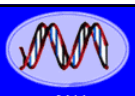


DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

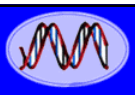
Plants of Tomorrow

HC70A, SAS70A, & PLSS599
Winter 2022
Genetic Engineering in Medicine, Agriculture, and Law

Professors Bob Goldberg, John Harada, & Channapatna Prakash

Lecture 5 - Part Two
How Are Genes Cloned & Engineered?
The Factor XIII Story

1



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

DNA Fingerprinting

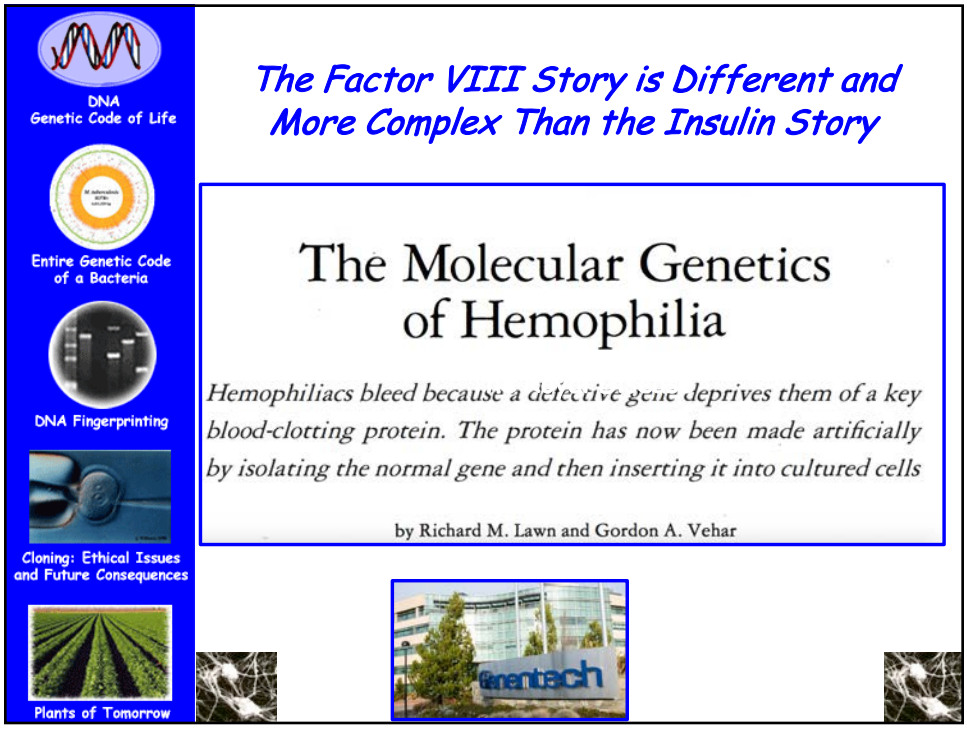
Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

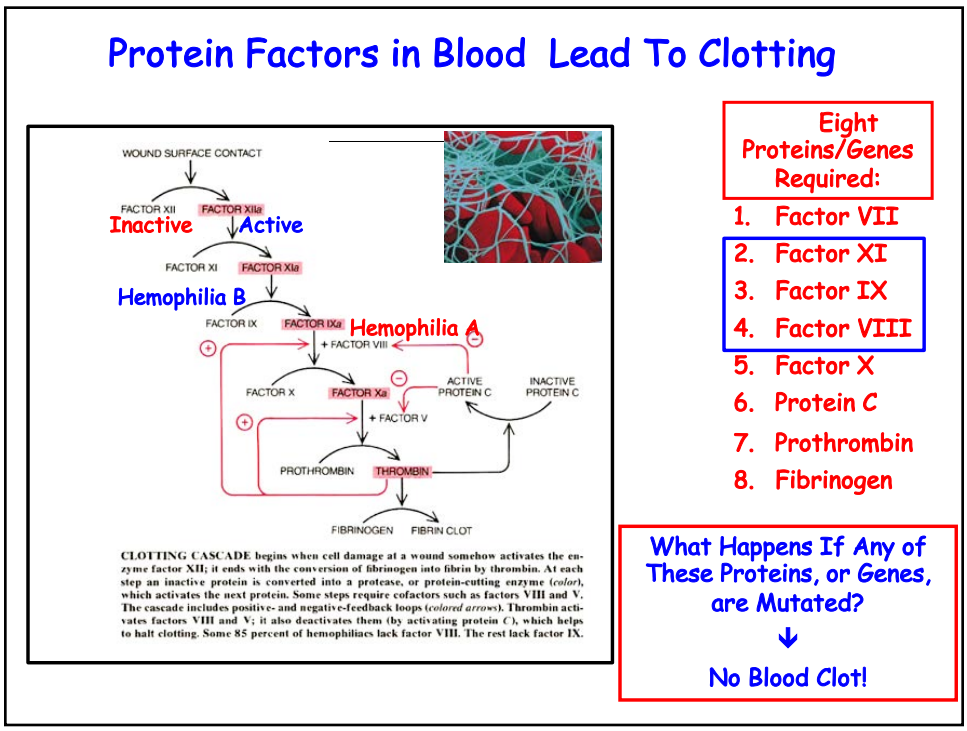
THEMES

1. How Did the Supreme Court Indirectly Give Rise to the Biotechnology Industry?
2. What Strategies Were Developed For Cloning Insulin mRNA and Expressing Insulin in Bacterial Cells? What Strategy "Won" Out?
3. What is Hemophilia and How is it Inherited?
4. How Can a Disease Gene Be Found When It is Not Known Where the Gene is Expressed?
5. What Vectors Can Be Used For Cloning DNA?
6. What is the Advantage of Using a Virus Vector For Constructing Genome Libraries?
7. How To Make a Library of the Human Genome?
8. How Find a Gene With Only a Knowledge of the Protein Sequence?
9. How Use DNA Testing to Detect Factor VIII Disease Alleles?
10. How Isolate a Factor VIII cDNA Clone?
11. Genomic vs. cDNA Libraries
12. How Produce Factor VIII Protein For Use as a Drug

2


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 is Different and More Complex Than the Insulin Story
The Molecular Genetics of Hemophilia
Hemophiliacs bleed because a defective gene deprives them of a key blood-clotting protein. The protein has now been made artificially by isolating the normal gene and then inserting it into cultured cells
by Richard M. Lawn and Gordon A. Vehar
Genentech

3


Protein Factors in Blood Lead To Clotting
WOUND SURFACE CONTACT
FACTOR XII
FACTOR XIIa
Inactive
Active
FACTOR XI
FACTOR XIa
Hemophilia B
FACTOR IX
FACTOR IXa
Hemophilia A
+ FACTOR VIII
FACTOR X
FACTOR Xa
ACTIVE PROTEIN C
INACTIVE PROTEIN C
+ FACTOR V
PROTHROMBIN
THROMBIN
FIBRINOGEN
FIBRIN CLOT
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.
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
What Happens If Any of These Proteins, or Genes, are Mutated?
No Blood Clot!

4

Hemophiliacs Have Mutations in Factor VIII, Factor IX, or Factor XI Genes

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

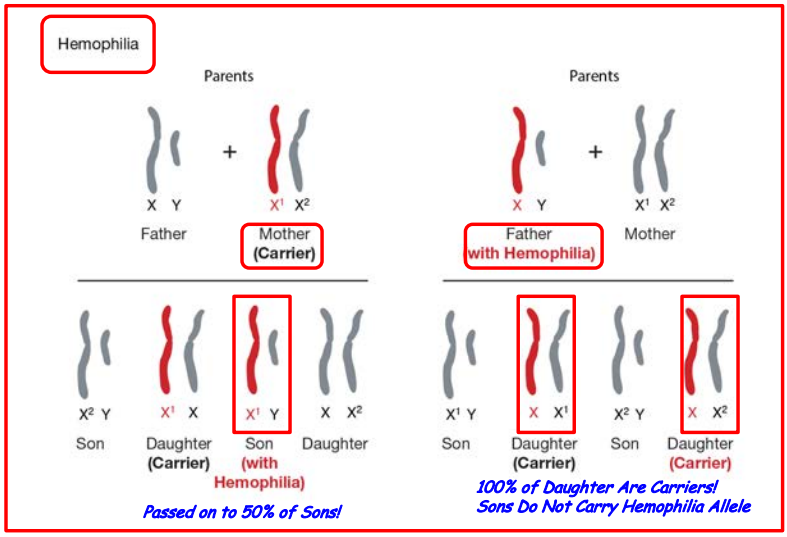
18,000 People in US Have Hemophilia & 400 Babies/Year Are Born With Disorder
Prior to 1960s – Average Life Span Was 11 Years

Hemophilia A	Defective Factor VIII Gene	1/10,000 males	80%
Hemophilia B	Defective Factor IX Gene	1/30,000 males	20%
Hemophilia C	Defective Factor XI Gene	Autosomal	<1%

Both Factor VIII & IX Genes
on X-Chromosome (♀→♂'s)

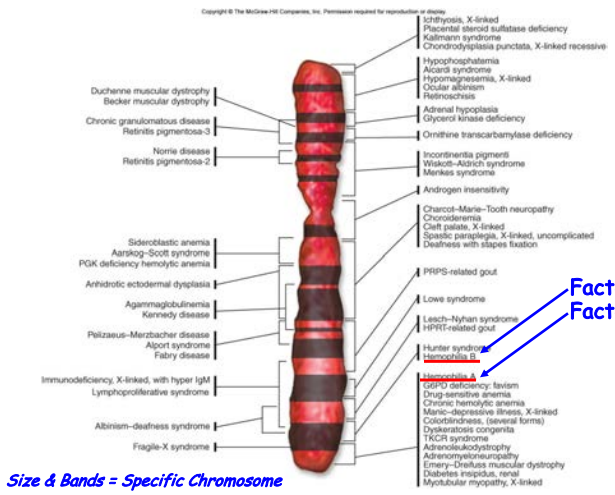
5

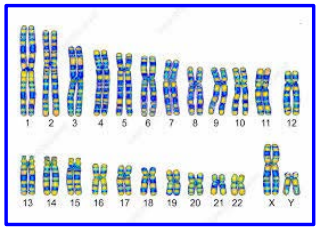
Factor VIII and Factor IX Genes Are on X-Chromosome and Show Sex-Linked Inheritance



6

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





Nature, March, 2005

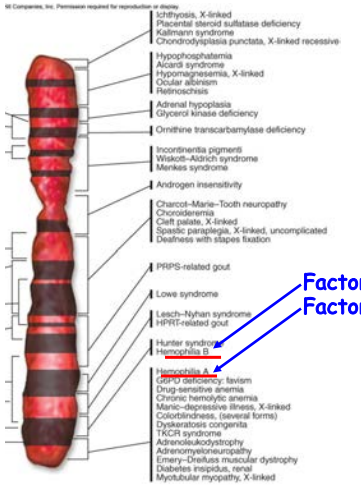
Size & Bands = Specific Chromosome

The X chromosome has ~1098 Genes and 150,000,000 bp (150 Mb).
168 Mendelian Diseases Explained by 113 X-Linked Genes

7

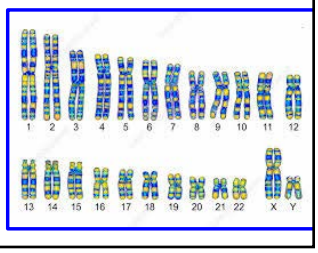
How to Find a Clone For Every Sequence on the X-Chromosome and Be Able to Order the Clones ?

Make a Genome Library To Order All Clones in the Library – Connect Clones & Genes to Each Other & Find Sequence of Entire Chromosome



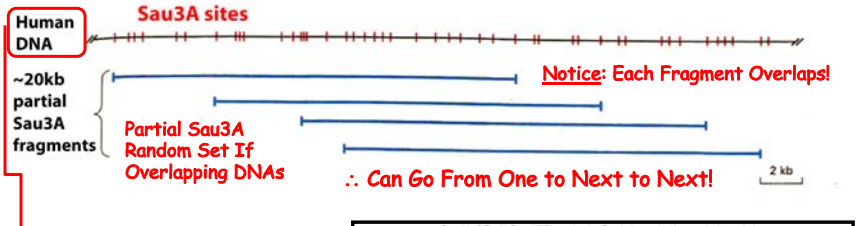
How Many Groups of Overlapping Clones in the Human Genome?

24 – One Group For Each Unique Chromosome



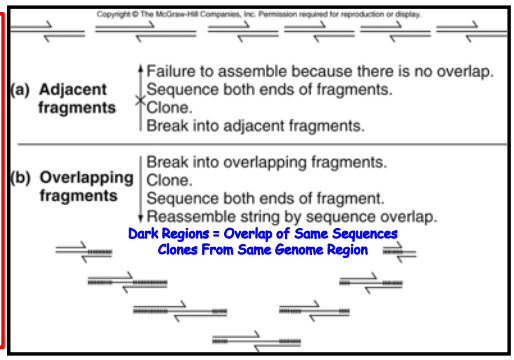
8

How to Construct a Human Genome Library That Creates Sets of Overlapping DNA Fragments/ Clones



Genes Connected to Each Other in Long Linear DNA Molecules
How Find Genes Contiguous to Each Other in Cells?

"Walking"



9

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

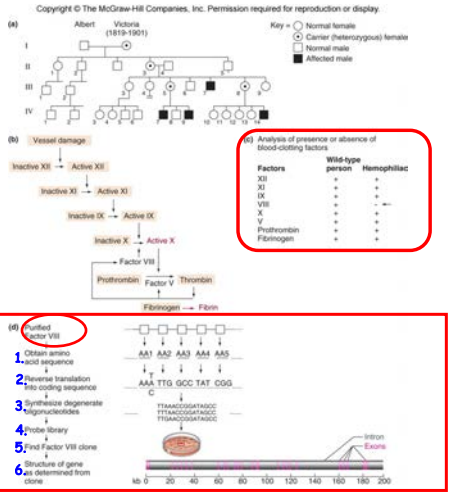
- Blood Protein (But Perhaps Synthesized Elsewhere!)
- Not Known Where Site of Synthesis Was
- Could Be Purified In Small Amounts From >20 Liters Of Human Blood + Cow Blood + Pig Blood
- Short Stretch Of Protein Sequenced = Known Protein Sequence!
- Hemophilia A Could Be Treated By *Blood Transfusions* From Normal Individuals, ∴ Clotting Factor In Blood
- 1980s Aids Epidemic Caused Many Hemophiliacs To Get HIV/AIDs (~50% Of Hemophiliacs Got Aids In 1985)
  - ∴ How To Go From Protein To Gene

10

The Problem!!

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

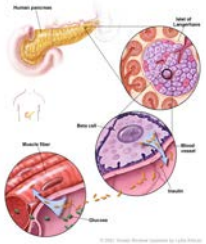
Early 1980's



Key Concept

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

Different Than Insulin Knew Where Protein Made!



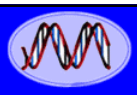
Key: Protein Sequence Known

How Find Gene & cDNA?

Protein → Gene → mRNA → Drug !

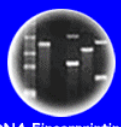
mRNA ⟶ Drug

11



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

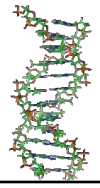
Plants of Tomorrow

Step One

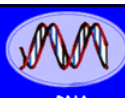
If It is Not Known Where Gene is Active Can "Look" to Genome Instead of mRNA to Find + Clone Gene!

How to Construct a Human Genome Library to:

a. Find the Factor VIII Gene
b. Find Genes Linked to the Factor VIII Gene
c. Be Able to Find Clones For All Contiguous Chromosome Regions
d. Sequence Every Human Chromosome or the Human Genome

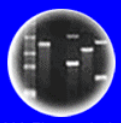


12



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

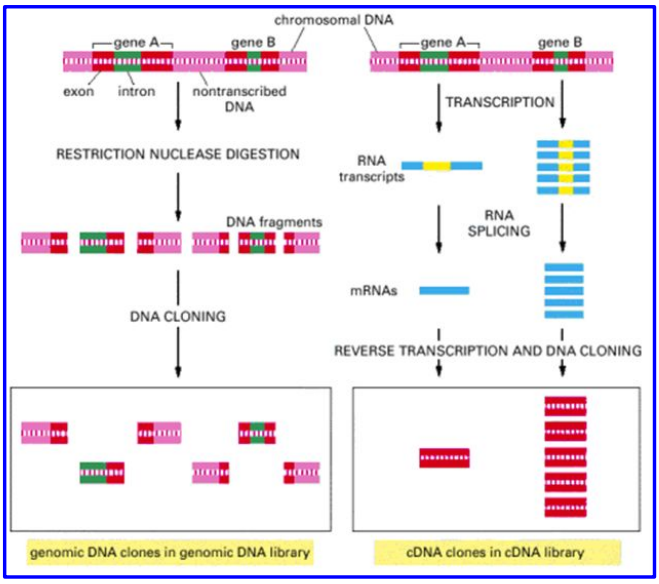


DNA Fingerprinting

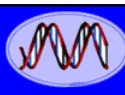
Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

Genomic Libraries vs. cDNA Libraries

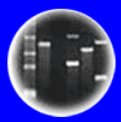


13



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

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

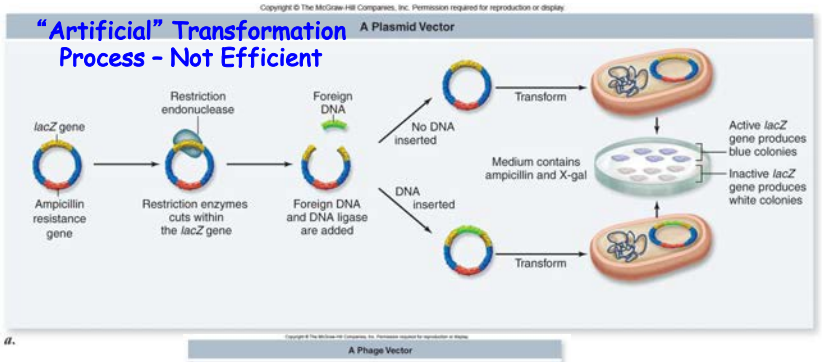
Properties of All Vectors

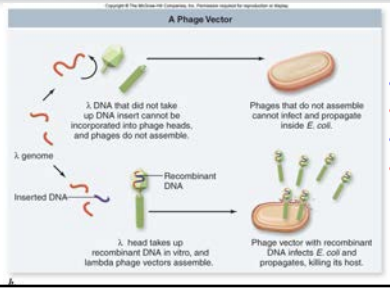
1. Replicate
2. Selectable
3. Can Be Used To Insert Foreign Genes/Restriction Sites
4. Easily Isolated + Transferred Back To Cells

14

Plasmid vs. Bacteriophage Vectors for Cloning DNA Fragments

"Artificial" Transformation Process – Not Efficient





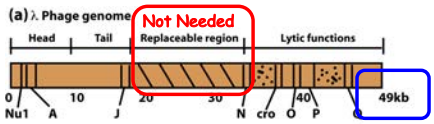
"Natural" Infection Process

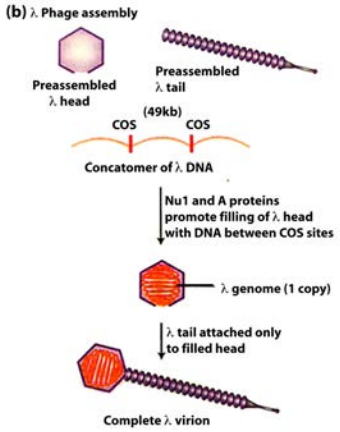
Advantages

- Much More Efficient
- Can Use Less DNA
- Get Lots More Clones
- Need Lots of Clones For Large Genome

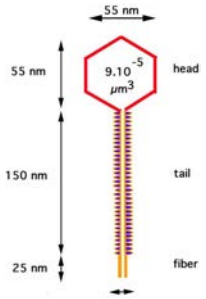
15

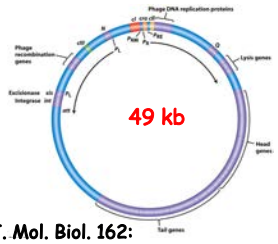
Structure of the λ Phage and Its Genome





Can Be Assembled From Parts In Vitro

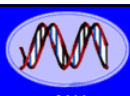




First Genome Sequence

Sanger et al. 1982 J. Mol. Biol. 162: 729-773.

16



DNA
Genetic Code of Life

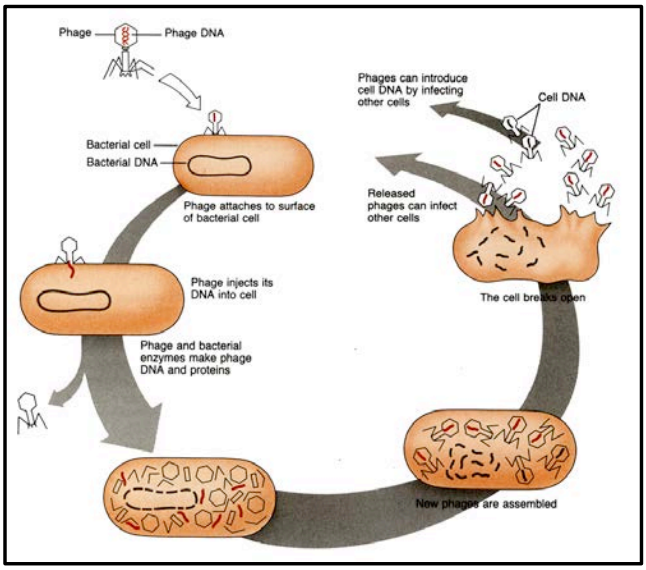
Entire Genetic Code of a Bacteria

DNA Fingerprinting

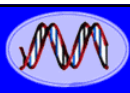
Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

λ Phage Infects *E.coli* & Destroys (Lyses) Cells

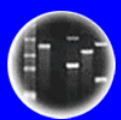


17



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

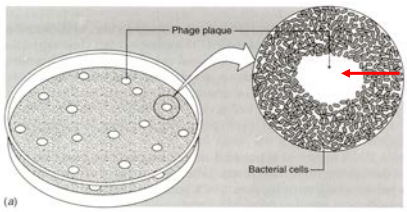


DNA Fingerprinting

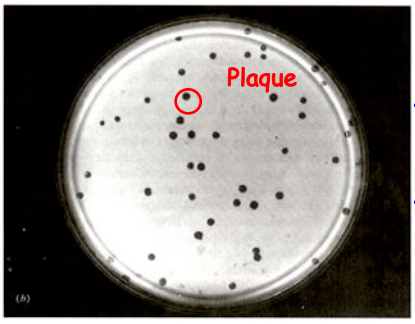
Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

Lysed Cells Can Be Seen as Clear Plaques on Agar Plates



Clear Plaque
Virus Particles
+ Dead Bacteria Cells



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

2. Selectable Marker is Bacterial Cell Destruction & Plaque Formation

18

Advantages of λ 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!

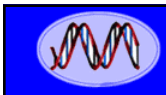
∴ Make Recombinant Viruses In The Lab!

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.

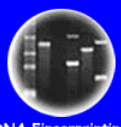
∴ Set More Clones Per Amount Of Recombinant DNA!

19



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

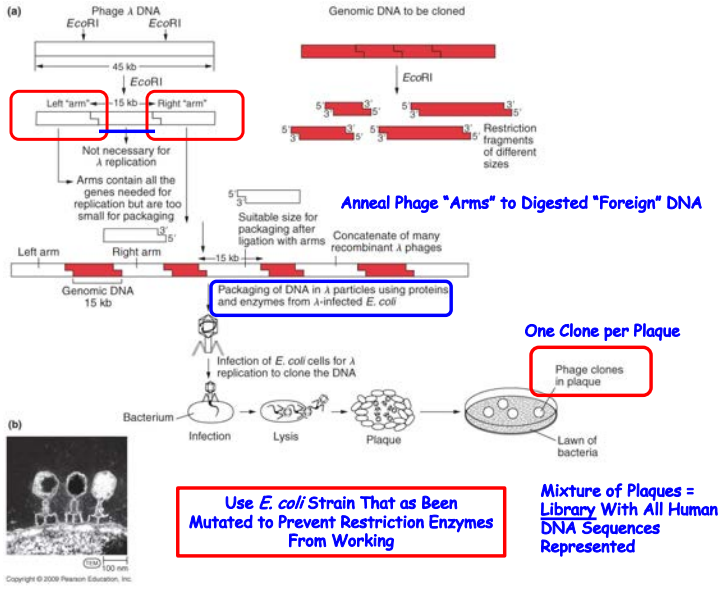


DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

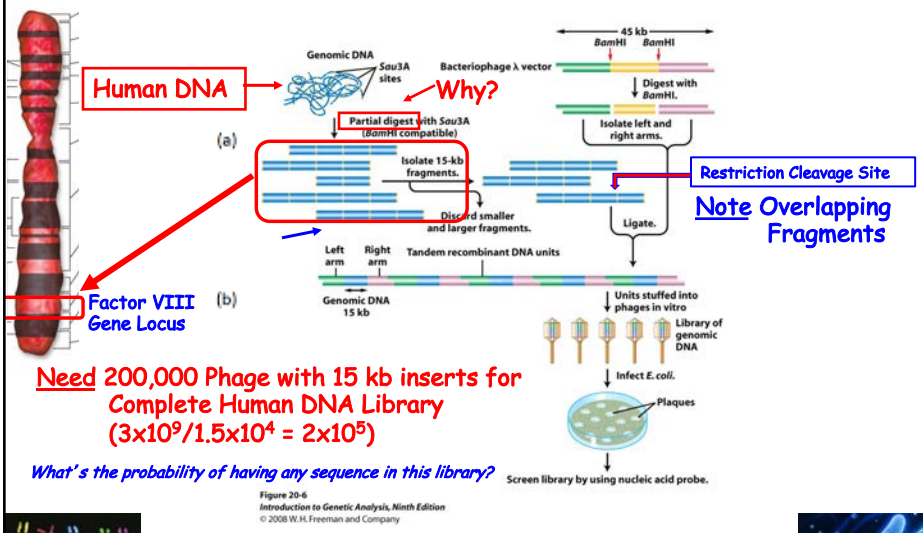
Plants of Tomorrow

Using a Bacterial Virus To Clone the Human Genome



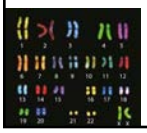
20

Cloning the Human Genome and Screening for the Factor VIII Gene



Need 200,000 Phage with 15 kb inserts for Complete Human DNA Library
($3x10^9/1.5x10^4 = 2x10^5$)

What's the probability of having any sequence in this library?



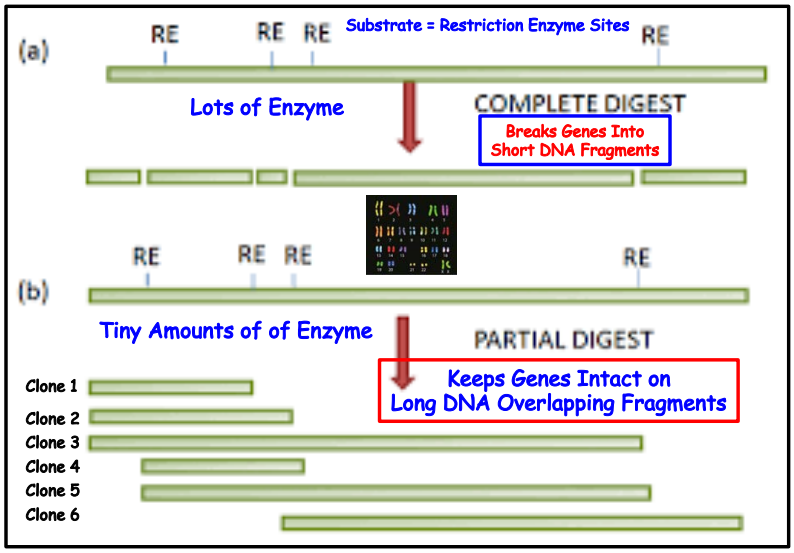
Why Partial Digestion? An Important Concept!

What is Complete & Partial Digestion?

21

Partial Restriction Enzyme Digestion Permits "Walking" From One DNA Region to the Next

Iterative Process of Screening & Rescreening Human Genome Library



22

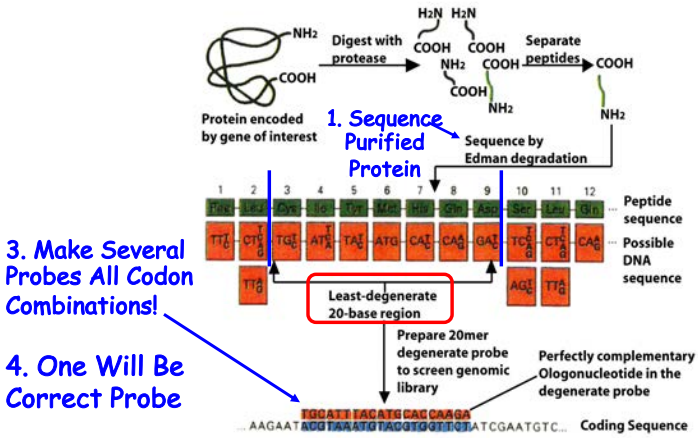
Step Two

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

23

Factor VIII Protein → Gene

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



3. Make Several Probes All Codon Combinations!

4. One Will Be Correct Probe

2. Use Genetic Code

How many Combinations of Synthetic Probes?

2x3x2x1x2x2x2=96

24

Finding The Factor VIII Gene Or Part of Gene!!

Mike Grunstein-UCLA

25

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 ∴ 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!

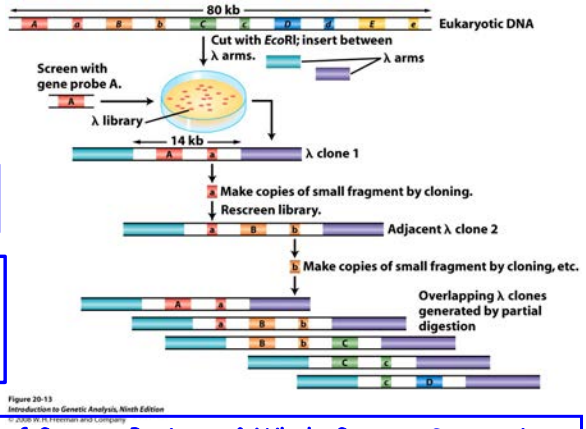
Sequence ----------> GenBank

26

Step Three

Finding the Entire Factor VIII Gene? Walking & Sequencing

Walking Up and Down Genes and Chromosomes



Iterative Library Screening Process

Find Overlapping Clones By Restriction Site Mapping

Basis of Genome Projects & Whole Genome Sequencing

Key Concepts → How know Find Complete Factor VIII Gene?

→ Compare Protein & DNA Sequences

27

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

28

Factor VIII SNP Mutations Occur Throughout the Gene

[*Haemophilia* 11, 481-491 (2005)] *Larger the Gene – Larger Number of Mutations!*

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

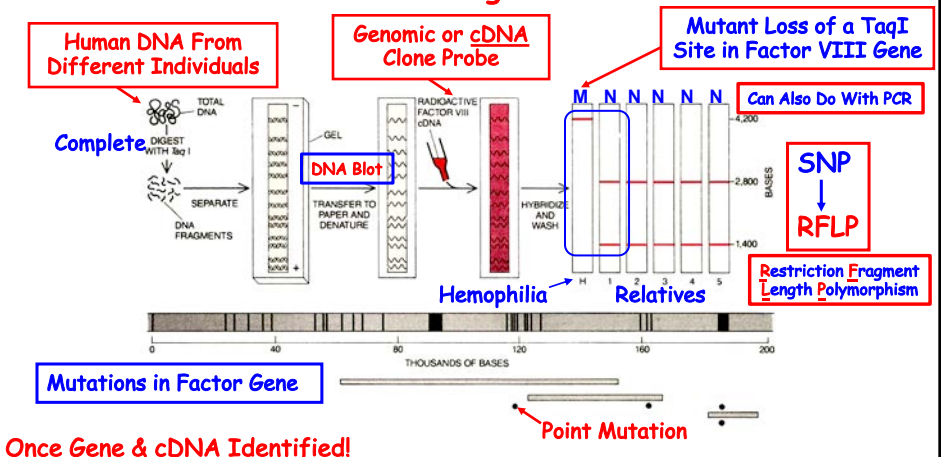
VIII:C (%)	Family history	Consanguinity*	Inversion	Codon†	Mutation	Amino acid change	Exon	Conservation‡
1	Sporadic	NC	Normal	51	TTT → TCT§	Phe → Ser	2	FFFF, identical
1.20	Sporadic	NC	Normal	80	GTT → GAT	Val → Asp	3	VVVV, identical
1	Sporadic	NC	Normal	102	GGT → GTT§	Gly → Val	3	GGGG, identical
2	Sporadic	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	Inerited	C	Normal	1392	1392del1418§	Pro → fs (TAG-1446)	14	
1	Sporadic	NC	Normal	1441	insA§		14	
1	Inerited	C	Normal	1441	insA§			
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
hibitor 96 BU								
1	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	1833	TGG → TGA§	Trp → Stop	16	WWWW, identical
2.60	Sporadic	C	Normal	1882	ATC → ATA§	Ile → Ile	17	IIII, identical
1	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

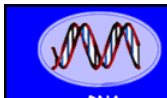
Need To Screen Across the Gene for Markers -- Family Specific

29

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



30


DNA
Genetic Code of Life

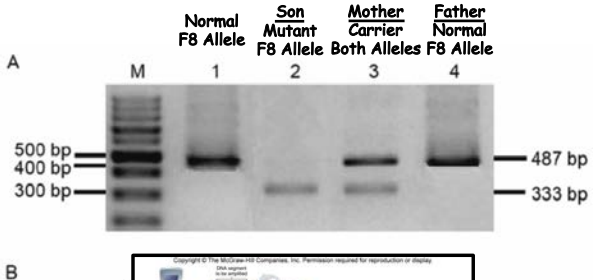

Entire Genetic Code of a Bacteria


DNA Fingerprinting

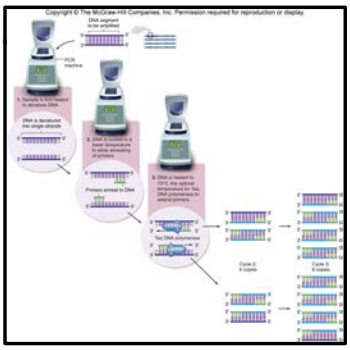

Cloning: Ethical Issues and Future Consequences


Plants of Tomorrow

Using PCR to Screen For Hemophilia Alleles

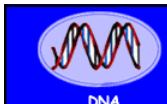


B

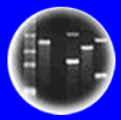


Mutations Family Specific
Different Test For Each Mutation

31


DNA
Genetic Code of Life


Entire Genetic Code of a Bacteria


DNA Fingerprinting


Cloning: Ethical Issues and Future Consequences


Plants of Tomorrow

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!

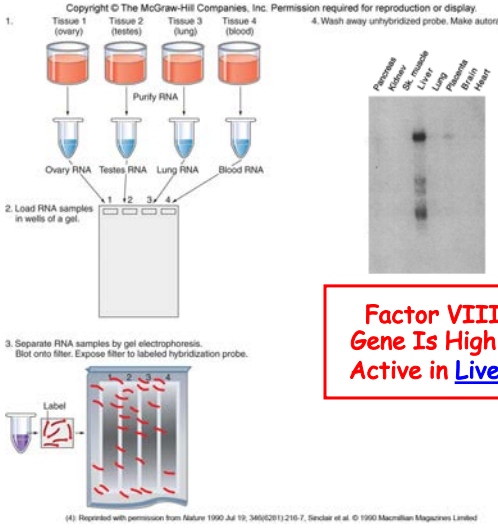
Introns! Switches!

32

Making the Drug

Need cDNA Not Gene

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



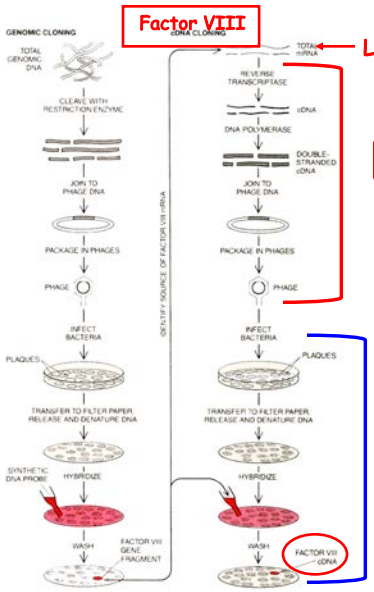
RNA Blot Is Like a DNA Blot Except That RNA is on Gel & Blotted

Factor VIII Gene Is Highly Active in Liver!

Can Also Use PCR (RT-PCR)

33

Using Factor VIII Gene Probe to Identify Factor VIII cDNA clone



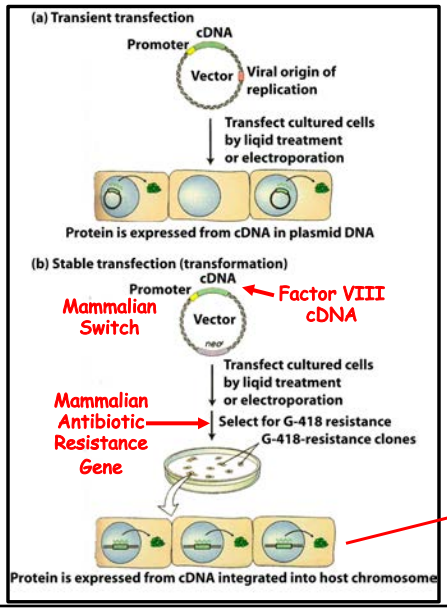
Construct Liver cDNA Library

Factor VIII cDNA → Used to Make Drug! Not Gene!!

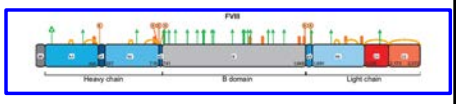
Screen Liver cDNA Library

34

Engineer Factor VIII cDNA to Produce Protein in Host Cell & Synthesize Factor VIII in Mammalian Cells



Why Mammalian Cells?



Purify Factor VIII Protein!

35

Recombinant Factor VIII

Prophylactic Treatment Costs $300,000/Year! Most Hemophiliacs Use "On Demand" or As Needed

Factor VIII Gene Cloned In 1983

Factor VIII (Recombinant) Approved As Drug In 1993! Ten Years From Gene → Drug! (Off Patent In 2011)

36

The Future: Gene Therapy – A Permanent "Cure"

December 10, 2011

Treatment for Blood Disease Is Gene Therapy Landmark

By NICHOLAS WADE

TIME Partners with CNN

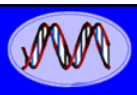
Gene Therapy Shows Promise for Treating Hemophilia

By ALICE PARK Monday, December 12, 2011

The First Ever In-Human Gene Editing Will Try and Combat Hemophilia

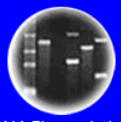
Factor IX – Hemoglobin B
FDA-Approved Clinical Trial
2016

37



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)

38