. sylvestris	1	H_2N-MET.....ILE-ASN.....

Figure 4. A region in the primary structure of the small subunit polypeptides of the amphidiploid *N. tabacum* Fraction I protein whose polymorphism could not be explained by a point mutation in the sequence of either of its diploid parental species.

EVOLUTION OF THE FRACTION I PROTEIN CONTAINED IN *N. digluta*

A further step in the evolution of the primary structure of Fraction I protein during speciation of Nicotiana has been seen in the case of *N. digluta* (n=36). CLAUSEN and GOODSPEED (1925) made the hybrid *N. glutinosa* ♀ x *N. tabacum* ♂. The seeds of the F_1 hybrid germinated and produced plants which, though vigorous, remained self-infertile and also infertile to either *N. glutinosa* or *N. tabacum* pollen. Evidently, infertility was due to failure of chromosome pairing during meiosis. A spontaneous doubling of somatic chromosomes occurred in one individual in the population. Doubling the chromosomes produced a plant which could self-fertilize and survive alternation of generations; consequently, it was eligible to be called the new species, *N. digluta*. The large subunit polypeptides of *N. glutinosa* have different isoelectric points than those of *N. tabacum*. *N. digluta* has those corresponding to *N. glutinosa* as demanded by maternal inheritance. The Fraction I protein of *N. glutinosa* has two small subunit

polypeptides, but their isoelectric points are different from the two in the protein from *N. tabacum*. The small subunit of *N. digluta* contains four polypeptides corresponding in isoelectric points to the two in *N. glutinosa* and the two in *N. tabacum* (KUNG *et al.* 1975). By extrapolation from the *N. tabacum* analysis, we would also suspect that the coding information for each of the *N. glutinosa* small subunit polypeptides is contained on different chromosomes. If not, the two genes must be at different loci on the same chromosome and are therefore non-allelic. There is no segregation of the two polypeptides following self-fertilization of *N. glutinosa* or cross-pollination by another member of the population. We make the further presumption that the two chromosomes in *N. glutinosa* are heterologous with respect to the two chromosomes containing the genetic code for small subunit polypeptides in *N. tabacum*. Extending the argument, the genetic information for the four small subunit polypeptides in *N. digluta* is thought to be distributed on four, heterologous chromosomes. It is clear, however, that further genetic analysis is badly needed. Otherwise we cannot be confident that linkage is not operative in this system, although our preference favors the four-chromosome hypothesis.

In tracing the origin of the *N. digluta* protein from the first hybridization of *N. sylvestris* and *N. tomentosiformis* to produce *N. tabacum*, the striking feature is that there is no evidence that mutations in the coding information for the small subunit polypeptides sufficient to change polypeptide isoelectric points accompanied the evolutionary process which increased the polypeptide number from one to four. It is equally striking that maternal inheritance of the coding information for the large subunit polypeptides was adhered to strictly during the two stages of speciation which resulted in the appearance of *N. digluta*. In our unpublished experiments one backcross of *N. digluta* with *N. tabacum* pollen was enough to remove two of the four small subunit polypeptides from Fraction I protein. The lost polypeptides corresponded in isoelectric points with the two found in *N. glutinosa* protein. This seems to be further evidence in favor of the view that genetic information for the lost polypeptides had been contained on chromosome(s) originally derived from the *N. glutinosa* genome.

FORMATION OF NEW SPECIES OF FRACTION I PROTEIN BY PARASEXUAL HYBRIDIZATION OF PLANT PROTOPLASTS

CARLSON *et al.* (1972) discovered that a new species of *Nicotiana* which is self-fertile could be produced by the initial act of fusing protoplasts of *N. glauca* (n=12) mesophyll tissue with those of *N. langsdorffii* (n=9). Fifteen *N. glauca* + *N. langsdorffii* parasexual hybrids have been created (SMITH *et al.* 1976). Eight of the fifteen parasexual hybrids contained large subunit polypeptides of the *N. langsdorffii* type and seven of the *N. glauca* type. All fifteen parasexual hybrids contained both *N. glauca* and *N. langsdorffii* nuclear genomes whose DNA was translated into three different small subunit polypeptides.

Why parasexually produced plants cannot tolerate the presence of two chloroplast DNA genomes operating in the same cell remains a mystifying question. The Fraction I protein small subunit of *N. glauca* consists of a single polypeptide whose isoelectric point is distinguishable from the two polypeptides of the *N. langsdorffii* small subunit. Since all parasexual hybrids have Fraction I proteins containing three small subunit polypeptides, we reason by analogy that the genetic information for the *N. glauca* polypeptide is on one chromosome and information for the *N. langsdorffii* polypeptides on two other chromosomes. In all hybrids the amounts of each of the three small subunit polypeptides appear to be equal, so that dominance is not a factor in the expression of the genetic information coming from different chromosomes.

The Fraction I protein in the selfed F_2 progeny from eleven of the parasexual hybrids has been found to contain the same three small subunit polypeptides. In two instances a selfed F_3 generation Fraction I protein was found to contain the same three small subunit polypeptides. Therefore the analysis of the Fraction I protein polypeptide composition of parasexual hybrid plants provides additional evidence in support of the argument that the genetic information for the small subunit amino acid sequence is located on heterologous chromosomes in the case of Fraction I proteins composed of more than one small subunit polypeptide.

The probable separation of information on different chromosomes provides the potential for a plant protein to undergo a relatively enormous change in its composition by a single evolutionary event, as illustrated by Fraction I protein. The two polypeptides of the small subunit of *N. tabacum* protein have been partially sequenced. In the first forty amino acids, two differences have been found (Figure 4). From amino acid analyses, it is predicted that two additional differences will be found in the sixty-two amino acids remaining to be sequenced (STRØBAEK *et al.* 1976). Thus the evolutionary event which produced *N. tabacum* was accompanied by four changes in amino acid sequence of the small subunit. If we assume that differences of similar magnitude are responsible for separation of the two polypeptides of *N. glutinosa*, then in reaching the evolutionary state of Fraction I protein in *N. digluta* eight differences could have occurred in the amino acid composition without any of them being attributable to point mutations in the genetic code. What is seen in this example at the molecular biological level is stark testimony to the accuracy of the view, so frequently advanced by plant evolutionists, that amphidiploids do not produce anything new in the genetical sense, but simply add together or combine genetic information already pre-existing in the parents. But evolution by amphidiploidy has produced, according to the estimates of SOLBRIG (1966), 40% of the world's total plant species.

Isozymes also arise by amphidiploidy, ferredoxin being an example. Two species of ferredoxin have become a part of the phenotype of *N. digluta*, whereas only one form exists in *N. tabacum* and *N. glutinosa*, the parents of *N. digluta* (KWANYUEN and

WILDMAN 1975). It can be imagined that species already containing a number of forms of an enzyme (like peroxidase, for example) could dramatically increase that number if speciation were to follow hybridization and amphidiploidy. As in Fraction I protein, the number could change dramatically and still be unaccompanied by any mutations in the genetic information coding for the individual proteins in the mixture.

HYPOTHESIS ON ORIGIN OF SMALL SUBUNIT POLYPEPTIDES DURING EVOLUTION OF FRACTION I PROTEIN IN THE GENUS NICOTIANA

The genus Nicotiana consists of 65 species (GOODSPEED 1954; BURBIDGE 1960). The polypeptide composition of Fraction I protein is known for sixty-two of these species (CHEN *et al.* 1976b). Only four differences are found in the isoelectric points of the three polypeptides of the large subunit coded by chloroplast DNA. During the evolution of the sixty-two species, thirteen differences have appeared in the isoelectric points of the small subunit polypeptides. Of the sixty-two species, forty-three contain two or more small subunit polypeptides. An understanding of how a Fraction I protein in an amphiploid comes to contain more than one small subunit polypeptide is apparent from the previous analysis of how the two polypeptides in *N. tabacum* and the four in *N. digluta* arose. However, nineteen species of Fraction I protein contain only a single small subunit polypeptide. These are all contained in diploids. Moreover, the isoelectric point of the single small subunit polypeptide of one species of a diploid Nicotiana Fraction I protein can be the same as, or different from, the isoelectric point of the protein isolated from another species of Nicotiana. There are five types of species of Fraction I protein containing a single small subunit polypeptide.

The origin of Fraction I proteins containing single small subunit polypeptides of differing isoelectric points is most likely the consequence of point mutations in the genetic code. The probable manner of how a small subunit polypeptide changes its isoelectric point is illustrated by the diagram in Figure 5. The basic premise is that the genetic information is contained in the alleles on a single chromosome. We presume that in a huge population of a single species of Nicotiana, designated as SS1 in Figure 5, a point mutation in the genetic code for one of the 100 amino acids of the small subunit occurs in one member of the population. Such an occurrence might be at the level of one in 10^6 individual plants. The mutation results in the replacement of a single amino acid in the primary structure of the small subunit. If the new amino acid has a charge that is different from the former, this will be reflected in a change in the isoelectric point of the polypeptide. The plant in which the mutation occurred now contains a chromosome in which the unmutated allele produces one polypeptide of the original isoelectric point and the mutated allele produces a polypeptide of different isoelectric point. If this plant, now containing two polypeptides, reproduces either by self- or by cross-fertilization with

another individual in the population, the two alleles will segregate, resulting in restoration of the previous species of Fraction I protein together with production of an entirely new species of protein as indicated by SS2 in Figure 5. Self-fertilization will perpetuate this new species of Fraction I protein containing a single small subunit polypeptide of an isoelectric point different from the single polypeptides in the SS1 population. Point mutation in the SS2 population could produce an alteration in the coding information contained in the single chromosome controlling the small subunit, leading to a single polypeptide with another difference in isoelectric point. Such events may have produced all of the thirteen differences that have evolved.

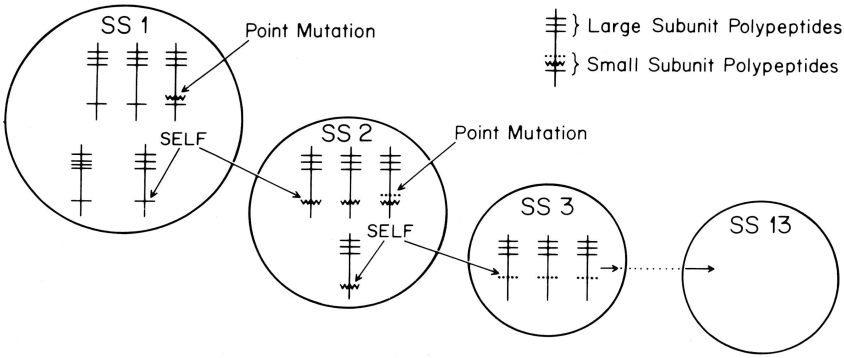


Figure 5. Probable manner in which the single small subunit polypeptides of Fraction I proteins undergo a change in isoelectric points.

In this long period of speciation the genetic information has presumably found its way by translocation onto as many as four different chromosomes in some polyploid species of Nicotiana which have Fraction I proteins displaying four small subunit polypeptides.

SPECIAL FEATURES OF FRACTION I PROTEIN EVOLUTION OBSERVED IN NICOTIANA WHICH ALSO APPLY TO OTHER GENERA

There has been no deviation encountered in the isoelectric points of large and small subunits of Fraction I proteins isolated from seventeen self-fertile cultivars of *N. tabacum* involving numerous repetitions of the analyses of different populations and individual plants within the populations. No changes have been found in *N. glauca* Fraction I protein isolated from plants whose self-fertile parents were growing on each continent of the world. The composition of Fraction I protein within a species therefore appears to be unusually stable since it has not gratuitously changed during centuries of plant breeding by

man or by the vicissitudes of Nature as affected by geography and edaphic factors. The large subunit polypeptide composition, because of its extraordinary stability during geological time, can be used as an estimate of the relative age of different plants.

Comparative Age of Genera Judged by Number of Differences in Composition of the Large Subunit

Twenty species of Nicotiana are unique to Australia, whereas the remaining 45 species are confined to the Western Hemisphere. GOODSPEED (1954) proposed that separation and isolation of the Australian species were the consequence of continental drift which commenced in the Cretaceous period. RAVEN and AXELROD (1974) have other ideas on the possible age of angiosperms and the origin of species. However, their view was advanced before it was known that two species of Nicotiana located at the tip of South America contain Fraction I protein compositions which have large subunit polypeptides of the type found in no other species except those in Australia (CHEN *et al.*1976b). Furthermore, these proteins have one small subunit polypeptide found elsewhere only in Australian species and another found only in Western Hemisphere species. The Fraction I proteins of these two species therefore appear to be relics of the proteins in those species which evolved into the Australian species. Their geographical location is supportive of the Goodspeed hypothesis, which proposed that the Australian species have been evolving in isolation from the Western Hemisphere species for 75-100 million years.

In the entire genus Nicotiana only four types of Fraction I protein large subunit polypeptides have been encountered (CHEN *et al.* 1976b). It was therefore estimated that $2^x=4$ is an indication that only two mutations in chloroplast DNA have survived during the entire life span of the genus. No differences were encountered among the Australian Fraction I protein large subunits. Therefore, if Goodspeed's contention is plausible, it follows that no surviving mutations in chloroplast DNA coding for the isoelectric points of the three Fraction I protein large subunit polypeptides have occurred in the last 75-100 million years among the Australian species. This would place the maximum surviving mutation rate for chloroplast DNA coding for Fraction I protein at 0.22 mutations/100 amino acids/10^8 years. The rate could be less. Pursuing this line of reasoning further, it would appear that, if more than 75 million years elapse between surviving mutations, the genus Nicotiana must have been evolving for at least 150 million years for the two surviving mutations to have occurred. GOODSPEED (1954) proposed that the Australian Nicotianas are of younger vintage than Western Hemisphere species because the former are all amphiploids. The Fraction I protein analyses agree with this notion. No Australian species has a Fraction I protein small subunit with less than two polypeptides, and 90% of the species have three or four polypeptides. However, six of the thirteen small subunit poly-

peptides are unique to the twenty Australian species. One of these six is found in those two Nicotiana species located at the tip of South America. Amphiploidy may have been a factor in the more rapid evolution of the polypeptides in Australian Fraction I proteins since amphiploidy provides a greater chance for point mutations to occur.

Another manner of looking at the problem of the age of Nicotiana is to consider the probable mutation rate of the small subunit polypeptide. Thirteen polypeptides are now existent among 62 species of Nicotiana. Comparing the single polypeptide of *N. tomentosiformis* with an isoelectric point different from the single polypeptide of *N. sylvestris*, the difference in isoelectric points resides in four changes in amino acids out of ca. 100 amino acids. On the presumption that each of the additional eleven polypeptides also underwent a change of four amino acids, $4 \times 12 = 48$ amino acid differences among thirteen polypeptides. Allowing a minimum of 75×10^6 years since the Nicotiana Fraction I proteins have existed, 75×10^6 years/48 differences = ca. 1.5 million years per amino acid sequence mutation. In comparison, a mutation in the sequence of cytochrome C, a protein also containing approximately the same number of amino acids, occurs about every twenty million years (DICKERSON 1972).

Table 1 is a compilation of recently acquired data for the Fraction I protein composition in species belonging to different genera of plants, providing an opportunity for comparison to Nicotiana. On the presumption that the greater the number of differences in the composition of the large subunit the greater the age of the genus, the genus Nicotiana with four differences is older than genera belonging to the Cruciferae with three, Chenopodiacea with two, etc. In the family Solanaceae, the genus Nicotiana is significantly older than the genus Lycopersicon, which has no differences. This estimation is thus in conformity with the previously advanced view on cytological, genetical and geographical grounds that the species constituting the genus Lycopersicon are of recent origin on the geological time scale (RICK *et al.* 1976).

Continuing the argument, genera capable of intergeneric hybridization in the family Cruciferae are older than Lycopersicon but younger than Nicotiana. Genera in the Chenopodiaceae are older than Lycopersicon but younger than Nicotiana and genera of the Cruciferae. Gossypium, with its two types, is a relatively old genus, whereas Oenothera is of more recent origin since no mutations have survived in its Fraction I protein large subunit.

Among the grain-producing plants belonging to the Monocotyledons, wheat and its relatives are older than barley or corn and its relatives, teosinte and Tripsacum, as judged by the differences in the polypeptide composition of the large subunit of their Fraction I proteins.

Age of One Species Relative to Another as Judged By the Number of Small Subunit Polypeptides

In the *N. sylvestris* x *N. tomentosiformis* → *N. glutinosa* x *N. tabacum* → *N. digluta* example previously considered, the presence of multiple polypeptides can be an indication of relative youth of the species compared to species containing single peptides in their Fraction I protein small subunits. Four polypeptides comprising the small subunit of *N. digluta* protein places the species as being younger than *N. tabacum* with its two polypeptides. In the same light *N. tabacum*, with two, is younger than either *N. sylvestris* or *N. tomentosiformis*, each having one polypeptide. On the basis of present knowledge, no distinction in age between species can be made if they contain only a single subunit polypeptide, even though the single polypeptides may differ in composition as shown by differences in the isoelectric points.

Corn and teosinte contain two small subunit polypeptides (Table 1) which place their origin as having occurred after Tripsacum with its single small subunit polypeptide.

In unpublished work, two species of Lycopersicon, as well as *Solanum pennelli*, were found to contain only a single small subunit polypeptide. Two other species contain small subunits with two polypeptides, which indicates that these species have evolved more recently than species with one polypeptide, but earlier than five other species, including the domestic tomato *L. esculentum*, all of which contain small subunits with three polypeptides. Nine different cultivars of the domestic tomato displayed identical compositions in regard to the isoelectric points of the large and small subunits of their Fraction I proteins. Thus domestication of this plant, followed by centuries of breeding activity by humans, has not produced alterations in the structural genes coding for the large or small subunits of Fraction I protein, a situation entirely reminiscent of the domestication of *N. tabacum*.

The presence of mainly single small subunit polypeptides in the Fraction I proteins of the grasses listed in Table 1 is probably the consequence of close pollination of these species. Mutations which occur in the nuclear DNA coding for the small subunit would seem to be eliminated soon after appearance by crossing over, segregation and further cross-hybridization with parental species.

Small Subunit Compositions as an Aid in Deciding On Designating a New Species of Plant

N. suaveolens is an example of a plant species whose Fraction I protein contains multiple small subunit polypeptides. A population of *N. suaveolens* plants presents considerable polymorphism in phenotypic characters between individuals (BURBIDGE 1960). For example, one individual may have a purple pigment in

Table 1. Number of polypeptides of differing isoelectric points comprising the large and small subunits of Fraction I proteins isolated from different species of plants.

FAMILY	GENERA	NUMBER OF SPECIES ANALYZED	NUMBER OF POLYPEPTIDES	
			LARGE	SMALL
DICOTS				
Solanaceae	Nicotiana	60	4	13
	Lycopersicon +Solanum	9	1	3
Cruciferae	Brassica Sinapis Rhaphanus	8	3	4
Chenopodiaceae	Beta Spinacia	2	2	2
Onagraceae	Oenothera	12	1	1
Malvaceae	Gossypium	3	2	2
MONOCOTS				
Gramineae	Zea	3	1	2
	Sorghum	7	1	1
	Hordeum	4	1	1
	Triticum +Aegilops	8	2	1

its anthers, another brown, and another green in plants which otherwise look alike. Similarly, we have found polymorphism in the electrofocusing pattern of Fraction I proteins isolated from individual plants in a population as shown by the diagram in Figure 6. All individuals have Fraction I proteins with identical large subunit polypeptides. However, some plants contain a Fraction I protein with a pattern of three small subunit polypeptides designated as A in Figure 6. Other plants have a three-polypeptide pattern designated as B, whereas still other plants have the four-polypeptide pattern shown as AB. Plants with these three different kinds of Fraction I protein are otherwise indistinguishable on morphological grounds, at least to the untutored eyes of biochemists. However, when plants of the A or B patterns self-fertilize, they breed true. The Fraction I protein small subunit polypeptide in their progeny will be either A or B. Plants containing the AB patterns can be generated in the laboratory by crossing plants of pattern A with plants of pattern B. Since there are no barriers to fertility, this F_1 hybrid plant containing four small subunit polypeptides is self-fertile. Although experiments are in progress, it is not yet known whether the small subunit polypeptide pattern in *N. suaveolens* AB breeds true or segregates into the A and B patterns after self-fertilization. The external phenotypes of these three kinds of plants are too similar to distinguish the plants from each other. However, the internal Fraction I protein patterns

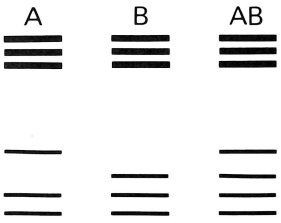


Figure 6. Example of how *N. suaveolens* plants could be reclassified into different species on the basis of differences in Fraction I protein composition.

could be used as stable, reproducible phenotypic markers to split *N. suaveolens* into two distinct subspecies. The analysis also makes clear that subspecies *N. suaveolens* AB is a derivative by hybridization between *N. suaveolens* A and *N. suaveolens* B.

Lycopersicon minutum, once considered to be a single species, has been split by RICK *et al.* (1976) into two species: *L. parviflorum* and *L. chiemelewski*. The argument for splitting was based on hybridization behavior and other grounds. Of interest here is the analysis of the Fraction I proteins contained in the two species. *L. parviflorum* contains a small subunit with a single polypeptide, while *L. chiemelewski* has two small subunit polypeptides. In addition to the fact that these differences in protein composition are consistent with the view of RICK *et al.* that they are distinct species, the existence of the two small subunit polypeptides in *L. chiemelewski* strongly indicates that it evolved later than *L. parviflorum*.

Origin of Genomes

Following the successful identification of *N. sylvestris* as having been the female donor of the chroloplast DNA genes coding for the large subunit polypeptides of *N. tabacum* Fraction I protein, a similar analysis was applied to the donor of the B genome of hexaploid wheat (CHEN *et al.* 1975). The outcome of the analysis is reproduced in Figure 7A. *Triticum dicoccum* was found to have chloroplast DNA genes corresponding to those in *T. aestivum* coding for the large subunit of Fraction I protein, and therefore *T. dicoccum* was the maternal parent in the cross. Furthermore, the analysis indicated that certain species such as *T. boeticum* and *T. urartu* could not have been direct participants as the B genome donor in the cross with *T. monococcum* which gave rise to *T. dicoccum*. However, the objection was raised that the single accessions used might not have been entirely representative of invariance in Fraction I protein composition of the species in question (JOHNSON 1976). Since then, Fraction I proteins from 38 additional accessions of *T. boeticum* and four of *Ae. squarrosa* have been analyzed with no deviation in polypeptide patterns on which the pedigrees in Figure 7A are based.

Among nine species of Brassica, to which the vegetables kale, cauliflower, Brussel sprouts and cabbage belong, one difference in the Fraction I protein large subunit polypeptide pattern is found. In the same nine species, three different small subunit polypeptides are found. Using these phenotypic markers, the evolutionary progression of these plants of commercial interest could be deduced with the results shown in Figure 7B.

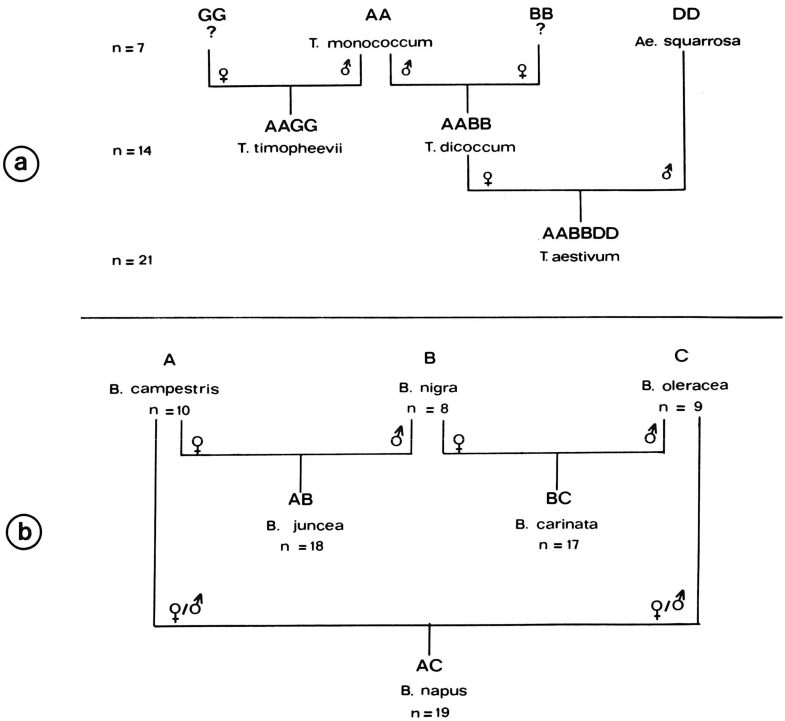


Figure 7. Diagrams indicating the sex of the plant giving rise to a particular genome during the origin of polyploid plants. a: plants belonging to wheat (Triticum) family; b: plants belonging to cabbage, kale, radish (Brassica) family.

The ancestor of modern day corn (*Zea mays*) is considered by BEADLE (1939) to have been teosinte (*Zea mexicana*), whereas MANGELSDORFF and REEVES (1939) have proposed Tripsacum as a probable parent. Analysis of the Fraction I proteins from the three species shows no difference in large subunit composition. However, both corn and teosinte have small subunits consisting of two polypeptides (CHEN *et al.* 1977), while an unpublished analysis has shown Tripsacum to have a single, small subunit polypeptide. It seems likely, therefore, that Tripsacum antedates both corn and teosinte and that the latter is more directly related to corn on the basis of having an identical Fraction I protein polypeptide composition.

ACKNOWLEDGMENTS

Research supported by Contract E (04-3) 34, Project 8, U. S. Energy Research and Development Administration; U. S. National Science Foundation Grant 75-07368 and a grant from Monsanto Commercial Porducts Corporation. The authors are specially in-

debted to Prof. H. J. Thompson of this department for advice and constructive criticism.

LITERATURE CITED

BAKER, T. S., D. EISENBERG, and F. A. EISERLING 1977 Ribulose diphosphate carboxylase is a two-layered, square-shaped molecule of symmetry 422. Science 196:293-295.

BEADLE, G. W. 1939 Teosinte and the origin of maize. J. Hered. 30:245-247

BURBIDGE, N. T. 1960 The Australian species of Nicotiana L. (Solanaceae). Aust. J. Bot. 8:342-380.

CARLSON, P. S., H. H. SMITH and R. D. DEARING 1972 Parasexual interspecific plant hybridization. Proc. Natl. Acad. Sci. USA 69:2292-2294.

CHAN, P. H., and S. G. WILDMAN 1972 Chloroplast DNA codes for the primary structure of the large subunit of Fraction I protein. Biochim. Biophys. Acta. 277:677-680.

CHAN, P. H., K. SAKANO, S. SINGH, and S. G. WILDMAN 1972 Crystalline Fraction I protein: preparation in large yield. Science 176:1145-1146.

CHEN, K., J. C. GRAY and S. G. WILDMAN 1975 Fraction I protein and the origin of polyploid wheats. Science 190:1304-1306.

CHEN, K., S. D. KUNG, J. C. GRAY, and S. G. WILDMAN 1976a Subunit polypeptide composition of Fraction I protein from various plant species. Plant Sci. Let. 7:429-434.

CHEN, K., S. JOHAL, and S. G. WILDMAN 1976b Role of chloroplast and nuclear DNA genes during evolution of Fraction I protein. In Genetics and Biogenesis of Chloroplasts and Mitochondria. (T. Bücher, W. Neupert, W. Sebald, S. Werner, Eds.) North Holland Publ. pp. 3-11.

CHEN, K., S. JOHAL, and S. G. WILDMAN 1977 Phenotypic markers for chloroplast DNA genes in higher plants and their use in biochemical genetics. In Nucleic Acid and Protein Synthesis in Plants. (J. W. Weil and L. Bogorad, Eds.) Plenum Publishing Co. pp. 183-194.

CLAUSEN, R. E. 1932 Interspecific hybridization in Nicotiana XIII. Further data as to the origin and constitution of Nicotiana tabacum. Svensk Bot. Tidskr. 26:123-136.

DICKERSON, R. E. 1972 The structure and history of an ancient protein. Sci. American 226:58-72.

GRAY, J. C., S. D. KUNG, S. G. WILDMAN and S. J. SHEEN 1974 Origin of Nicotiana tabacum L. detected by polypeptide composition of Fraction I protein. Nature 252:226-227.

GIBBONS, G. C., S. STRØBAEK, B. HASLETT, and D. BOULTER 1975 The N-terminal amino acid sequence of the small subunit of ribulose 1,5-diphosphate carboxylase from Nicotiana tabacum. Experientia 31:1040-1041.

GOODSPEED, T. H. 1954 The Genus Nicotiana. Chronica Botanica; Waltham, Massachusetts.

IWAI, S., Y. TANABE, AND N. KAWASHIMA 1976 Origin of sequence heterogenity of the small subunit of Fraction I protein from Nicotiana tabacum. Biochem. Biophys. Res. Commun. 73:993-996.

JOHNSON, B. L. 1976 Polyploid wheat and Fraction I protein. Science 192:1252.

KAWASHIMA, N., Y. TANABE, and S. IWAI 1976 Origin of Nicotiana tabacum detected by primary structure of Fraction I protein. Biochim. Biophys. Acta. 427:70-77.

KUNG, S. D., K. SAKANO, J. C. GRAY and S. G. WILDMAN 1975 The evolution of Fraction I protein during the origin of a new species of Nicotiana. J. Mol. Evol. 7:59-64.

KOSTOFF, D. 1938 Studies on polyploid plants. XVIII. Cytogenetic studies on Nicotiana sylvestris x N. tomentosiformis hybrids and amphidiploids and their bearings on the problem of the origin of N. tabacum. Doklady Akad. Nauk. SSSR 18: 459-462.

KWANYUEN, P., and S. G. WILDMAN 1975 Nuclear DNA codes for Nicotiana ferredoxin. Biochim. Biophys. Acta. 405:167-174.

MANGELSDORF, P. C., and R. G. REEVES 1939 The origin of Indian corn and its relatives. Tex. Agr. Exp. Sta. Bull. 574:1-315.

RAVEN, P. H. and D. L. AXELROD 1974 Angiosperm biogeography and past continental movements. Ann. Missouri Bot. Gard. 61:539-673.

RICK, C. M., E. KESICKI, J. F. FOBES, and M. HOLLE 1976 Genetic biosystematic studies on two new sibling species of Lycopersicon from Interandean Peru. Theo. and Appl. Genet. 47:55-68.

SAKANO, K., S. D. KUNG, and S. G. WILDMAN 1974 Identification of several chloroplast DNA genes which code for the large subunit of Nicotiana Fraction I proteins. Molec. Gen. Genet. 130:91-97.

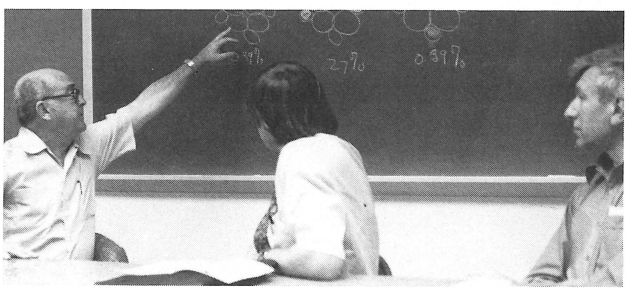
SMITH, H. H., K. N. KAO, and N. C. COMBATTI 1976 Interspecific hybridization by protoplast fusion in Nicotiana. J. Hered. 67:123-128.

SOLBRIG, O. T. 1966 Evolution and systematics. The MacMillian Company.

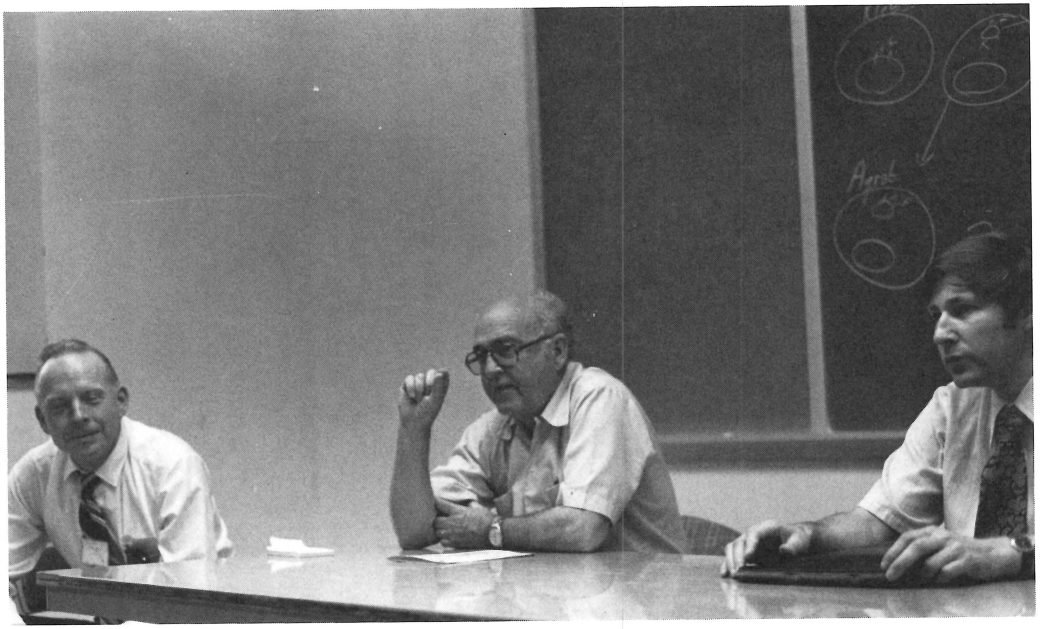
STEER, M. W. 1975 Evolution in the genus Avena: Inheritance of different forms of ribulose diphosphate carboxylase. Can. J. Genet. Cytol. 17:337-344.

STRØBAEK, S., G. C. GIBBONS, B. HASLETT, D. BOULTER and S. G. WILDMAN 1976 On the nature of the polymorphism of the small subunit of ribulose-1,5-diphosphate carboxylase in the amphidiploid Nicotiana tabacum. Carlsberg Res. Commun. 41:335-343.

WILDMAN, S. G., C. JOPE, and B. A. ATCHISON 1974 Role of mitochondria in the origin of chloroplast starch grains. Plant Physiol. 54:231-237.



From left to right: Drs. Wildman, Brill, Fincham



From left to right: Drs. Coe, Wildman and Brill