



The nature of gene

Deependra Dhakal

Agriculture and Forestry University

ddhakal.rookie@gmail.com

<https://rookie.rbind.io>

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

Genetic analysis begins with mutants

- Genetic approach to understanding any biological property is through gene discovery
- One simple and useful analytical approaches to gene discovery is detection of single-gene inheritance pattern.
- Variants in biological system represent “ tweaks ” in the biological machine.
- Common form of any property of an organism is called **wild type**
- Heritable variants different from wild types are called **mutants**
- Alternative forms of the property are called **phenotypes**
- **Polymorphism:** coexistence of two or more reasonably common phenotypes of a biological property

Approaches to gene identification

- This type of approach to gene discovery is sometimes called **forward genetics**, a strategy to understanding biological function starting with random single-gene mutants and ending with their DNA sequence and biochemical function.
- **Reverse genetics** starts with genomic analysis at the DNA level to identify a set of genes as candidates for encoding the biological property of interest¹.

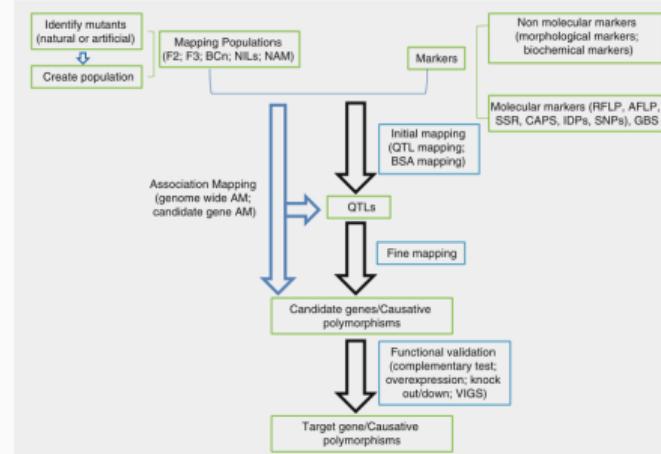


Figure 1: Flowchart of forward genetics approach in gene identification. For an entire chapter dedicated to the approach, refer to Ji (2013)

¹For a book chapter dedicated to Reverse genetics approaches, refer to [Gilchrist and Haughn \(2013\)](#).

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

Physical nature

- Are genes composed of protein, nucleic acid, or some other substance?
- In 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty offered the first compelling experimental evidence that genes are made of deoxyribonucleic acid (DNA).
- They showed that DNA extracted from a virulent strain of bacteria carried the necessary genetic information to transform a nonvirulent strain into a virulent one.

How can DNA molecules store information ?

- In the 1950s, there was something of a race among several groups of geneticists and chemists to answer this question. In 1953, James Watson and Francis Crick working at Cambridge University in England won that race.
- They determined that the molecular structure of DNA was in the form of a double helix – two strands of DNA wound side-by-side in a spiral. Their structure of the double helix is like a twisted ladder (Figure 2).
- The sides of the ladder are made of sugar and phosphate groups. The rungs of the ladder are made of four bases: adenine (A), thymine (T), guanine (G), and cytosine (C).

- The bases face the center, and each base is hydrogen bonded to the base facing it in the opposite strand.
- Nucleotide base pairing is highly specific. The bonding specificity is based on the complementary shapes and charges of the bases.
- The sequence of A, T, G, and C represents the coded information carried by the DNA molecule.

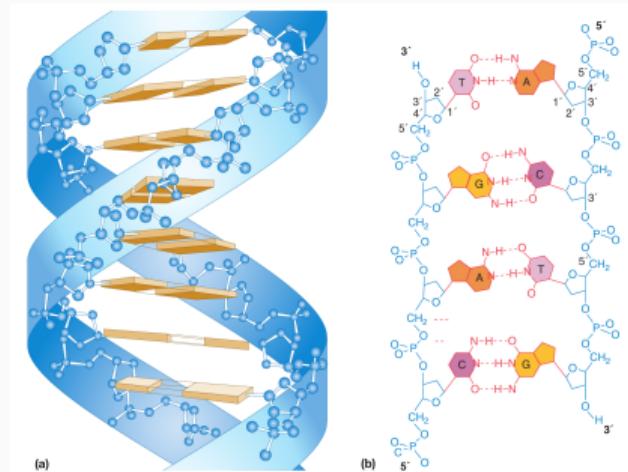


Figure 2: Twisted ladder structure of DNA

How are genes regulated?

- Cells need mechanisms to turn genes on or off in specific cell and tissue types and at specific times during development.
- In 1961, Francois Jacob and Jacques Monod made a conceptual breakthrough on this question.
- Working the genes necessary to metabolize the sugar lactose in the bacterium *Escherichia coli*, they demonstrated that genes have regulatory elements that regulate gene expression – that is, whether a gene is turned on or off.
- The regulatory elements are specific DNA sequences to which a regulatory protein binds and acts as either an activator or repressor of the expression of the gene.

How is the information stored in DNA decoded to synthesize protein ?

- While the discovery of the double-helical structure of DNA was a watershed for biology, many details were still unknown.
- Precisely how information was encoded into DNA and how it was decoded to form the enzymes that Tatum and Beadle had shown to be the workhorses of gene action remained unknown.
- Over the years 1961 through 1967, teams of molecular geneticists and chemists working in several countries answered these questions when they “cracked the genetic code.”

- What this means is that they deduced how a string of DNA nucleotides, each with one of four different bases (A, T, C, or G), encodes the set of 20 different amino acids that are the building blocks of proteins.
- They also discovered that there is a messenger molecule made of ribonucleic acid (RNA) that carries information in the DNA in the nucleus to the cytoplasm where proteins are synthesized.
- By 1967, the basic flowchart for information transmission – central dogma of biology – in cells was known.

Central dogma of biology

- The flow of genetic information in cells from DNA to mRNA to protein is described by the Central Dogma.
- The decoding of one molecule to another is performed by specific proteins and RNAs.
- Because the information stored in DNA is so central to cellular function, it makes sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected.

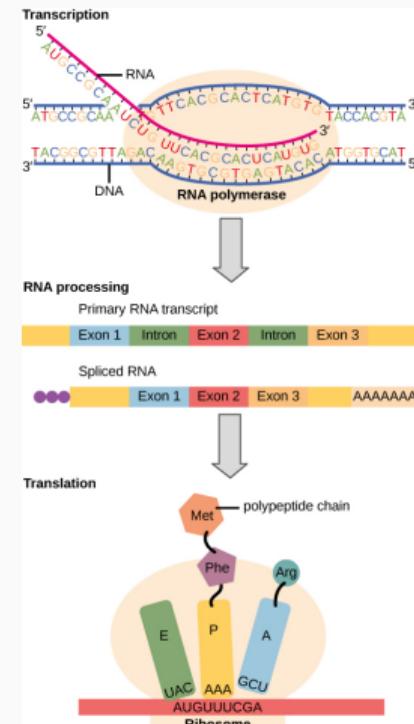


Figure 3: The central dogma of biology

Terminologies

Table 1: Some terms applicable in the context of genetic code

Term	Meaning
Code letter	A nucleotide e.g., A, U, G, C (in mRNA) or A, T, G, C (in DNA)
Code word	A sequence of nucleotides specifying an amino acid. e.g., the RNA codon for leucine = CUG (of GAC for DNA)
Anticodon	A sequence of nucleotides on tRNA that complements the codon, e.g., GAC = anticodon for leucine
Word size of codon	The number of letters in a code word, e.g., three letters in a triplet code. These are the same as coding ratio in a nonoverlapping code.
Non overlapping code	A code in which only as many amino acids are coded as there are code words in end-to-end sequence, e.g., for a triplet code UUUCCC = phenylalanine (UUU) + proline (CCC)
Overlapping code	When more amino acids are coded for in a single reading frame than there are code words present in end-to-end sequence, e.g., UUUCCC = phenylalanine (UUU) + phenylalanine (UUU) + serine (UCC) + proline (CCC)
Nondegenerate code	When there is only one codon for each amino acid, e.g., 20 different amino acids have a total of 20 codons.
Degenerate code	When there is more than one codon for a particular amino acid, e.g., UUU, UUC = phenylalanine, or 20 different amino acids have a total of more than 20 codons.

Table 2: Some terms applicable in the context of genetic code (...continued)

Term	Meaning
Synonymous codons	Different codons that specify the same amino acid in a degenerate code, e.g., UUU = UUC = phenylalanine
Ambiguous code	When one codon can code for more than one amino acid, e.g., GGA = glycine, glutamic acid
Commaless code	When there are no intermediary nucleotides (spacers) between words, e.g., UUUCCC = two amino acids in triplet nonoverlapping code
Reading frame	The particular nucleotide sequence, coding for a polypeptide, that starts at a specific point and is then partitioned into codon until the final code word of that sequence is reached.
Frameshift mutation	A change in the reading frame because of the insertion or deletion of nucleotides in numbers other than multiples of the codon length. This modifies the previous partitioning of codons in the reading frame, and causes a new sequence of codons to be read.
Sense word	A codon that specifies an amino acid normally present at that position in a protein
Missense mutation	A change in nucleotide sequence, either by deletion, insertion, or substitution, resulting in the appearance of a codon that produces a different amino acid in a particular protein, e.g., UUU (phenylalanine) mutates to UGU (Cysteine)
Nonsense mutation	A mutation that results in a codon that does not produce an amino acid, e.g., UAG (also called a "stop" mutation or "chain termination" codon)
Universality	Utilization of the same genetic code in all organisms, e.g. UUU = phenylalanine in bacteria, mouse, man and tobacco (with some exceptions in mitochondria)

Genetic code: The search

- Evidence first presented on its nature in 1961 by Francis Crick and Sydney Brenner used the chemical mutagen proflavin to insert (+) or delete (-) one, two, or three nucleotides into the gene of a virus.
 - ▶ When one or two nucleotides were inserted/deleted, protein synthesis was completely abolished. When three nucleotides were inserted/deleted, the protein was synthesized and functional.
 - ▶ The insertion/deletion of one or two nucleotides completely changed the triplet reading frame, thereby altering the message for every subsequent amino acid.
 - ▶ Insertion/deletion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

Nature: Code triplet

- Given the different numbers of “letters” in the mRNA and protein “alphabets,” scientists theorized that combinations of nucleotides corresponded to single amino acids.
- These nucleotide triplets are called codons².
- Nucleotide doublets would not be sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations.

$$4^2 = 16$$

$$4^3 = 64$$

- If there were 200 commonly occurring amino acids instead of 20. Given what you know about the genetic code, what would be the shortest possible codon length? Explain.

²Codons: Three consecutive nucleotides in mRNA that specify the insertion of an amino acid or the release of a polypeptide chain during translation

Nature: Degeneracy

- Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was degenerate³.
- Scientists solved the genetic code by translating synthetic mRNAs in vitro and sequencing the proteins they specified
- In addition to instructing the addition of a specific amino acid to a polypeptide chain, three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called nonsense codons⁴, or stop codons.
- AUG codon has a special function.
 - ▶ Specifies the amino acid methionine,
 - ▶ Serves as the start codon to initiate translation. The reading frame for translation is set near the 5' end of the mRNA.

³(of the genetic code) describes that a given amino acid can be encoded by more than one nucleotide triplet

⁴one of the three mRNA codons that specifies termination of translation

Nature: Universality

- The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis.
- Conservation of codons means that a purified mRNA encoding the globin protein in horses could be transferred to a tulip cell, and the tulip would synthesize horse globin.
- That there is only one genetic code is powerful evidence that all of life on Earth shares a common origin, especially considering that there are about 10^{84} possible combinations of 20 amino acids and 64 triplet codons.

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } ACA } ACG }	AAU } Asn AAC } AAA Lys AAG }	AGU } Ser AGC } AGA Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } GCA } Ala GCG }	GAU } Asp GAC } GAA Glu GAG }	GGU } GGC } GGA } Gly GGG }	U C A G
Third letter						

Figure 4: The genetic code is universal for all life forms

Property: Overlapping/non-overlapping genetic code

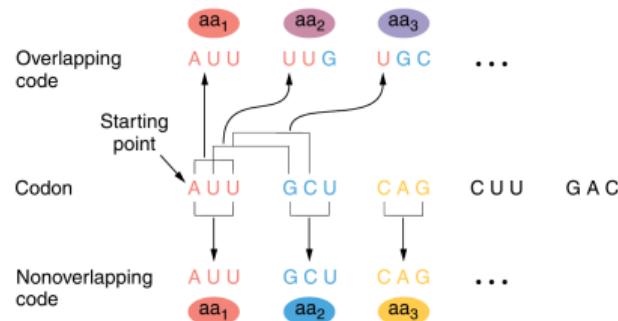


Figure 5: Overlapping versus non overlapping genetic codes

- An overlapping and a nonoverlapping genetic code would translate differently into an amino acidc sequence.
- In an overlapping code, single nucleotides occupy positions in multiple codons.
- In a nonoverlapping code, a protein is translated by reading nucleotides sequentially in sets of three.

Property

- With comma/commaless
- Non-ambiguous

Translational consequences of a mutation

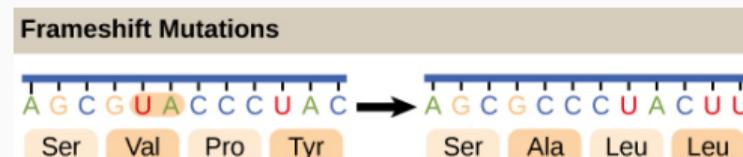


Figure 6: Frameshift mutation

- Degeneracy is believed to be a cellular mechanism to reduce the negative impact of random mutations.
- Codons that specify the same amino acid typically only differ by one nucleotide.
- Amino acids with chemically similar side chains are encoded by similar codons.
- So what does single nucleotide mutation do ?

Molecular phenotype

- The primary phenotype of a gene is the protein it produces.
- Two variants of a gene – wild and mutant, produce functionally different proteins to be called alleles.
- Human disease phenylketonuria (PKU) provides a hint to the topic.

Molecular biology of PKU

- Due to defective allele of a gene that encodes for liver enzyme phenylalanine hydroxylase (PAH).



- A mutation in the gene encoding this enzyme may alter the amino acid sequence in the vicinity of the enzyme's active site. In this case, the enzyme cannot bind phenylalanine or convert it into tyrosine.
- Phenylalanine builds up and is instead converted into phenylpyruvic acid, this interferes with the development of the nervous system.

- Mutations occur at different sites along the gene. Functional mutations are essentially those occurring in the protein-encoding regions, or the exons. They represent a range of DNA changes, but most are small changes affecting only one nucleotide pair among the thousands that constitute the gene.
- By changing one or more amino acids, (non-synonymous) mutations inactivate some essential part of the protein encoded by the gene. The effect of the mutation on the function of the gene depends on where within the gene the mutation occurs.

- An important functional region of the gene is that encoding an enzyme's active site. In addition, a minority of mutations are found to be in introns, and these mutations often prevent the normal processing of the primary RNA transcript.
- Many of the mutant alleles are of a type generally called null alleles: the proteins encoded by them completely lack PAH function. Other mutant alleles reduce the level of enzyme function; they are sometimes called leaky mutations, because some wild-type function seems to “leak” into the mutant phenotype.

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

Functionality of genes

- Genes don't always encode protein
- Some genes encode RNAs that have special functions, however all genes are transcribed to make RNA.
- RNAs that do not end up being mRNAs are functional RNAs (ribosomal RNAs, small cytoplasmic RNAs).

Study in *Neurospora* fungus

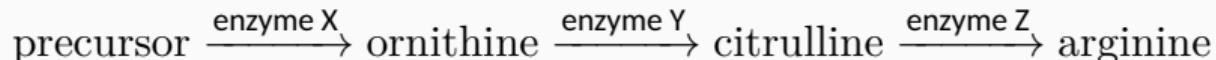
- Genes act by controlling cellular chemistry.
- Beadle and Tatum received a Nobel Prize for their study on the haploid fungus *Neurospora*.
- They used forward genetic approach by first inducing mutations and then tested cultures grown from ascospores for interesting mutant phenotypes.
- Specifically auxotrophic mutants were the subject; Auxotrophic mutants cannot synthesize all its cellular components from the inorganic nutrients and a carbon source in the medium.
- Thus mutant is defective for some normal synthesizing step.

- It was first confirmed that each mutation that generated a nutrient requirement was inherited as a single-gene mutation because each gave a 1:1 ratio when crossed with a wild type. Letting *aux* represent auxotrophic mutation and wild type as “+”,

$$\begin{array}{c} + \times aux \\ \downarrow \\ \text{progeny : } \frac{1}{2} + \text{ and } \frac{1}{2} aux \end{array}$$

- The second step was to classify the specific nutritional requirement of each auxotroph.
- Some grew only if proline was supplied, others methionine, others pyridoxine, others arginine. Focus was set on arginine types.
- Three separate genes different loci were identified on separate chromosomes.
- It was discovered that auxotrophs for each of the three loci differed in their response to the structurally related compounds ornithine and citrulline.

- *arg-1* mutants grew when supplied with any one of the chemicals ornithine, citrulline, or arginine.
- *arg-2* mutants grew when either citrulline or arginine was supplied.
- *arg-3* mutants grew only when arginine was supplied.
- Enzymes were already known to interconvert such related compounds.
- B & P proposed a biochemical pathway for such conversion in the fungus:

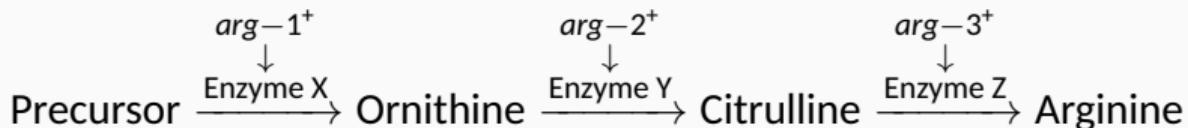


Mutant	Supplement		
	Ornithine	Citrulline	Arginine
<i>arg-1</i>	+	+	+
<i>arg-2</i>	-	+	+
<i>arg-3</i>	-	-	+

Figure 7: Growth of *arg* Mutants in response to supplements.

■ The model:

- ▶ *arg-1* mutants have a defective enzyme X, and so they are unable to convert the precursor into ornithine as the first step in producing arginine. However they have normal enzymes Y and Z, and so the *arg-1* mutants are able to produce arginine if supplied with either ornithine or citrulline.
- ▶ *arg-2* mutants lack enzyme Y, and the *arg-3* mutants lack enzyme Z.
- ▶ A mutation at a particular gene is assumed to interfere with the production of a single enzyme.
- ▶ The defective enzyme creates a block in biosynthetic pathway. The block can be prevented by supplying to the cells any compound that normally comes after the block in the pathway.



- This model was initially known as the *one-gene-one-enzyme hypothesis*.
- Genes somehow were responsible for the function of enzymes, and each gene apparently controlled one specific enzyme in a series of interconnected steps in a bio-chemical pathway
- All **Proteins** are infact encoded by gene hence the phrase one-gene-one-polypeptide hypothesis.
- Hence Beadle and Tatum's hypothesis became the great unifying theory that brought together areas of genetics and biochemistry together.

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

- An enzyme is a product of gene/s.
 - ▶ A tertiary or quarternary form of polypeptide derived from amino-acid units.
- Enzyme controls/catalyzes biochemical reaction and the reaction produces biochemical product which is responsible to final phenotype expression.
- So, phenotypic ratio is determined by the enzyme.

The 9:3:3:1 ratio; No interaction

- In inheritance of skin coloration in corn snakes, the phenotype is produced by two separate pigments. The snake's natural color is a repeating black and orange camouflage pattern, as shown in Figure 8 (a).
- The phenotype is produced by two separate pigments, both of which are under genetic control.
- One gene determines orange pigment (o^+ and o) and another gene determines black pigment (b^+ and b).
- Two genes are unlinked.
- Natural pattern is produced by the genotype $o^+/_$; $b^+/_$.
- What genotypes leads to completely black pigmentation?

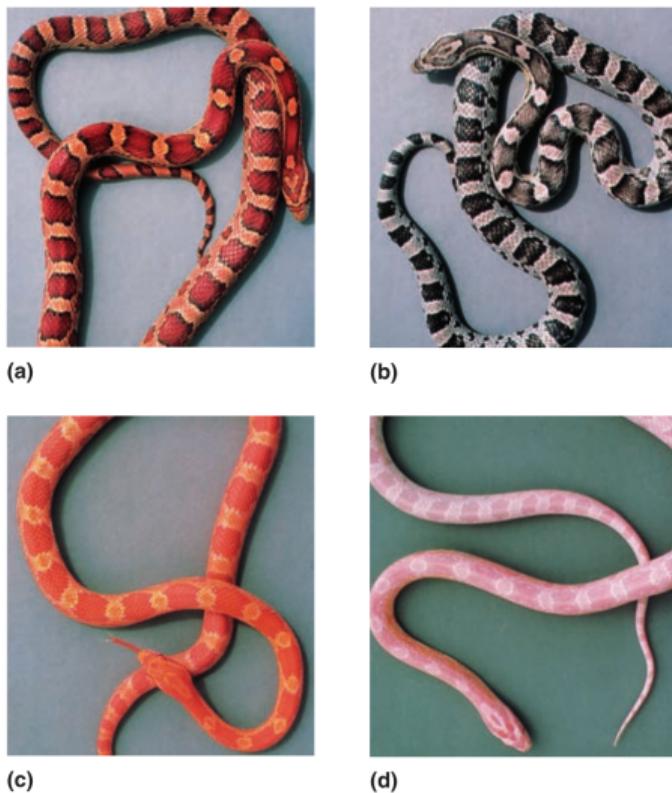
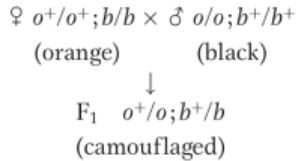
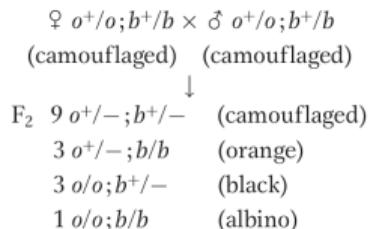


Figure 8: Phenotypes of snake skin pigmentation

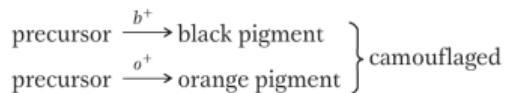
- If homozygous orange and a homozygous black snake are crossed, the F_1 is wild type (camouflaged), demonstrating complementation.



Here, however, an F_2 shows a standard 9:3:3:1 ratio:



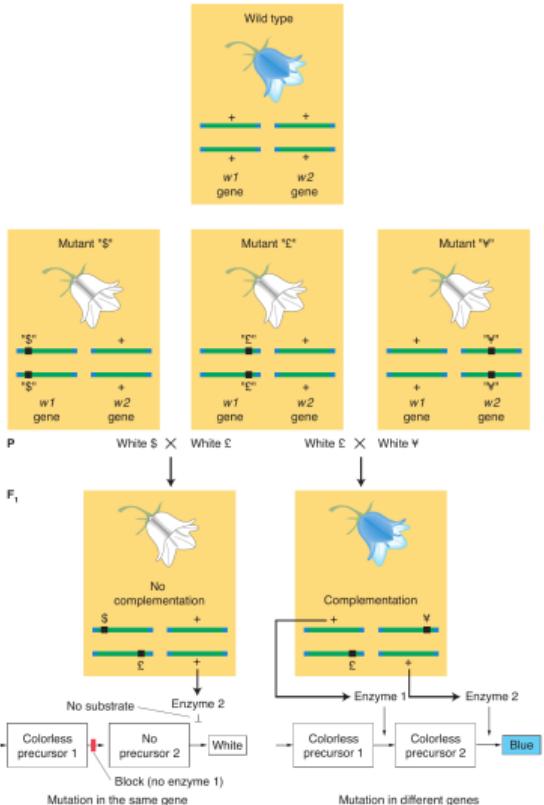
The 9:3:3:1 ratio is produced because the two pigment genes act independently at the cellular level.



The 9:7 ratio: gene complementation (genes acting on same pathway)

- The F_2 ratio from the harbell dihybrid cross shows both blue and white plants in ratio of 9:7. How can such results be explained ?
- The 9:7 ratio is a modification of the dihybrid 9:3:3:1 ratio with the 3:3:1 combined to make 7.

Enzymatic explanation of genetic ratios



$w1/w1; w2^+/w2^+$ (white) × $w1^+/w1^+$; $w2/w2$ (white)	
↓	
F ₁	$w1^+/w1$; $w2^+/w2$ (blue)
	$w1^+/w1$; $w2^+/w2$ × $w1^+/w1$; $w2^+/w2$
↓	
F ₂	9 $w1^+/-; w2^+/-$ (blue) 9
	3 $w1^+/-; w2/w2$ (white) }
	3 $w1/w1; w2^+/-$ (white) 7
	1 $w1/w1; w2/w2$ (white)

- Only way in which a 9:7 ratio is possible is if the double mutant has the same phenotypes as the two single mutants.
- Hence, the modified ratio constitutes a way of identifying the double mutant's phenotype.
- Furthermore, the identical phenotypes of the single and double mutants suggest that each mutant allele controls a different step in the same pathway.
- The results show that a plant will have white petals if it is homozygous for the recessive mutant allele of either gene or both genes.
- To have the blue phenotype, a plant must have at least one copy of the dominant allele of both genes because both are needed to complete the sequential steps in the pathway.
- No matter which is absent, the same pathway fails, producing the same phenotype. The enzymatic pathways allowing for complementation is through regulatory gene products (Shown in 11).
- Thus, three of the genotypic classes will produce the same phenotype, and so, overall, only two phenotypes result.

The 9:3:4 ratio: recessive epistasis

- One important type of functional interaction between different genes occurs when an allele of genotype at one locus “masks” or inhibits the expression of a allele or genotype at a different locus, such an interaction is called epistasis.
- Epistasis means stand upon.
- Double mutants show the phenotype of one mutation but not the other.
- Overriding mutation is epistatic, whereas the overridden one is hypostatic.
- Also results from the genes in same pathway.
- In a simple model, if epistatic mutation is carried by a gene that is farther upstream than the one being overridden, the mutant phenotype of the upstream gene takes precedence, no matter what is taking place later in the pathway.

	$w/w; m^+/m^+$ (white) \times $w^+/w^+; m/m$ (magenta)	
F ₁	$w^+/w; m^+/m$ (blue)	
	↓	
	$w^+/w; m^+/m \times w^+/w; m^+/m$	
	↓	
F ₂	$9w^+/_; m^+/_$ (blue)	9
	$3w^+/_; m/m$ (magenta)	3
	$\left\{ \begin{array}{l} 3w/w; m^+/_ \text{(white)} \\ 1w/w; m/m \text{(white)} \end{array} \right\}$	4

Enzymatic explanation of genetic ratios

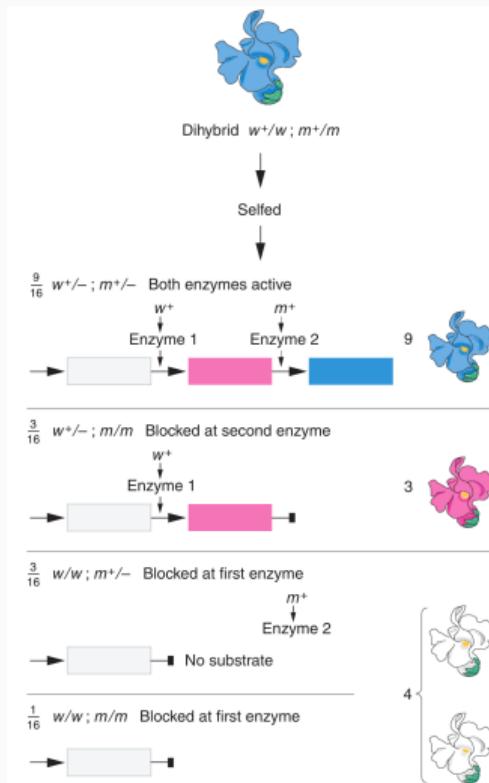


Figure 9: A model for recessive epistasis

Epistasis is inferred when a mutant allele of one gene masks the expression of a mutant allele of another gene and expresses its own phenotype instead.

- This interaction is called recessive epistasis because a recessive phenotype (white) overrides the other phenotype.
- At the cellular level, we can account for the recessive epistasis in *Collinsia* by the following type of pathway.



- Notice that the epistatic mutation occurs in a step in the pathway leading to blue pigment; this step is upstream of the step that is blocked by the masked mutation.

The 12:3:1 ratio: dominant epistasis

- In foxgloves (*Digitalis purpurea*), two genes interact in the pathway that determines petal coloration.
- The two genes are unlinked.
- One gene affects the intensity of red pigment in the petal,
 - ▶ *d* results in the light red color seen in natural populations
 - ▶ *D* is a mutant allele that produces dark red color.
- Other gene determines in which cells the pigment is synthesized,
 - ▶ *w* allows synthesis of the pigment throughout the petals as in the wild type,
 - ▶ *W* is a mutant allele that confines pigment synthesis to the small throat spots.
- Selfing of dihybrid shows:

$\left\{ \begin{array}{l} 9D/_; W/_ \text{ (white with spots)} \\ 3d/d; W/_ \text{ (white with spots)} \end{array} \right\}$	12
$3D/_; w/w$ (dark with red)	3
$1d/d; w/w$ (light red)	1



Figure 10: Dominant epistasis due to white mutation

Regulators

	Regulatory gene	Gene for protein A	Protein product of gene a
(a) Normal			Wild-type protein A produced
(b) Mutation in the gene that encodes the regulatory protein			No protein A produced
(c) Mutation in the gene that encodes the structural protein			Mutant protein A produced
(d) Mutation in both genes			No protein A produced

Figure 11: Molecular mechanisms of genes acting on same pathway. Interaction between a regulatory protein and its target

Suppressors

- A suppressor is a mutant allele of a gene that reverses the effect of a mutation of another gene, resulting in a wild-type or near-wild-type phenotype.
- Suppression implies that the target gene and the suppressor gene normally interact at some functional level in their wild-type states.
- Assume that an allele a^+ produces the normal phenotype, whereas a recessive mutant allele a results in abnormality.
- Screening for suppressors involves:
 - ▶ Assemble mutants in some process of interest,
 - ▶ Expose these mutants to mutation causing agents such as high energy radiation, and screen the descendants for wild types.
 - ▶ Most wild types arising in this way are merely reversals of the original mutation event and are called revertants. However, some will be “pseudorevertants”, double mutants in which one of the mutations is a suppressor.

- revertants and suppressed states can be distinguished by appropriate crossing. For example, in yeast, two results would be distinguished as,

Progeny	true revertants $a^+ \times$ standard wild-type a^+	
	↓	
Progeny	all a^+	
Progeny	supressed mutant $a.s \times$ standard wild-type $a^+.s^+$	
	$a^+.s^+$	wild type
	$a^+.s$	wild type
	$a.s^+$	original mutant
	$a.s$	wild type (supressed)

- The appearance of the original mutant phenotype identifies the parent as a suppressed mutant.

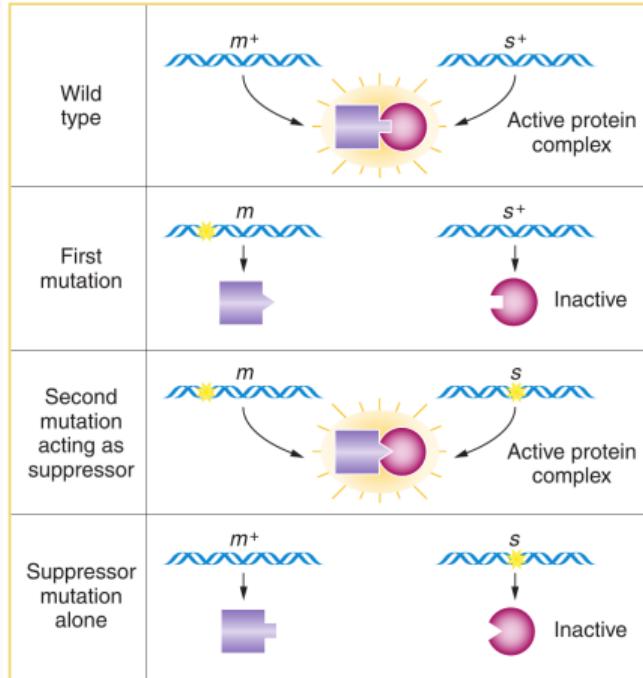


Figure 12: Molecular mechanism for suppression. A first mutation alters the binding site of one protein so that it can no longer bind to a partner. A suppressor mutation in the partner alters the binding site so that both proteins are able to bind once again.

Modifiers

- Modifier mutation at a second locus changes the degree of expression of the mutated gene at the first locus. Regulatory genes are an example.
- Regulatory proteins bind to the sequence of the DNA upstream of the start site for transcription. These proteins regulate the level of transcription.
- As opposed to null mutation of a regulatory gene (in complementation), which completely (almost) prevented transcription, some regulatory mutations change the level of transcription of the target gene so that either more (upregulation) or less protein (downregulation) is produced.
- A popular downregulating regulatory mutation is that in gene “B” affecting gene A in a fungus such as yeast.
- We cross a leaky mutation a with the regulatory mutation b ;

leaky mutant $a.b^+ \times$ inefficient regulator $a^+.b$

Table 3: Progeny characterization

Progeny	Phenotype
$a^+.b^+$	wild type
$a^+.b$	defective (low transcription)
$a.b^+$	defective (defective protein A)
$a.b$	extremely defective (low transcription of defective protein)

- Hence the action of modifiers is seen in the appearance of two grades of mutant phenotypes *within* the a progeny.

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

- There are two concepts of gene.
 - ▶ Older theory aka bead theory
 - ▶ Recent theory aka molecular theory

Bead theory

- A gene is a unit of inheritance.
- Its existence is in the form of an allele.
- Each of the alleles affects a single phenotypic character. All alleles map to one chromosome locus.
- A gene gives a mutant phenotype when they are in pairs.
- Genes produce mendelian phenotypic ratio when intercrossed.
- Old concept speaks of indivisible nature of gene (even upon occurrence of crossing over).
- Crossing over occurs between genes, but crossing over does not occur in the same gene. Thus, a gene is fundamental unit of change or mutation.

- A gene changes from one allelic form to another. There are no smaller units inside the gene that can change.
- A gene is basic unit of function. Parts of a gene if they exist, cannot function.
- In the same way as the beads are arranged in an unfastened necklace, genes are arranged linearly in a chromosome. So chromosome is a linear array of genes according to the bead theory.
- There is no complementation between the allele pair of a gene.
- The old theory of the gene says that there is only one site of mutation in a gene. In other words: one gene ~ one site of mutation.

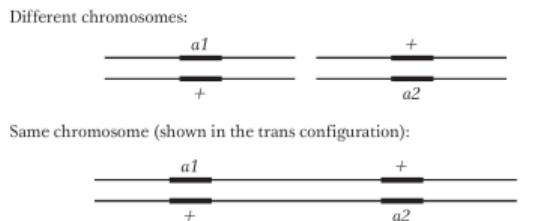
Molecular theory

- Genes and cistrons are synonymous.
- A cistron or a gene is a fundamental unit to produce a kind of a polypeptide.
- But the gene is most commonly used.

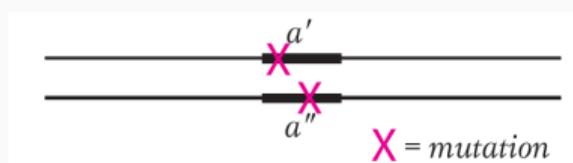
Complementation test

- Determines whether two mutations belong to the same gene.
- Quicker approach when linkage mapping is time consuming.
- The decision whether members of a multiple allelic series are located in a single gene or in two or more separately but closely-linked genes is based on complementation test.
- When the trans-heterozygotes for two mutant alleles (affecting the same trait) have the mutant phenotype, they are placed in the same gene. But if they have the wild type phenotype, the mutant alleles are considered to be located in two different genes.

- In a diploid, the complementation test is performed by intercrossing two individuals that are homozygous for different recessive mutations.
- Next step is to observe whether the progeny have the wild-type phenotype. If the progeny are wild type, two recessive mutations must be in *different* genes because the respective wild-type alleles provide wild-type function. In this case, the two mutations are said to have *complemented*.
- Suppose $a1$ and $a2$ are mutant alleles of two genes.
- Heterozygotes can be represented as follows, depending on whether genes are on the same or different chromosomes.



- However, if the progeny are not wild type, then the recessive mutant must be alleles of the same gene. Because both alleles of the gene are mutants, there is no wild-type allele to help distinguish between two different mutant alleles of a gene whose wild-type allele is a^+ .
- These alleles could have different mutant sites within the same gene, but they would both be non functional.
- The heterozygote a'/a'' would be



- At operational level, complementation is defined as the production of a wild type phenotype when two haploid genomes bearing different recessive mutations are united in the same cell.

- Let us take an example on harebell plants (genus *Campanula*). The wild-type flower color of this plant is blue. Let's assume that, from a mutant hunt, we have obtained three white-petaled mutants and that they are available as homozygous pure-breeding strains.
- They all look the same, and so we do not know whether they are genetically identical.
- We will call the mutant strains \$, £, and ¥ to avoid any symbolism using letters, which might imply dominance. When crossed with wild type, each mutant gives the same results in the F_1 and F_2 as follows:

white \$ × blue → F_1 , all blue → F_2 , $\frac{3}{4}$ blue, $\frac{1}{4}$ white

white £ × blue → F_1 , all blue → F_2 , $\frac{3}{4}$ blue, $\frac{1}{4}$ white

white ¥ × blue → F_1 , all blue → F_2 , $\frac{3}{4}$ blue, $\frac{1}{4}$ white

- In each case, results show that the mutant condition is determined by the recessive allele of a single gene.
- However, are they three alleles of one gene, of two genes, or of three genes?
- Because the mutants are recessive, the question can be answered by the complementation test, which asks if the mutants complement one another.
- Let us intercross the mutants and assume following results are obtained:

white \$ × white £ → F₁, all white

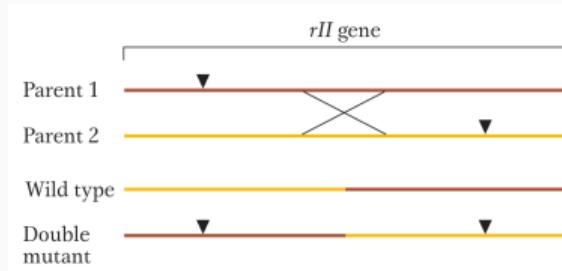
white \$ × white ¥ → F₁, all blue

white £ × white ¥ → F₁, all blue

- From this set of results, we can conclude that mutants \$ and £ must be caused by alleles of one gene (say, w1) because they do not complement, but ¥ must be caused by a mutant allele of another gene (w2) because ¥ complements both \$ and £.

Benzer's experiment on rII locus of T4 with *E. coli*

- Because astronomically large numbers of phages can be used in phage-recombination analyses, very rare crossover events can be detected.
- In the 1950s, Seymour Benzer made use of such rare crossover events to map the mutant sites within the rII gene of phage T4, a gene that controls lysis.
- For different rII mutant alleles arising spontaneously, the mutant site is usually at different positions within the gene.
- Therefore, when two different rII mutants are crossed, a few rare crossovers may take place between the mutant sites, producing wild-type recombinants, as shown here:



As distance between two mutant sites increases, such a crossover event is more likely. Thus, the frequency of rII^+ recombinants is a measure of that distance within the gene. (The reciprocal product is a double mutant and indistinguishable from the parentals.)

Benzer used a clever approach to detect the very rare rII^+ recombinants. He made use of the fact that rII mutants will not infect a strain of *E. coli* called K. Therefore, he made the $rII \times rII$ cross on another strain and then plated the phage lysate on a lawn of strain K. Only rII^+ recombinants will form plaques on this lawn. This way of finding a rare genetic event (in this case, a recombinant) is a selective system: only the desired rare event can produce a certain visible outcome.

- The T4 has A and B cistrons.
- According to Benzer, a gene is a functional unit to produce a particular polypeptide.
- Benzer isolated 3000 mutants of rII locus of T4 phage based on frequency of crossing over.
- Benzer developed a map of the mutant sites of rII locus of T4 phage. The mutation sites were found spread along the entire length of the cistron A and B.
- He concluded that crossing over occurs inside the gene or cistron. The gene is divisible.
- There are many sites of recombination in a gene. There are many sites of change or mutation in a gene.
- A unit of mutation in a gene is a muton. The muton is a single nucleotide pair
- Benzer termed sites that are more mutable than other sites as hot spots.
- The number of mutational sites was determined to be approximately one-fifth (1/5) of the number of nucleotide pairs.
- In other words, the smallest mutable site was five nucleotide pair or less.

- From the experiment following conclusions could be drawn:
 - ▶ A gene is a distinct chromosomal region responsible for a single cellular function. It consists of a linear array of potentially mutable sites between which recombination can occur.
 - ▶ A gene is subdivisible into several units called nucleotide pairs (recombination is possible within a gene) among these nucleotide pairs.

Outline

- 1 Gene
- 2 Nature of genetic code
- 3 One gene one polypeptide hypothesis
- 4 Enzymatic explanation of genetic ratios
- 5 Fine structure of gene
- 6 Bibliography

See also

- For a discourse on types of gene interaction, refer to lecture notes on 'Introductory genetics', 3rd semester, BScAg.

References

Bibliography

Erin Gilchrist and George Haughn. Gene identification: Reverse genetics. *Diagnostics in Plant Breeding*, pages 61–89, 2013.

Qing Ji. Gene identification: forward genetics. *Diagnostics in Plant Breeding*, pages 41–60, 2013.