Genes in nonexperimental herds have been identified by several procedures. Some procedures, like the search for ancestors common to both parents in the case of recessive inheritance, afford presumptive evidence. Others, like segregation analysis, are more rigorous and require fewer restrictive assumptions. The usual analytical approach is to examine sequentially familial data for conformance with various hypothesis, from simplest to more complex when the simpler is untenable. Often the simplest later proves to have been tenable only because the data were scanty, particularly when a trait is rare or difficult or costly to measure or observe.

Modes of inheritance and other genetic hypothesis

The simplest modes involve only two alleles at a single locus, most likely on the 29 pairs of autosomes but some on sex chromosomes and their patterns of transmission and of recognition altered accordingly. Multiple alleles are not often hypothesized yet at least two thirds of the blood group loci, the most thoroughly studied in cattle, have more than two alleles.

The expression of anatomically or physiologically complex traits that may result from a single allelic difference may be regulated somewhat by other loci. Expression of traits also may be modified by environmental causes (including internal ones operating during development) or by an interplay of both genetic and environmental factors. Such traits vary in expressivity and may fail to appear (have incomplete penetrance) when genetic hypothesis indicates they should. This
statistical description of a trait’s failure to behave as regularly expected under a simple mode of inheritance represents a second order of complexity of hypothesis. Hypotheses involving variable expressivity or incomplete penetrance emphasize the importance of a single locus in regulating a trait despite unexplained irregularities in its behavior. It is not clear when a third order of complexity of hypothesis, the interactions of two or more loci, should be introduced to account for the irregularities. However, data usually are not available to provide convincing evidence of major effects controlled by two or more loci.

A principle trait often is accompanied regularly or irregularly by one or several other traits. All may originate from the same genic alteration during development or one may simply be a consequence of alteration produced in development by another. The most easily perceived, visible, or regular one may be chosen as the major effect of the allele while the others are said to be pleiotropic effects. That often is an analytically convenient way to handle a constellation of disorders, or syndrome, produced by a single locus. However, each element in the constellation also might be viewed as resulting from effects of different loci acting simultaneously (3). The rarity of several unusual independent or «loosely linked» traits appearing simultaneously in several individuals usually dispels such hypothesis (3, 11). Comparative embryology often provides unifying insights.

Mendelian hypothesis and their phenotypic systems

Each common Mendelian hypothesis, or mode of inheritance, requires a specific number of phenotypes. No matter how little or how much variability is observed in one trait or how many traits are involved in a syndrome, they must be classified into the required number of phenotypes before analysis can proceed. COTTERMAN (4) classified one locus genotypic systems by the number of phenotypes that could be produced; some two-locus systems are given in SNYDER (16).

Common two-phenotype systems (e.g., normal versus affected; normal versus mutant) in order of genetic simplicity include 1) those produced by two alleles at one locus: the dominant mode and its converse, the recessive mode, 2) those produced by three or more alleles at one locus, and 3) those produced by two alleles at each of two loci: duplicate dominant, duplicate recessive, and dominant and recessive epistasis. Three-phenotype systems (e.g., normal, weak, dead; red, roan, white) include: 1) the incompletely dominant mode produced by two alleles at one locus, 2) those produced by three or more alleles at one locus, and 3) those produced by two alleles at each of two loci: recessive epistasis, incompletely duplicate epistasis and dominant epistasis. COTTERMAN (4) lists 9 two-phenotype systems and 21 three phenotype systems for a single locus with three alleles. An alternative way of treating two-phenotype systems not conforming to the expectations of a two-allele: one-locus system is to measure irregularity in terms of expressivity and penetrance.

Segregation analysis

The most rigorous proof of inheritance is furnished by segregation analysis, an operation based on the phenomenon that progenies of certain parents specified
genotypically or phenotypically segregate into two or more phenotypic groups with predictable and mathematical regularity (15). For example, the offspring of two normal parents heterozygous for a deleterious recessive trait will segregate into two phenotypic groups, averaging 3/4 normal, 1/4 with the recessive trait. The segregation frequency (of the mutant form) is \( p = 1/4 \). Another example is the progeny by a normal heterozygote mated to a recessive homozygote. Such progeny also segregate into two phenotypic groups 1/2 normal; 1/2 recessive. Segregation frequency is 1/2. These two examples are full sib families. Other families also segregate.

For example, a 3/4 sib family produced by mating a normal heterozygous male to daughters of another heterozygous male, whose mothers were unrelated normal homozygous females, should contain 7/8 normal sibs and 1/8 recessive sibs; segregation frequency is 1/8, on the average.

Segregation analysis compares statistically a segregation frequency estimate derived from certain families chosen because they should display a segregation frequency characteristic of a specified mode of inheritance. If that estimate is consistent with the theory, it is evident that the trait being investigated is inherited in that specified mode. A concrete example will make the operation clear but before it can be given, certain complications caused by dominance or by epistasis must be considered.

It is easy to write segregation frequencies expected in families whose parents are of known genotypes or where each genotype has a unique phenotype. Unfortunately dominance and epistasis cause individuals of different genotypes to look alike. And genotypes often are unknown. In the case of a simple recessive lethal trait (a one-locus: two allele: two phenotype system), the dominant homozygote and heterozygote, by definition, have the same normal phenotype. Segregation analysis might be based on a segregation frequency, expected to be 1/4, in full-sib families produced by two heterozygous parents, if the parents could be identified some way as heterozygotes. That happens when normal parents produce one or more recessive sibs, the first of which is usually the proband, because it usually is the sib that identifies the segregating family first. Clearly, if recessive inheritance is involved, both parents must be heterozygous. But estimates of segregation frequency derived from such families (those containing at least one recessive sib) are larger than the expected 1/4 because other families that contain no recessive sibs, which also normally are produced by two heterozygous parents, have been excluded. Statistically, the restriction of the analysis to families that contain one or more affected sibs alters the distribution of the number of affected sibs in these families from a full binomial distribution to a truncated one with corresponding changes in mean and variance. Two alternate procedures have been used to accommodate truncate selection of families.

One, like the post-proband procedure of Lush and Hazel (14), attempts to restore the original frequency observed in the full binomial distribution. The post-proband procedure accomplishes this by restricting analysis to the younger sibs of probands which, as they are independent, form a full binomial distribution with mean of 1/4 as expected in mating two known heterozygotes.

The usual technique is to retain the truncated binomial distribution and obtain from it an estimate of segregation frequency by allowing for truncation algebraically.
The proband procedure also may be viewed broadly as a simple way to divide familial data into two or more groups in a way that one group provides an estimate or segregation frequency characteristic of a particular mode of inheritance.

**Herd data and biases**

Data usually include information on the afflicted or mutant and normal animals, including such relatives as full or half sibs, parents, aunts, nieces, and cousins of various degrees (Figures 1 and 2). Naturally bred herds often consist of several large half-sib families composed of smaller 3/4, 5/8, full-sib, and other families because few breeding males are kept. Artificial breeding often reduces family size.

Investigators often rely on herd owners' records or memories of animal phenotypes and genealogies; they differ in acuity and memory. Often only families or herds with several afflicted or mutant animals are reported. Morton (15) has reviewed other concerns in human data; parallels exist in cattle data.

**Segregation analysis in certain families**

*Full-sib families*, unlike other sib families of lesser relationship, involve no restrictive analytic assumptions. Unfortunately, cattle full-sib families are rare and small—both features reducing their value. Li (13) outlined the procedures that we (11) used to analyze the data of Cole and Moore (3) on hydrocephalus. Figure 1 shows 6 full-sib families containing at least one or more hydrocephalic calves, 4 with 2 sibs and 2 with 3. We derived an estimate of 0.39 for the segregation frequency there. Its standard error is ±0.19. The estimate exceeds the segregation
frequency of 0.25 expected under recessive inheritance by +0.14, which is 14/19 of a standard error too large, an insignificant deviation statistically.

In limited data, standard errors are very large. In the example, using ±2 standard errors, we would not have rejected any other theoretical segregation frequency from 0.11 to 0.67. We don't know yet what other genetic or nongenetic etiologic agents might have produced such fractions of mutants because we have not solved the consequences of other hypotheses. Because the segregation frequency estimate is a ratio of mutants of family size, both integers, it may be biassed when the integers are small. For example, from a full-sib family of two individuals (containing at least one recessive) only two estimates of segregation frequency can be obtained: 0.0 and 1.0, corresponding to 1 and 2 segregants. In families of three, 0.0, 0.6, and 1.0. Dr. Kemp and Chase (1) have estimated these biases for families of various sizes. With single truncation they are very important in families of less than 10, particularly for segregation frequencies less than 0.20.

Again following Li (13) we (11) also derived another estimate based on the assumption that the probability of ascertaining families was proportional to the number of recessive sibs contained (single selection). That estimate was 0.25±0.15, precisely the value expected under recessive inheritance.

**Figure 2.—Epitheliogenesis imperfecta in a Holstein-Friesian herd. [Courtesy of the authors (12).]**
Estimates less than the suspected 1/4 also might have aroused suspicions of incomplete penetrance.

*Three-quarter sib families* are produced in four ways (2). The commonest is produced by one male from otherwise unrelated daughters of another male. Only recessive inheritance in the commonest family type has been studied. There a segregation frequency of 1/8 is customarily cited for families that descend from ancestors with the dominant (and usually normal) phenotype (6, 8). That is true only under restrictive assumptions. Other 3/4 sib families descend from one or more recessive grandparents or parents whose phenotype immediately identifies their genotype and requires their progeny to be at least heterozygous for the recessive gene. Segregation frequencies will be larger then \( p = 1/4 \) or 1/2.

Two other complex situations also may be encountered. Often mutant recessives reproduce (e.g., brachygathia) but are not identified. So some grandparental or even parental phenotypes may be unknown. Another situation occurs when both full-sib and 3/4 sib families are recorded simultaneously in the data so one serves as *proband* for the other because parents are related in some remote irregular degree.

For our 3/4 sib family descended from normal ancestors to contain one or more recessive individuals, the sire must be heterozygous as must one or more of the dams. Each dam might have the recessive allele from her sire, the usual assumption, or from her dam.

We assumed equilibrium in the maternal grandparental generation and various frequencies of the recessive gene. Two kinds of normal sub-families of mates can be drawn from that grandparental generation. One includes a heterozygous male and his mates, a random sample of the normal dominant genotypes. The other includes a homozygous normal male and a random sample of normal females including at least one or more heterozygotes. In the first sub-family, the segregation frequency is from 1/8 to 1/6; in the second, 0 to 1/8, but hovering around 0.02 for the usual range of frequencies expected from populations in which recessive individuals do not reproduce. Unfortunately about 60% of families containing one or more lethal sibs are of the second type; and about 40%, of the first type. That almost precludes any test of hypothesis except a one-sided test excluding values > 1/6. The notable exception is the 3/4 sib family produced by sire-daughter mating, which is of the first type.

Fortunately virtually all families containing two or more recessives are of the first type. That means that 3/4 sib families with two or more recessive sibs provide an approximate test of recessive inheritance when the recessive allele is at usual low frequencies.

In figure 2, from the family of 10 (VI: 2, 4, 5-7, 9, 10, 12-14) we obtained an estimate of 0.24 ± 0.29, which exceeds the expected segregation frequency of 0.125 by less than 1/2 standard error. This approach seems particularly useful because it clearly pictures the amount of data at hand, the tenability of the hypothesis, and ideas about other hypotheses.

Half-sib families’ segregation ratios (7) depend completely on gene frequency. There is no procedure for using half-sibs to identify unestablished loci. Perhaps Lauvergne and Lefort’s rationale (9) may be modified to this end.
**Other Analyses**

*Penetrance.*—Ericksson (5) has derived algebraic equations for matings involving incomplete penetrance. Lauvergne and LeFort (9) have proposed a half-sib method for recessives in bull testing schemes. Full-sib and 3/4 sib analyses may be extended to include incomplete penetrance.

*Multiple alleles and two or more Loci.*—We know of no published methods. Huston (7) has examined full sib segregation ratios in three two-allele: two phenotype: two locus systems. All are gene frequency dependent and sometimes in the range for simpler modes.

*Other families.*—Rigorous analysis of complex genealogic data has been attempted for dominant inheritance (10). As yet there seems to be none for recessives, though comparing pedigrees of mutant calves with those of a random sample of contemporaries may be informative.

*Heritability.*—Methods of quantitative genetics can estimate heritability in both the broad and narrow senses and where appropriate, genetic correlations.

**SUMMARY**

Genes with visible effects are identified by showing that traits they regulate conform to intragenerational familial frequencies or intergenerational patterns produced by segregation characteristic of certain modes of inheritance.

Modes of inheritance can be grouped according to number of phenotypes, genotypes, alleles, and loci involved. Analysis proceeds from simplest to more complex modes. Segregation analysis in full-sib families offers the most rigorous proof but its use has been limited because full-sib families are small and those containing mutant animals, infrequent. A recently proposed 3/4 sib method may be more useful for certain traits. For most rare disorders, familial data often are adequate only for tentative acceptance of a genetic hypothesis. Irregularities in phenotypic expression for certain modes also may be described statistically as variable expressivity or incomplete penetrance.

**RESUMEN**

Genes con efectos visibles se identifican demostrando que las características que estos regulan están de acuerdo con las frecuencias dentro de la misma generación familiar o están de acuerdo con los patrones en diferentes generaciones producidos por la dispersión característica de ciertas clases de herencia.

Las clases de herencia pueden ser agrupadas de acuerdo con el número de fenotipos, genotipos, formas variantes de un mismo gene y su locus respectivo en el cromosoma. El análisis va de formas simples a formas más complejas. El análisis de dispersion en familias de dobles hermanos da la prueba más rigurosa, pero su uso es limitado debido a que las familias dobles de hermanos son pequeñas y aquéllas que contienen animales mutantes no son frecuentes. Se ha propuesto un método de familias de 3/4 de hermanos que puede ser más útil para el estudio de ciertas características. Para las anomalías menos frecuentes la información sobre la familia es a menudo adecuada sólo para la aceptación tenta-
tiva de una hipótesis genética. Las irregularidades en la manifestación fenotípica de ciertas formas también se puede describir estadísticamente como manifestación variable o como penetración incompleta.

ZUSAMMENFASSUNG

Gene mit sichtbarem Effekt auf Merkmale zeigen sich innerhalb einer Generation in typischen Frequenzen und segregieren zwischen Generationen charakteristisch für den Vererbungsmodus.


REFERENCES

1. CHASE, R. G. (1973): *Inference Techniques For Truncated Binomial Distributions of the Lower Type*. Kansas State University Library, Ph.D.