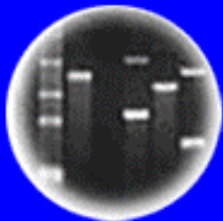


DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

## HC70A & SAS70A Winter 2011 Genetic Engineering in Medicine, Agriculture, and Law

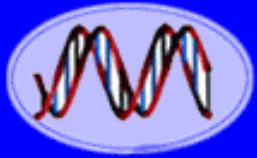
Professors Bob Goldberg & John Harada

### Lecture 4

The Nuts & Bolts of Genetic  
Engineering: The Factor VIII Story -  
From Gene To Drug

**UCLA**

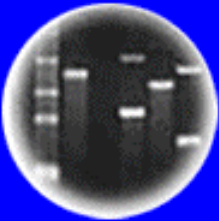
**UC DAVIS**  
UNIVERSITY OF CALIFORNIA



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



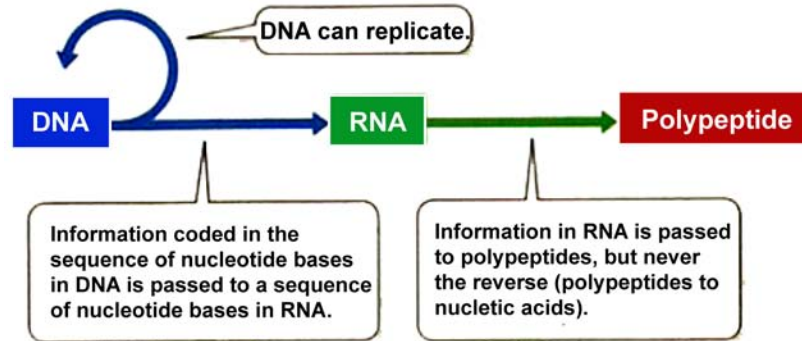
Plants of Tomorrow

# THEMES

1. Continuation of How Do Genes Work - Parts One and Two
2. What is Hemophilia?
3. How Is Hemophilia Inherited?
4. What is the Pedigree Pattern of a Sex-Linked Gene?
5. How Find a Disease Gene When It is Not Known Where the Gene is Expressed?
6. What Vectors Can Be Used For Cloning DNA?
7. What Are the Advantage of Using a Virus Vector For Constructing Genome Libraries?
8. How Make a Library of the Human Genome?
9. How Find a Gene With Only a Knowledge of the Protein Sequence?
10. What is Chromosome Walking & What Role Did it Play in Cloning the Factor VIII Gene?
11. How Use DNA Testing to Detect Factor VIII Disease Alleles?
12. How Isolate a Factor VIII cDNA Clone?
13. How Produce Factor VIII Protein For Use as a Drug?

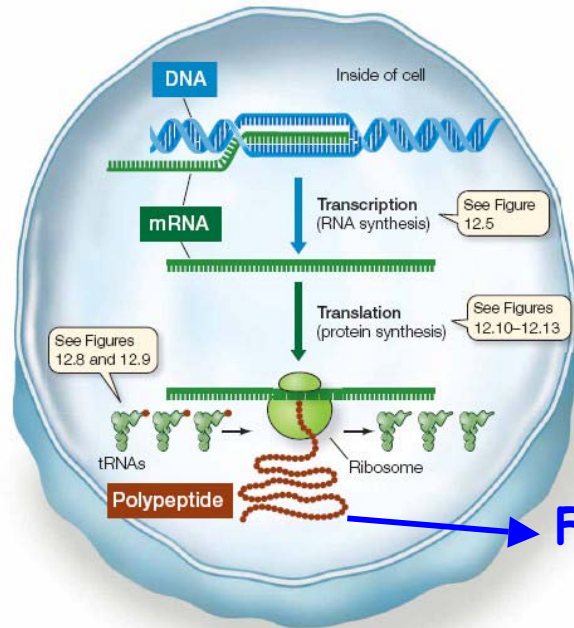
# How Do Genes Work (Lecture 3 Continued)

## ① Replication

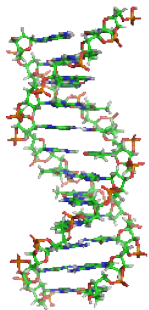


## ② Gene Activity to Function & Phenotype

Gene Activity  
↓  
Protein  
↓  
Function  
↓  
Phenotype (Trait)



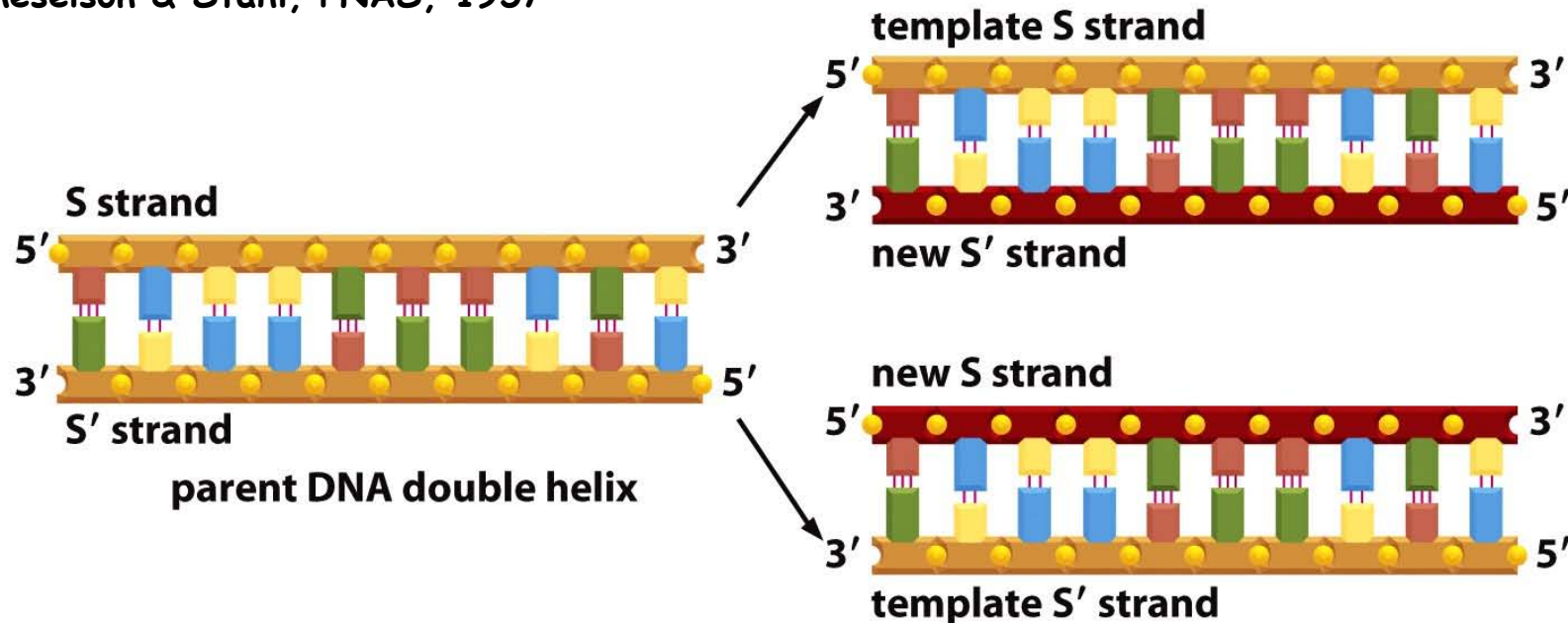
Function →



**A Gene is NOT Expressed Unless A Functional Protein Produced!**

# DNA Replication Occurs Semi-Conservatively

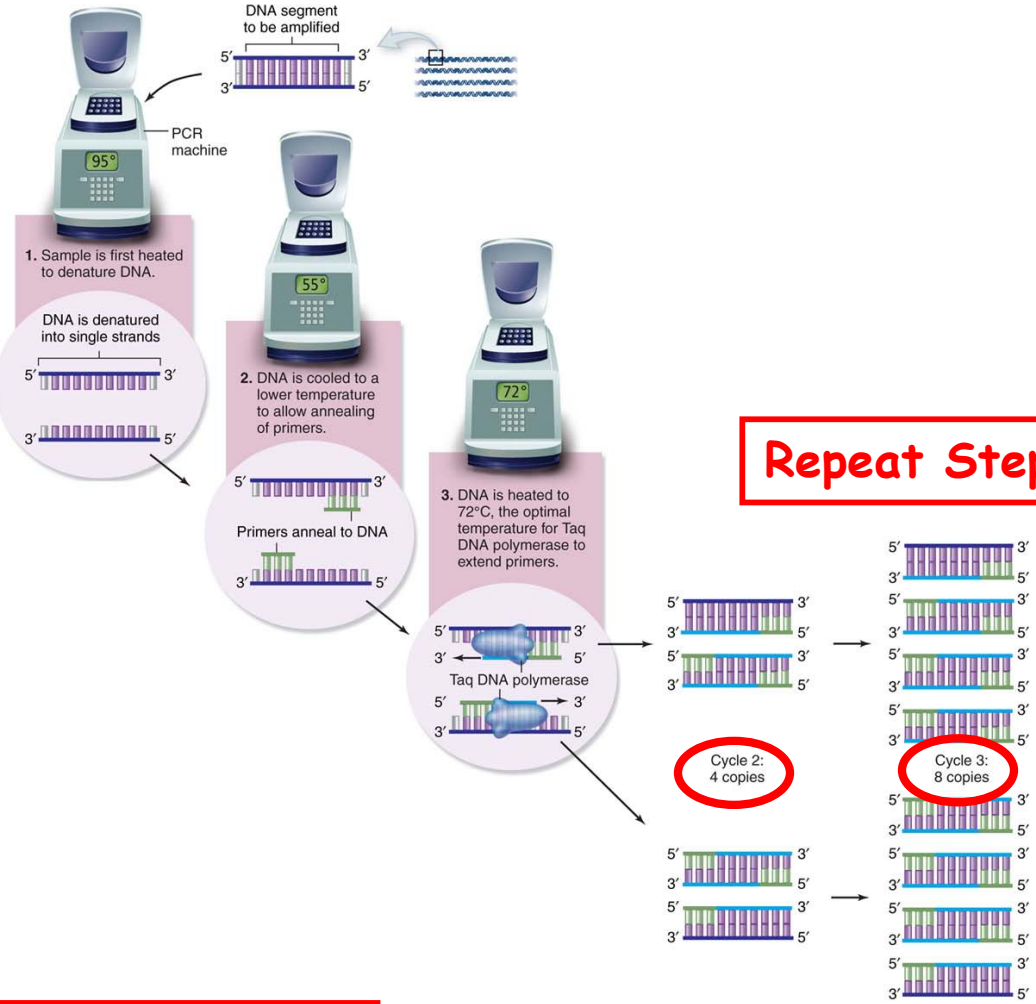
Meselson & Stahl, PNAS, 1957



1. DNA Structure Allows DNA Sequence to Be Maintained by Complementary Base Pairing
2. Each Strand Serves as a Template for the Synthesis of a Complementary Strand
3. New DNA Molecules are Precise Copies of Parental DNA - Each Containing One Newly Synthesized Complementary Strand

# PCR is A Cyclical Process of DNA Replication

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



- Requires**
1. Template
  2. Primers
  3. Knowledge of Specific Sequence
  4. Nucleotides
  5. Heat-Stable DNA Polymerase
  6. Cycler

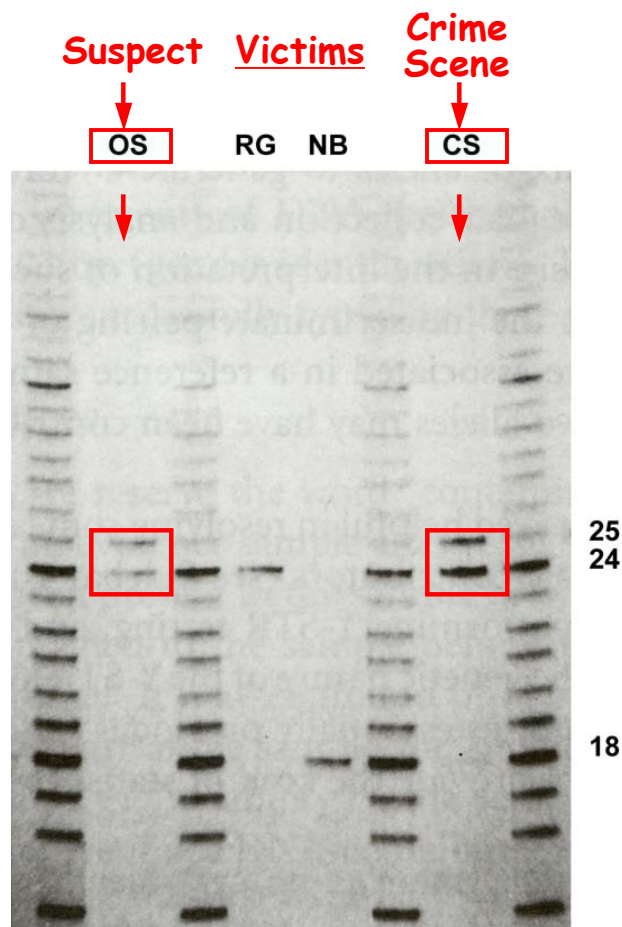
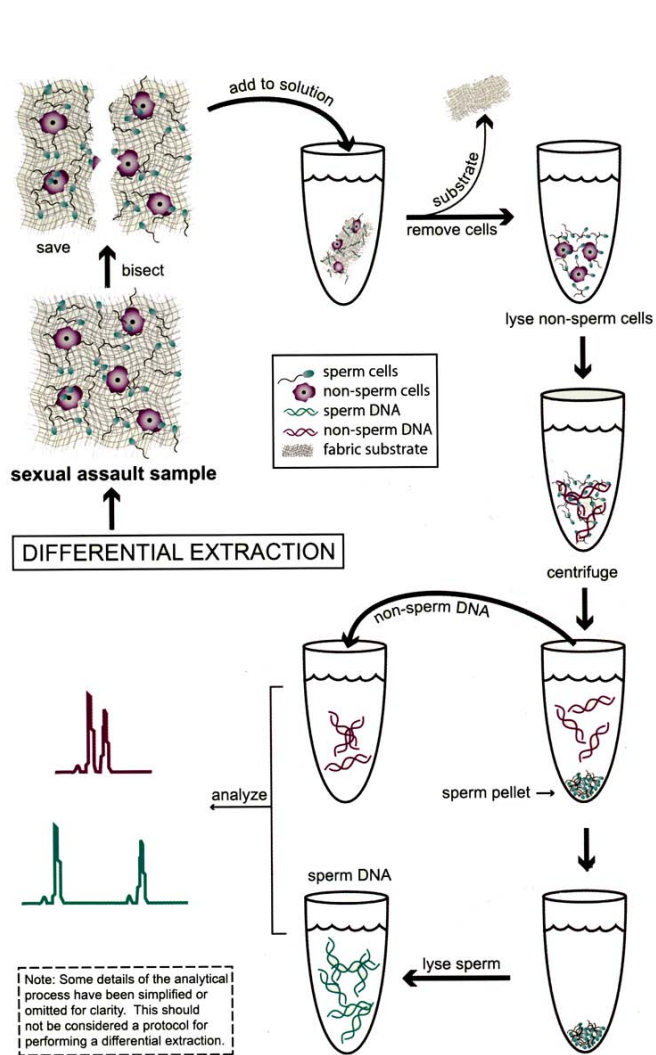
**Repeat Steps or Cycle**

**2<sup>n</sup> Molecules of DNA where n = Number of Cycles**

**Diagnostic For Amplified DNA Sequence**

**DNA Fragments All The Same Size Primer-Sequence-Primer**

# Using PCR in Crime Scenes



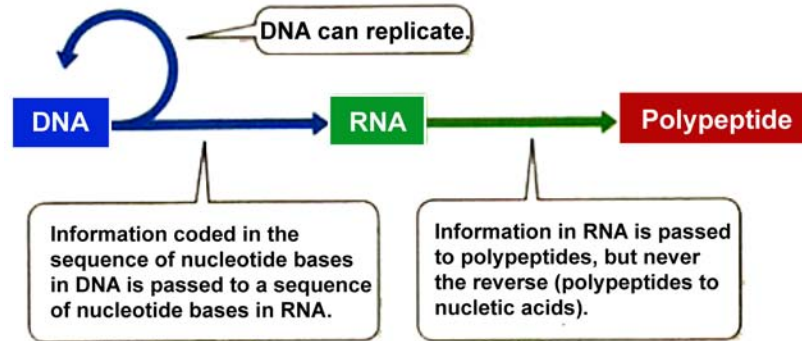
OS = Suspect  
 CS = Crime Scene  
 RG & NB = Victims

**“Match”**  
 What is Probability  
 That This  
 Will Occur  
 by Chance?

**DNA Doesn't “Lie” !!**

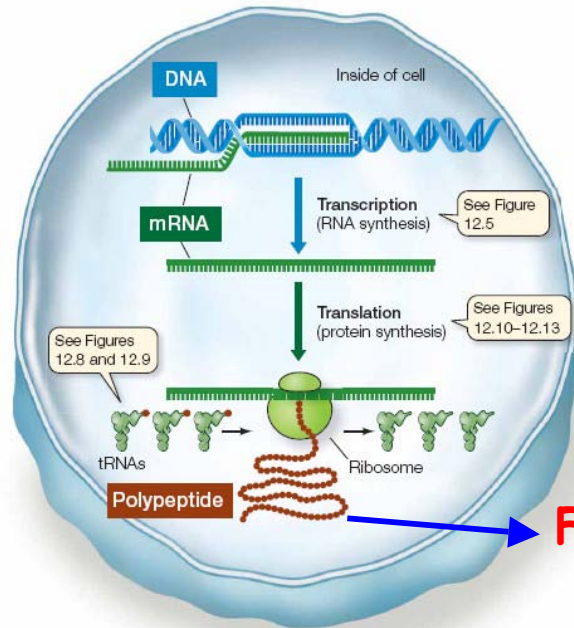
# How Do Genes Work - Mutations!

## ① Replication Mistakes?

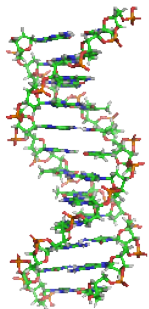


## ② Gene Activity to Function & Phenotype

Gene Activity  
↓  
Protein  
↓  
Function  
↓  
Phenotype (Trait)



Function →

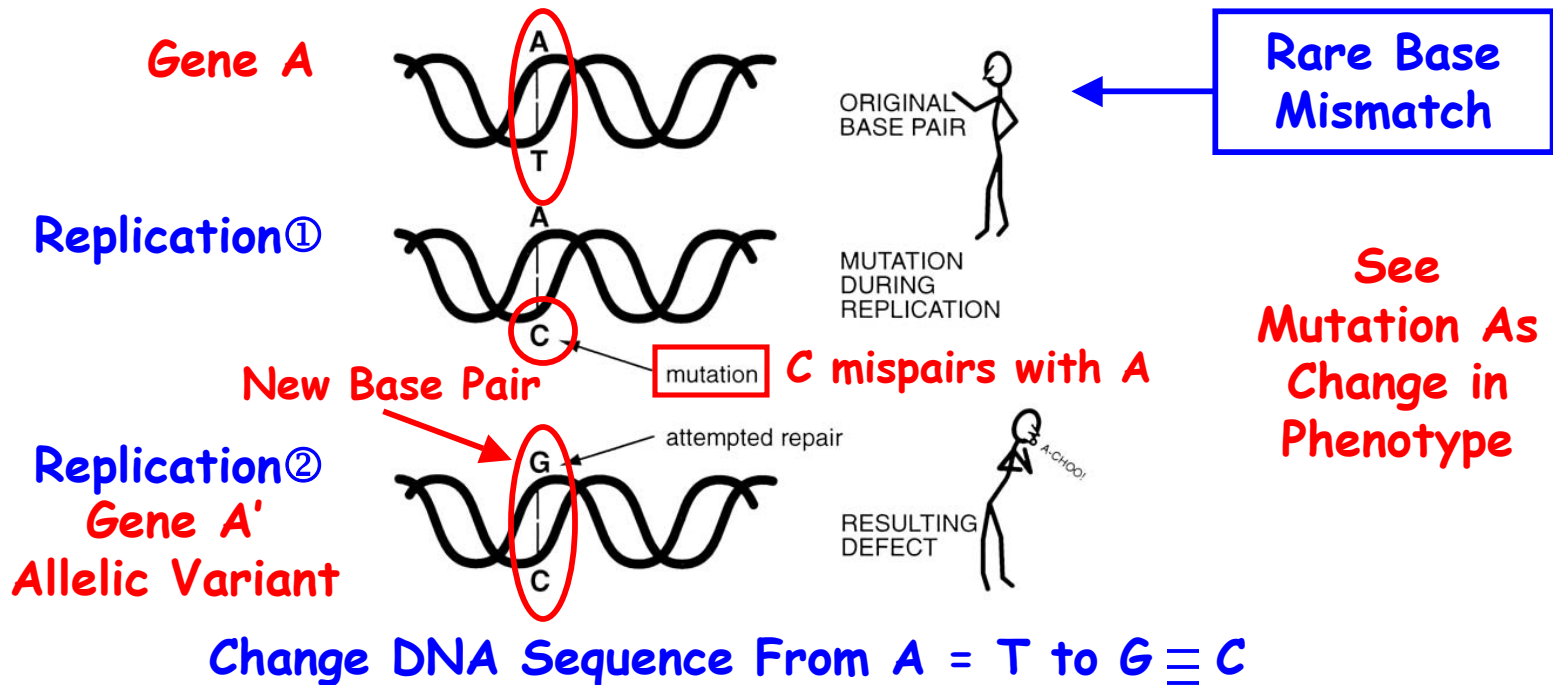


**A Gene is NOT Expressed Unless A Functional Protein Produced!**

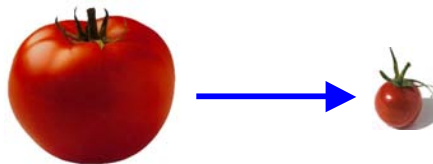
# DNA Replication is Precise But Mistakes or Mutations Can Occur!!

	DNA	RNA	
pair	A	A	} pair
	T	U	
pair	G	G	} pair
	C	C	

BASE PAIR RULES



**∴ Change Protein Amino Acid Sequence ⇨ Alter Function!**



**Big Tomato to Small Tomato**

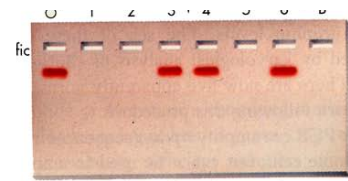
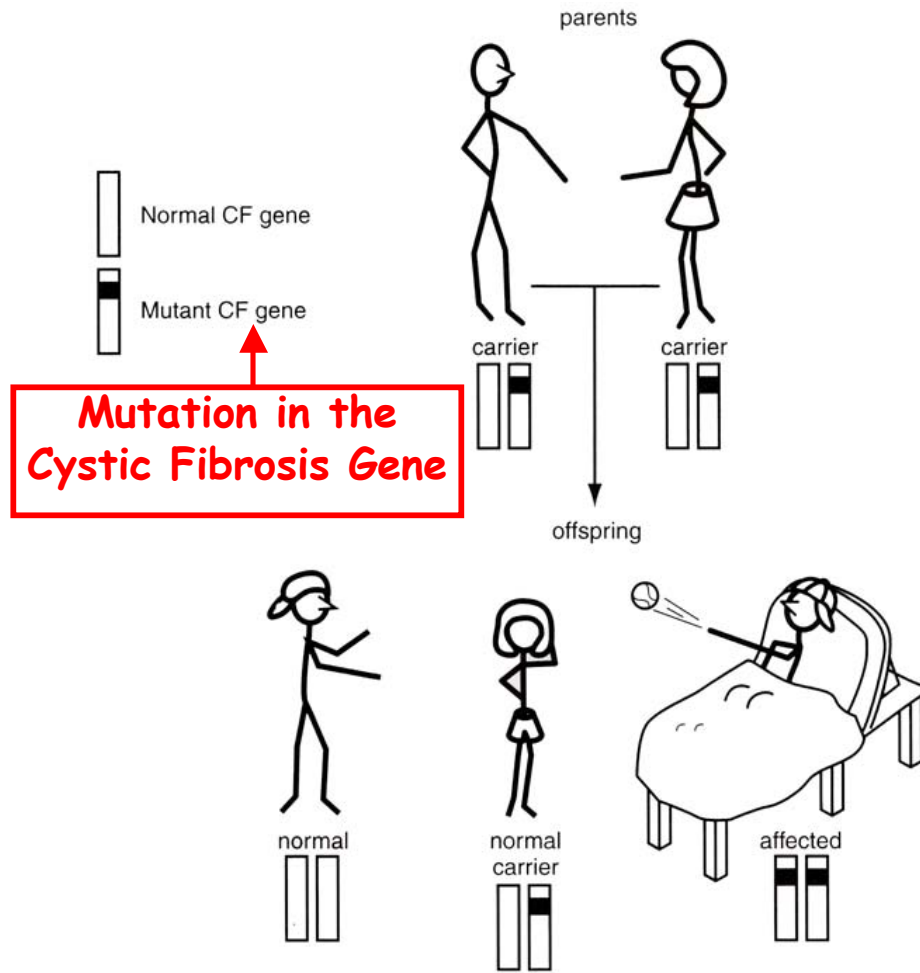


# Mutation in Genes Are Rare But Are Inherited (1 out of $10^7$ replications)

**One Gene Per Gamete**

♀ + ♂

**Two Genes per Somatic Cells**

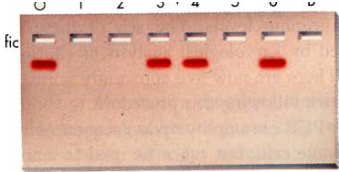
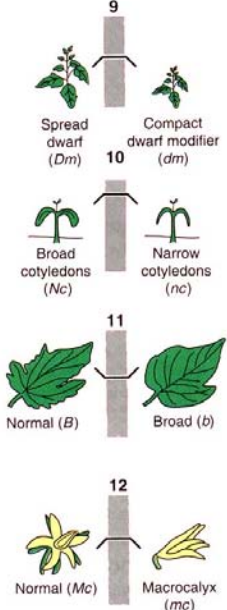
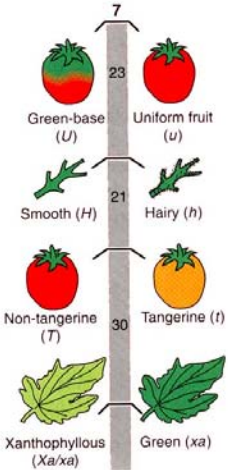
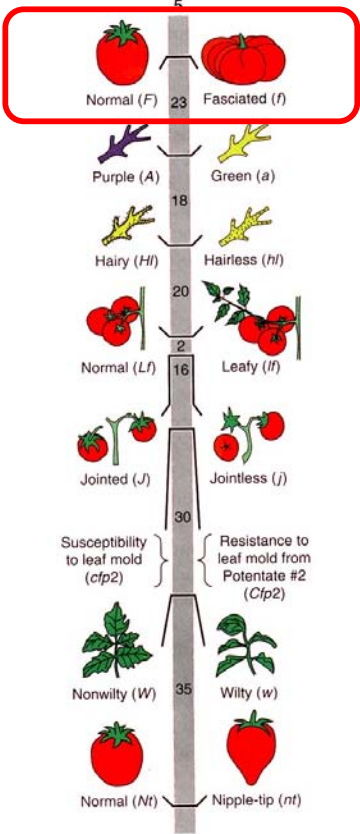


**How Follow Inheritance?  
What Allows Disease To Be Followed?**

**DNA Marker or Fingerprint!**

# Alternative Forms of the Same Gene Lead to Genetic Diversity

*Alleles*



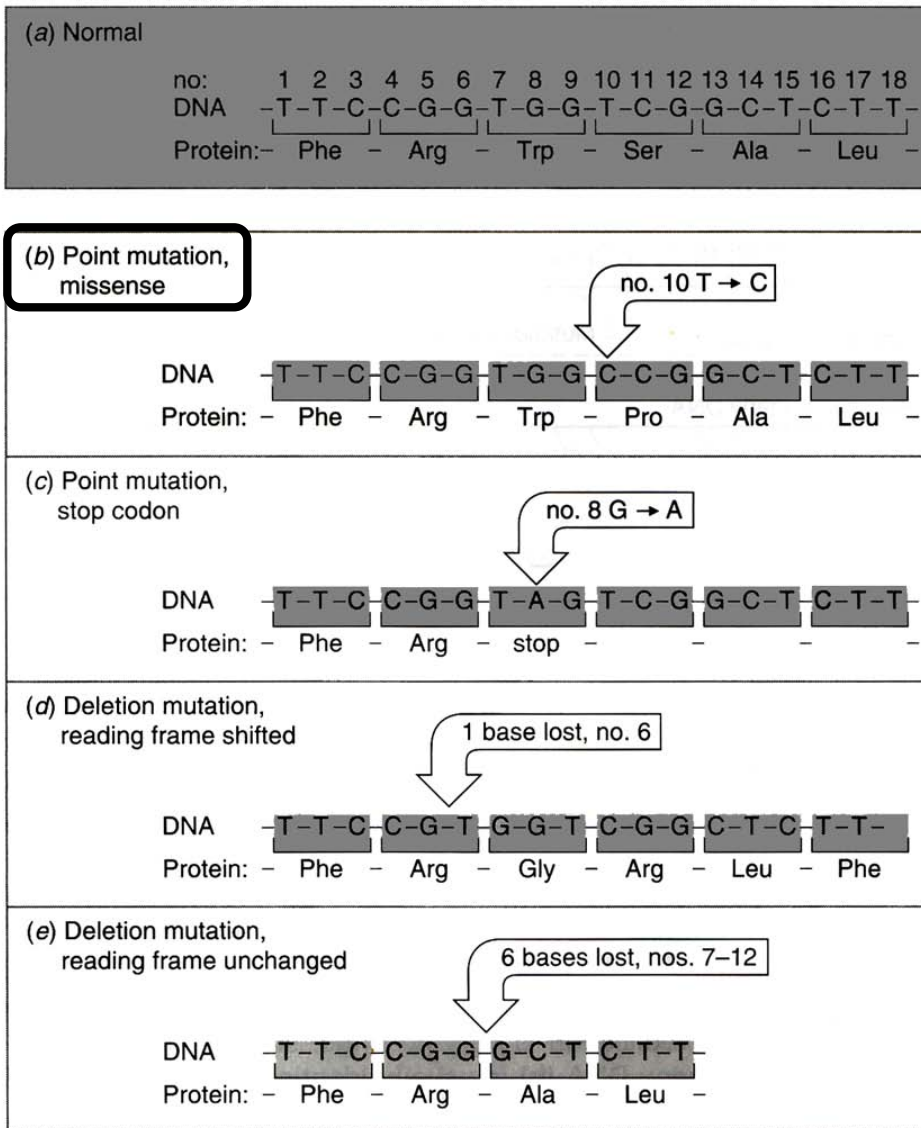
Analyze PCR products on gel

Can Follow These Traits With DNA Markers As Well

*mutations result in genetic diversity!!!*

*Spontaneous Mutations Give Rise To Alleles, or Different Forms of the Same Gene, And result in Small DNA Sequence Changes (e.g., SNPs or Single Nucleotide Polymorphisms)*

# Mutations Can Occur Different Ways



1. Base-Pair Change
2. Insert or Delete Base (Indel)
3. Move Gene, or Part of Gene, to New Location (Switches Change)!

Function of Protein Lost and/or Changed

∴

Phenotype Changes

# Human Genetic Disorders Occur As a Result of Mutations

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

TABLE 13.2		Some Important Genetic Disorders		
Disorder	Symptom	Defect	Dominant/ Recessive	Frequency Among Human Births
Hemophilia	Blood fails to clot	Defective blood-clotting factor VIII	X-linked recessive	1/10,000 (Caucasian males)
Huntington disease	Brain tissue gradually deteriorates in middle age	Production of an inhibitor of brain cell metabolism	Dominant	1/24,000
Muscular dystrophy (Duchenne)	Muscles waste away	Degradation of myelin coating of nerves stimulating muscles	X-linked recessive	1/3700 (males)
Hypercholesterolemia	Excessive cholesterol levels in blood lead to heart disease	Abnormal form of cholesterol cell surface receptor	Dominant	1/500

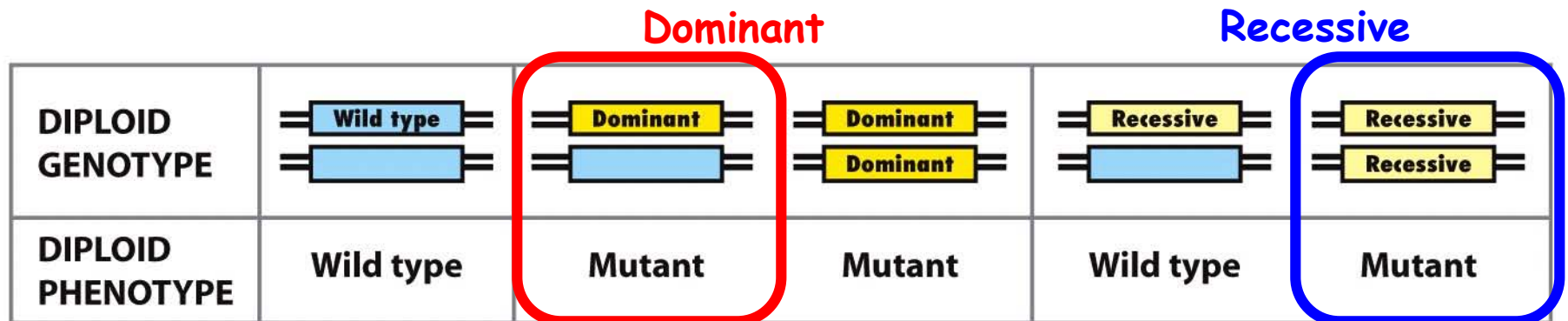
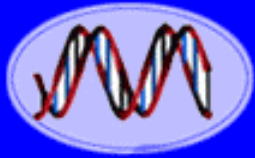


Figure 5-2  
*Molecular Cell Biology, Sixth Edition*  
 © 2008 W. H. Freeman and Company

Need One Allele

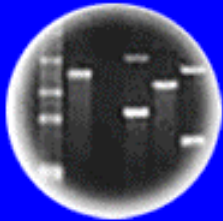
Need Two Alleles



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

# ARTICLE

Nature, October 10, 2010

doi:10.1038/nature09534

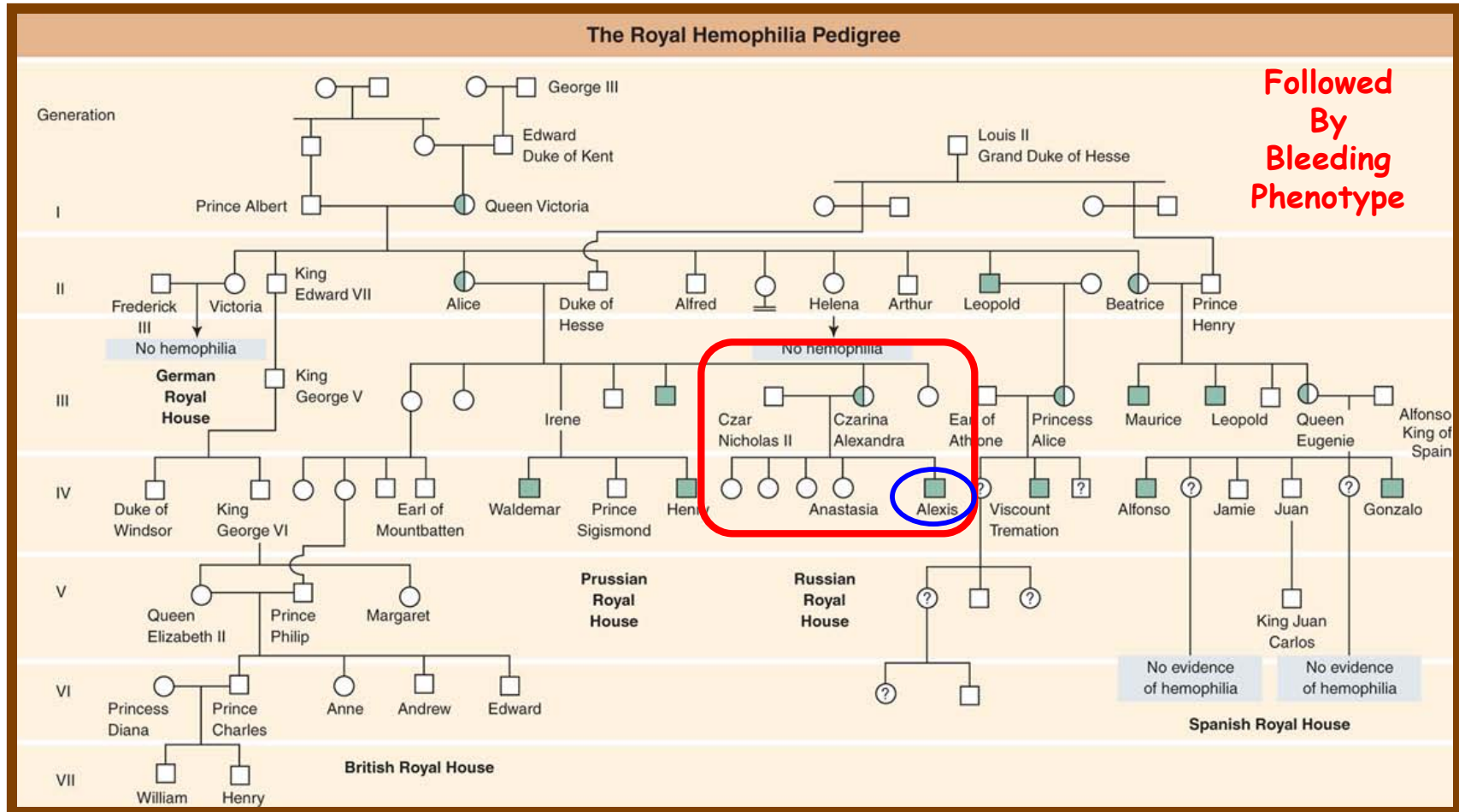
## A map of human genome variation from population-scale sequencing

The 1000 Genomes Project Consortium\*

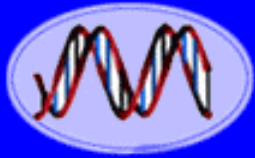
The 1000 Genomes Project aims to provide a deep characterization of human genome sequence variation as a foundation for investigating the relationship between genotype and phenotype. Here we present results of the pilot phase of the project, designed to develop and compare different strategies for genome-wide sequencing with high-throughput platforms. We undertook three projects: low-coverage whole-genome sequencing of 179 individuals from four populations; high-coverage sequencing of two mother-father-child trios; and exon-targeted sequencing of 697 individuals from seven populations. We describe the location, allele frequency and local haplotype structure of approximately 15 million single nucleotide polymorphisms, 1 million short insertions and deletions, and 20,000 structural variants, most of which were previously undescribed. We show that, because we have catalogued the vast majority of common variation, over 95% of the currently accessible variants found in any individual are present in this data set. On average, each person is found to carry approximately 250 to 300 loss-of-function variants in annotated genes and 50 to 100 variants previously implicated in inherited disorders. We demonstrate how these results can be used to inform association and functional studies. From the two trios, we directly estimate the rate of *de novo* germline base substitution mutations to be approximately  $10^{-8}$  per base pair per generation. We explore the data with regard to signatures of natural selection, and identify a marked reduction of genetic variation in the neighbourhood of genes, due to selection at linked sites. These methods and public data will support the next phase of human genetic research.

- Sequenced Genomes of ~900 individuals
- From Seven Different Global Populations
- Found 250-300 Loss-Of-Function Mutations (KOs) Per Person
- $10^{-8}$  bp Mutations per Generation (30 per Genome)

# Pedigrees Can Be Used To Follow Disease Genes in Human Families



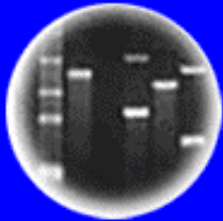
**Recessive Sex Linked**



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



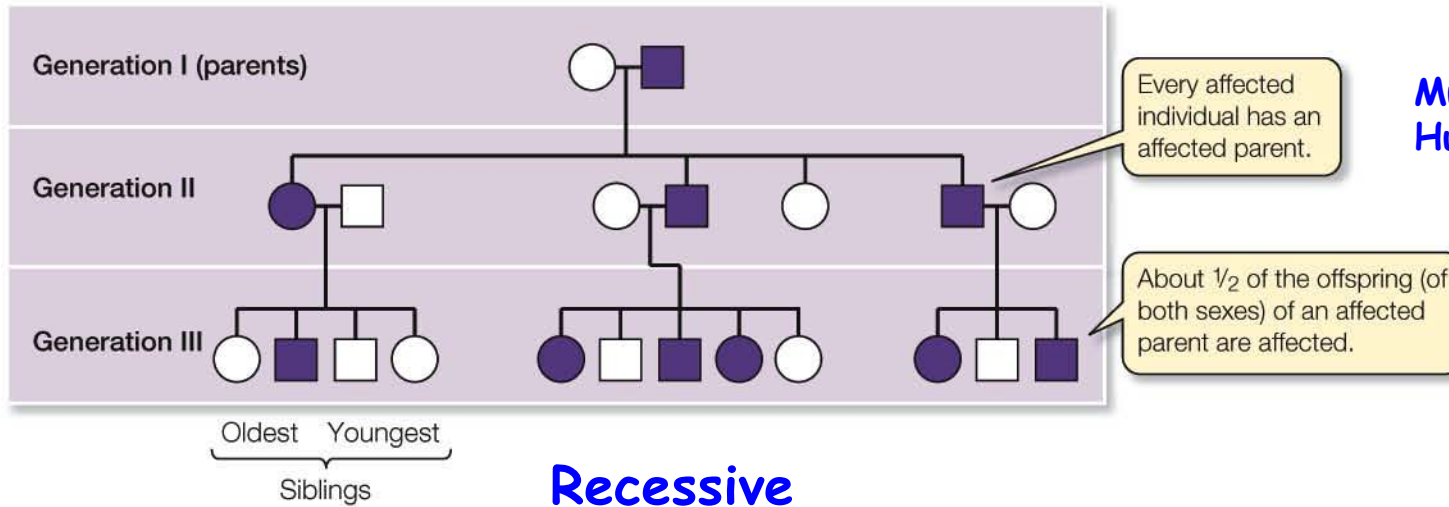
Plants of Tomorrow

**Pedigrees Can Be Used To Determine If  
a Trait is Dominant or Recessive**

**Each Type of Inheritance Predicts  
Specific Results in Each Generation**

# Dominant

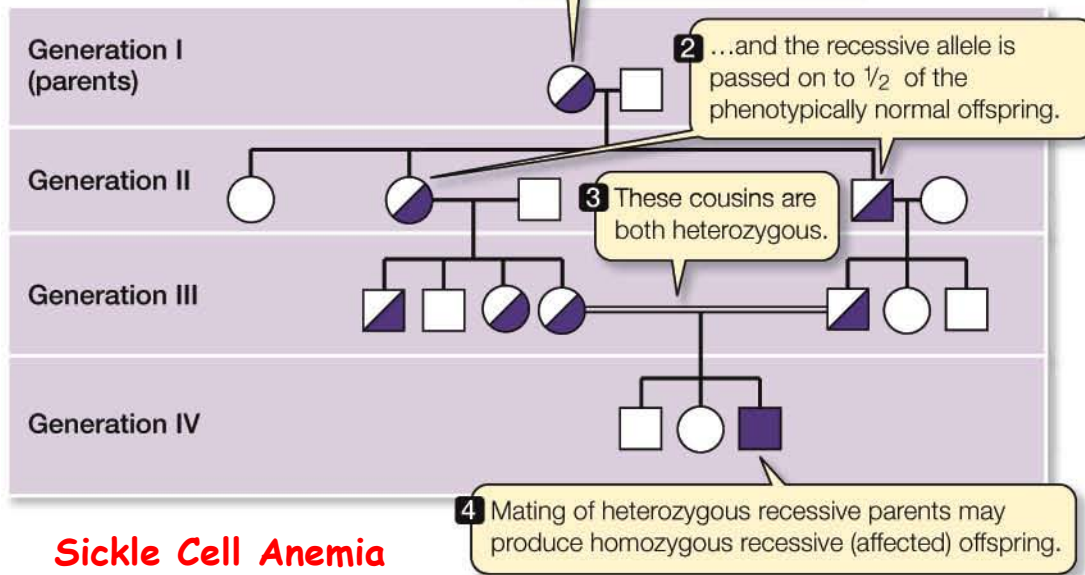
(A) Dominant inheritance



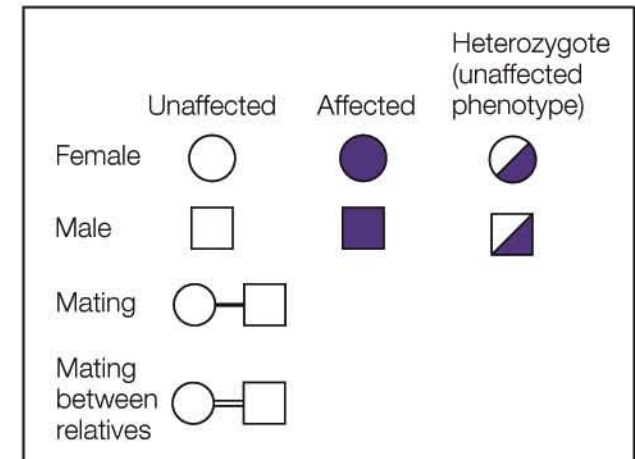
**Muscular Dystrophy**  
**Huntington Disease**

# Recessive

(B) Recessive inheritance



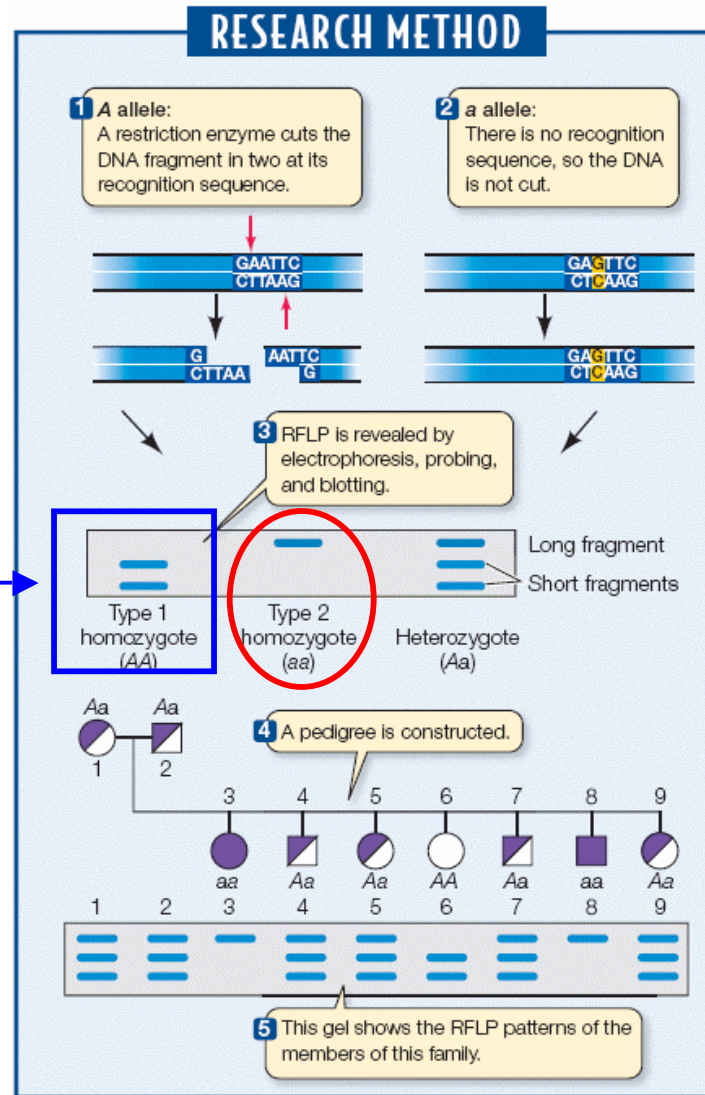
**Sickle Cell Anemia**  
**Cystic Fibrosis**  
**Tay-Sachs Disease**





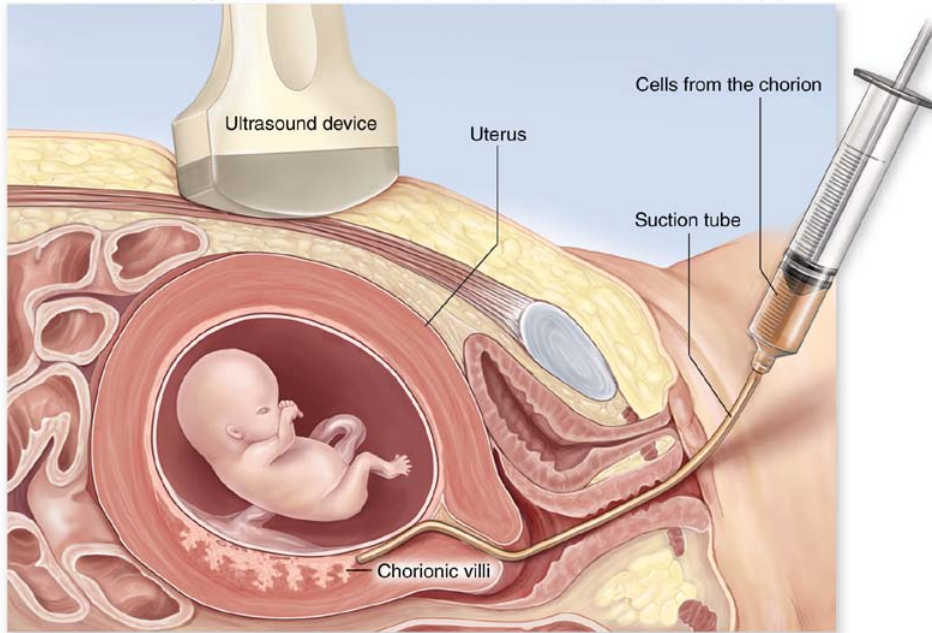
# Genetic Diseases Can Be Followed in Families Using Molecular Methods (e.g., DNA Blots or PCR)

DNA Fingerprints →

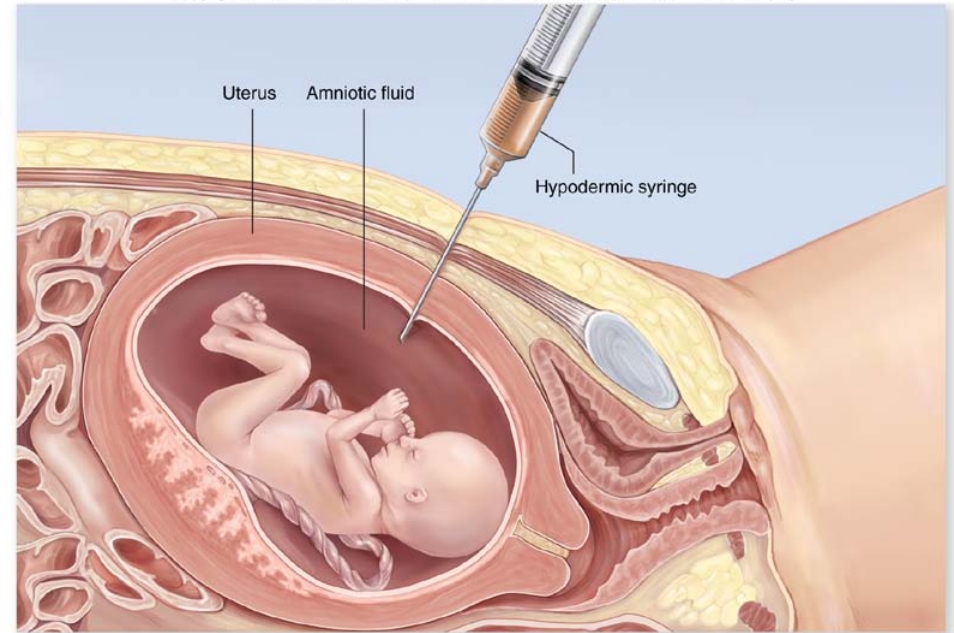


# PCR Can Be Used To Analyze Genes During Pregnancy

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

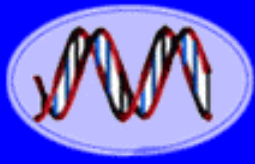


Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

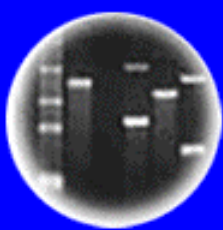
Disorder	Symptom	Defect	Dominant/Recessive	Frequency Among Human Births
Hemophilia	Blood fails to clot	Defective blood-clotting factor VIII	X-linked recessive	1/10,000 (Caucasian males)
Huntington disease	Brain tissue gradually deteriorates in middle age	Production of an inhibitor of brain cell metabolism	Dominant	1/24,000
Muscular dystrophy (Duchenne)	Muscles waste away	Degradation of myelin coating of nerves stimulating muscles	X-linked recessive	1/3700 (males)
Hypercholesterolemia	Excessive cholesterol levels in blood lead to heart disease	Abnormal form of cholesterol cell surface receptor	Dominant	1/500



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



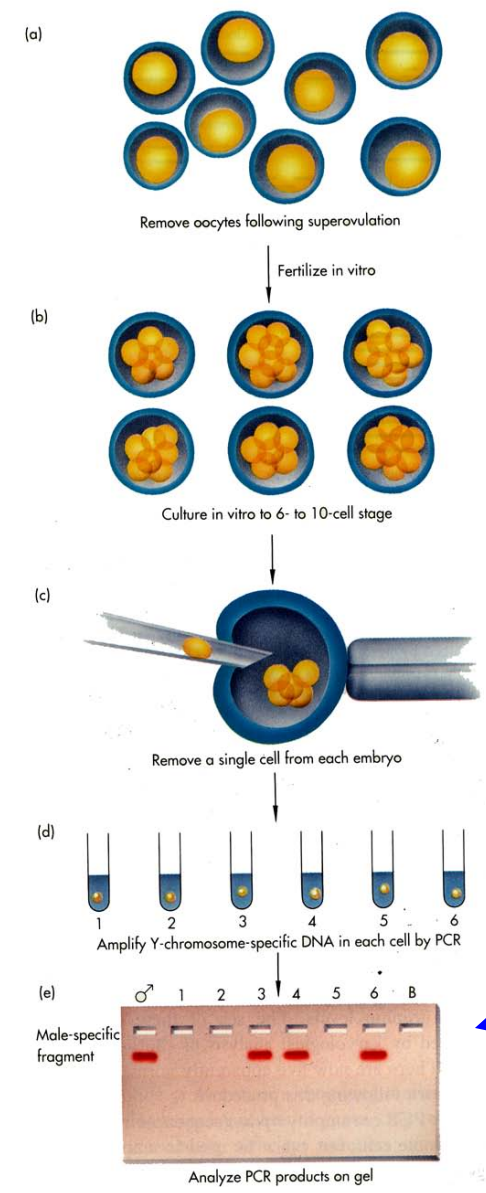
Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

# PCR Can Be Used To Analyze Mutant Gene in A Single Embryo Cell Before Pregnancy

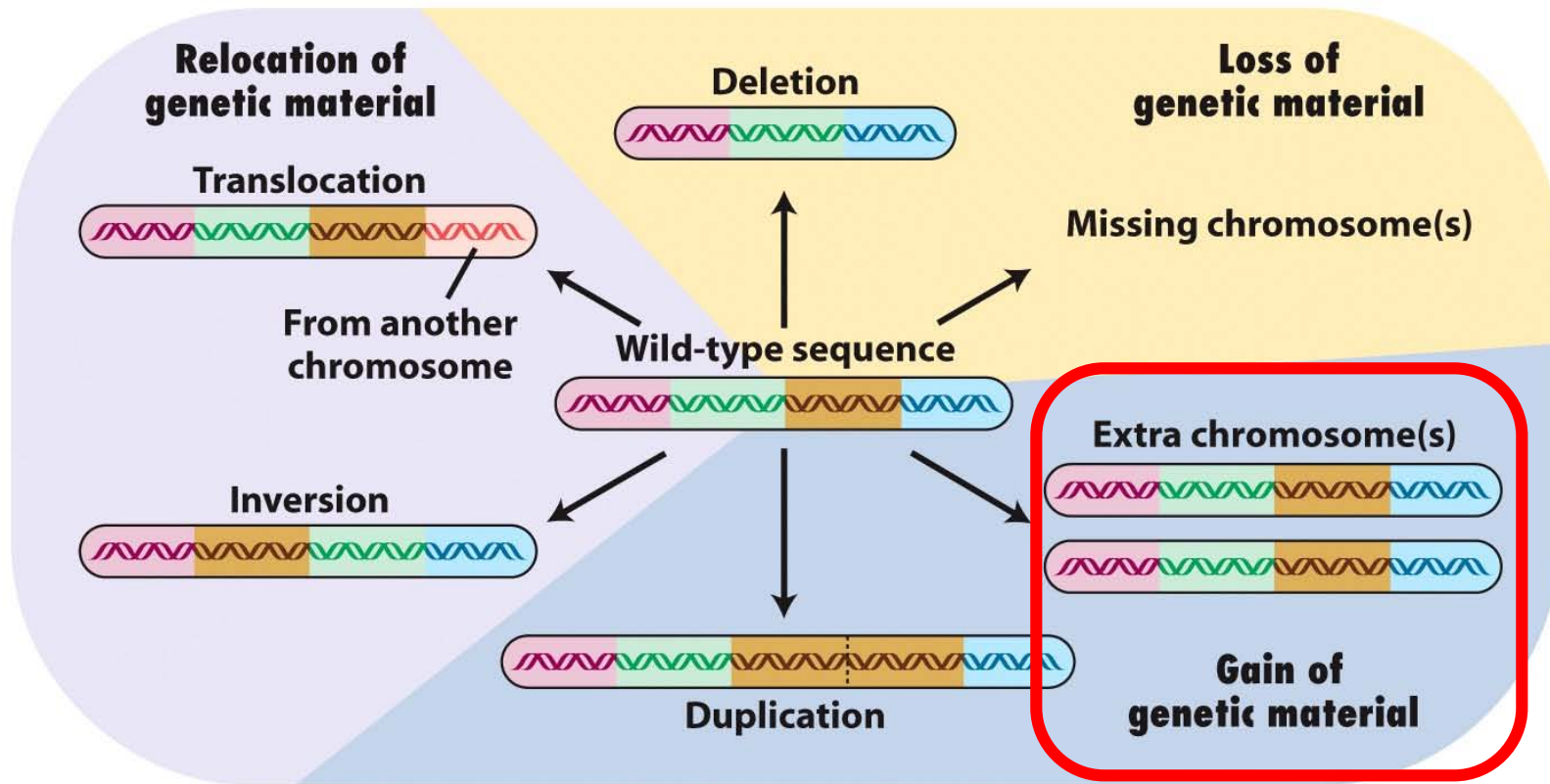
## PGD Pre- Implantation Genetic Diagnosis



**What is The Implication of This Procedure Considering That The Human Genome Has Been Sequenced?**

**Sex Determination in 8-cell Embryo!**

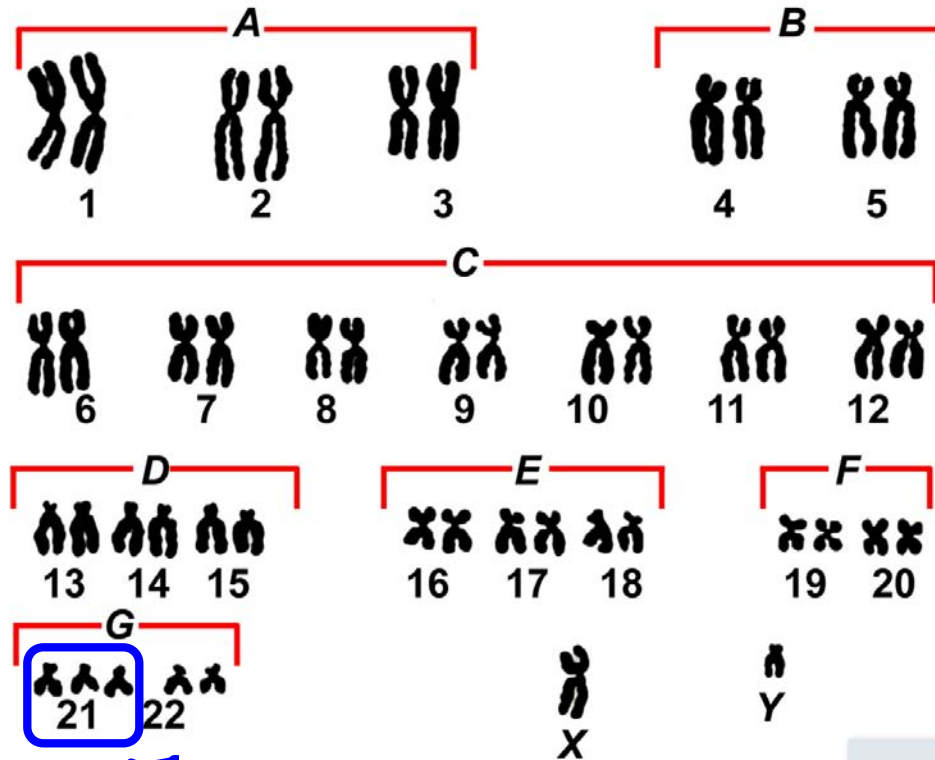
# "Mutations" Can Also Occur By Large Chromosomal Changes



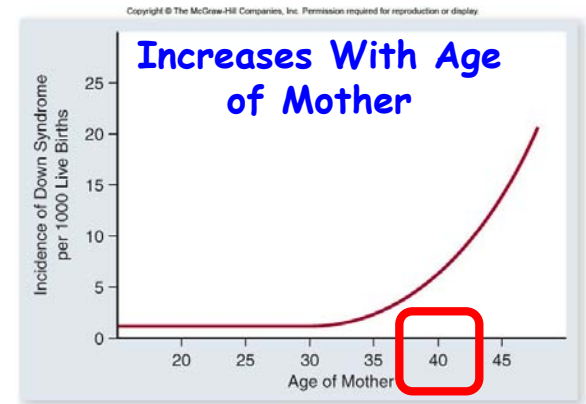
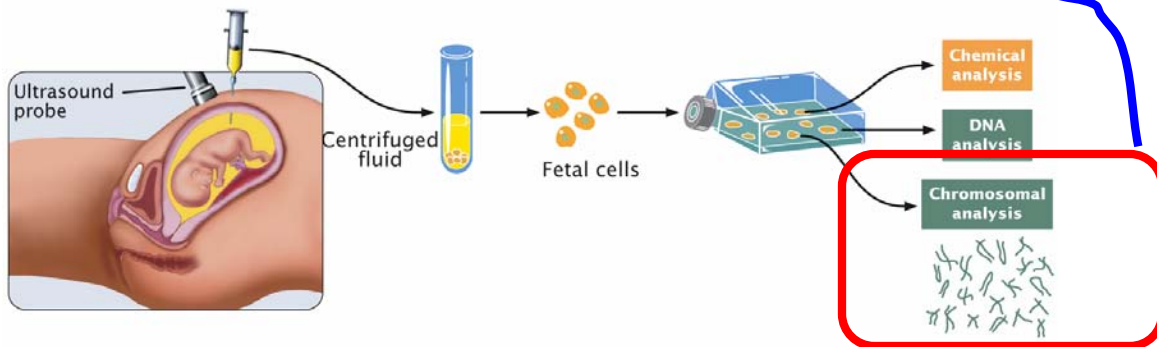
**These changes affect many genes!**

e.g. Down's Syndrome (3 Chromosome #21s)

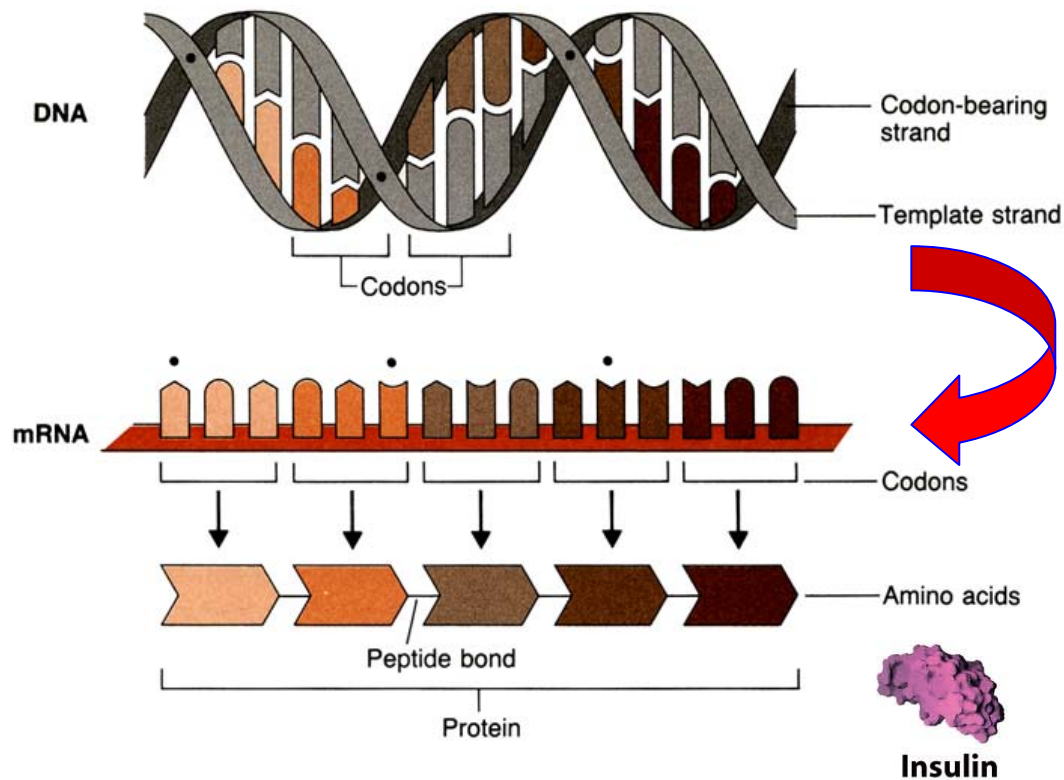
# A Down's Syndrome Karyotype



Three Chromosome # 21s



## ② How Does A Gene Lead To A Phenotype?



### ① mRNA Synthesized by Transcription

- Complementary to Transcribed, Non-Sense Strand
- Same Sequence As Sense Strand

### ② mRNA Translated into Protein by Translation of The Genetic Code

Genetic Code on mRNA Translated to Protein Sequence

∴ Sequence of Gene  
↓  
Sequence of mRNA  
↓  
Sequence of Protein

Know Sequence  
Know Protein

Engineer New Protein

# The Genetic Code is Universal!



DNA codons	Amino acid
GCA GCG GCT GCC	Ala
AGA AGG CGA CGG CGT CGC	Arg
GAT GAC	Asp
AAT AAC	Asn
TGT TGC	Cys
GAA GAG	Glu
CAA CAG	Gln
GGA GGG GGT GGC	Gly
CAT CAC	His
ATA ATT ATC	Ile
TTA TTG CTA CTG CTT CTC	Leu
AAA AAG	Lys
ATG	Met
TTT TTC	Phe
CCA CCG CCT CCC	Pro
AGT AGC TCA TCG TCT TCC	Ser
ACA ACG ACT ACC	Thr
TGG	Trp
TAT TAC	Tyr
GTA GTG GTT GTC	Val
TAA TAG TGA	Stop

Start

For RNA, The Ts are replaced by Us.

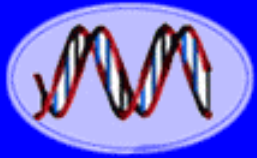
How Know?

1. Universal
2. Triplet
3. Punctuation
4. Degenerate

Know Sequence of Gene-Know Sequence of Protein  
Using Genetic Code

Big Implication For Genetic Engineering! Can Make Genes,  
Genomes & Specify Proteins Wanted! Can Express Genes  
From One Organism in Another!

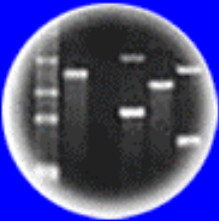
Design An Experiment to Show Code is Universal!



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting

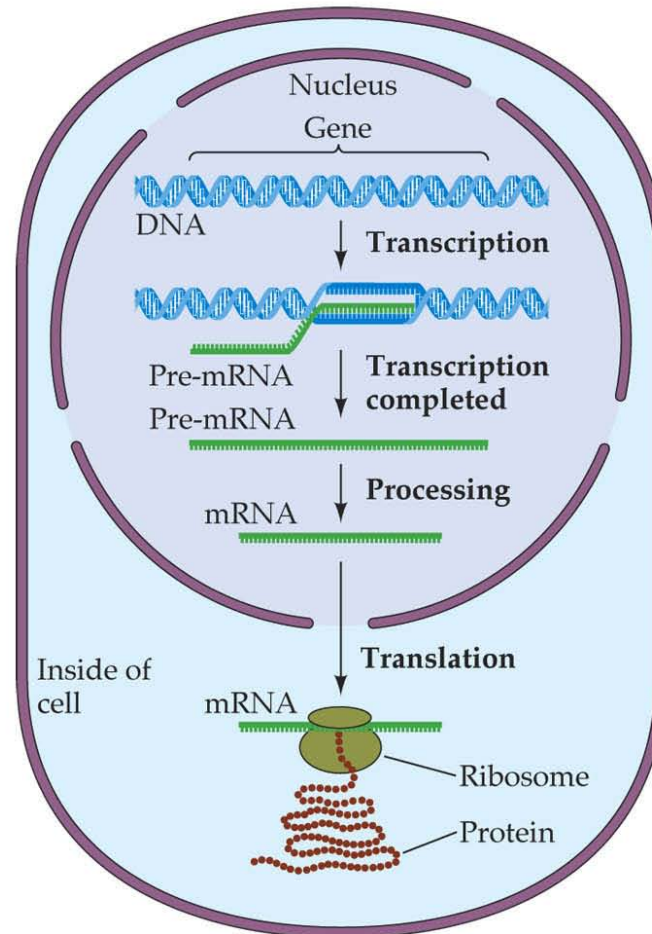


Cloning: Ethical Issues  
and Future Consequences



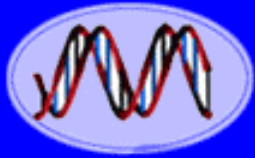
Plants of Tomorrow

# An Elaborate Cellular Machinery Requiring Thousands Of Genes is Required To Produce Proteins Encoded By Specific Genes!!



It takes Genes  
to Express  
(and Replicate)  
A GENE!!!

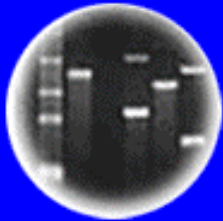




DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences

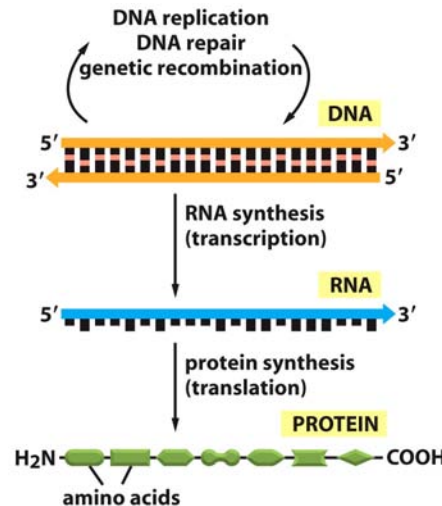


Plants of Tomorrow

# How Do Genes Work & What are Genes in Context of...



## Thinking About The Consequences of GMOs



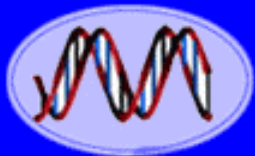
**Need Science-Based Questions & Science-Based Solutions-NOT OPINIONS!**

1. What is a Gene?
2. What is the Anatomy of a gene?
3. How Does the Gene Replicate?
4. How Does the Gene Direct Synthesis of a Protein?
5. Does the Gene Work Independently of other Genes?
6. What is the Sequence & Structure of the Protein?
7. How does it work in cell?
8. Does the Protein Structure imply any Potential "Harm"?
9. Does the Gene Change the organism? Fitness?

**There's NO HOCUS POCUS all hypothesis are testable!!**

**"Behind" All Traits!**

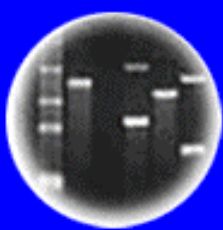
**Same Processes!**



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



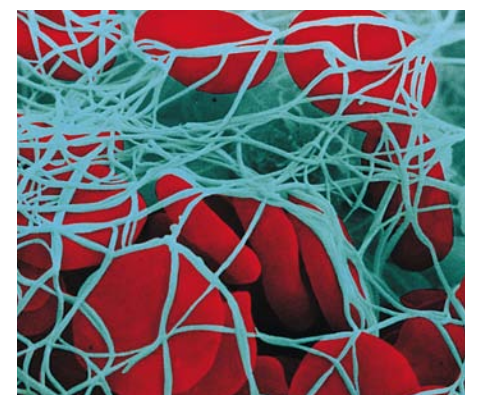
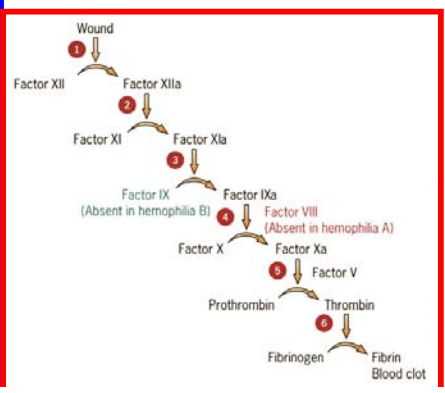
Cloning: Ethical Issues  
and Future Consequences



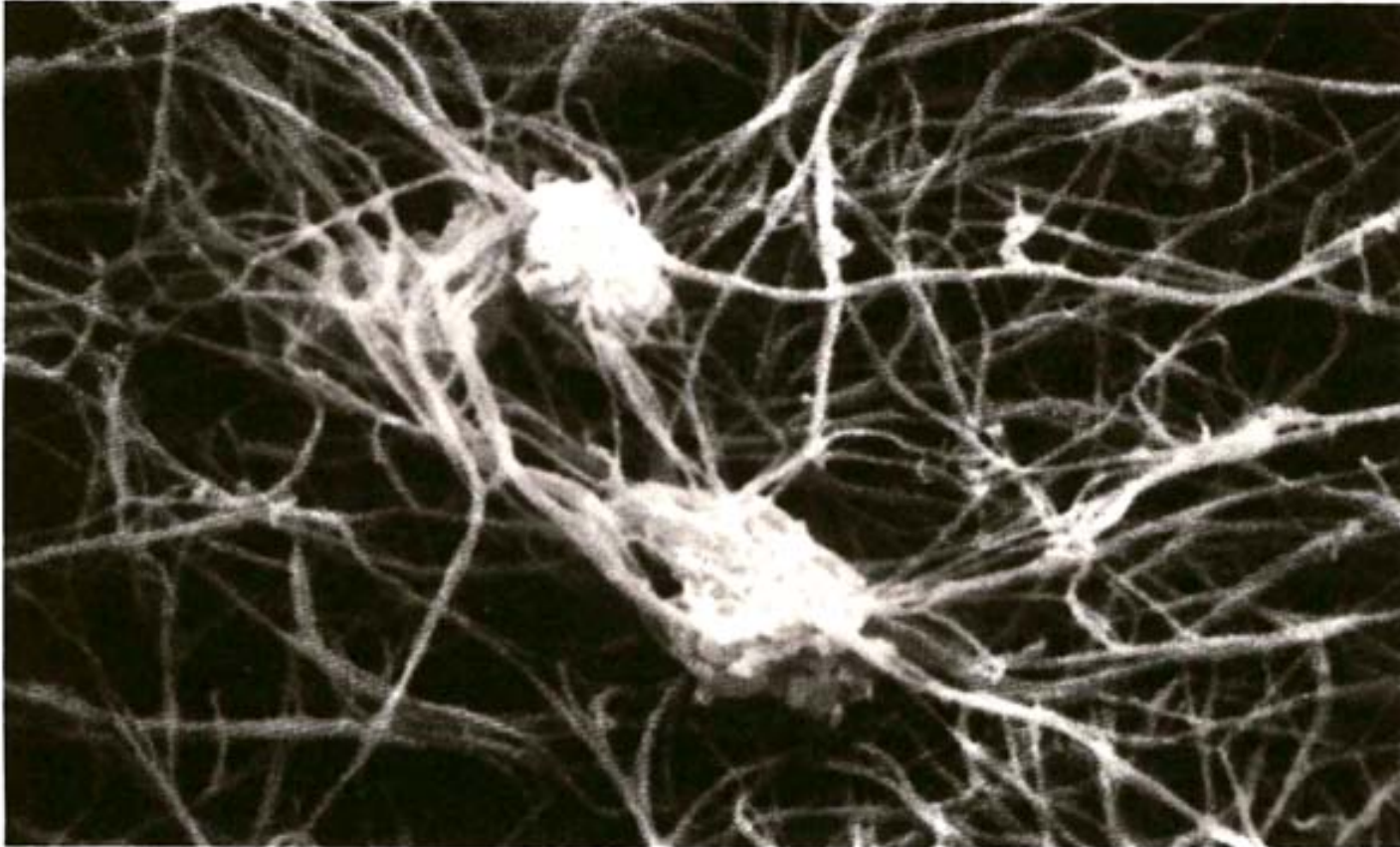
Plants of Tomorrow

# The Nuts & Bolts of Genetic Engineering

## The Factor VIII Story: From Gene To Genetic Screening To Drug



# The Molecular Genetics of Hemophilia (Potentially Lethal Disease)



FIBRIN STRANDS stabilize a blood clot at the site of a wound by trapping the platelets that form the bulk of the clot. The electron micrograph, which was made by Jon C. Lewis of Wake Forest University, shows a clot formed in a suspension of platelets and fibrin.

A clot in the bloodstream is the result of a complex cascade of enzymatic reactions culminating in the conversion of fibrinogen, a soluble protein, into insoluble fibrin strands. In hemophiliacs a crucial protein in the blood-clotting cascade is either missing or defective.

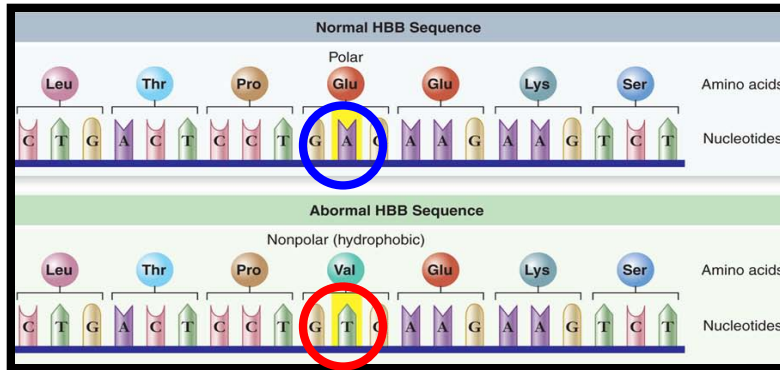
## A Case Study of Cloning Genes and mRNAs

Reference: Lawn & Vehar, *Sci. Amer.*, January, 1986

# Human Genetic Disorders Occur As A Result of Mutations: *Change Code-Alter Protein*



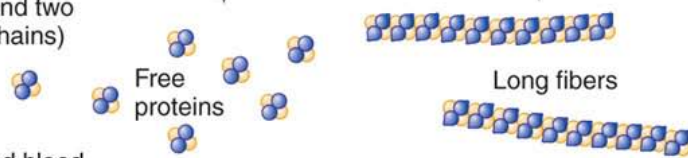
Chromosome 11



1. The polypeptide: the  $\beta$  chain of hemoglobin



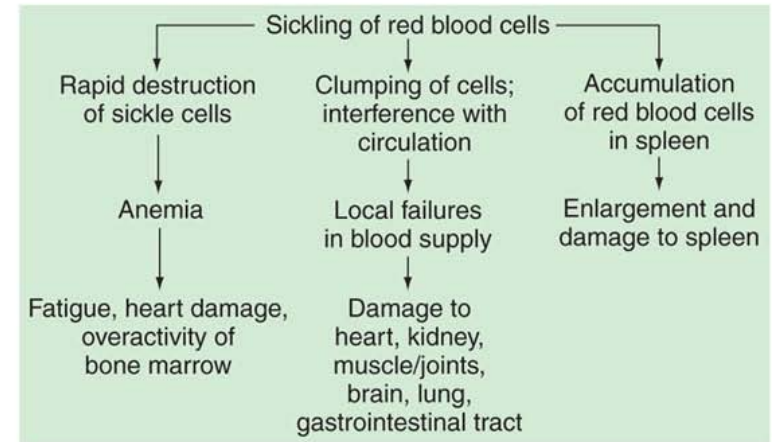
2. The protein: (made of two  $\alpha$  and two  $\beta$  chains)



3. Red blood cell making thousands of hemoglobin molecules



## (b) Sickle-cell anemia is pleiotropic



## (c) $\beta$ -chain substitutions/variants

	Amino-acid position															
	1	2	3	...	6	7	...	26	...	63	...	67	...	125	...	146
Normal (HbA)	Val	His	Leu	Glu	Glu	Glu	His	Val	Glu	His						
HbS	Val	His	Leu	Val	Glu	Glu	His	Val	Glu	His						
HbC	Val	His	Leu	Lys	Glu	Glu	His	Val	Glu	His						
HbG San Jose	Val	His	Leu	Glu	Gly	Glu	His	Val	Glu	His						
HbE	Val	His	Leu	Glu	Glu	Lys	His	Val	Glu	His						
HbM Saskatoon	Val	His	Leu	Glu	Glu	Glu	Tyr	Val	Glu	His						
Hb Zurich	Val	His	Leu	Glu	Glu	Glu	Arg	Val	Glu	His						
HbM Milwaukee 1	Val	His	Leu	Glu	Glu	Glu	His	Glu	Glu	His						
HbD $\beta$ Punjab	Val	His	Leu	Glu	Glu	Glu	His	Val	Gln	His						

## Sickle-Cell Anemia

# Hemophilia Has Been Known As An Inherited Disease For >2500 Years!

Old Testament-Circumcisions  
Royal Family-Europe



a = activated form



# First Reference to Hemophilia is in the Old Testament

## *Genesis 17:10-14*

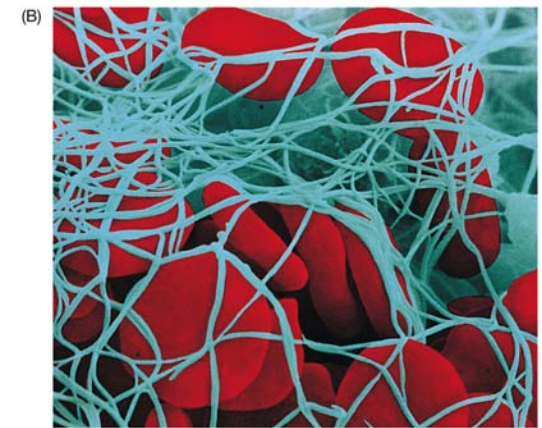
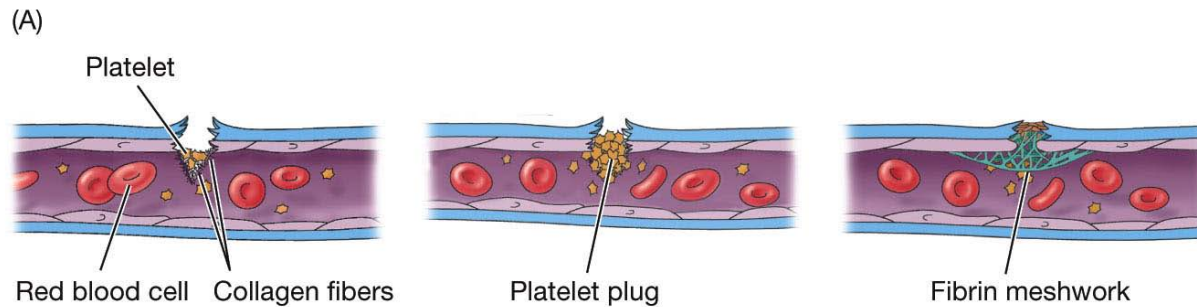
'This is My covenant that you shall keep between Me and you and your descendants after you: every male among you shall be circumcised. You shall circumcise the flesh of the foreskin.....At the age of eight days every male among you shall be circumcised throughout your generations.....an uncircumcised male...that soul shall be cut off from its people, he has invalidated My covenant.'



The Talmud also makes reference to families in whom children have died as a result of circumcision (Babylonian Talmud, Chapter Yevamoth p64b) [6]. Should a mother lose two children or should two sisters lose a child each after circumcision, subsequent children of the woman, the two sisters or of any other sisters of the same family should not be circumcised until they are older, or possibly not at all. This is thought to be the earliest reference to haemophilia; it was recognized in the Talmud that this condition was transmitted by the mother.

*Abraham was circumcised at 93 and gave birth to Isaac at 99.  
His wife - Sarah - was 90!*

# A Cascade Of Events After Wounding Leads to A Fibrin Clot



*LIFE 8e*, Figure 49.10 (Part 2)

*LIFE: THE SCIENCE OF BIOLOGY*, Eighth Edition © 2007 Sinauer Associates, Inc. and W. H. Freeman & Co.

Clotting factors:

1. Released from platelets and injured tissue
2. Plasma proteins synthesized in liver and circulated in inactive form

Prothrombin  
circulating  
in plasma

Thrombin

Fibrinogen  
circulating  
in plasma

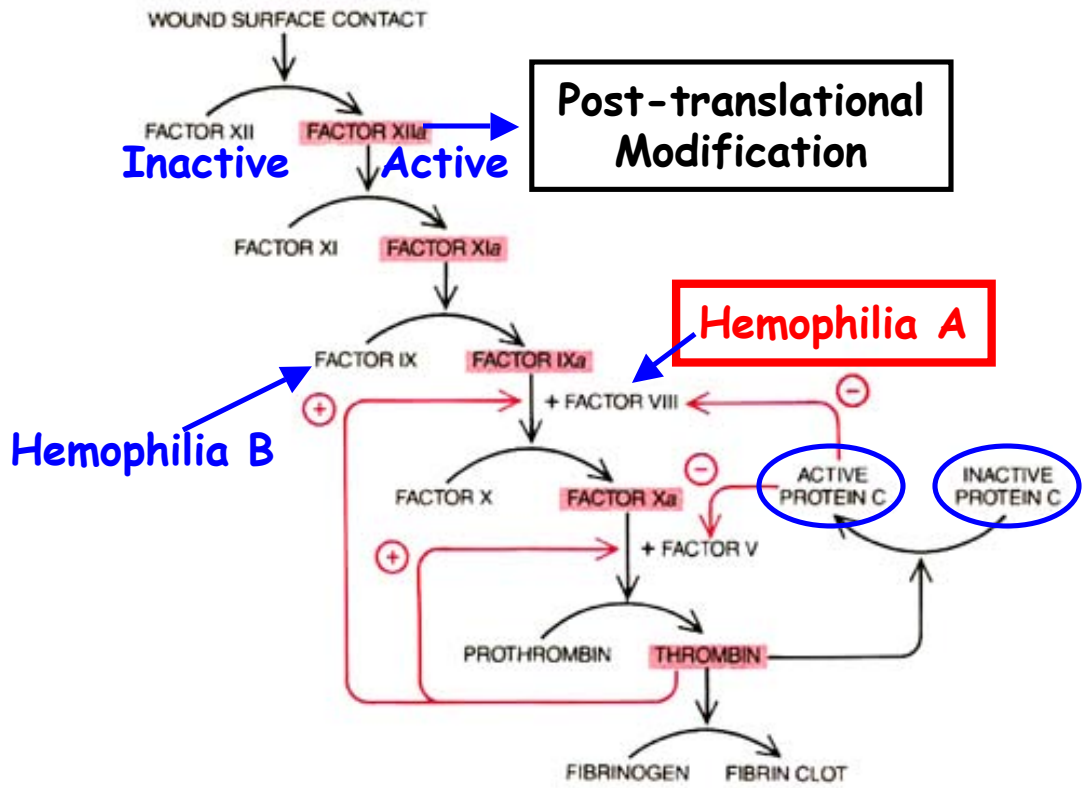
Fibrin

*LIFE 8e*, Figure 49.10 (Part 1)

*LIFE: THE SCIENCE OF BIOLOGY*, Eighth Edition © 2007 Sinauer Associates, Inc. and W. H. Freeman & Co.

**Clotting Factors Such As Factor VIII Play A Critical Role in This Process**

# How Does Blood Clot After Wounding?



**Eight Proteins/Genes Required:**

1. Factor VII
2. Factor XI
3. Factor IX
4. **Factor VIII**
5. Factor X
6. **Protein C**
7. Prothrombin
8. Fibrinogen

**CLOTTING CASCADE** begins when cell damage at a wound somehow activates the enzyme factor XII; it ends with the conversion of fibrinogen into fibrin by thrombin. At each step an inactive protein is converted into a protease, or protein-cutting enzyme (color), which activates the next protein. Some steps require cofactors such as factors VIII and V. The cascade includes positive- and negative-feedback loops (colored arrows). Thrombin activates factors VIII and V; it also deactivates them (by activating protein C), which helps to halt clotting. Some 85 percent of hemophiliacs lack factor VIII. The rest lack factor IX.

ATryn® 2009

**Anti-Thrombin??**

**Cascade**

→ **Anti-Thrombin Deficiency (At-III) genetic disease**

**What Happens If Any Of These Proteins Or Genes Are Mutated?**

↓  
**No Blood Clot!**



# Hemophiliacs Have Mutations In Either Factor VIII or Factor IX Genes

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

TABLE 13.2		Some Important Genetic Disorders		
Disorder	Symptom	Defect	Dominant/ Recessive	Frequency Among Human Births
Cystic fibrosis	Mucus clogs lungs, liver, and pancreas	Failure of chloride ion transport mechanism	Recessive	1/2500 (Caucasians)
Sickle cell anemia	Blood circulation is poor	Abnormal hemoglobin molecules	Recessive	1/600 (African Americans)
Tay-Sachs disease	Central nervous system deteriorates in infancy	Defective enzyme (hexosaminidase A)	Recessive	1/3500 (Ashkenazi Jews)
Phenylketonuria	Brain fails to develop in infancy	Defective enzyme (phenylalanine hydroxylase)	Recessive	1/12,000
Hemophilia	Blood fails to clot	Defective blood-clotting factor VIII	X-linked recessive	1/10,000 (Caucasian males)
Huntington disease	Brain tissue gradually deteriorates in middle age	Production of an inhibitor of brain cell metabolism	Dominant	1/24,000
Muscular dystrophy (Duchenne)	Muscles waste away	Degradation of myelin coating of nerves stimulating muscles	X-linked recessive	1/3700 (males)
Hypercholesterolemia	Excessive cholesterol levels in blood lead to heart disease	Abnormal form of cholesterol cell surface receptor	Dominant	1/500

<b>Hemophilia A</b>	<b>Defective Factor VIII Gene</b>	<b>1/10,000 males</b>
<b>Hemophilia B</b>	<b>Defective Factor IX Gene</b>	<b>1/30,000 males</b>

**Hypothesis For High Frequency in Males?**

**Both Genes On X-Chromosome ♀ → ♂'s**

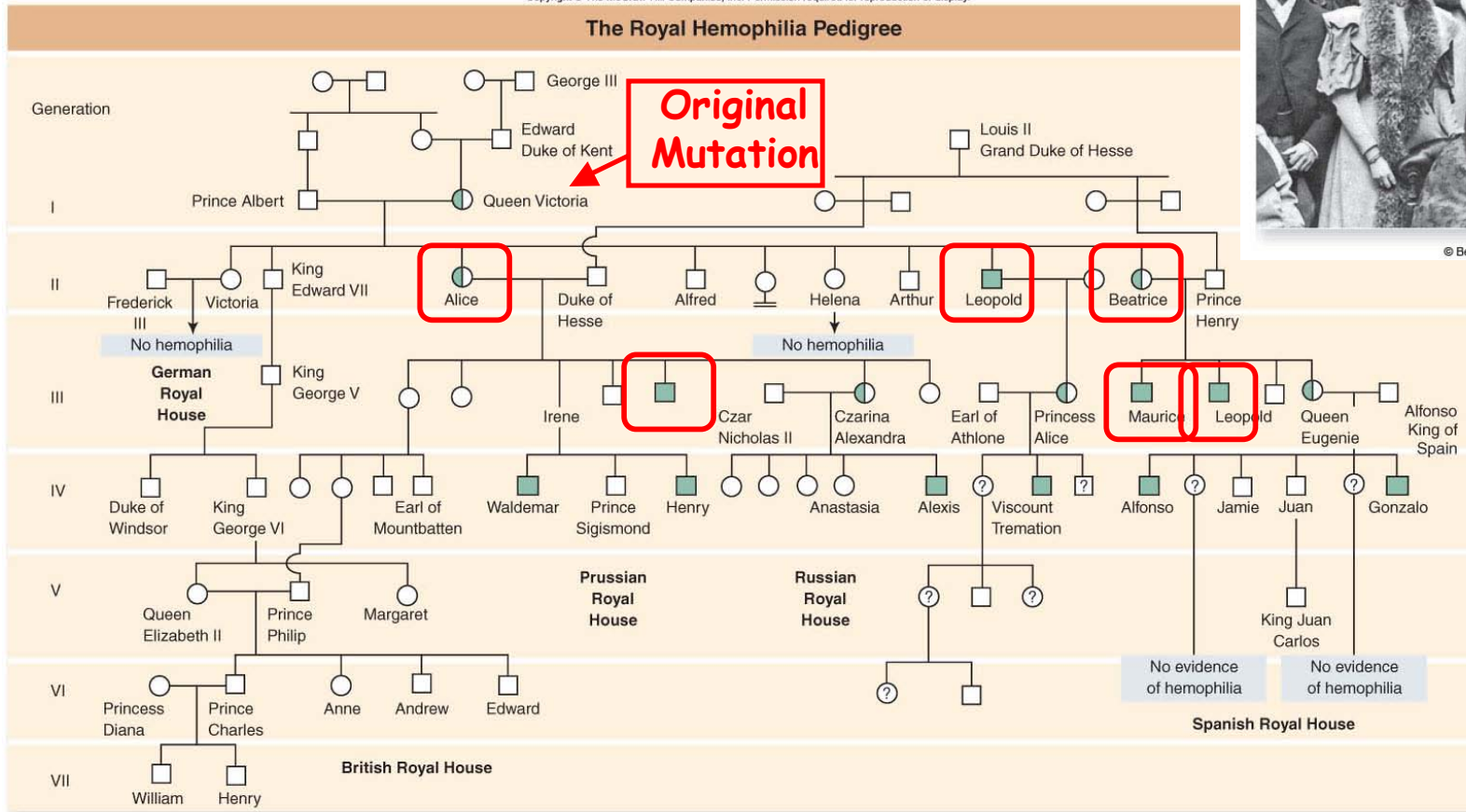
# Hemophilia A and B Genes (Traits) Are Sex Linked

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



© Bettmann/Corbis

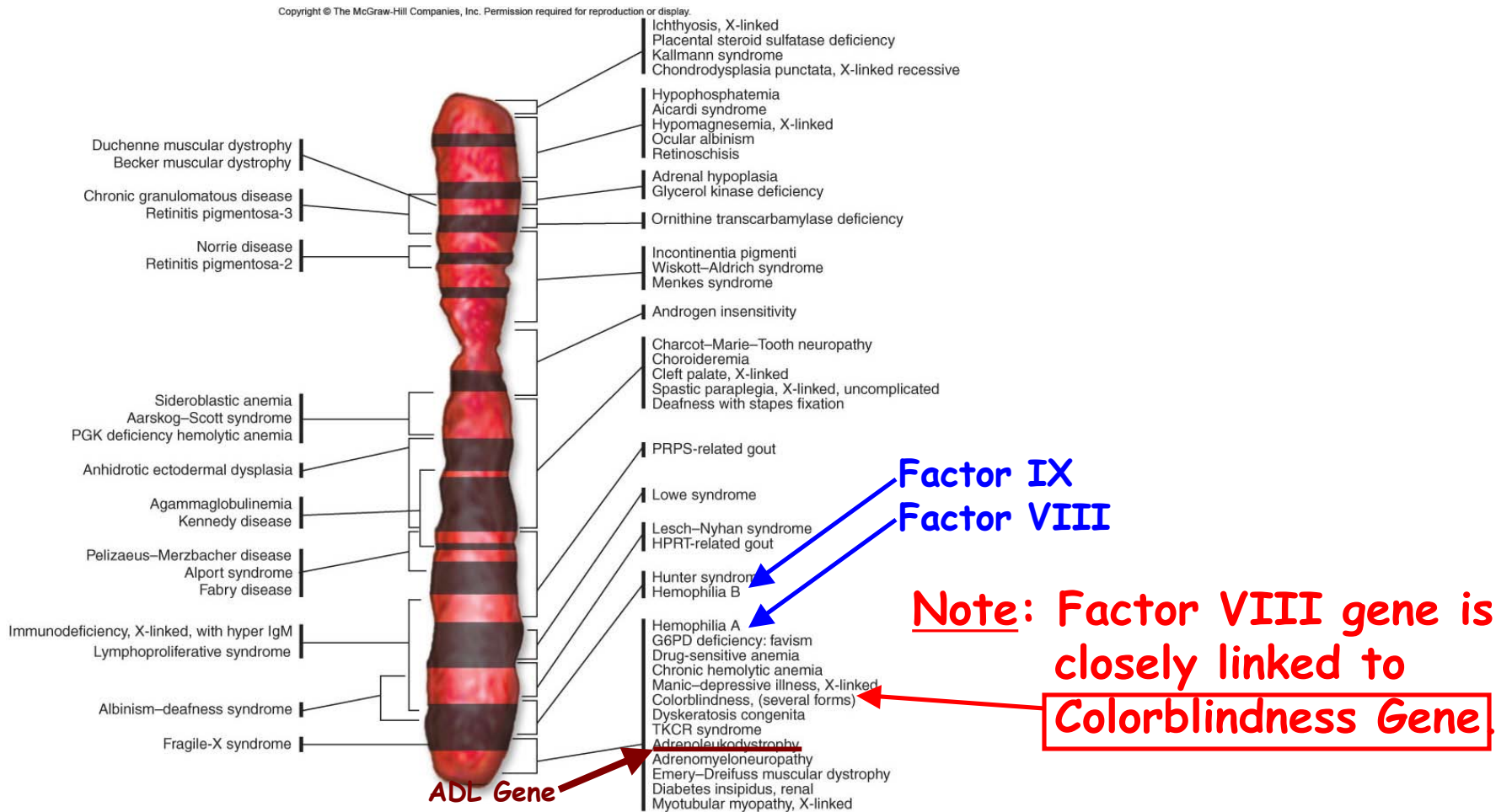
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



**Note: 1. Males Obtain Defective Gene From Mothers**

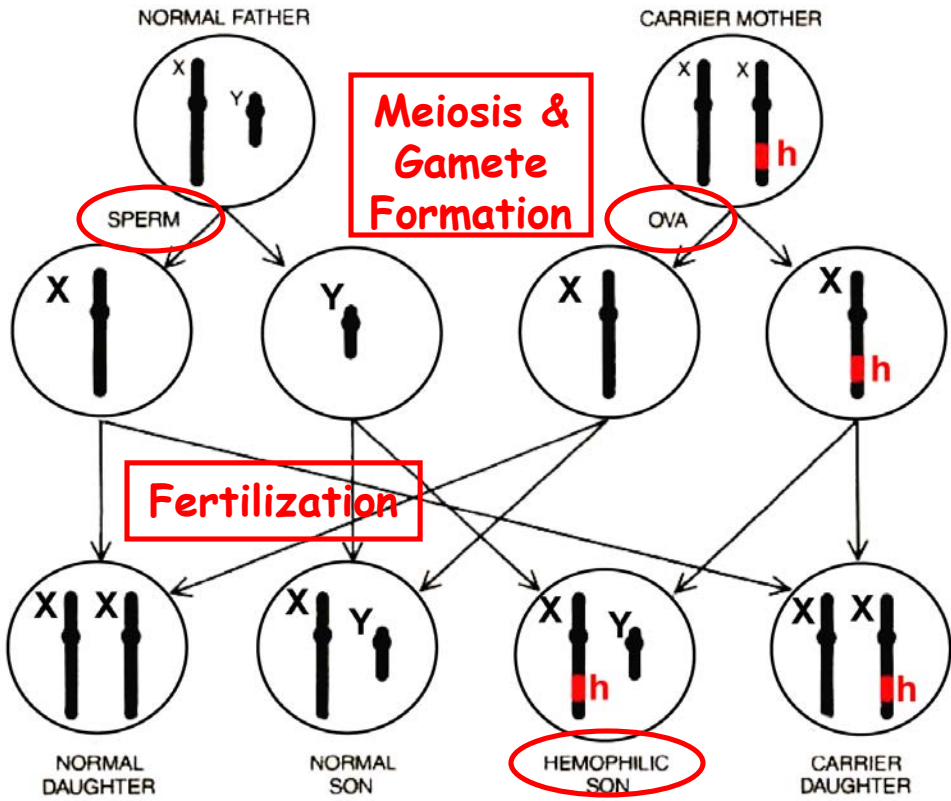
**2. 50% of Sons Of A Maternal Carrier Have The Defective Gene**

# Factor VIII and Factor IX Genes are Closely Linked on the X Chromosome

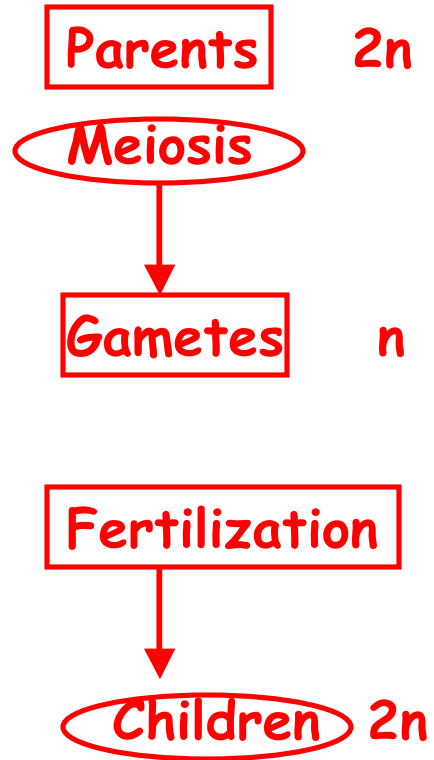


The X chromosome has ~1500 Genes (2008) and 150,000,000 bp (150 Mb)

# Hemophilia A and B Inheritance



**SEX-LINKED INHERITANCE** of hemophilia results from the location of the factor VIII gene on the X chromosome. A male carrying a mutant factor VIII gene lacks normal factor VIII and is hemophilic. A female carrier is protected by the normal gene on her second X chromosome, but half of her daughters will be carriers and half of her sons will be hemophilic. In the case of a hemophilic father (not shown), his sons will not be hemophilic, because they receive his Y (not his X) chromosome, but his daughters will be carriers.



## Sex-Linked Inheritance

♀ Carriers → 1/2 Sons + No Daughters!

Only One X-Chromosome is ♂

## What Was Known About Factor VIII *Before Gene Cloned?*

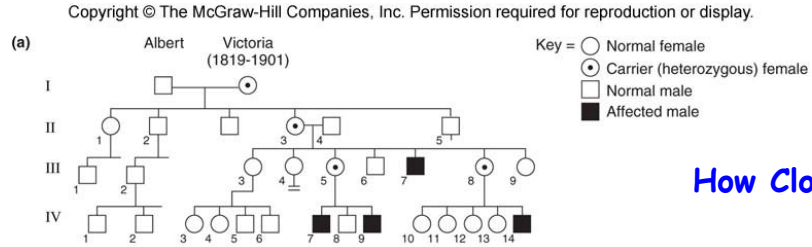
1. Blood Protein (But Perhaps Synthesized Elsewhere!)
2. Could be purified in small amounts from >20 Liters of human blood + cow blood + pig blood
3. Short Stretch of Proteins Sequenced = Known Protein Sequence!
4. Hemophilia A could be treated by blood transfusions from normal individuals, ∴ clotting factor in blood.

∴ How to go From Protein to Gene

# The Problem

# For Factor VIII- Not Known Where Gene is Expressed ∴ **Must Use Genome Library**

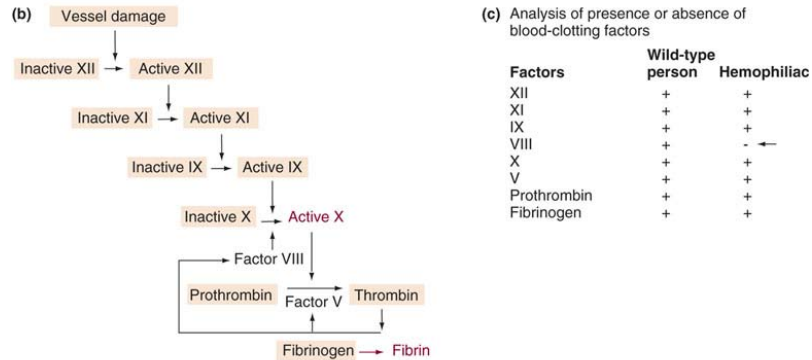
Early 1980's



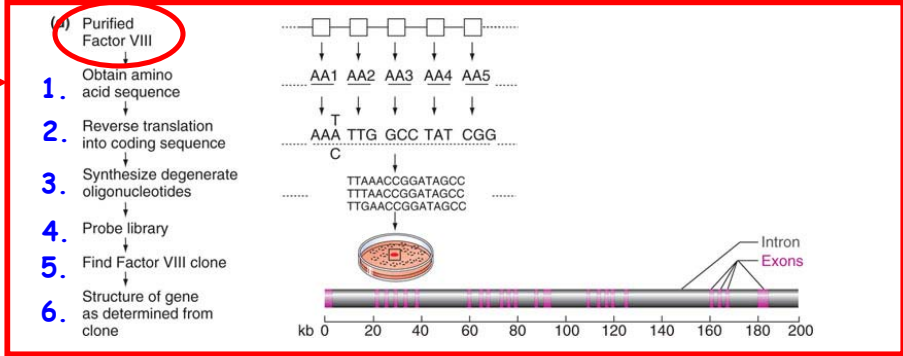
Key Concept



How Clone A Gene When You Don't Know Where it is Expressed !



**Key:**  
**Protein Sequence Known**



## How Find Gene & cDNA?

Protein → Gene → mRNA → Drug !

# Knowledge of the Protein Sequence and the Genetic Code Makes it Possible to Identify a Gene

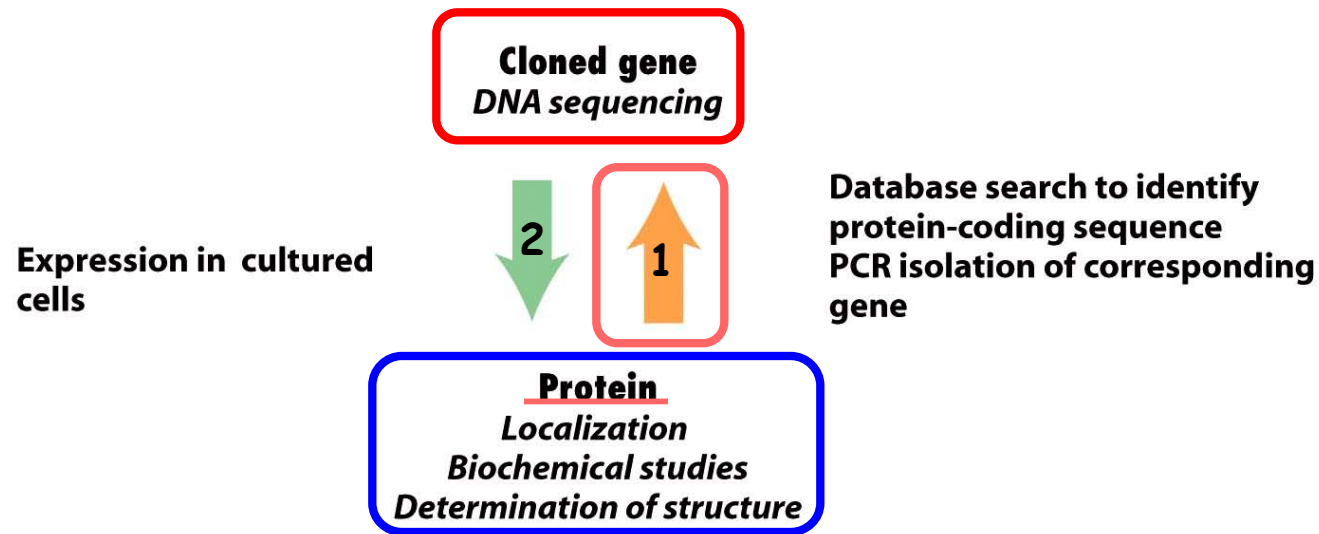


Figure 5-1  
*Molecular Cell Biology, Sixth Edition*  
© 2008 W.H. Freeman and Company

∴ 1. Protein → Gene → Drug

or

Factor VIII Strategy (1985)

Genomics

2. Gene → Protein Using Sequencing and Genetic Code

GenBank

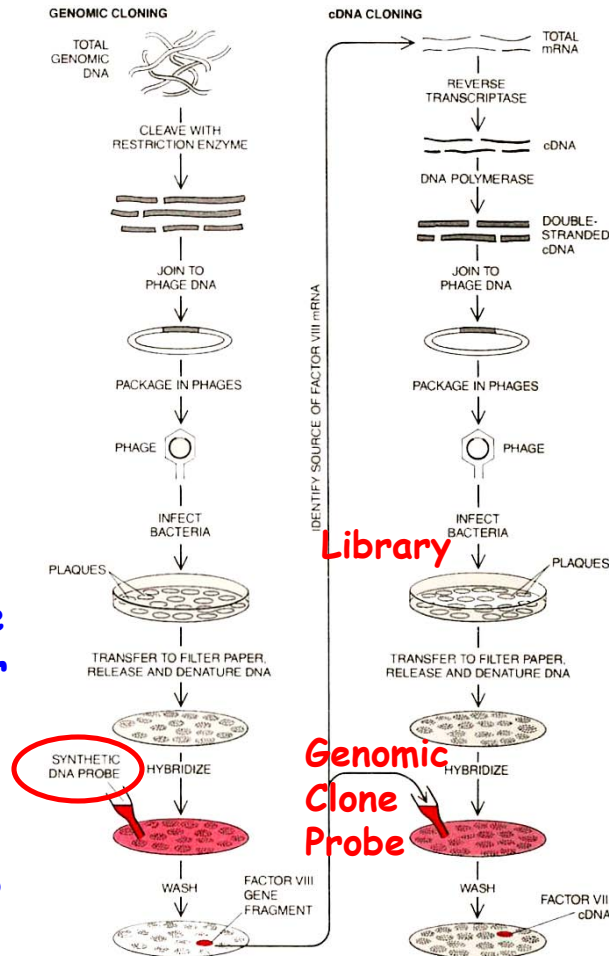
2011

3. Just Sequence Everything + Identify Protein- GenBank Huge

# Steps Required to Clone Factor VIII Gene and cDNA

## Gene

1. Make Genome Library Because Factor VIII Gene in Genome!
2. Purify Protein from Blood- that's where it works (wasn't known where made)
3. Reverse Translate using the genetic code a portion of the protein sequence
4. Synthesize a DNA probe complementary to Factor VIII gene corresponding to protein sequence
5. Screen Genome Library Entire Gene on The Clone?



## cDNA

1. Use Gene probe to screen cDNA library for Factor VIII cDNA clone
2. How know what mRNA to use to make cDNA library?
3. Use gene probe to probe RNA blots containing mRNA from all major organs (liver, kidney, blood, etc.)
4. Find Factor VIII mRNA in liver- male, liver- secrete into blood

Why Need cDNA?  
Story continued

Want cDNA to Manufacture Factor VIII as a Drug to Treat Hemophilia A!



## Step One

How to Construct a Human Genome Library to Find the Factor VIII Gene?

If It is Not Known Where Gene is Active  
Can "Look" to Genome Instead of mRNA to  
Find + Clone Gene!

# Vectors Used in Genetic Engineering Have Similar Conceptual Properties But are Used in Different Situations

**Table 3.2 A COMPARISON OF DNA VECTORS AND THEIR APPLICATIONS**

Vector Type	Maximum Insert Size (kb)	Applications	Limitations
Bacterial plasmid vectors (circular)	~6-12	DNA cloning, protein expression, subcloning, direct sequencing of insert	Restricted insert size; limited expression of proteins; copy number problems; replication restricted to bacteria
DNA		DNA	
Bacteriophage vectors (linear)	~25	cDNA, genomic and expression libraries	Packaging limits DNA insert size; host replication problems
Cosmid (circular)	~35	cDNA and genomic libraries, cloning large DNA fragments	Phage packaging restrictions; not ideal for protein expression; cannot be replicated in mammalian cells
Bacterial artificial chromosome (BAC, circular)	~300	Genomic libraries, cloning large DNA fragments	Replication restricted to bacteria; cannot be used for protein expression
Yeast artificial chromosome (YAC, circular)	200-2,000	Genomic libraries, cloning large DNA fragments	Must be grown in yeast; cannot be used in bacteria
Ti vector (circular)	Varies depending on type of Ti vector used	Gene transfer in plants	Limited to use in plant cells only; number of restriction sites randomly distributed; large size of vector not easily manipulated

Copyright © 2009 Pearson Education, Inc.

## Plasmids vs. Bacteriophage Vectors

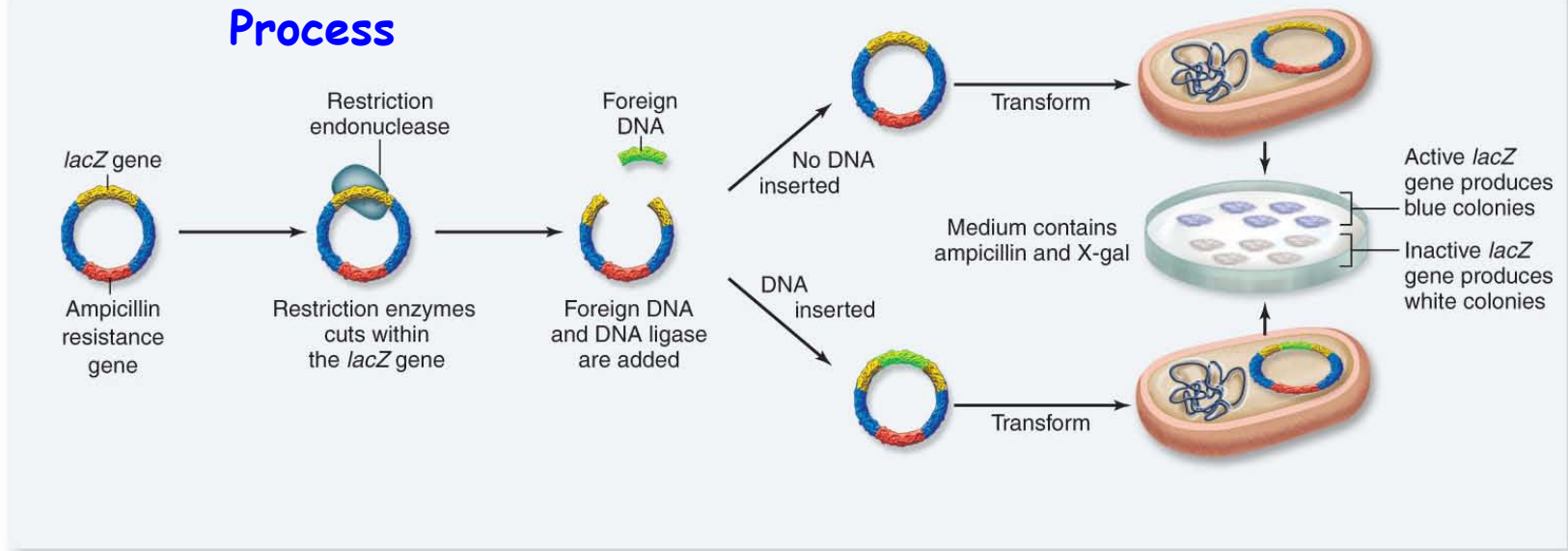
1. Replicate
2. Selectable
3. Can be used to insert foreign genes/restriction sites
4. Easily isolated + transferred back to cells

# Plasmid vs. Bacteriophage Vectors for Cloning DNA Fragments

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

## "Artificial" Transformation Process

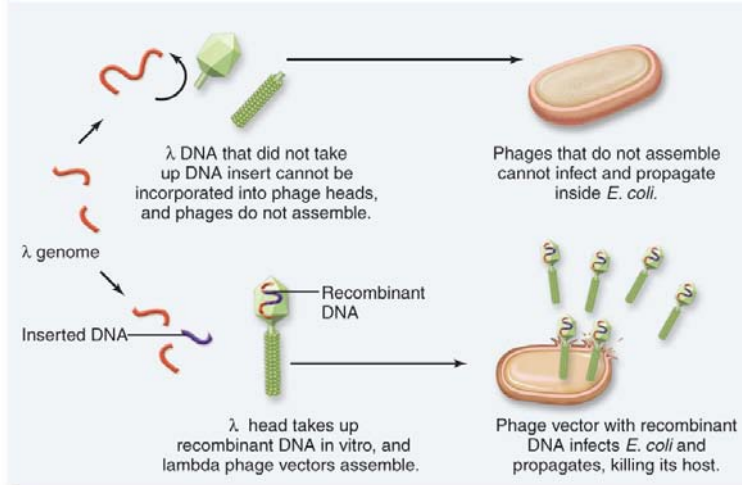
### A Plasmid Vector



a.

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

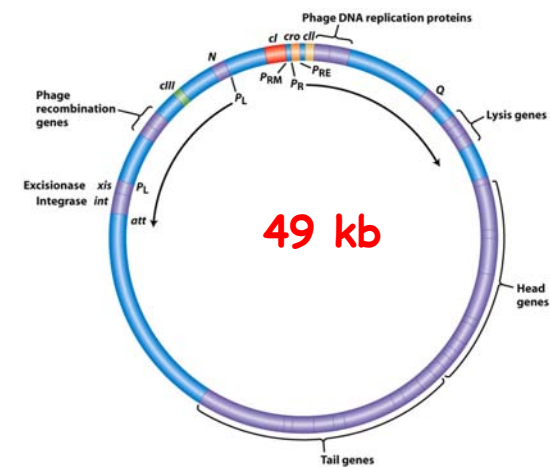
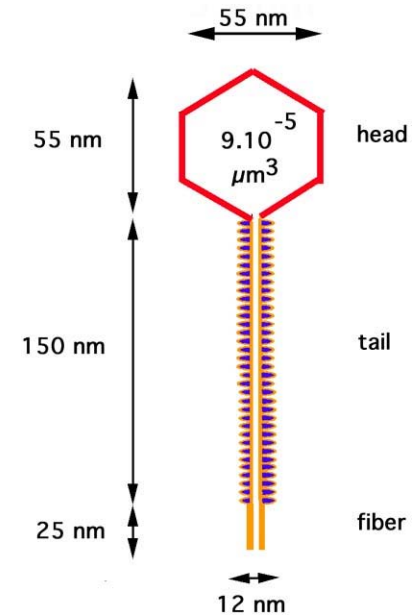
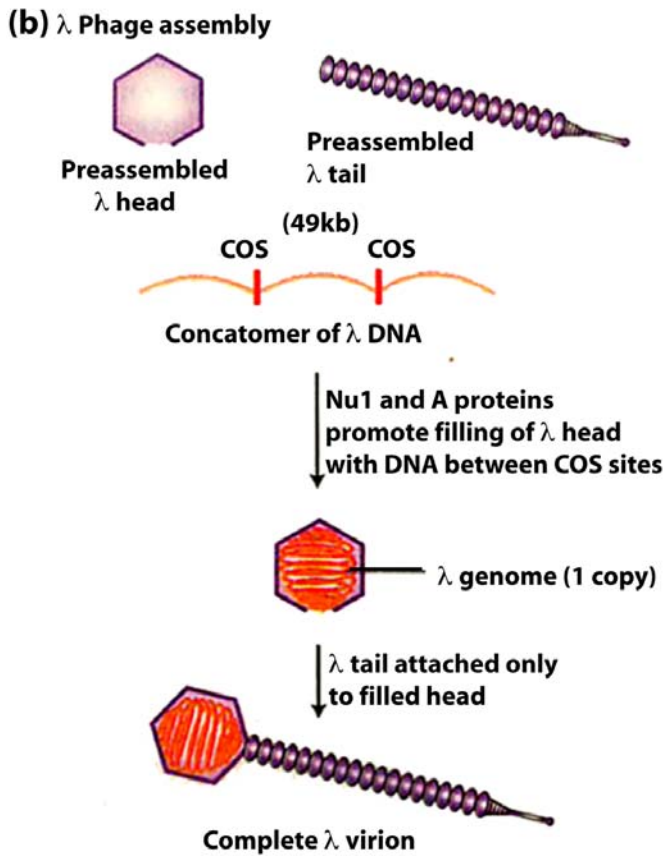
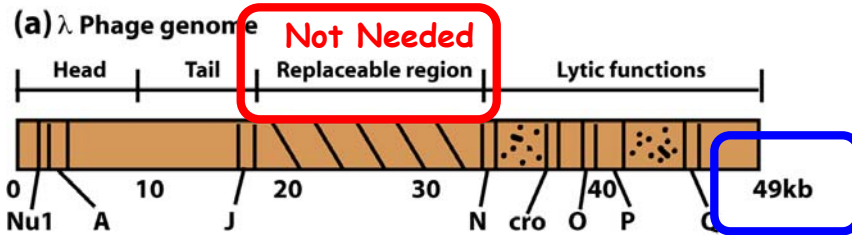
### A Phage Vector



b.

**Natural  
Infection Process**

# Structure of the $\lambda$ Phage and Its Genome

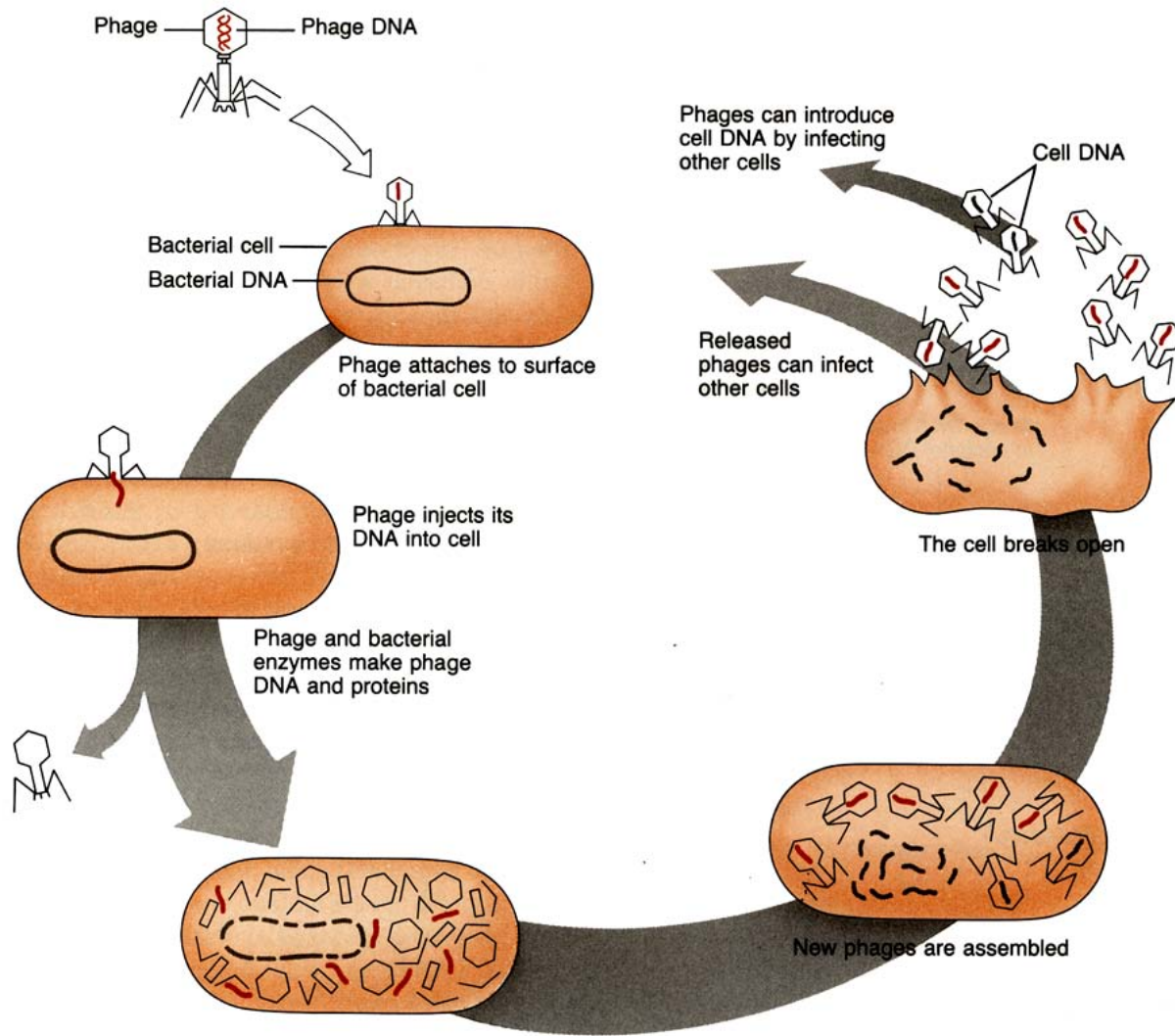


One of First Genome Sequences

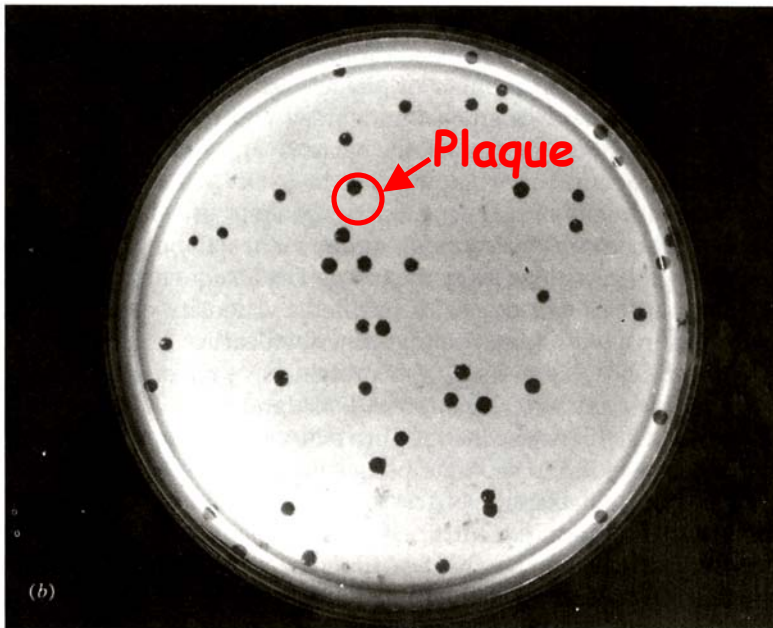
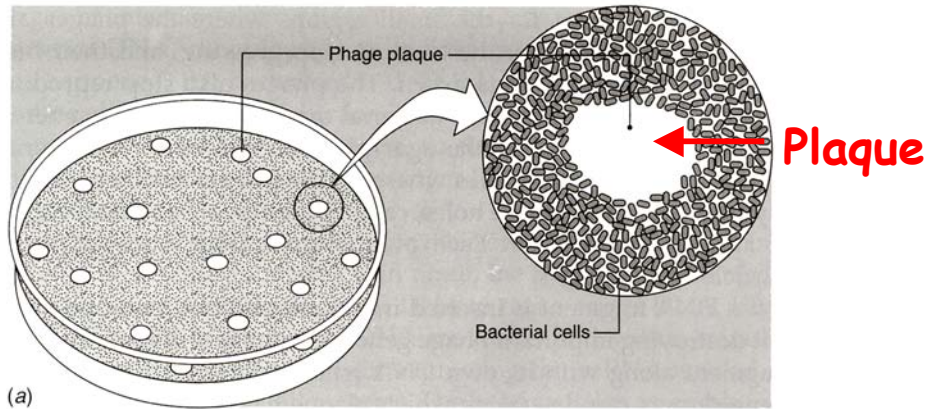
How Can the Genome "fit" into the Head?

Figure 10-26  
© 2004 Garland Science, an imprint of Macmillan, a division of The McGraw-Hill Companies

# $\lambda$ Phage Infects E.coli & Destroys (Lyses) cells



# Lysed Cells Can Be Seen as Clear Plaques on Agar Plates



1. Each Plaque is a  
Virus Clone  
Representing One  
Viral Infection!

2. Selectable Marker is  
Bacterial Cell Destruction  
& Plaque Formation

## Advantages of $\lambda$ Virus as a Vector for Cloning DNA

1. Long DNA Segments can be Cloned (~20kb) Need fewer clones for whole Genome!

2. Can clone DNA Segments in Viral Genome & Self-Assemble with viral proteins into virus in a test tube!

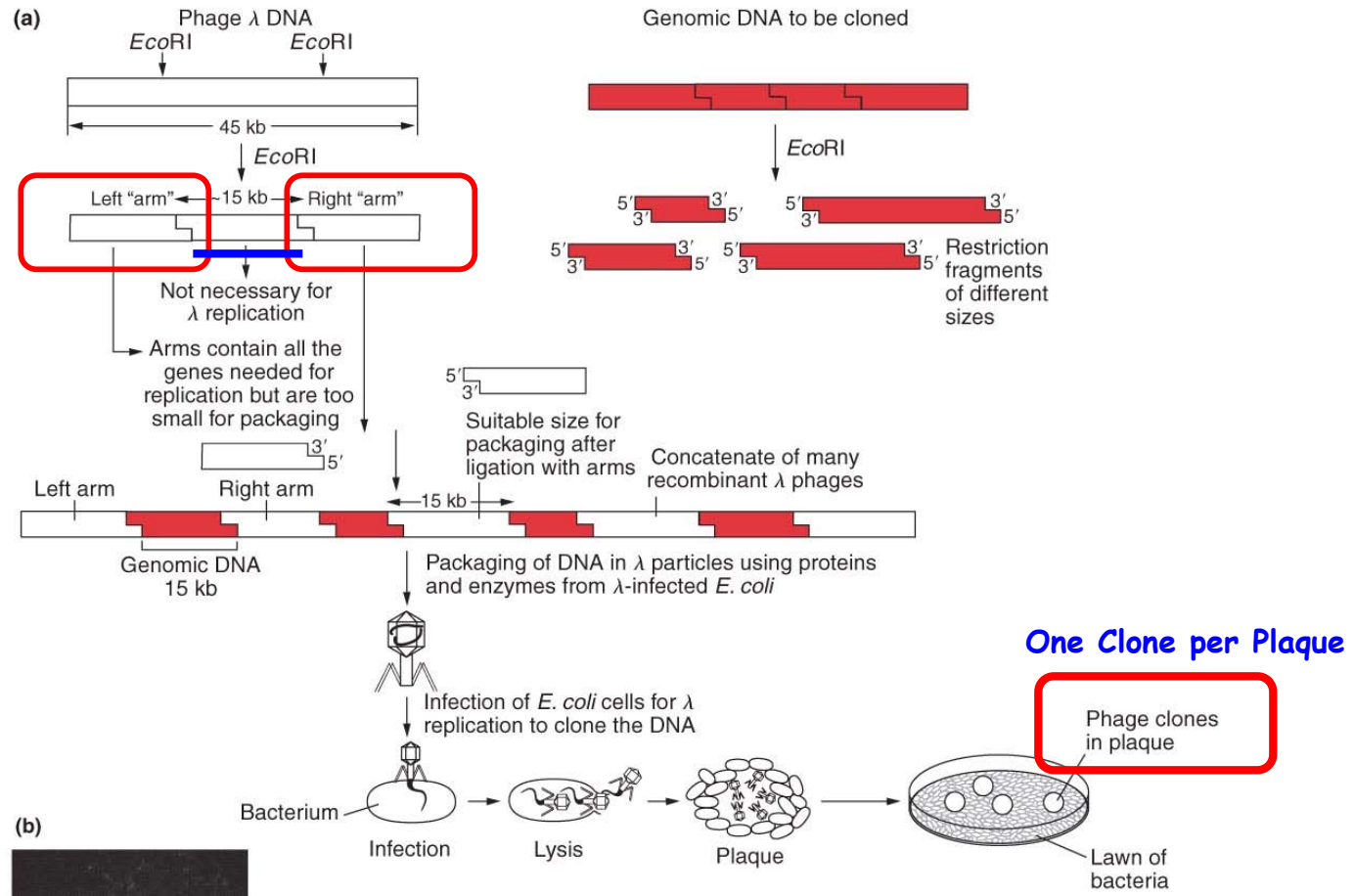
∴ Make Recombinant Viruses in the Lab!

3. Use "Natural" Infection process to Generate Large Number of Clones for a Eukaryotic Genome Library.

Much higher efficiency for getting recombinant DNA → bacterial cells compared with DNA transformation.

∴ set more clones per amount of recombinant DNA!

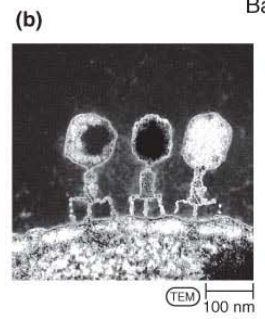
# Using a Bacterial Virus To Clone the Human Genome



One Clone per Plaque

Phage clones in plaque

Mixture of Plaques = Library With All Human DNA Sequences Represented



Copyright © 2009 Pearson Education, Inc.



# Cloning the Human Genome and Screening for the Factor VIII Gene

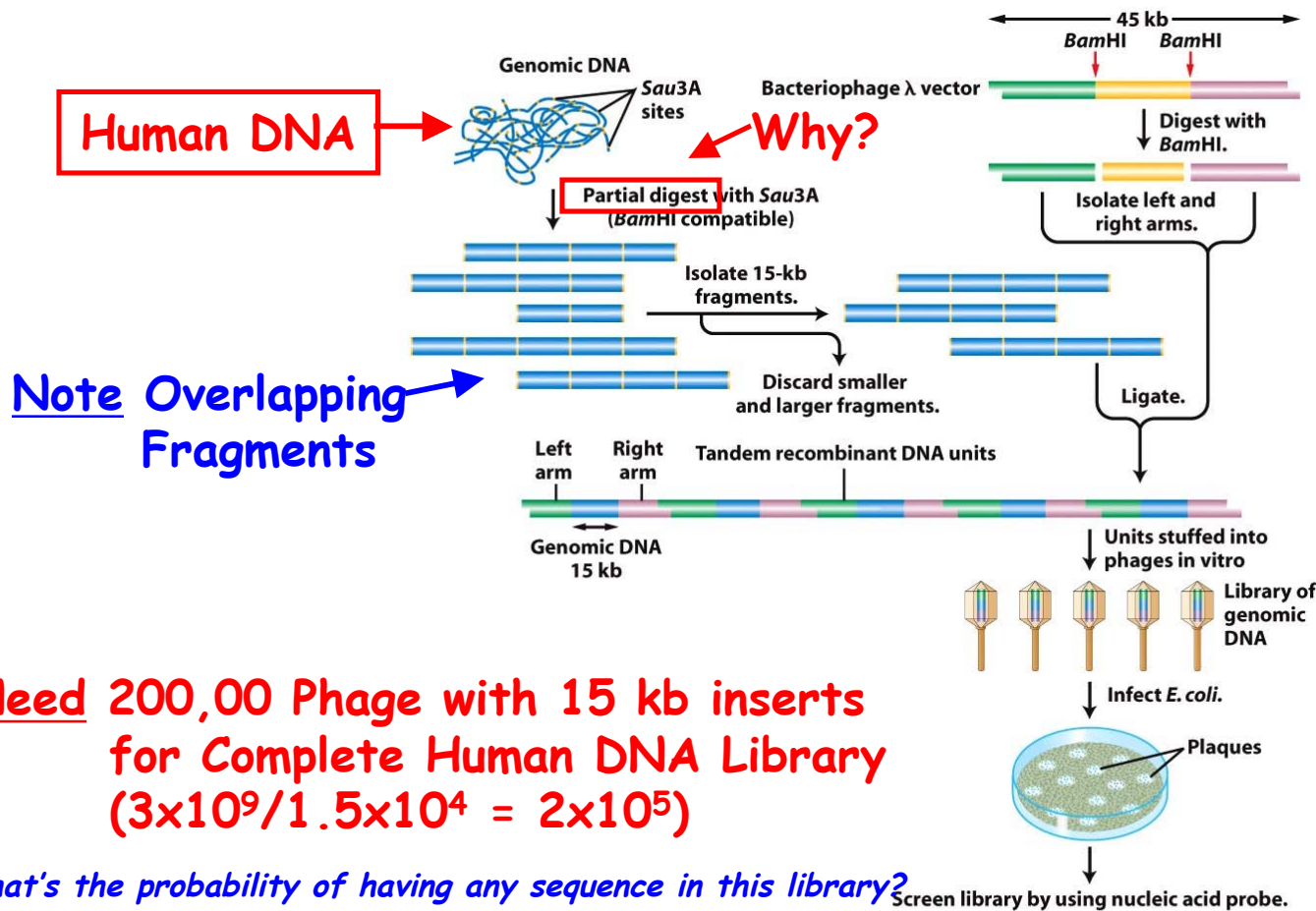


Figure 20-6  
*Introduction to Genetic Analysis, Ninth Edition*  
 © 2008 W. H. Freeman and Company

Why Partial Digestion? An Important Concept!

What is Complete & Partial Digestion?

An EcoRI Restriction Enzyme Site is Found Only Once in the Human Genome:

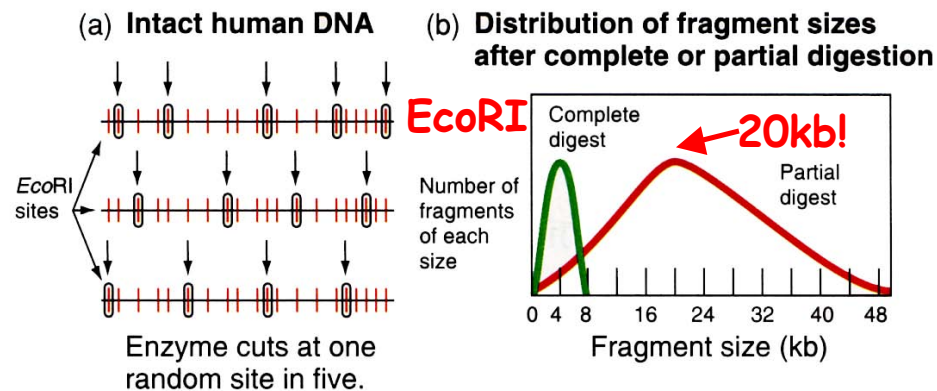
- a. Yes
- b. No

# What is the Purpose of Partial Digestion of Human DNA?

Sau 3A= 4bp= 5'GATC3' ∴ 1 site every 280bp if digest to completion =  $1 \times 10^7$  DNA fragments

Eco RI= 6bp= 5'GAATTC3' ∴ 1 site every 3100 bp if digest to completion (cleaves every site) = 972,000 DNA fragments

1. Complete Digestion Produces fragments that are too small to clone in  $\lambda$  virus (need 20Kb)
2. Complete Digestion would create huge genome libraries with large # clones to screen
3. Complete Digestion would break up genes of different DNA fragments- particularly if human genes big- ∴ would have one gene on many different clones- parts separated !
4. Complete Digestion provides no way to find neighbors of clones in genome- what's next to gene in chromosome!

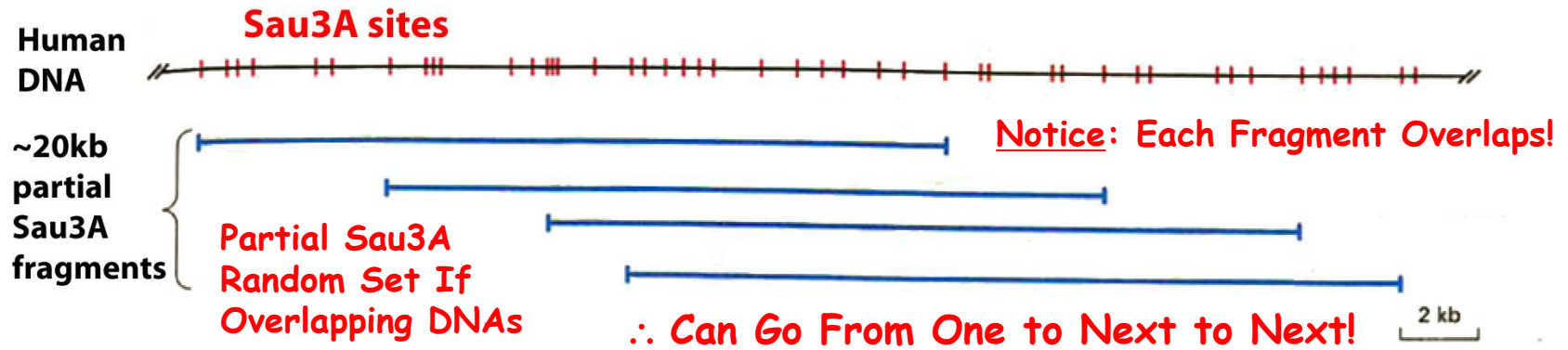


Principle of  
Genome  
Sequencing  
Too!!

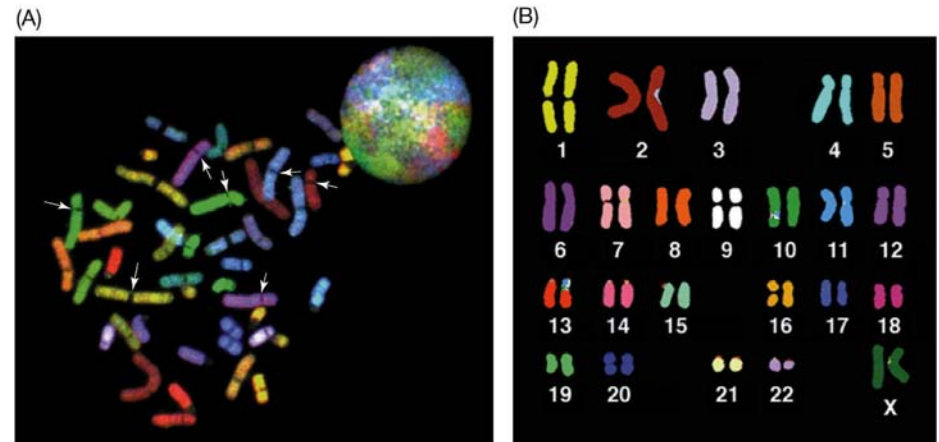
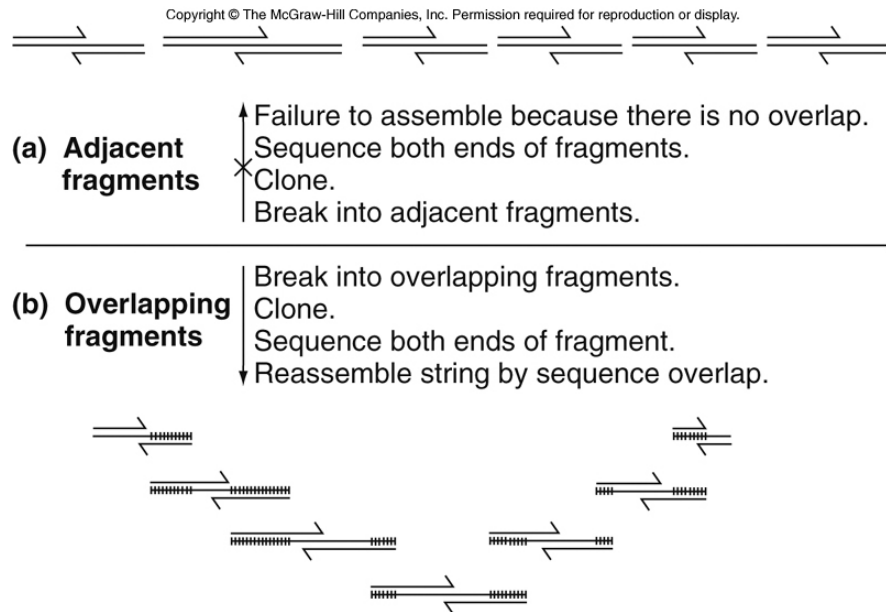
Partial Digestion Produces A series of Large, Overlapping DNA Fragments/ Clones

Can connect one clone with another!! Build up clones of each chromosome!!

# Constructing a Human Genome Library by Partial Digestion Creates a Set of Overlapping DNA Fragments/ Clones



**∴ An overlapping set for each of the 24 chromosomes would allow clones to be ordered from beginning to end by restriction mapping because each chromosome contains one DNA molecule !**



## Step Two

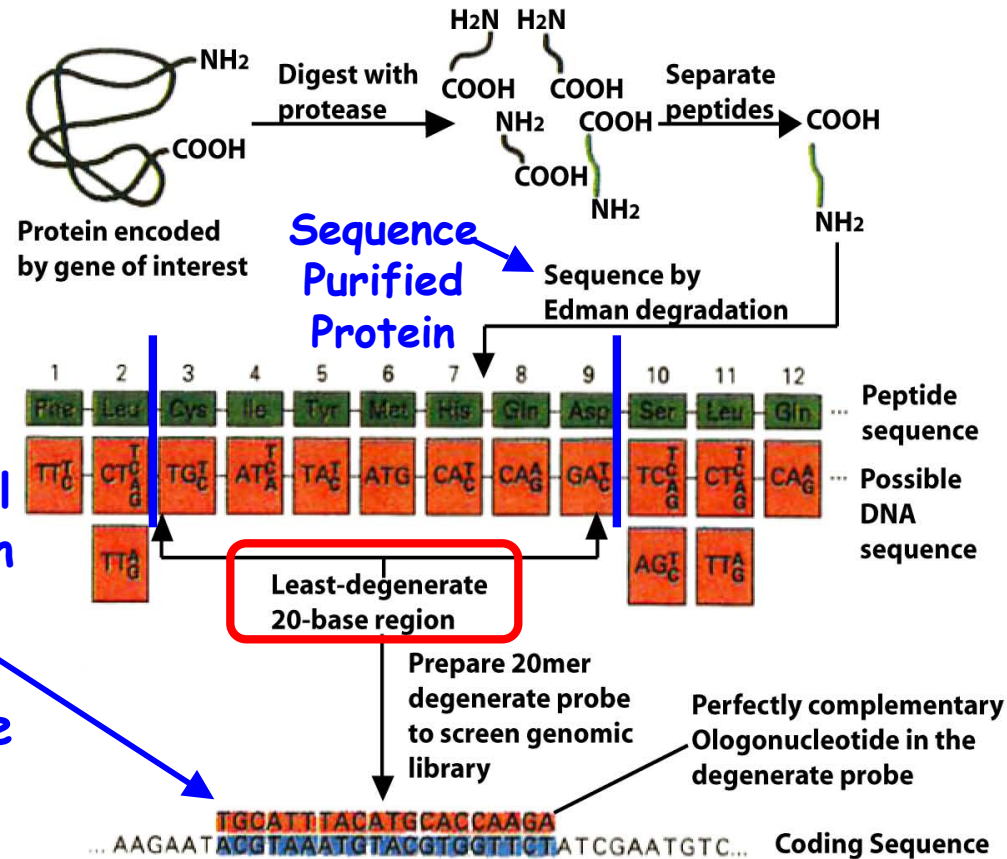
**How Find the Factor VIII  
Gene in a Human  
Genome Library?**

**A Specific Gene Can Be Identified in a Genome Library  
if the Amino Acid Sequence of its Protein is Known  
Because of the :**

- a. Double Helical Structure of DNA**
- b. Antisense Strand DNA Sequence**
- c. Genetic Code**
- d. Mutant Gene Phenotype**

# Factor VIII Protein → Gene

Using the Factor VIII Protein Sequence and Genetic Code as a Guide to Synthesize a Factor VIII Probe



2. Make Several Probes All Codon Combinations!

3. One Will Be Correct Probe

1. Use Genetic Code

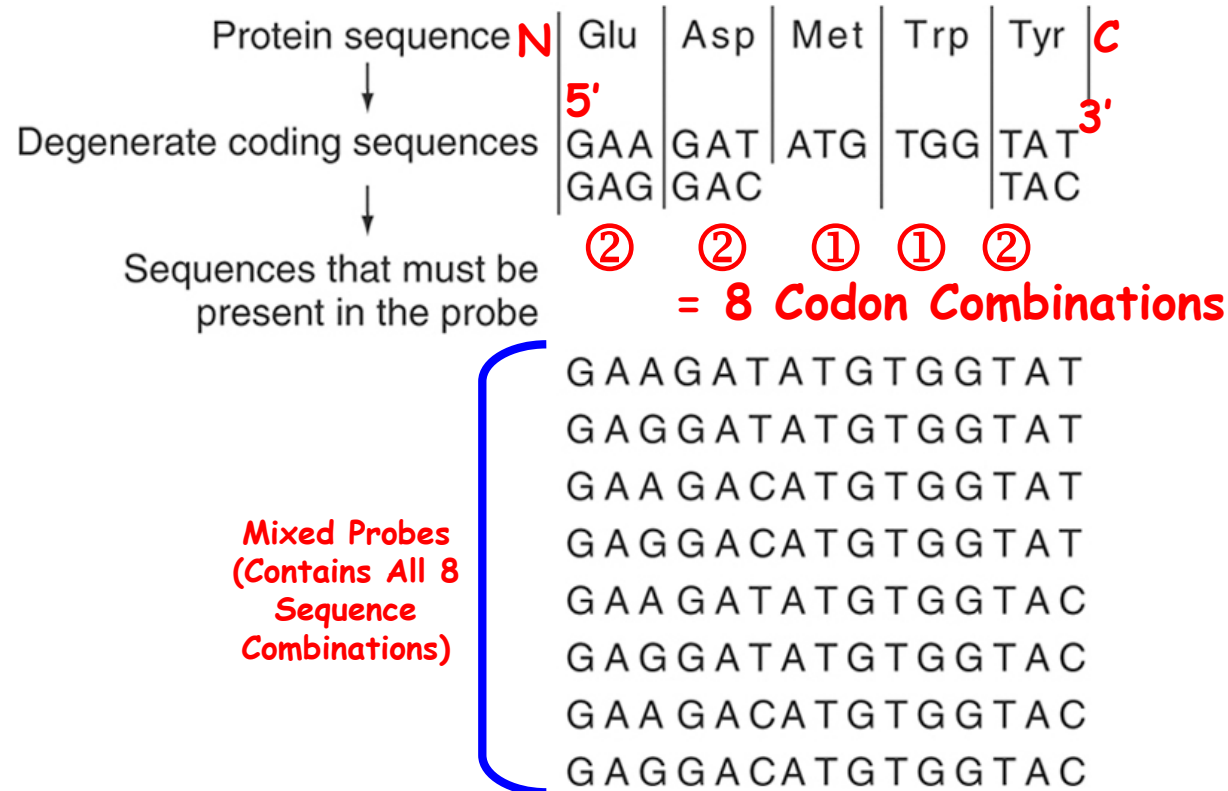
How many Combinations of Synthetic Probes?

$$2 \times 3 \times 2 \times 1 \times 2 \times 2 \times 2 = 96$$

# Using the Genetic Code to go From Protein Sequence to Gene Sequence

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

## (b) Synthesizing DNA probes based on reverse translation



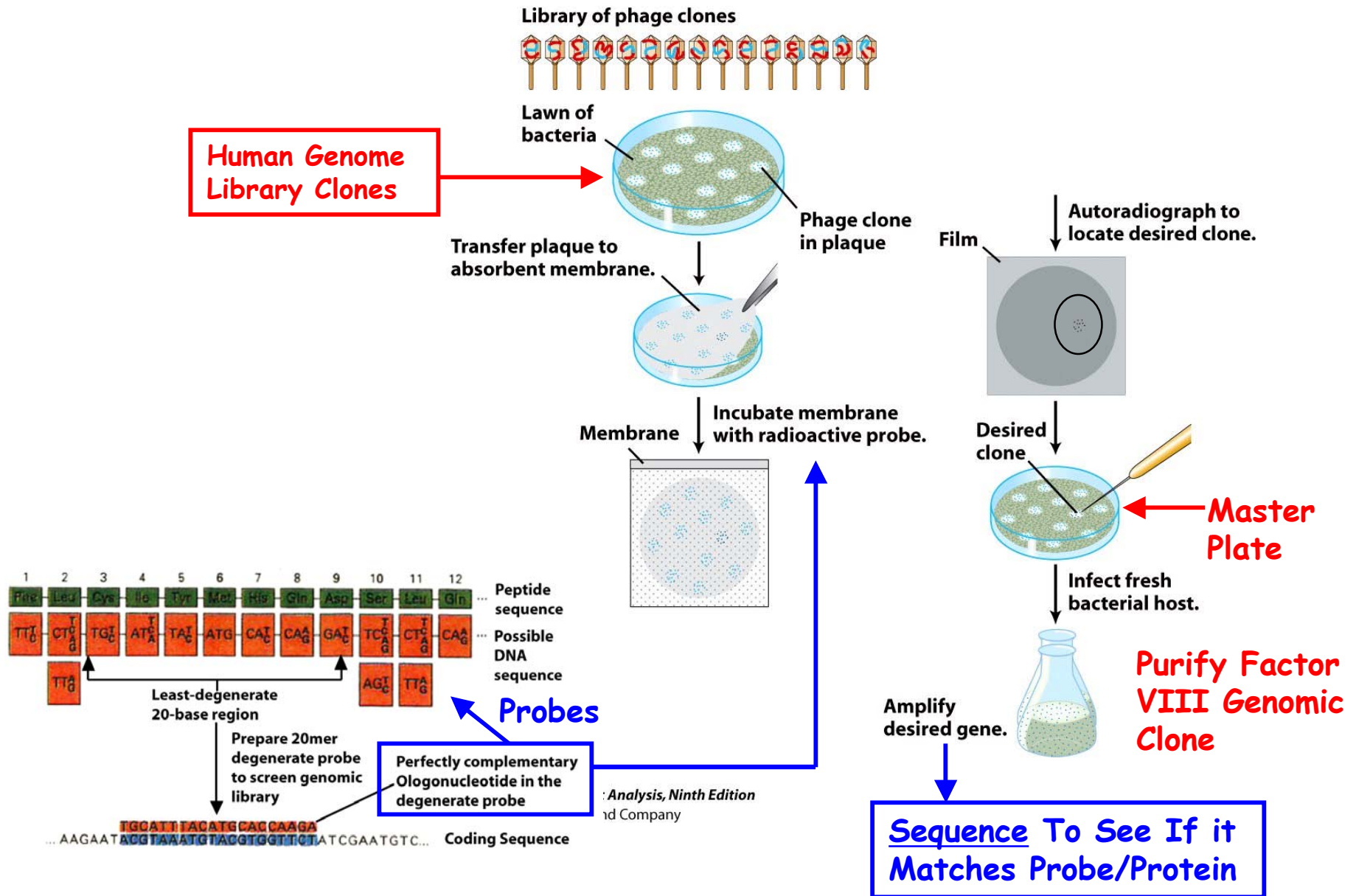
1. Need Amino Acid Sequence of Part of the Protein
2. Need DNA Sequences Representing all Codon Combinations
3. Synthesize DNA Sequence Probes!



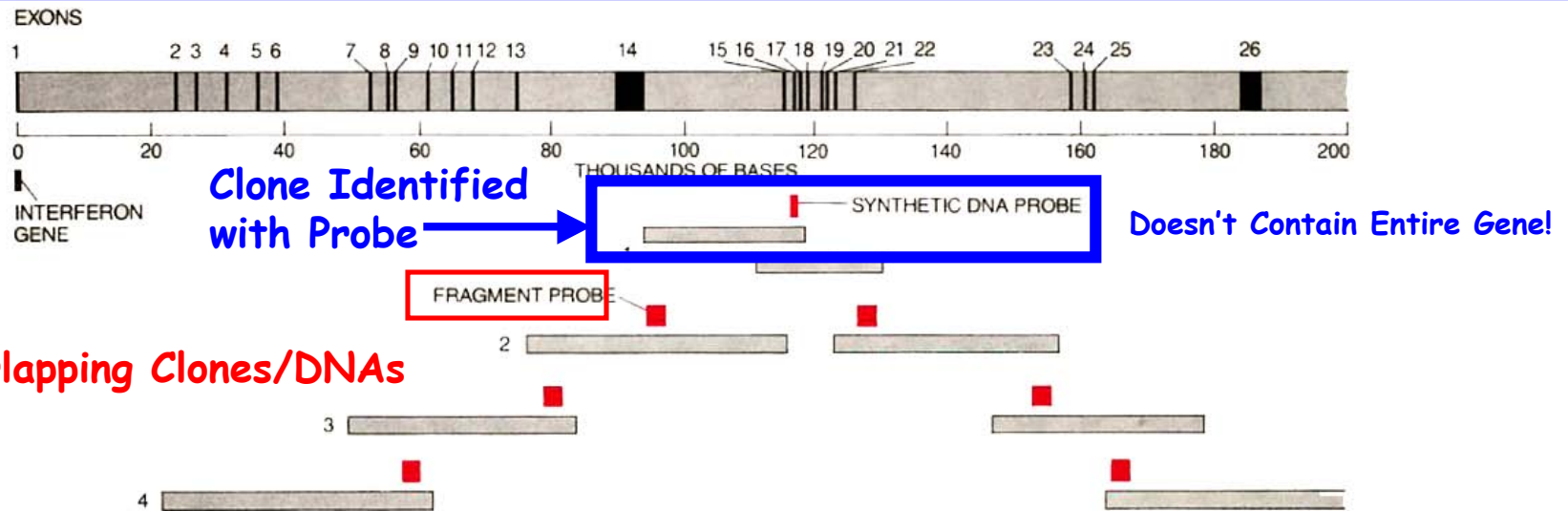
Probes Can Identify Genes in a Genome Library  
Because They Are: ?

- a. Synthetic
- b. Complementary to Specific DNA Sequences
- c. Contain the Correct Amino Acid Sequence
- d. Are Non-Radioactive

# Finding The Factor VIII Gene Or Part of Gene!!



The Result-The Factor VIII Gene is Huge- 186,000 bp- The Probe Identified a Clone Containing **Only One Part of Gene !!!** Why?



How Find Clones with Rest of Gene?

Key Question !

Remember - the library contains overlapping DNA clones  $\therefore$  can use one part of first clone to re-screen library & "walk" to other gene regions- using restriction maps & sequencing (compare with protein sequence) as guides!

Sequence -----> GenBank

## Step Three

# Finding the Entire Factor VIII Gene? Walking & Sequencing

## Walking up and down Genes and Chromosomes

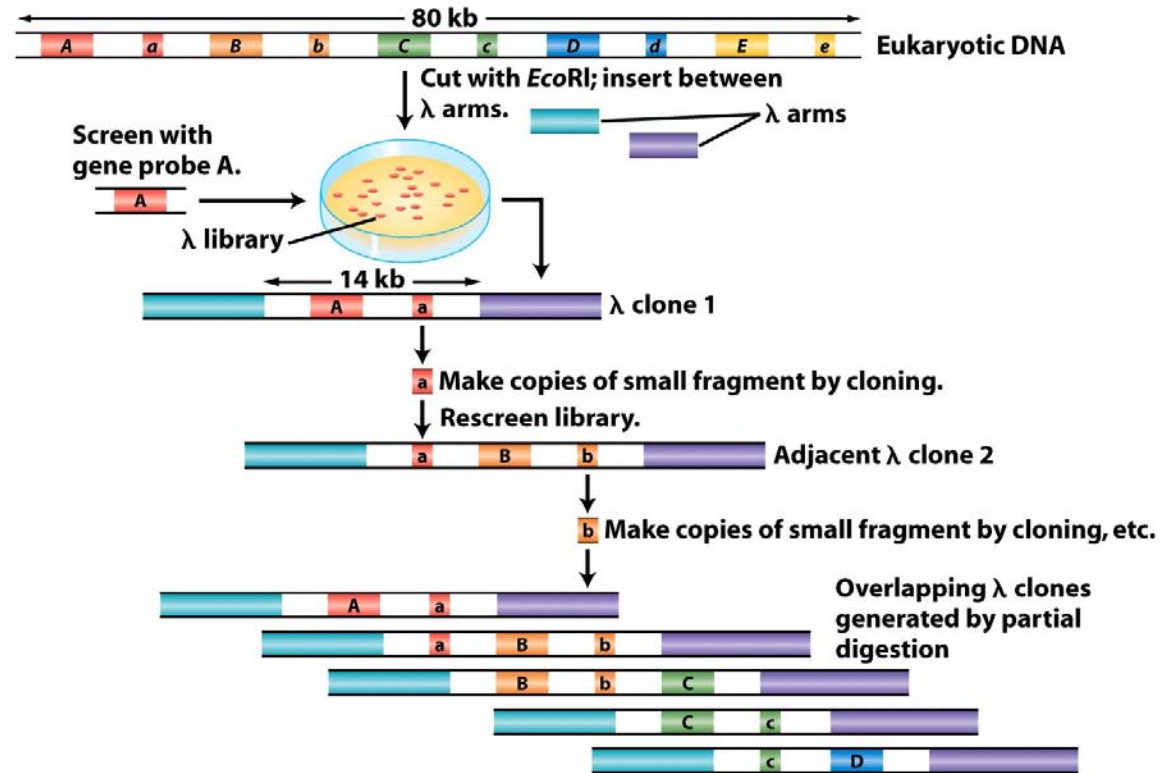


Figure 20-13  
Introduction to Genetic Analysis, Ninth Edition  
© 2005 W. H. Freeman and Company

## Basis of Genome Projects & Whole Genome Sequencing

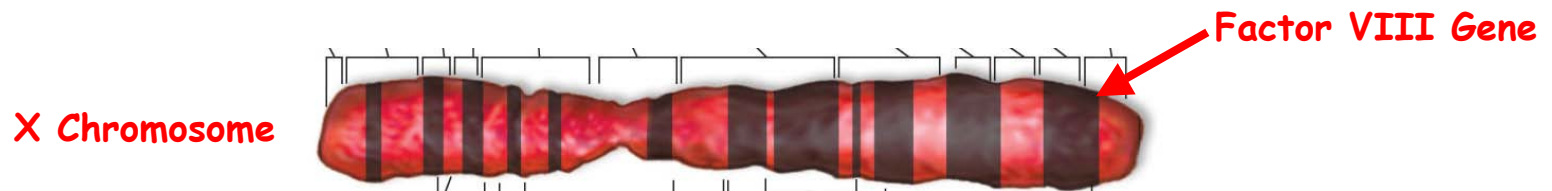
Key Concepts

How know Find Complete Factor VIII Gene?

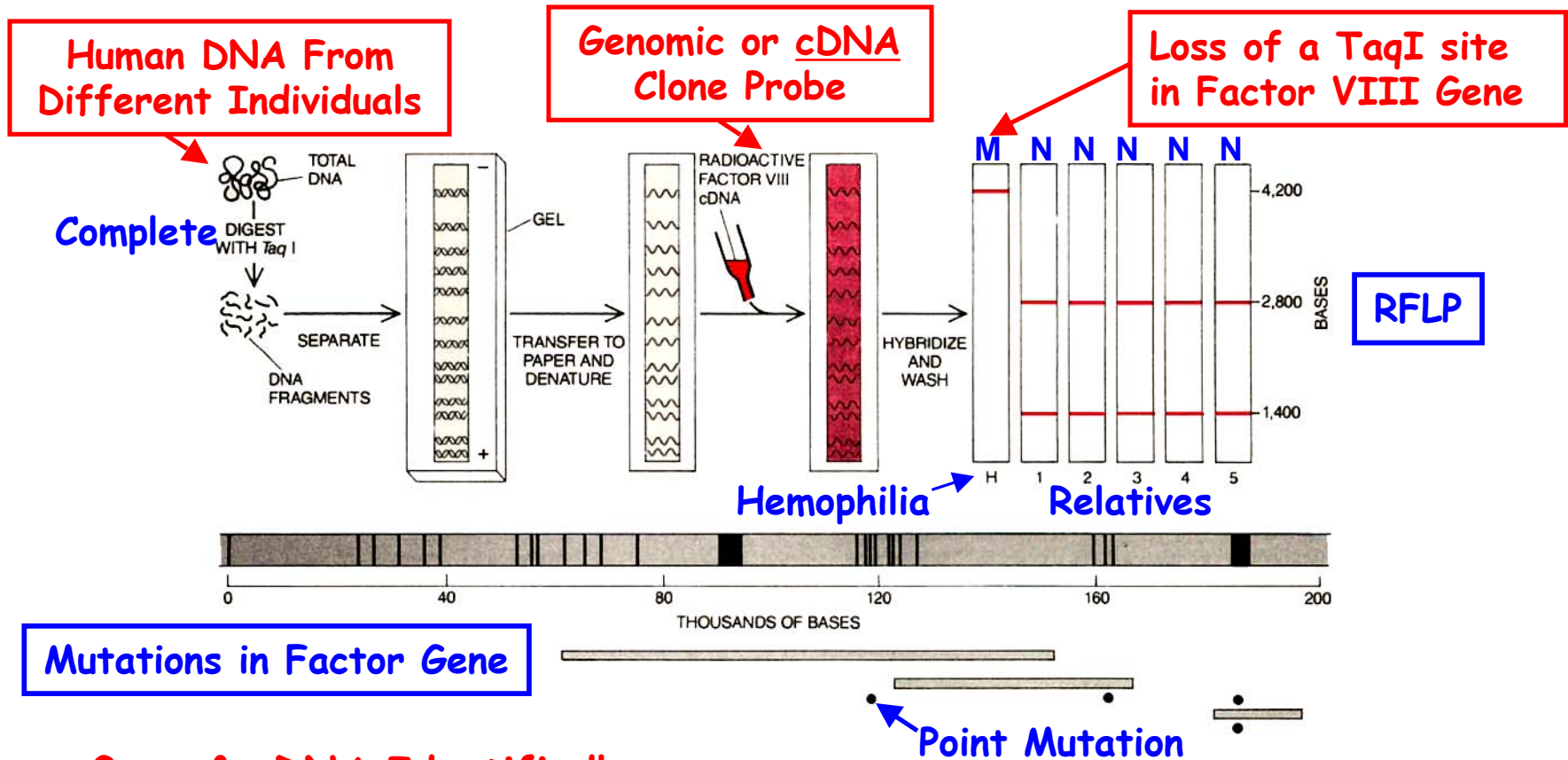
Compare Protein & DNA Sequences

## The Factor VIII Gene Was Found To Be Very Large

- **186,000 Nucleotides in Length** (Won't Fit in One Phage Clone)
- **25 Introns**
- **9,000 Nucleotide Coding Sequence (cDNA)**
- **2,351 Amino Acids in Protein**



# Factor VIII Gene Probes/ Sequence Can Be Used to Characterize Mutant Genes & Do DNA Testing for Carriers



Once Gene & cDNA Identified!

Use DNA Gel Blots (or PCR) & Factor VIII Probes to Investigate Presence of Mutant Alleles in Families (carriers)

Mutations Arise Independently in Families

# Factor VIII Mutations Occur Throughout the Gene

[*Haemophilia* 11, 481-491 (2005)]

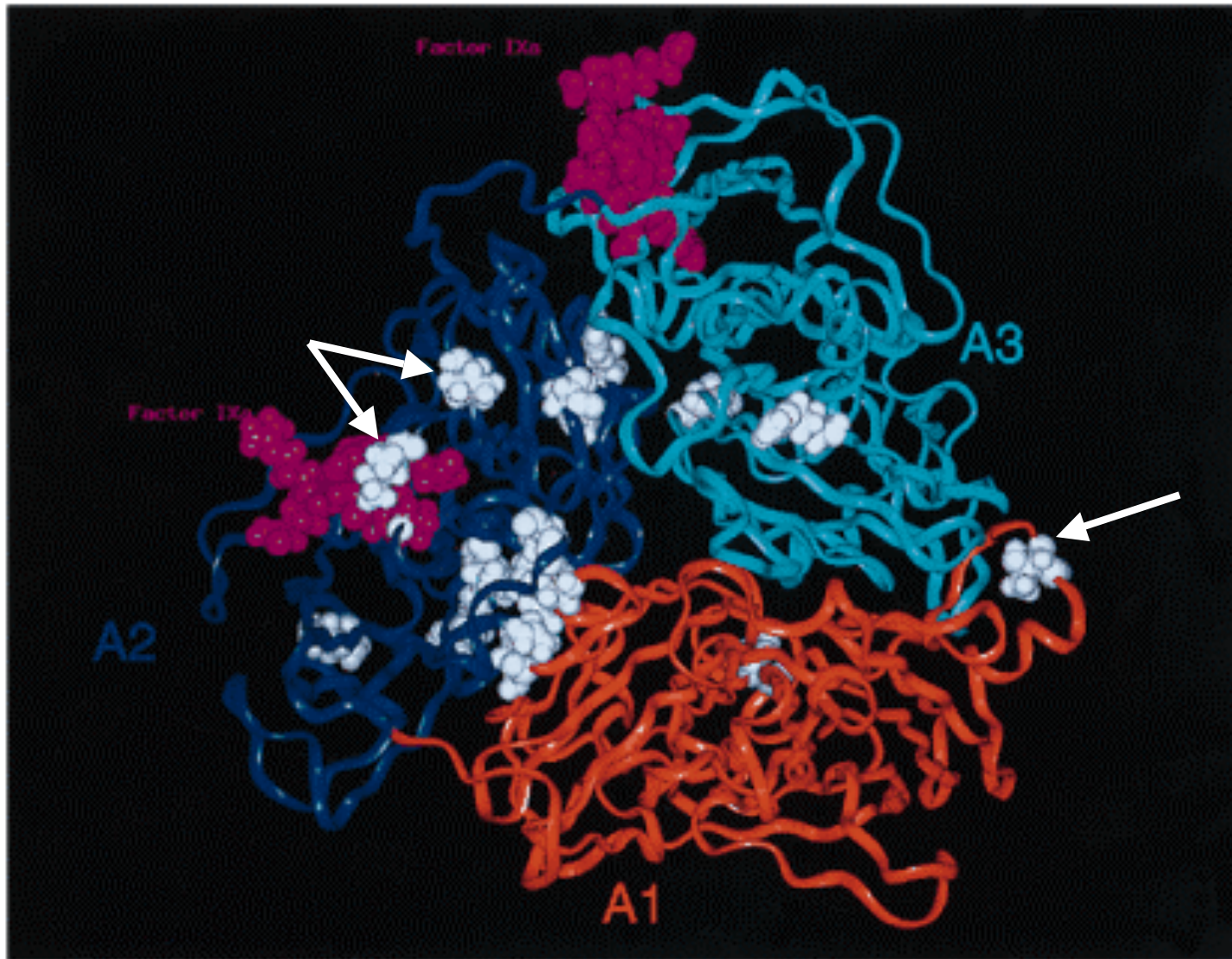
Factor VIII gene mutations in haemophilia A patients without intron 22 inversion.

VIII:C (%)	Family history	Consanguinity*	Inversion	Codon†	Mutation	Amino acid change	Exon	Conservation‡
1	Sporadic	NC	Normal	51	TTT → TCT§	Phe → Ser	2	FFFF, identical
1.20	Sporadic	NC	Normal	80	GTT → GAT	Val → Asp	3	VVVV, identical
1	Sporadic	NC	Normal	102	GGT → GTT§	Gly → Val	3	GGGG, identical
2	Sporadic	NC	Normal	104	TCC → CCC§	Ser → Pro	3	SSSS, identical
6	Sporadic	NC	Normal	143	GAG → AAG§	Glu → Lys	4	EEEE, identical
1	Sporadic	NC	Normal	233	delCA§	Thr → fs (TGA-264)	6	
2.70	Inherited	NC	Normal	321	GAA → AAA	Glu → Lys	8	EEEE, identical
0	Sporadic	NC	Normal	372	CGC → CAC	Arg → His	8	RRRR, identical
3	Inherited	NC	Normal	527	CGG → TGG	Arg → Trp	11	RRRR, identical
1	Sporadic	NC	Normal	528	TGC → TAC§	Cys → Tyr	11	CCCC, identical
1	Inherited	NC	Normal	592	CAA → TAA	Gln → Stop	12	QQQQ, identical
1	Inherited	NC	Normal	864	delGACA insCAATTAAATGAGAA§	Gly → fs [TAA-867]	14	
1	Sporadic	NC	Normal	948	insA§	Lys → fs (TGA-984)	14	
1	Sporadic	NC	Intron 1	1107	AGG → TGG§	Arg → Trp	14	RGKK, dissimilar
1	Sporadic	NC	Normal	1107	AGG → TGG§	Arg → Trp	14	RGKK, dissimilar
1	Inherited	NC	Normal	1191-1194	delA	Ile → fs (TAG-1198)	14	
1.40	Sporadic	NC	Normal	1191-1194	insA	Ile → fs (TAA-1220)	14	
1	Sporadic	C	Normal	1227	delC§	Leu → fs (TGA-1231)	14	
2.10	Sporadic	NC	Normal	1241	GAC → GAG	Asp → Glu	14	DGGE, similar
1	Sporadic	NC	Normal	1392	1392del1418§	Pro → fs (TAG-1446)	14	
1	Inherited	C	Normal	1392	1392del1418§	Pro → fs (TAG-1446)	14	
1	Sporadic	NC	Normal	1441	insA§		14	
1	Inherited	C	Normal	1441	insA§		14	
1	Inherited	NC	Normal	1502	CAG → TAG§	Gln → Stop	14	QREQ, dissimilar
1	Inherited	NC	Normal	1504	delGT§	Val → fs (TGA-1517)	14	
1	Sporadic	NC	Normal	1535	TGG → TGA	Trp → Stop	14	WLWM, dissimilar
inhibitor 96 BU	Sporadic	NC	Normal	1571	TAT → TAA§	Tyr → Stop	14	Y-YY, dissimilar
1	Sporadic	NC	Normal	1581	AAA → TAA§	Lys → Stop	14	KEKK, dissimilar
0.20	Sporadic	NC	Normal	1696	CGA → GGA	Arg → Gly	14	RRRR, identical
1.80	Sporadic	NC	Normal	1729	delA§	Gln → fs (TAA-1752)	15	
1	Inherited	NC	Normal	1751	GAA → AAA§	Glu → Lys	15	EEEE, identical
1	Sporadic	NC	Normal	1775	TTC → TCC§	Phe → Pro	16	FFFF, identical
1	Sporadic	NC	Normal	1835	TGG → TGA§	Trp → Stop	16	WWWW, identical
7.60	Sporadic	C	Normal	1882	ATC → ATA§	Ile → Ile	17	III, identical
3	Inherited	C	Normal	1966	CGA → CAA	Arg → Glu	18	RRRR, identical
1	Sporadic	NC	Normal	1966	CGA → TGA	Arg → Stop	18	RRRR, identical

FVIII GENE MUTATIONS IN INDIAN PATIENTS

**Need To Screen Across the Gene for Markers -- Family Specific**

# Factor VIII Protein Structure & Positions Where Mutations Disrupt Protein Function and Lead to Hemophilia





# How is a Specific Gene Detected in Genome?

DNA can be Transferred "in situ" to paper & annealed with radioactive probes

## DNA Blots!

Probe Represents a Cloned Fragment from Genome with a Unique Sequence!

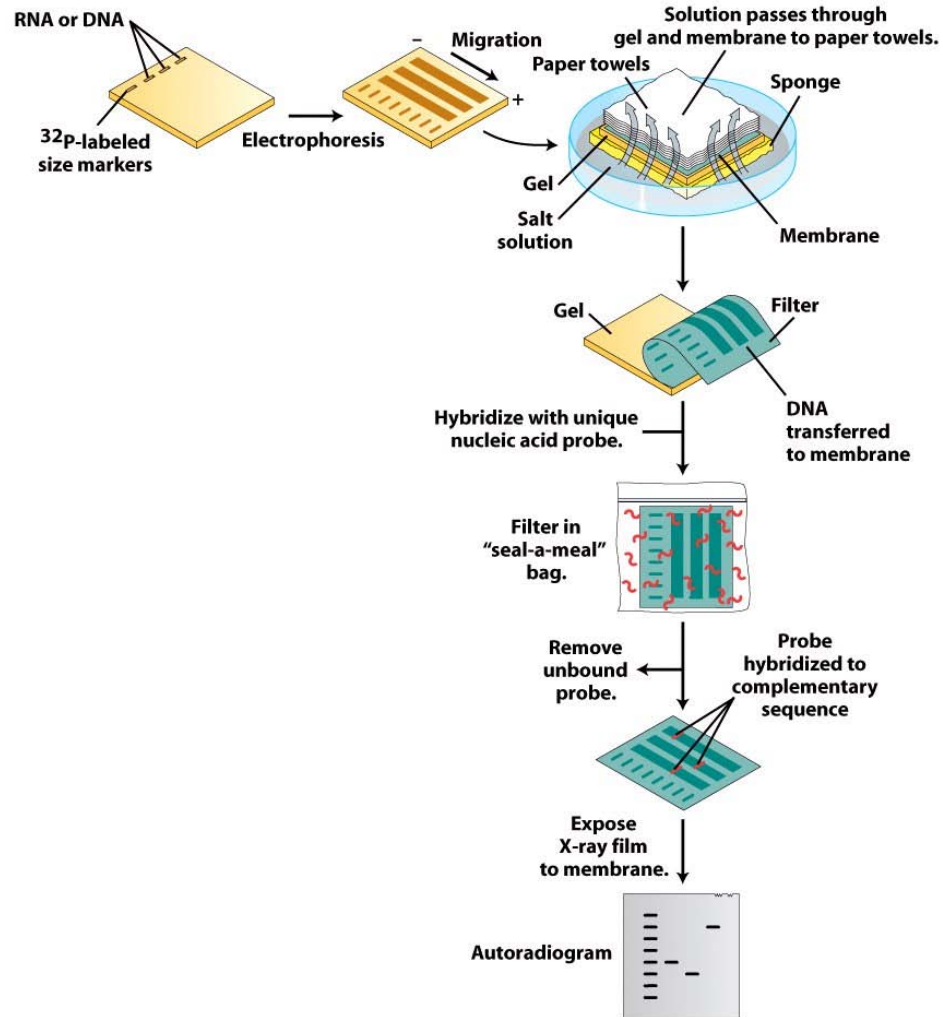
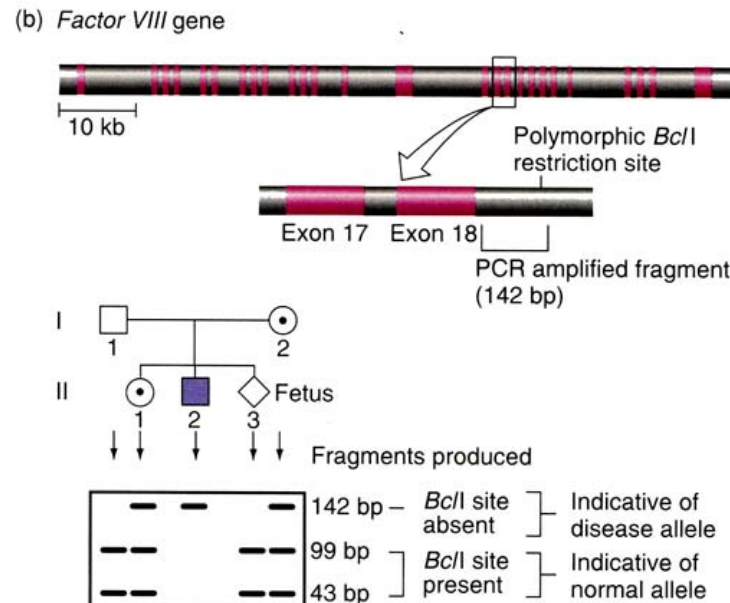
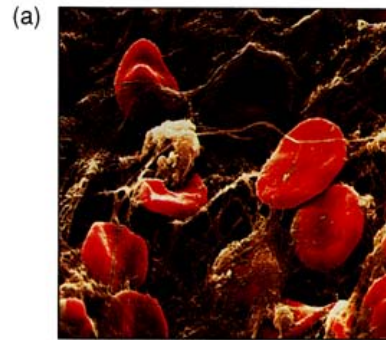


Figure 20-12  
*Introduction to Genetic Analysis, Ninth Edition*  
© 2008 W. H. Freeman and Company

# Using PCR and RFLPs (Markers) to Detect the Hemophilia A Disease Allele/ Gene

1. Use PCR to amplify a specific Factor VIII gene region
2. Use restriction enzyme (Bcl I) to distinguish between normal allele (1 site) & disease allele (no site)

= = Normal allele  
- = Disease allele

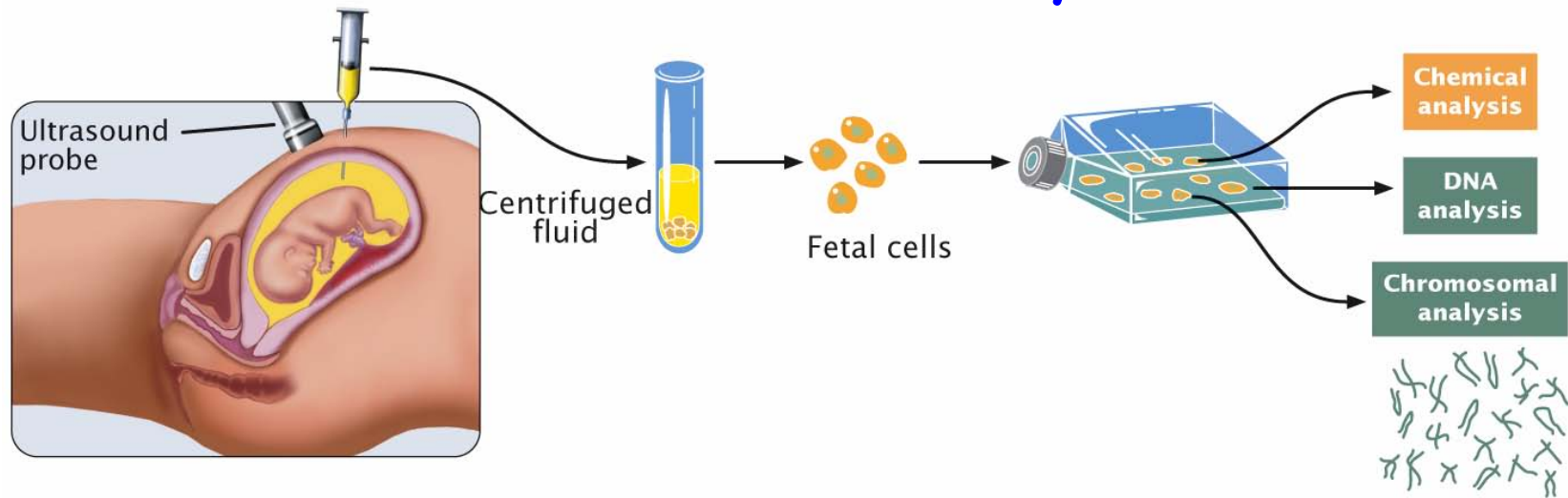


The 21<sup>st</sup> Century Approach!

1. Sequence the Entire Gene & Find Mutation
2. Then Synthesize Probes to Test Family Members Using PCR

Only Can Do This With a Knowledge of DNA Sequence of Wild-type (Normal) and Disease Genes (Can Vary family to Family)

# Use Gene Probe to Test for Disease Gene Prenatally



## Ultrasound Picture

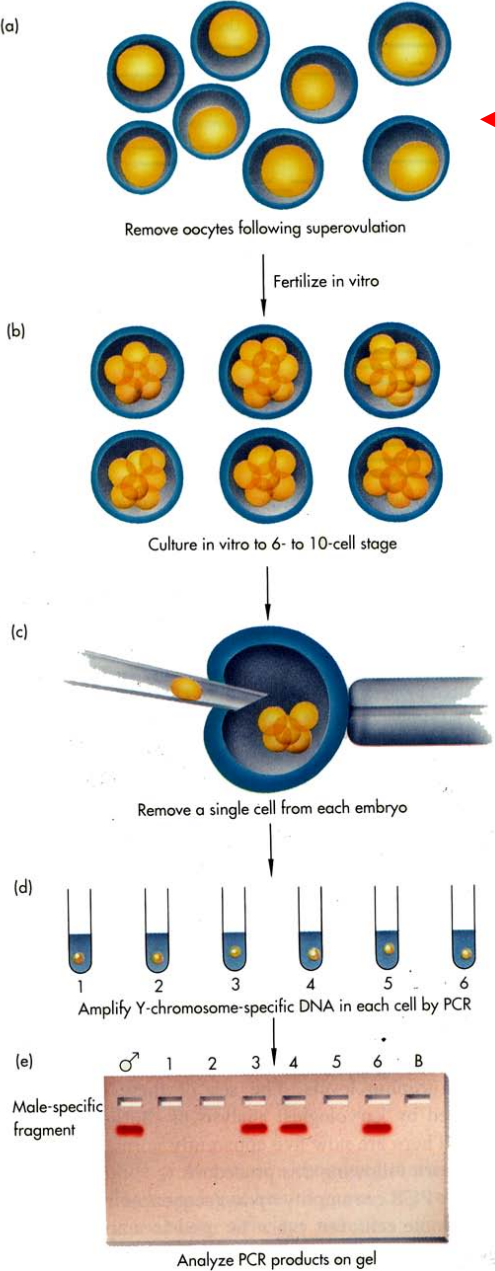


Fig. 06-15 Genetics, Second Edition © 2005 W.H. Freeman and Company

**Table 6.5** Examples of genetic diseases and disorders that can be detected prenatally and the techniques used in their detection

Disorder	Method of Detection
Chromosome abnormalities	Examination of a karyotype from cells obtained by amniocentesis or CVS
Cleft lip and palate	Ultrasound
Cystic fibrosis	DNA analysis of cells obtained by amniocentesis or CVS
Dwarfism	Ultrasound or X-ray; some forms can be detected by DNA analysis of cells obtained by amniocentesis or CVS
Hemophilia	Fetal blood sampling* or DNA analysis of cells obtained by amniocentesis or CVS
Lesch-Nyhan syndrome (deficiency of purine metabolism leading to spasms, seizures, and compulsory self-mutilation)	Biochemical tests on cells obtained by amniocentesis or CVS
Neural-tube defects	Initial screening with maternal blood test, followed by biochemical tests on amniotic fluid obtained by amniocentesis and ultrasound
Osteogenesis imperfecta (brittle bones)	Ultrasound or X-ray
Phenylketonuria	DNA analysis of cells obtained by amniocentesis or CVS
Sickle-cell anemia	Fetal blood sampling or DNA analysis of cells obtained by amniocentesis or CVS
Tay-Sachs disease	Biochemical tests on cells obtained by amniocentesis or CVS

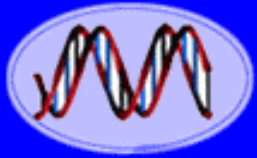
# Using PGD to Detect Hemophilia A Disease Alleles



Mother is a Carrier  $X^HX^h$

- 1. Test for Male Embryos
- 2. Test for Presence of Hemophilia A Disease Alleles!

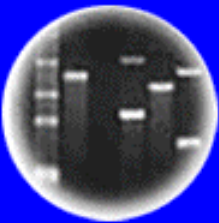
$X^hY$



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

# Genetic Screening Issues

- Why Screen For Genes?
- When is a Test Accurate Enough?
- Mandatory or Voluntary Screening?
- Who Should Be Tested?
- Employer & Insurance Company Testing?
- Protection From Genotype Discrimination?
- Testing for Genetic Diseases With No Cures??
  - How Ensure Privacy & Confidentiality?
- Obligations to Inform Others (Spouse/Sibling) of Genetic Disorder Knowledge?
- Genetic Databases??

## Step Four

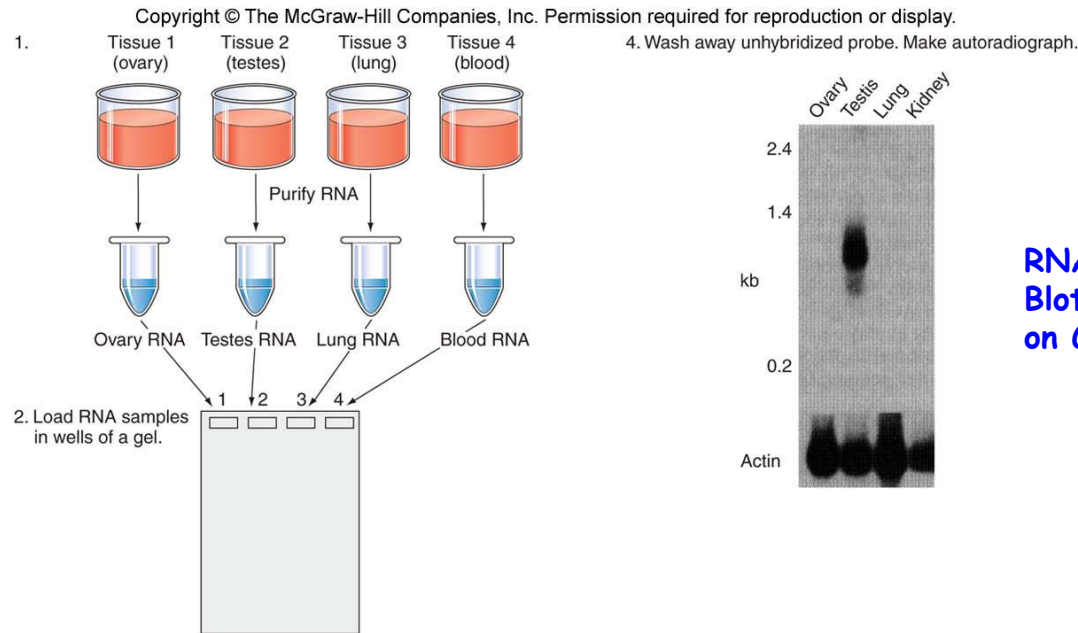
How Find Factor VIII mRNA to  
Generate a cDNA for Protein  
Production in Host Cells?

Recall: Eukaryotic Genes Provide  
Obstacles for Efficient Protein  
Production in Genetically  
Engineered Cells! Reasons???

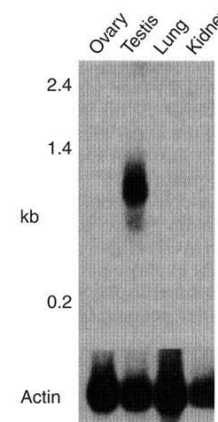
# Making the Drug

Need cDNA Not Gene

Factor VIII Gene Can Be Used to Find Out Where It is Active Using RNA Blots



4. Wash away unhybridized probe. Make autoradiograph.

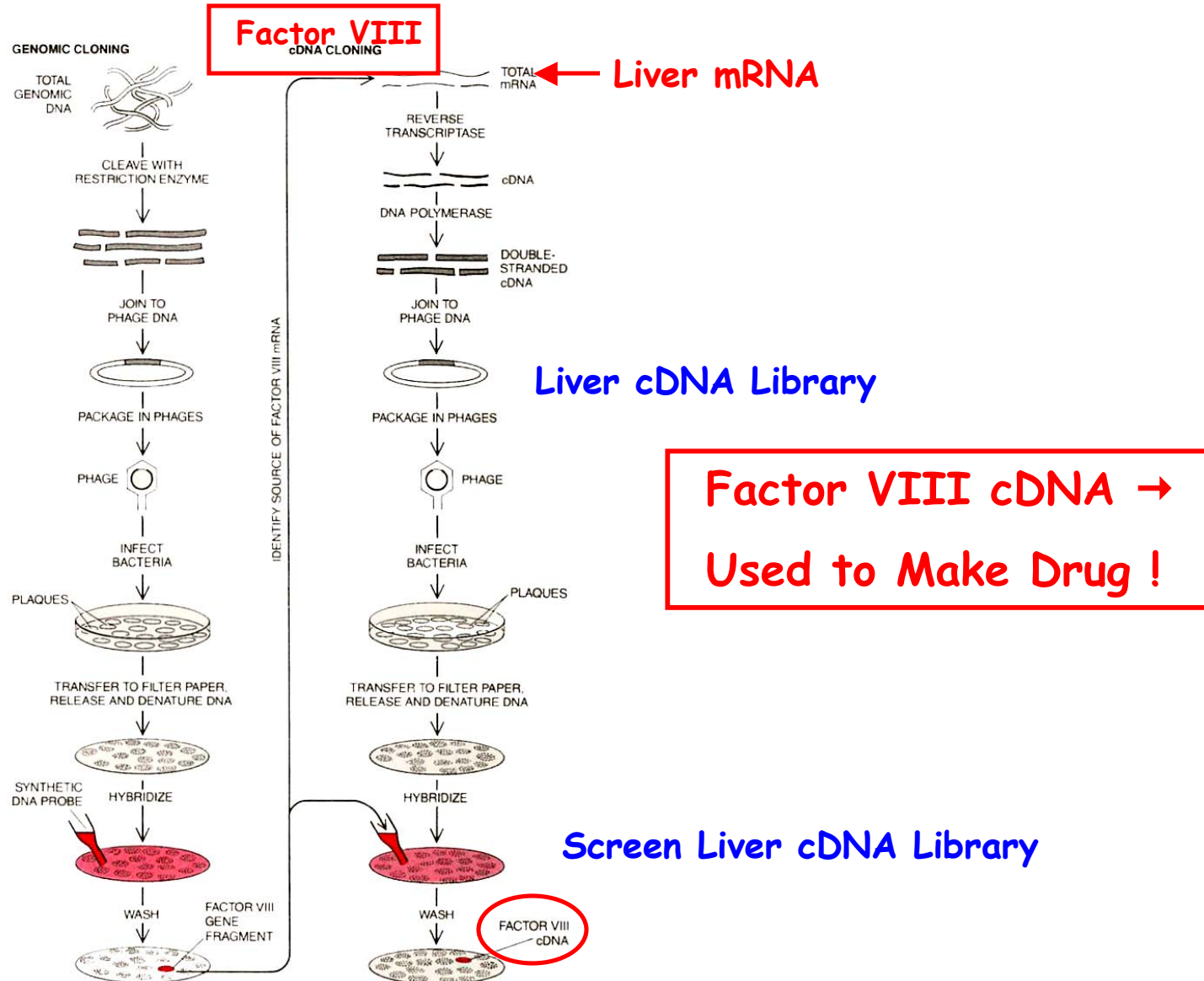


RNA Blot Is Like a DNA Blot Except That RNA is on Gel & Blotted

Factor VIII Gene Is Highly Active in Liver!

Could Also Use PCR (RT-PCR)

# Using Factor VIII Gene Probe to Identify Factor VIII cDNA clone





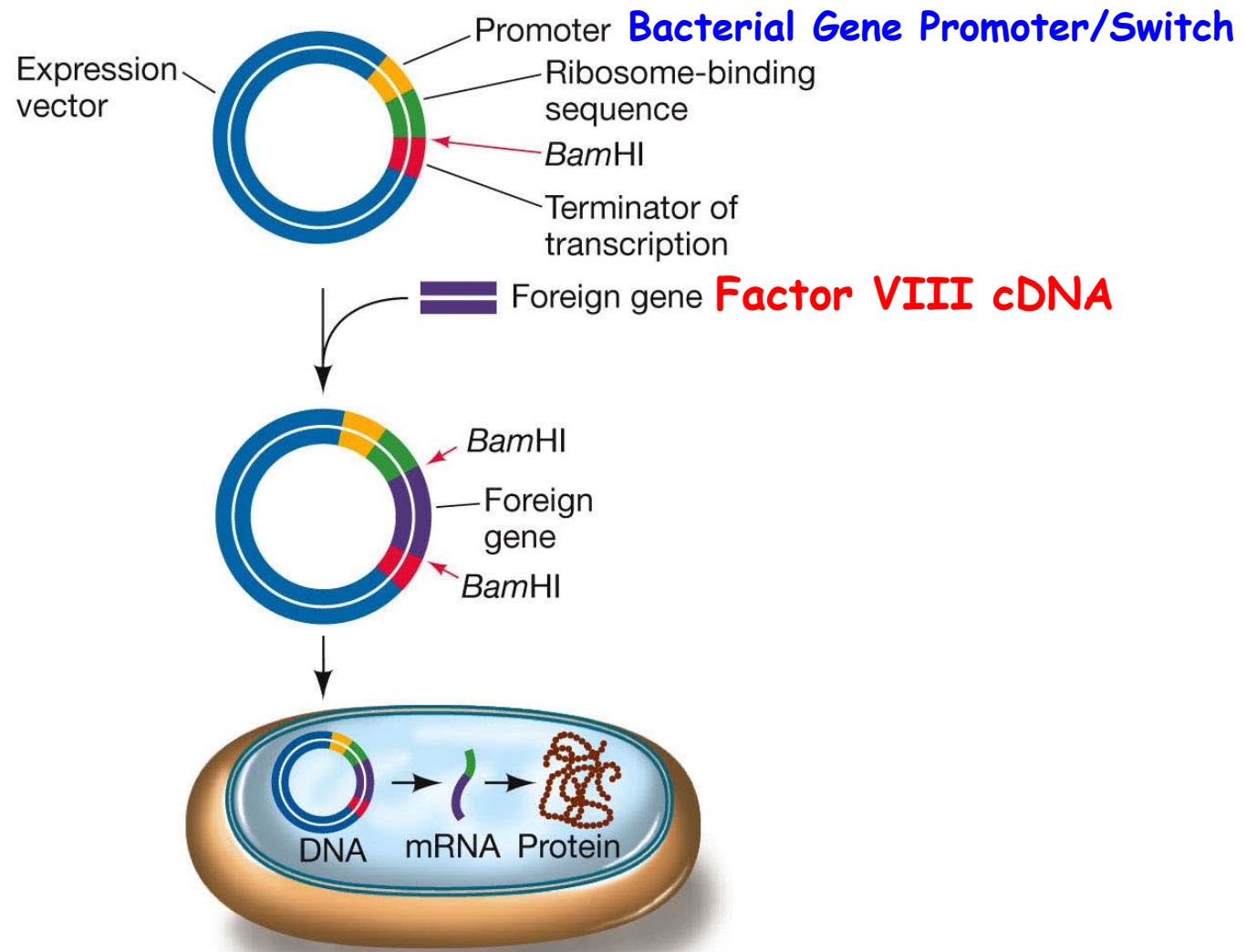
A cDNA is made by using:

- a. Primers
- b. mRNA
- c. DNA Polymerase
- d. All of above

The sequence of a cDNA clone is the same as:

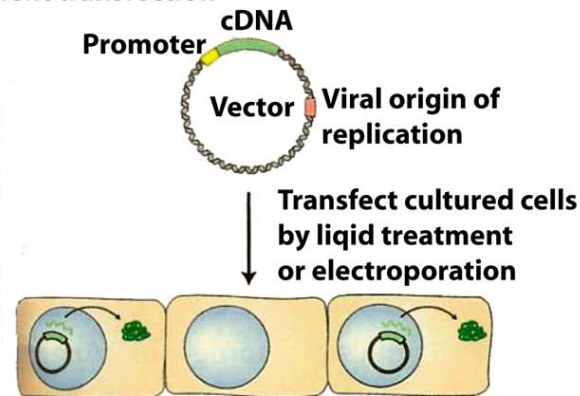
- a. The Sense Strand of the Corresponding Gene
- b. The mRNA Template
- c. The Antisense (Template Strand) of the Corresponding Gene
- d. The Sense and Antisense Strands of the Corresponding Gene Minus Introns

# Use Expression Vector to Allow cDNA to Produce Protein in Host Cell



# A Factor VIII Drug/"Cure" Making Factor VIII in Mammalian Cells

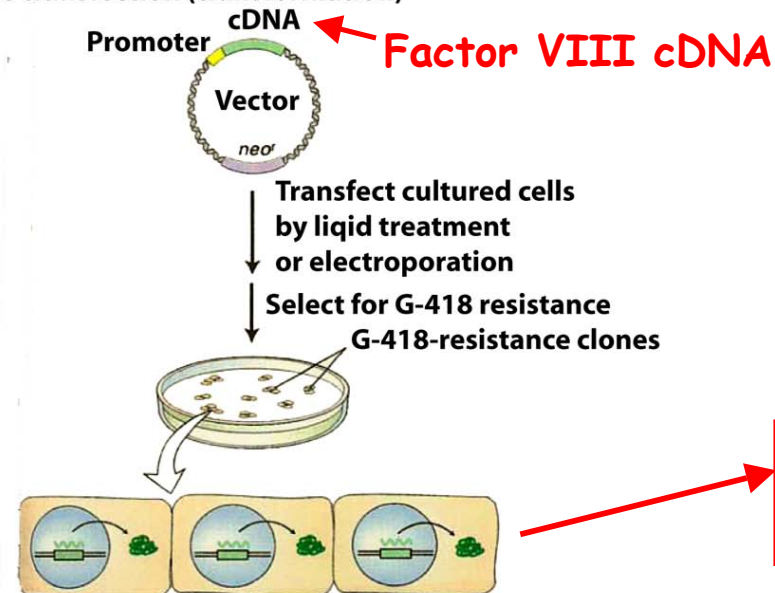
## (a) Transient transfection



Protein is expressed from cDNA in plasmid DNA

Why Mammalian Cells?

## (b) Stable transfection (transformation)



Protein is expressed from cDNA integrated into host chromosome

**Purify Factor VIII Protein!**

# Using Factor VIII to Treat Hemophilia

## Formation of a Blood Clot



FIBRIN STRANDS stabilize a blood clot at the site of a wound by trapping the platelets that form the bulk of the clot. The electron micrograph, which was made by Jon C. Lewis of Wake Forest University, shows a clot formed in a suspension of platelets and fibrin.

A clot in the bloodstream is the result of a complex cascade of enzymatic reactions culminating in the conversion of fibrinogen, a soluble protein, into insoluble fibrin strands. In hemophiliacs a crucial protein in the blood-clotting cascade is either missing or defective.

## A Triumph of Genetic Engineering

# Recombinant Factor VIII



Bayer Biological Products EU



Bayer HealthCare  
Biological Products Division  
Search | Sitemap

Home About Us **About Haemophilia** For Kids Research & Development Press Releases

→ Recombinant Factor VIII

More Resources

Haemophilia Centres in Europe

**Related Links**

Haemo-QoL Project  
Haemophilia Research Awards

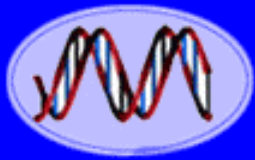
## Recombinant factor VIII

Recombinant factor VIII (rFVIII) is the antihemophilic factor A, obtained using recombinant DNA technology. With this technology, pure protein is synthesized in the laboratory instead of being extracted from blood. In the following pages, it will be explained in detail how the knowledge and analysis of DNA, using the new instruments of molecular genetics, have represented both the beginning



**Factor VIII gene cloned in 1983**

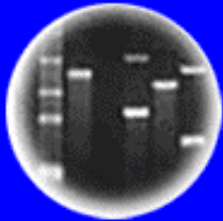
**Factor VIII (recombinant) approved as drug in 1993!**  
**Ten years from gene → drug! (Off Patent in 2011)**



DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

# A Patent on YOUR Factor VIII Gene!

United States Patent

5,618,788

Capon , et al.

April 8, 1997

Preparation of functional human factor VIII and pharmaceutical treatment therewith

### Abstract

Functional human factor VIII produced recombinantly is used in the treatment of human beings diagnosed to be deficient in factor VIII coagulant activity. Also provided are DNA solates and expression vehicles encoding functional human factor VIII, as well as transformed host cells and processes for producing human factor VIII by use of recombinant DNA technology.

Inventors: **Capon; Daniel J.** (San Mateo, CA), **Lawn; Richard M.** (San Francisco, CA), **Vehar; Gordon A.** (San Carlos, CA), **Wood; William I.** (San Mateo, CA)

Assignee: **Genentech, Inc.** (South San Francisco, CA)

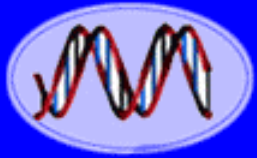
Appl. No.: **07/570,096**

Filed: **August 20, 1990**

**An Individual Should Be Allowed to Patent the  
Factor VIII DNA Sequence:**

- a. Yes**
- b. No**

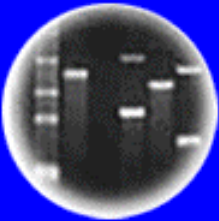




DNA  
Genetic Code of Life



Entire Genetic Code  
of a Bacteria



DNA Fingerprinting



Cloning: Ethical Issues  
and Future Consequences



Plants of Tomorrow

## The Factor VIII Story -- A Summary

1. Purify Small Amounts of Factor VIII
2. Obtain Partial or Complete Amino Acid Sequence
3. Use the Genetic Code to Synthesize Degenerate DNA Probes
4. Isolate Factor VIII DNA Clones Complementary to Probe in Genome Library
5. Determine if Factor VIII Clones Contain the Complete Gene By Sequencing and Comparing With Protein Sequence
6. If Not, "Walk" to Obtain Overlapping DNA Clones That Collectively Contain the Factor VIII Gene
7. Sequence Clones To Determine Where the Factor VIII Gene Starts and Stops
8. Use Factor VIII Genome Probe to Find Out What Body Organ/Tissue Expresses the Factor VIII Gene
9. Make a cDNA Library From the Target Organ/Tissue and Isolate a Factor VIII cDNA Clone
10. Sequence the Factor VIII cDNA Clone and Compare With Factor VIII Gene Sequence to Map its Anatomy (I.e., introns, exons, swtiches) and Ensure That it Contains the Complete Protein Coding Sequence
11. Use Factor VIII cDNA and/or Genome Fragments as a Probe to Find RFLP Markers For Disease Alleles -- Or Sequence Disease Alleles to Find Relevant RFLP Markers By Comparison With Wild-Type Sequence
12. Insert Factor VIII cDNA Into an Expression Vector and Synthesize Factor VIII Protein in Host Cells (e.g., Mammalian Cells)