Using the Biochemical Pathway Model to Teach the Concepts of Gene Interaction and Epistasis (*American Biology Teacher*, in press)

Joseph P. Chinnici
Department of Biology, Box 842012
Virginia Commonwealth University
Richmond, VA 23284-2012
Phone: 804-828-1562
e-mail: jpchinni@vcu.edu
FAX: 804-828-0503

Typical dihybrid Mendelian phenotypic ratios are 9:3:3:1 for *AaBb* x Aabb crosses. If the genes involved interact in their effects on the phenotype, altered phenotypic ratios such as 9:3:4, 9:7, 12:3:1, and others may result as a consequence of gene interactions called epistasis. The colors of onions found in grocery stores is a relevant example. When red onions (*AABB*) and crossed with white onions (*aabb*), the dihybrid offspring are red (*AaBb*), and the offspring of a mating between dihybrids (*AaBb* x *AaBb*) are 9 red: 3 yellow: 4 white onions.

Students learning genetics often have difficulty grasping the concept of gene interaction well enough for them to be comfortable predicting the outcomes of crosses or interpreting the genetic bases of such outcomes. A method I use successfully to teach gene interaction is to give examples where the genes involved affect various steps in biochemical pathways. A biochemical (metabolic) pathway is a sequential series of chemical reactions by which substrates are altered step-by-step to produce the desired end product. Once students grasp the notion that genes can interact because they control different steps within biochemical pathways, the reason why phenotypic ratios appear to be non-Mendelian becomes more obvious and comprehensible.

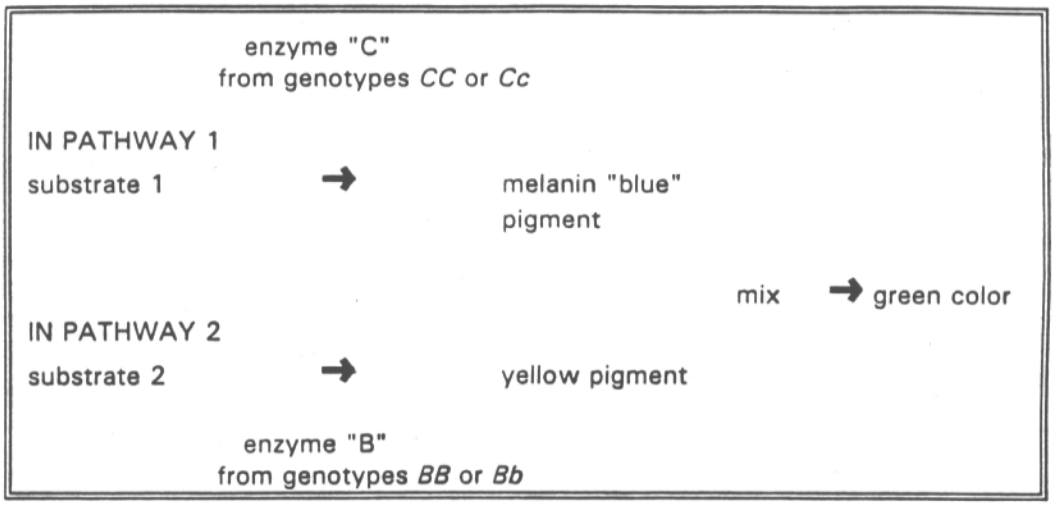
"Simple" Gene Interaction: Different Genes Affect Different Biochemical Pathways.

The simplest type of gene interaction involves two independently assorting pairs of genes controlling a single phenotypic characteristic with variations occurring in typical 9:3:3:1 Mendelian ratios among the offspring of dihybrids (*AaBb* x *AaBb*). A good example to use here is one found in many textbooks, plumage color in parakeets (budgies).

Matings between pure breeding "sky blue" birds (*bbCC*) and pure breeding "black-eyed yellow" ones (*BBcc*) produce only green (*BbCc*) offspring. When these green F_1 birds are interbred, the F_2 phenotype consists of 9/16 green (*B-C-*, the generalized genotype for *BBCC*, *BbCC*, *BbCC*, and *BbCc*): 3/16 blue (*bbC-*): 3/16 yellow (*B-cc*): 1/16 white-feathered (*bbcc*) birds. Because of the 9:3:3:1 ratio of the F_2 generation, we know this is a dihybrid situation. An explanation of these results, based on the biochemical pathway model, makes use of the principle that mixing blue and yellow colors produces green, and lack of either color is seen as white. Each gene pair affects the synthesis of a different pigment in a separate biochemical pathway.

In one pathway, gene *C* produces enzyme "C" which converts substrate 1 into melanin pigment responsible for black markings on the wing and head feathers and the "sky blue" feather color. In animals, blue color results from an optical illusion: the light brown color from scanty amounts of melanin are refracted and dispersed by light rays to give the illusion of blue color. The blue color of insects is a similar phenomenon, as is blue eye color in humans. Gene *C*'s recessive allele *c* produces a non-functional enzyme, but in this case the amount of enzyme in the heterozygote *Cc* is sufficient to fully catalyze the reaction. The homozygous genotype *cc*, however, results in no active enzyme "C" and a consequent metabolic block: substrate 1 is not

converted to melanin. Similarly, gene *B* in another pigment pathway is responsible for the synthesis of enzyme "B" which converts substrate 2 to yellow pigment. Homozygosity for the *b* allele (which produces no enzyme "B") results in a metabolic block in that reaction. So:



In the F_2 generation, a parakeet with genotype *C-bb* (3/16 of the total) makes only blue-appearing melanin pigment and black feather markings, since there is no active enzyme "B", while a *ccB-* bird (3/16 of the total) produces only yellow pigment (no activity for enzyme "C"). However, a *C-B-* bird 9/16 of the total) makes both pigments and has green plumage and black markings, while a *ccbb* individual (1/16 of the total), lacking activity for enzymes "C" and "B", makes neither pigment and consequently has white feathers without black markings. Here, the 9:3:3:1 ratio in the F_2 generation is not altered because the interacting genes *C,c* and *B,b* affect steps in different biochemical pathways, the end products of which affect a single phenotypic characteristic, plumage color.

Another example of simple gene interaction is eye color in the fruit fly Drosophila melanogaster. A cross between a fly with scarlet (bright red) eyes (*st st BW BW*) and one with brown eyes (*ST ST bw bw*) produce F_1 flies with normal (reddish-brown) eyes (*ST st BW bw*) eyes. Matings between F_1 flies produce an F_2 generation of 9/16 normal: 3/16 scarlet: 3/16 brown: 1/16 white-eyed (*st st bw bw*) flies. This occurs because the *ST/st* alleles control a step in the production of brown (ommochrome) eye pigment while the *BW/bw* alleles control a step in the production of red (drosopterin) eye pigment. A doubly recessive fly produces no eye pigment and has white eyes as a consequence. Again, the 9:3:3:1 ratio in the F_2 generation is not altered because the interacting genes *ST,st* and *BW,bw* affect steps in different biochemical pathways, the end products of which affect a single phenotypic characteristic, eye color.

"Epistatic" Gene Interaction: Different Genes Affect Different Steps Within a Single Biochemical Pathway.

The basic concept illustrated by a biochemical pathway

When two different pairs of genes control different steps in the same biochemical pathway, they will interact with each other to produce modifications of typical Mendelian phenotypic ratios. This phenomenon is known as **epistasis**, defined as the masking of the phenotypic effects of allele substitution at one locus by the presence of specific alleles at a second locus. A frequent result of

epistasis is that typical Mendelian phenotypic ratios are altered. Most often, the interactions among the non-allelic genes is non-reciprocal. For example, consider a portion of the pathway for anthocyanin flower pigment synthesis in some species of cape primroses (Streptocarpus). Gene pairs *R,r* and *D,d* show independent assortment but control separate steps in a biochemical pathway producing salmon, rose, and magenta colored forms of anthocyanin pigment:

	enzyme R from genotypes *RR* or *Rr*		enzyme D from genotypes *DD* or *Dd*	
salmon color	→	rose color	→	magenta color

In this example, epistasis occurs since rr masks the difference in phenotypic effects of *D-* and *dd*. The genotype *rrD-* produces salmon color since the *rr* condition "blocks" the action of *D-* by not producing the rose substrate needed by enzyme D. This effect of *rr* masking *D-* and *dd* causes an alteration of typical Mendelian phenotypic ratios. In the dihybrid mating *RrDd* x *RrDd*, instead of getting a 9:3:3:1 ratio of phenotypes among the offspring, a 9:3:4 ratio occurs:

P magenta flowers (*RRDD*) x salmon flowers (*rrdd*)

↓

F_1 all magenta flowers (*RrDd*)

↓

F_2

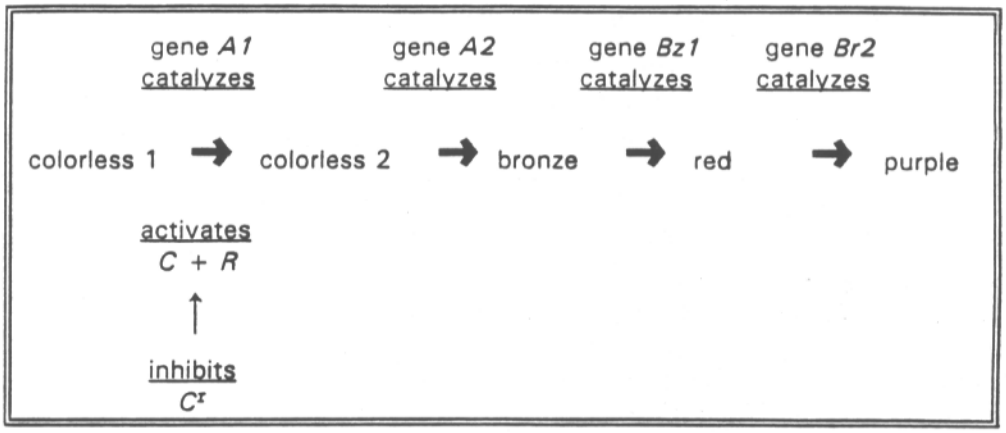
GENERALIZED RATIOS	GENOTYPES	PHENO-TYPES	OBSERVED PHENOTYPE RATIOS
9/16	*R-D-*	magenta	9/16
3/16	*R-dd*	rose	3/16
3/16	*rrD-*	salmon	
1/16	*rrdd*	salmon	4/16

The expected ratio of generalized genotypes (9:3:3:1) occurs but the phenotype ratio becomes 9:3:4 due to epistasis. This example illustrates what happens when two different non-allelic genes both affect the same characteristic: the expression of one hides or masks the expression of the other. The *rr* genotype prevents us from observing the phenotypic effect of the *D-* genotype. Here, the *rr* genotype is **epistatic** ("standing above"), masking the effect of alleles at the *D* locus which are **hypostatic** ("standing below" or being masked).

Using aleurone color in maize as an example of many types of epistasis

Corn (maize, Zea mays) has been more intensely studied than any other species of angiosperm, and much is known about the genetic interactions involving kernel (seed) coloration. A colorless layer of cells called the pericarp surrounds the outside of each kernel. The aleurone layer of cells lies just beneath the pericarp and surrounds the endosperm cell mass. Cells of the aleurone layer may accumulate bronze (purplish-brown), reddish, or purplish anthocyanin pigments, or remain colorless. Endosperm cells are either yellow (genotypes *YY* or *Yy*) or colorless (white, genotype *yy*). If the aleurone is bronze, reddish, or purplish, this becomes the kernel color since the endosperm cannot be seen through the colored aleurone. If the aleurone is colorless, the kernel will appear either yellow or white due to the endosperm color showing through the aleurone.

Various genes govern aleurone color. A simplified biochemical pathway for the actions of these genes is given below. More biochemical details about this pathway are given in the appendix.



The dominant structural genes *A1*, *A2*, *Bz1*, and *Br2* produce enzymes that perform biochemical steps in the anthocyanin pathway in maize kernel aleurone cells. The functioning of the structural gene *A1* (and possibly others) is transcriptionally regulated (i.e., turned on) in the aleurone tissue by the dominant gene *C* in combination with the dominant gene *R*. Gene C^I, an allele of *C*, has a dominant inhibitory effect on pigment biosynthesis by not allowing the gene products of *C* to polymerize into their active form in C/C^I heterozygotes.

Using this biochemical pathway for aleurone pigmentation, students can determine the basis for several types of epistasis involving interactions of different dihybrid gene combinations. These interactions produce phenotypic ratios of 9:3:4, 12:3:1, 13:3, and 9:7 in the F_2 generations from dihybrid crosses.

9:3:4 Epistasis. Consider the following dihybrid mating (genes C/C^I and *Yy* differ):

A1 A1 A2 A2 Bz1 Bz1 Br2 Br2 $C^I\ C^I$ *R R y y* (white kernel) x
A1 A1 A2 A2 Bz1 Bz1 Bz2 Bz2 C C R R Y Y (purple kernel)

↓

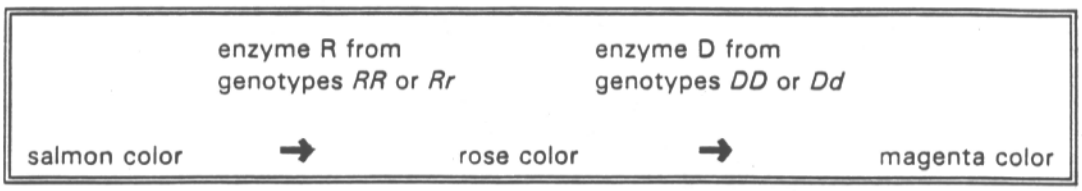
C C^I *Y y* (white kernel)

↓

9/16	C^I - *Y* -	yellow kernels 9/16
3/16	C^I - *y y*	white kernels 3/16
3/16	*C C Y* -	purple kernels
1/16	*C C y y*	purple kernels 4/16

Here, the homozygous genotype *CC* masks the effect of genetic variation at the *Y* locus since both *CCY*- and *CCyy* kernels are purple. Thus, *CC* is epistatic to alleles at the *Y* locus, which are hypostatic. Gene C^I inhibits the ability of the C protein to polymerize, not permitting colored pigment to be produced in the aleurone. This allows the phenotypic effect of *Y*- or *yy* in the endosperm to be observed as the kernel color phenotype.

epistasis is that typical Mendelian phenotypic ratios are altered. Most often, the interactions among the non-allelic genes is non-reciprocal. For example, consider a portion of the pathway for anthocyanin flower pigment synthesis in some species of cape primroses (Streptocarpus). Gene pairs *R,r* and *D,d* show independent assortment but control separate steps in a biochemical pathway producing salmon, rose, and magenta colored forms of anthocyanin pigment:



In this example, epistasis occurs since rr masks the difference in phenotypic effects of *D-* and *dd*. The genotype *rrD-* produces salmon color since the *rr* condition "blocks" the action of *D-* by not producing the rose substrate needed by enzyme D. This effect of *rr* masking *D-* and *dd* causes an alteration of typical Mendelian phenotypic ratios. In the dihybrid mating *RrDd* x *RrDd*, instead of getting a 9:3:3:1 ratio of phenotypes among the offspring, a 9:3:4 ratio occurs:

P magenta flowers (*RRDD*) x salmon flowers (*rrdd*)

↓

F_1 all magenta flowers (*RrDd*)

↓

F_2

GENERALIZED RATIOS	GENOTYPES	PHENOTYPES	OBSERVED PHENOTYPE RATIOS
9/16	*R-D-*	magenta	9/16
3/16	*R-dd*	rose	3/16
3/16	*rrD-*	salmon	
1/16	*rrdd*	salmon	4/16

The expected ratio of generalized genotypes (9:3:3:1) occurs but the phenotype ratio becomes 9:3:4 due to epistasis. This example illustrates what happens when two different non-allelic genes both affect the same characteristic: the expression of one hides or masks the expression of the other. The *rr* genotype prevents us from observing the phenotypic effect of the *D-* genotype. Here, the *rr* genotype is **epistatic** ("standing above"), masking the effect of alleles at the *D* locus which are **hypostatic** ("standing below" or being masked).

Using aleurone color in maize as an example of many types of epistasis

Corn (maize, Zea mays) has been more intensely studied than any other species of angiosperm, and much is known about the genetic interactions involving kernel (seed) coloration. A colorless layer of cells called the pericarp surrounds the outside of each kernel. The aleurone layer of cells lies just beneath the pericarp and surrounds the endosperm cell mass. Cells of the aleurone layer may accumulate bronze (purplish-brown), reddish, or purplish anthocyanin pigments, or remain colorless. Endosperm cells are either yellow (genotypes *YY* or *Yy*) or colorless (white, genotype *yy*). If the aleurone is bronze, reddish, or purplish, this becomes the kernel color since the endosperm cannot be seen through the colored aleurone. If the aleurone is colorless, the kernel will appear either yellow or white due to the endosperm color showing through the aleurone.

Other examples of 9:3:4 epistasis include: rat coat color (9 black: 3 cream: 4 albino), onion bulb color (9 red: 3 yellow: 4 white), gladiolus flower color (9 red: 3 cream: 4 white)

12:3:1 Epistasis. Consider the following dihybrid mating (genes *Rr* and *Yy* differ):

A1 A1 A2 A2 Bz1 Bz1 Bz2 Bz2 C C R R y y (purple kernel) x
A1 A1 A2 A2 Bz1 Bz1 Bz2 Bz2 C C r r Y Y (white kernel)

↓

R r Y y (purple kernel)

↓

9/16	*R - Y -*	purple kernels
3/16	*R - y y*	purple kernels 12/16
3/16	*r r Y -*	yellow kernels 3/16
1/16	*r r y y*	white kernels 1/16

Kernels that are *rr* produce colorless aleurone, allowing the endosperm color to be visible as the kernel phenotype. However, *R-* is epistatic to the phenotypic effect of genetic variation at the *Y* locus, since both *R-Y-* and *R-yy* kernels are purple.

Other examples of 12:3:1 epistasis include: summer squash fruit color (12 white: 3 yellow: 1 green), oat hull color (12 black: 3 gray: 1 white), pineapple leaf shape (12 piping: 3 spiny tip: 1 spiny), dog hair color (12 white: 3 black: 1 brown)

13:3 Epistasis. Consider the following dihybrid mating (genes C/C^I and *Rr* differ):

A1 A1 A2 A2 Bz1 Bz1 Bz2 Bz2 $C^I\ C^I\ r\ r$ *Y Y* (yellow kernel) x
A1 A1 A2 A2 Bz1 Bz1 Bz2 Bz2 C C R R Y Y (purple kernel)

↓

C C^I *R r* (yellow kernel)

↓

9/16	C^I *- R -*	yellow kernels
3/16	C^I *- r r*	yellow kernels
3/16	*C C R -*	purple kernels 3/16
1/16	*C C r r*	yellow kernels 13/16

In this case, C^I - is epistatic to allele differences at the *R* locus since both C^I*-R-* and C^I*-rr* kernels are yellow. Also, *rr* is epistatic to allele differences at the *C* locus since both C^I*-rr* and *CCrr* kernels are yellow. This occurs because a combination of gene products from the *C* and the *R* genes is necessary to activate the anthocyanin pathway. Thus, in this example only *CCR-* kernels can produce purple color.

Other examples of 13:3 epistasis include: chicken plumage color (13 white: 3 colored), fruit fly eye facet shape (13 normal: 3 star-shaped)

Bibliography: listed are various articles on gene interactions and epistasis in plants and animals, including humans.

Dooner H. K., Robbins, T. P., and Jorgensen, R. A. (1991). Genetic and developmental control of anthocyanin biosynthesis. *Annual Review of Genetics, 25*, 173-199.

Emmerling-Thompson, M., and Nawrocky, M. M. (1979). Genetic studies of flower color in *Tradescantia*. *Journal of Heredity, 70*, 115-122.

Frankel, J. S. (1985). Inheritance of trunk striping in the Sumatran tiger barb, *Barbus tetrazona*. *Journal of Heredity, 76*, 478-479.

Freeling, M., and Walbot, V. (Ed.). *The Maize Handbook* (1994). New York: Springer-Verlag. See, in particular, chapters 34 and 35.

Fulton, J. E., Juriloff, D. M., Cheng, K. M., and Nichols, C. R. (1983). Defective feathers in Japanese quail. A two-locus model for a new trait. *Journal of Heredity, 74*, 184-188.

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *The Plant Cell, 7*, 1071-1083.

Jaynes, R. A. (1981). Inheritance of ornamental traits in mountain laurel, *Kalmia latifolia*. *Journal of Heredity, 72*, 245-248.

King, R. A., Wirtschafter, J. D., Olds, D. A. and Brumbaugh, J. (1986). Minimal pigment: A new type of oculocutaneous albinism. *Clinical Genetics, 29*, 42-50.

King, R. A. and Olds, D. P. (1985). Hairbulb tyrosinase activity in oculocutaneous albinism: Suggestions for pathway control and block location. *American Journal of Human Genetics, 20*, 49-55.

Lawrence, W. J. C., and Sturgess, V. C. (1957). Studies on *Streptocarpus* III. Genetics and chemistry of flower colour in the garden forms, species and hybrids. *Heredity, 11*, 303-336.

Leiper, B. D., and Robinson, R. (1985). Gray mutant in the Mongolian gerbil. *Journal of Heredity, 76*, 473.

Long, C. A., and Hogan, A. (1988). Two independent loci for albinism in raccoons, *Procyon lotor*. *Journal of Heredity, 79*, 387-389.

Okonkwo, C. A., and Clayberg, C. D. (1984). Genetics of flower and pod color in *Phaseolus vulgaris*. *Journal of Heredity, 75*, 440-444.

Paris, H. S., and Nerson, H. (1986). Genes for intense fruit pigmentation of squash. *Journal of Heredity, 77*, 403-409.

Pawelek, J. M. and Korner, A. M. (1982). The biosynthesis of mammalian melanin. *American Scientist, 70*, 136-145.

Reese, P. F., and Boerma, H. R. (1989). Additional genes for green seed coat in soybeans. *Journal of Heredity, 80*, 86-88.

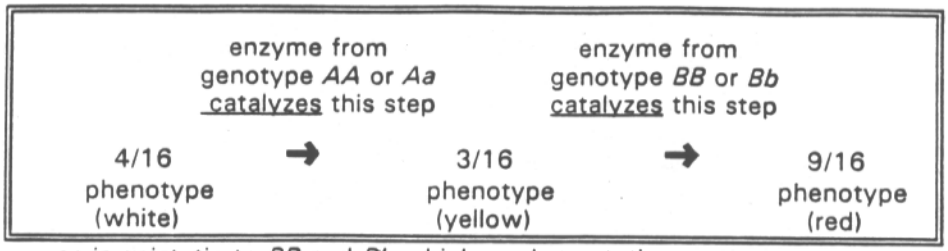
Rhodes, B. B. (1986). Genes affecting foliage color in watermelon. *Journal of Heredity, 77*, 134-135.

Witkop, C. J., Nance, W. E., Rawls R. F. and White, J. G. (1970). Autosomal recessive oculocutaneous albinism in man: Evidence for genetic heterogeneity. *American Journal of Human Genetics, 22*, 55-74.

Wolf, S. J., and Craig, R. (1988). Inheritance of flower and stem color in *Exacum affine* Balf. *Journal of Heredity, 79*, 303-306.

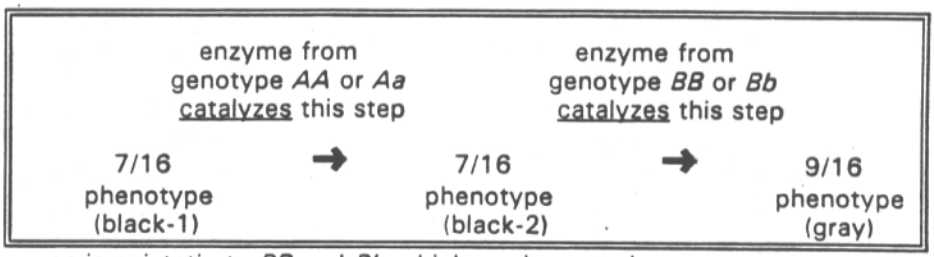
Table 1. Model biochemical pathways to explain the altered Mendelian ratios observed in the various types of epistasis

9:3:4 epistasis (onion bulb color)



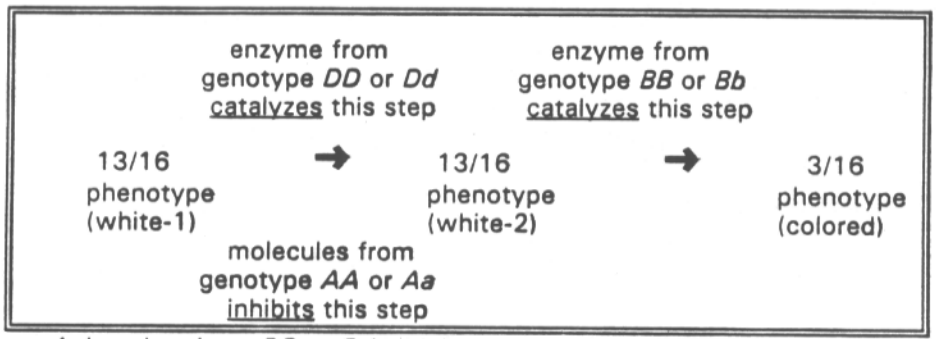
aa is epistatic to *BB* and *Bb* which are hypostatic

9:7 epistasis (fruit fly body color)



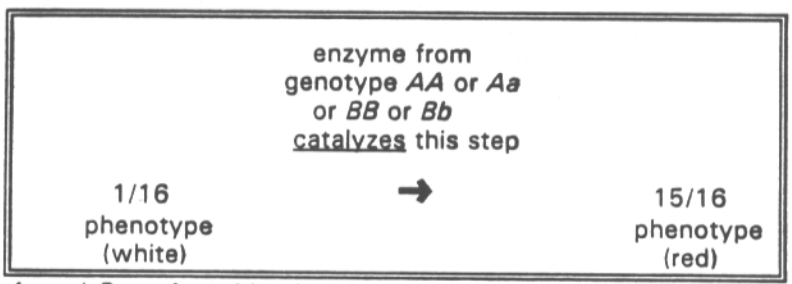
aa is epistatic to *BB* and *Bb* which are hypostatic

13:3 epistasis (chicken plumage color)



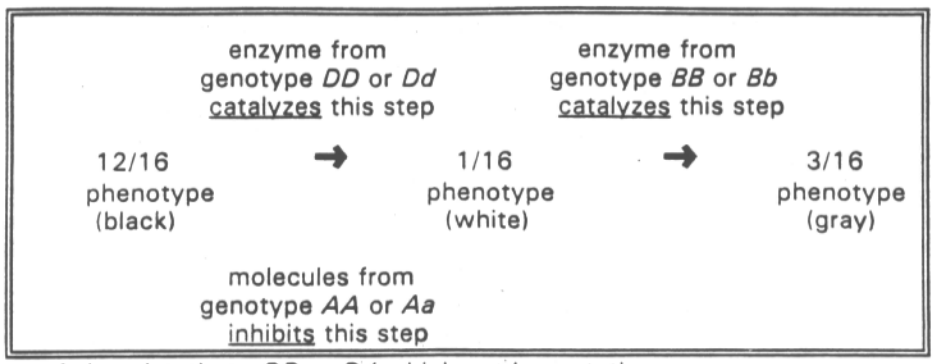
A- is epistatic to *DD* or *Dd* which are hypostatic
aadd is epistatic to *BB* or *Bb* which are hypostatic

15:1 epistasis (white clover flower color)



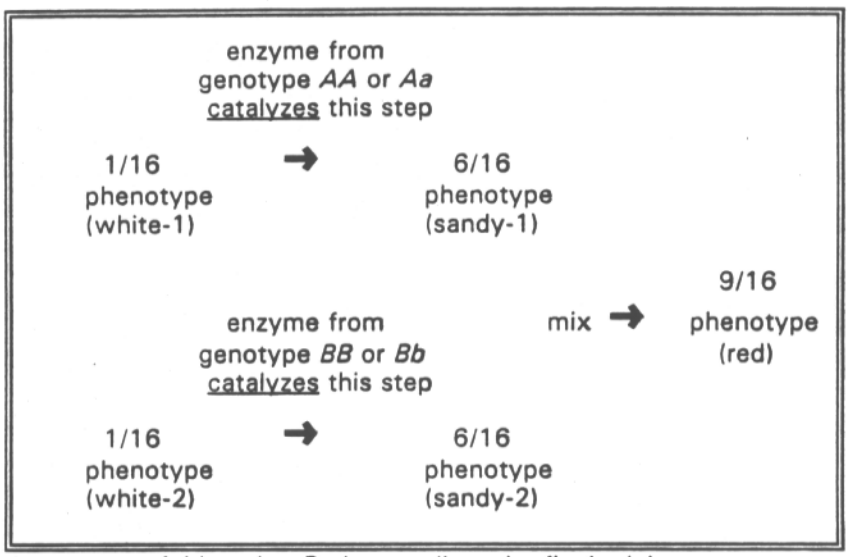
A- and *B-* perform identical functions; only the *aabb* genotype does not allow the reaction to occur

12:3:1 epistasis (oat hull color)



A- is epistatic to *DD* or *Dd* which are hypostatic
aadd is epistatic to *BB* or *Bb* which are hypostatic

9:6:1 epistasis (pig hair color)



genotypes *A-bb* and *aaB-* do not allow the final mixing to occur; thus, genes *A* and *B* complement each other's effect to produce the end product

Table 1a. Examples of various types of epistasis in plants and animals.

ORGANISM	CHARACTERISTICS	PROGENY OF DIHYBRID MATINGS (*AaBb* x *AaBb*)		
9:3:4 types		9 *A-B-*	3 *A-bb*	3 *aaB-* + 1 *aabb*
rats	coat color	9 black:	3 cream:	4 albino
onions	bulb color	9 red:	3 yellow:	4 white
gladiolus	flower color	9 red:	3 cream:	4 white
9:7 types		9 *A-B-*	3 *A-bb* + 3 *aaB-* + 1 *aabb*	
fruit flies	body color	9 gray:	7 black	
white clover	cyanide synthesis	9 high:	7 low cyanide	
daisies	flower center color	9 purple center:	7 yellow center	
minks	fur color	9 normal color:	7 platinum	
rabbits	fur length	9 normal length:	7 short hair	
mice	wavy hair	9 straight hair:	7 wavy hair	
chickens	behavior of hens	9 broodiness: (sit on eggs)	7 no broodiness towards eggs (ignore eggs)	
13:3 types		9 *A-B-* + 3 *A-bb* + 1 *aaB-*		3 *aaB-*
chickens	plumage color	13 white:		3 colored plumes
fruit flies	eye facet shape	13 normal:		3 star-shaped facets
12:3:1 types		9 *A-B-* + 3 *A-bb*	3 *aaB-*	1 *aabb*
summer squash	fruit color	12 white:	3 yellow:	1 green
oats	hull color	12 black:	3 gray:	1 white
pineapples	leaf shape	12 piping:	3 spiny tip:	1 spiny
dogs	hair color	12 white:	3 black:	1 brown
9:6:1 types		9 *A-B-*	3 *A-bb* + 3 *aaB-*	1 *aabb*
Sumatran tiger barb fish	trunk banding	9 complete:	6 incomplete:	1 half
summer squash	fruit shape	9 disk:	6 sphere:	1 long
wheat	kernel color	9 red:	6 brown:	1 white
pigs	hair color	9 red:	6 sandy:	1 white
15:1 types		9 *A-B-* + 3 *A-bb* + 3 *aaB-*		1 *aabb*
shephard's purse weed	fruit pod shape	15 heart shape:		1 ovoid
white clover	flower color	15 white:		1 red
chickens	shank feathers	15 with feathers		1 without

YOUR NAME:_____ CODE: _____

Observed Phenotype (F_2)	Observed Number (O)	Theoretical Proportion (TP)	Expected Number (E) ($E=TP \times O_{total}$)	deviation^2/E $(O-E)^2/E$
TOTALS:		1.00		X^2=

X^2 = _____, df = _____; P = _____

CONCLUSION:_____

1. Genotypes and Phenotypes of the parental (P) generation seeds which grew into plants that were crossed to produced the F_1 generation:

parent 1 Genotype _____ Phenotype_____ parent 2 Genotype _____ Phenotype_____

2. Genotypes and Phenotypes of the F_1 seeds which grew into plants that were crossed to produce the F_2 seeds/seedlings:

Genotype: _____ Phenotype: _____

3. Genotypes and theoretical proportions of all F_2 types

GENOTYPES	THEO. PROPORTIONS
_____	_____
_____	_____
_____	_____
_____	_____

YOUR NAME:_____ CODE: _____

Observed Phenotype (F_2)	Observed Number (O)	Theoretical Proportion (TP)	Expected Number (E) ($E=TP \times O_{total}$)	deviation^2/E $(O-E)^2/E$
TOTALS:		1.00		X^2=

X^2 = _____, df = _____; P = _____

CONCLUSION:_____

1. Genotypes and Phenotypes of the parental (P) generation seeds which grew into plants that were crossed to produced the F_1 generation:

parent 1 Genotype _____ Phenotype_____ parent 2 Genotype _____ Phenotype_____

2. Genotypes and Phenotypes of the F_1 seeds which grew into plants that were crossed to produce the F_2 seeds/seedlings:

Genotype: _____ Phenotype: _____

3. Genotypes and theoretical proportions of all F_2 types

GENOTYPES	THEO. PROPORTIONS
_____	_____
_____	_____
_____	_____
_____	_____

Degrees of Freedom. We toss a coin into the air. If it does not land heads-up, it must land tails-up. Although there are two sides to a coin, it has only one "choice" as to which side is up. Or, when we put our shoes on, if we put one shoe on the right foot first, then we must then put the other shoe on the left foot. Thus, given two possibilities, there is only one free choice or "degree of freedom." In general, we have n-1 degrees of freedom (df) in assigning numbers at random to n classes within an experiment. Thus, if there are four possible phenotypic combinations possible (say, in a 9:3:3:1 dihybrid ratio situation), there are three degrees of freedom in assigning an organism to one of these (if we choose not to place it in the first, second, or third group, we must place it in the fourth category). For most genetics situations, the number of degrees of freedom will be one less than the number of phenotypic classes.

Chi-Square Test (X^2). The chi-square test enables an experimenter to convert the amount of deviation from expected values into the probability of such differences occurring by chance. This test takes into account the size of the sample tested and the number of variables (degrees of freedom). The question we try to answer with the X^2 test is "How small can the deviations be to be attributed to chance alone?" The formula for X^2 is:

$$\text{chi-square} = \sum (O - E)^2 / E$$

where Σ = the grand total of the squared deviations [observed number minus expected number, $(O-E)^2$] divided by the expected number (E) for each class. The value of chi-square may then be converted into the probability (P) that the deviation is due to chance by using the table below for the proper number of degrees of freedom.

Chi-Square Distribution Table.

	Probability that deviation is due to chance alone	Numbers of Degrees of Freedom 1	2	3	4	5
these values	0.95 (95%)	0.004	0.10	0.35	0.71	1.15
support the	0.70 (70%)	0.15	0.71	1.42	2.20	3.00
hypothesis	0.50 (50%)	0.46	1.39	2.37	3.36	4.35
under	0.30 (30%)	1.07	2.41	3.66	4.88	6.06
consideration	0.10 (10%)	2.71	4.60	6.25	7.78	9.24
don't support	0.05 (5%) **	3.84	5.99	7.82	9.49	11.07
hypothesis	0.01 (1%) **	6.64	9.21	11.34	13.28	15.09
(P= or <.05)	0.001(0.1%) **	10.83	13.82	16.27	18.47	20.52

** Observed results are significantly different from the expected results.

The chi-square test has two important limitations. First, it must be used only for the numerical data itself, never on any percentages or ratios derived from the data. Second, it cannot be used for experiments where the expected number in any phenotypic class is less than 5.