Genetic Basics and Engineering

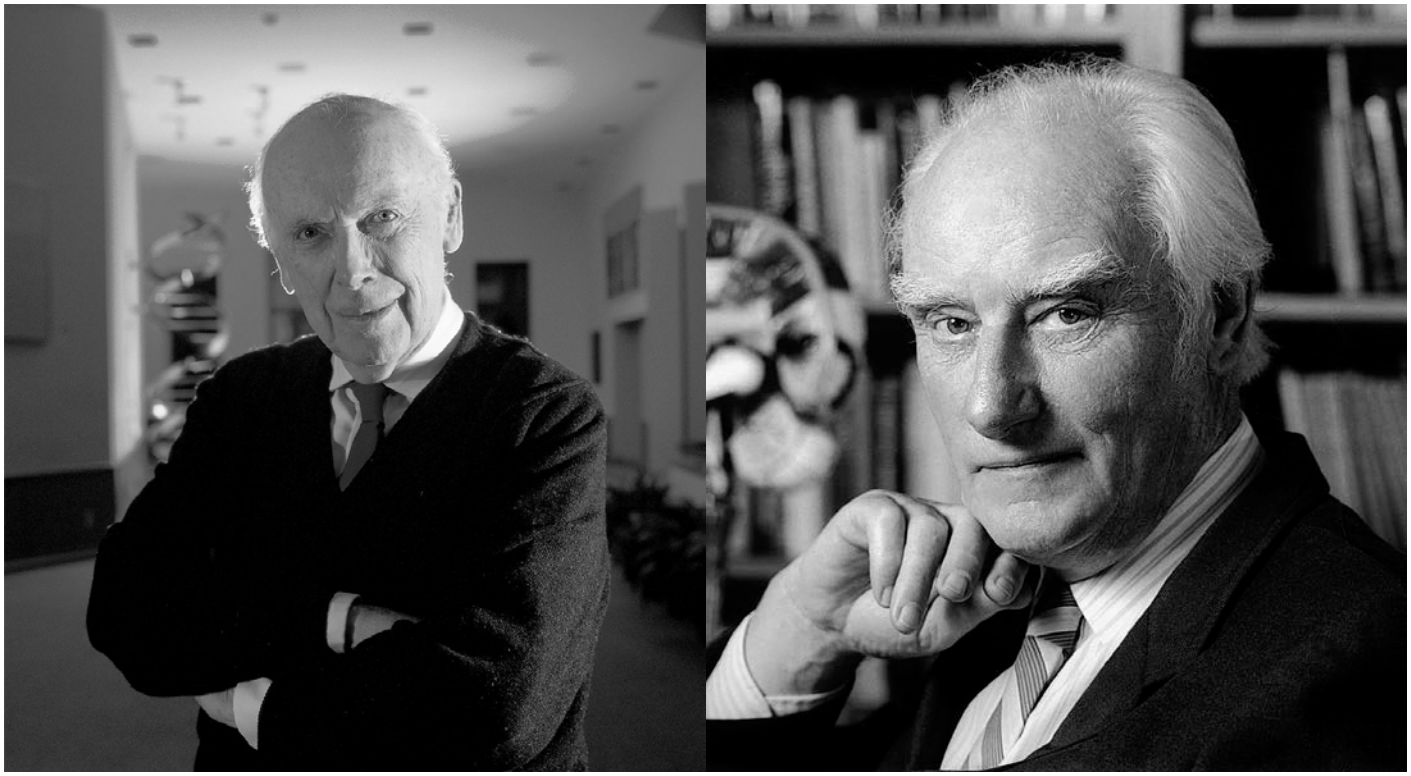
BME2105

The Discovery of structure of DNA

- **Deoxyribonucleic acid (DNA)**

The Discovery of structure of DNA

- **Deoxyribonucleic acid (DNA)**



The code of Life and Health

Disease	Location of Defect	Incidence	Treatment
	Enzyme involved		
Lesch–Nyhan syndrome	Hypoxanthine–guanine phosphoribosyltransferase	1 in 100,000–380,000 (males only)	
Adenosine deaminase deficiency	Adenosine deaminase		Enzyme infusion; gene therapy
Gaucher's disease	Lysosomal glucocerebrosidase		Enzyme infusion
Phenylketonuria	Phenylalanine hydroxylase		Diet
Sickle cell anemia	Hemoglobin	0.4% of African-American males	
Hemophilia	Factor VIII	1 in 10,000 males	Blood transfusion; protein infusion
	Membrane protein involved		
Cystic fibrosis	Cystic fibrosis transmembrane conductance receptor (CFTR)	1 in 2000 live births in Caucasians	Gene therapy not yet successful
Disaccharidase deficiency (lactose intolerance)	Lactase	Frequent	Diet
Duchenne's muscular dystrophy	Dystrophin gene product	1 in 3500 males	None
	Protein involved		
Huntington's disease	Huntington		None

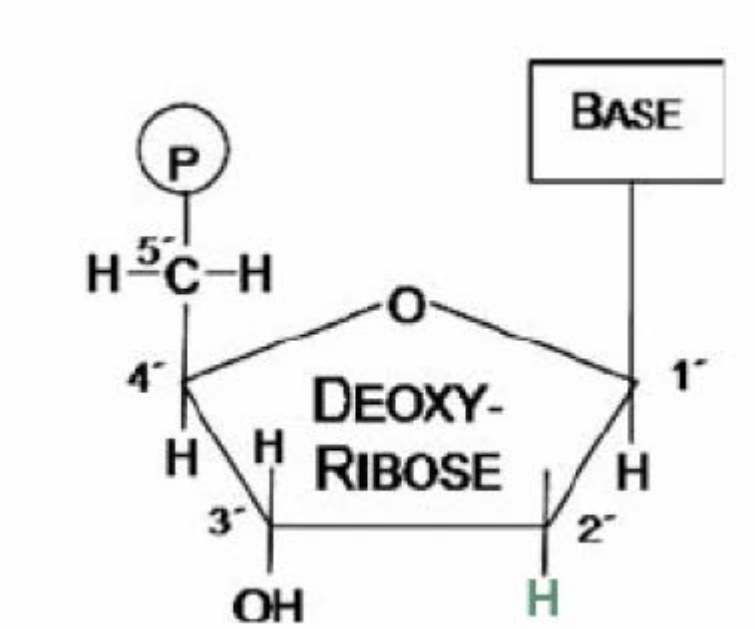
The code of Life and Health

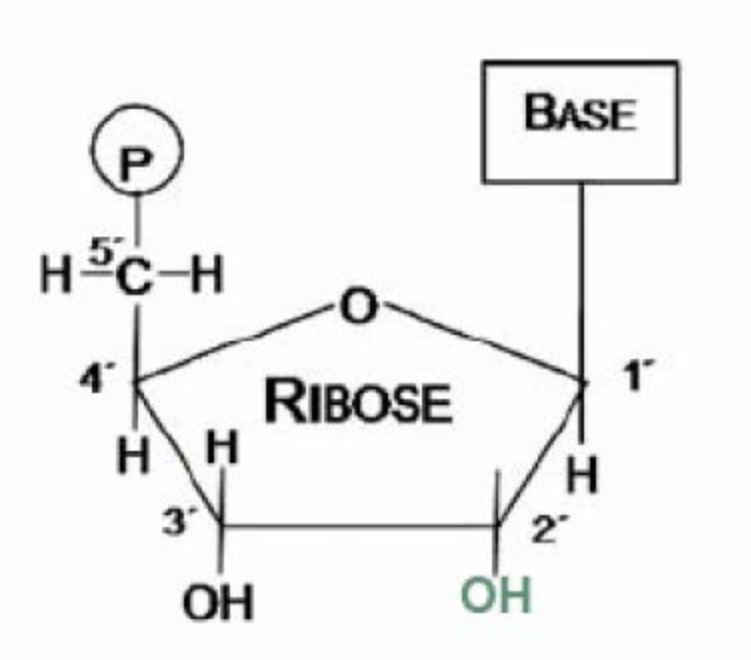
Disease	Location of Defect	Incidence	Treatment
	Enzyme involved		
Lesch–Nyhan syndrome	Hypoxanthine–guanine phosphoribosyltransferase	1 in 100,000–380,000 (males only)	
Adenosine deaminase deficiency	Adenosine deaminase		Enzyme infusion; gene therapy
Gaucher's disease	Lysosomal glucocerebrosidase		Enzyme infusion
Phenylketonuria	Phenylalanine hydroxylase		Diet
Sickle cell anemia	Hemoglobin	0.4% of African-American males	
Hemophilia	Factor VIII	1 in 10,000 males	Blood transfusion; protein infusion
	Membrane protein involved		
Cystic fibrosis	Cystic fibrosis transmembrane conductance receptor (CFTR)	1 in 2000 live births in Caucasians	Gene therapy not yet successful
Disaccharidase deficiency (lactose intolerance)	Lactase	Frequent	Diet
Duchenne's muscular dystrophy	Dystrophin gene product	1 in 3500 males	None
	Protein involved		
Huntington's disease	Huntington		None

DNA Structure

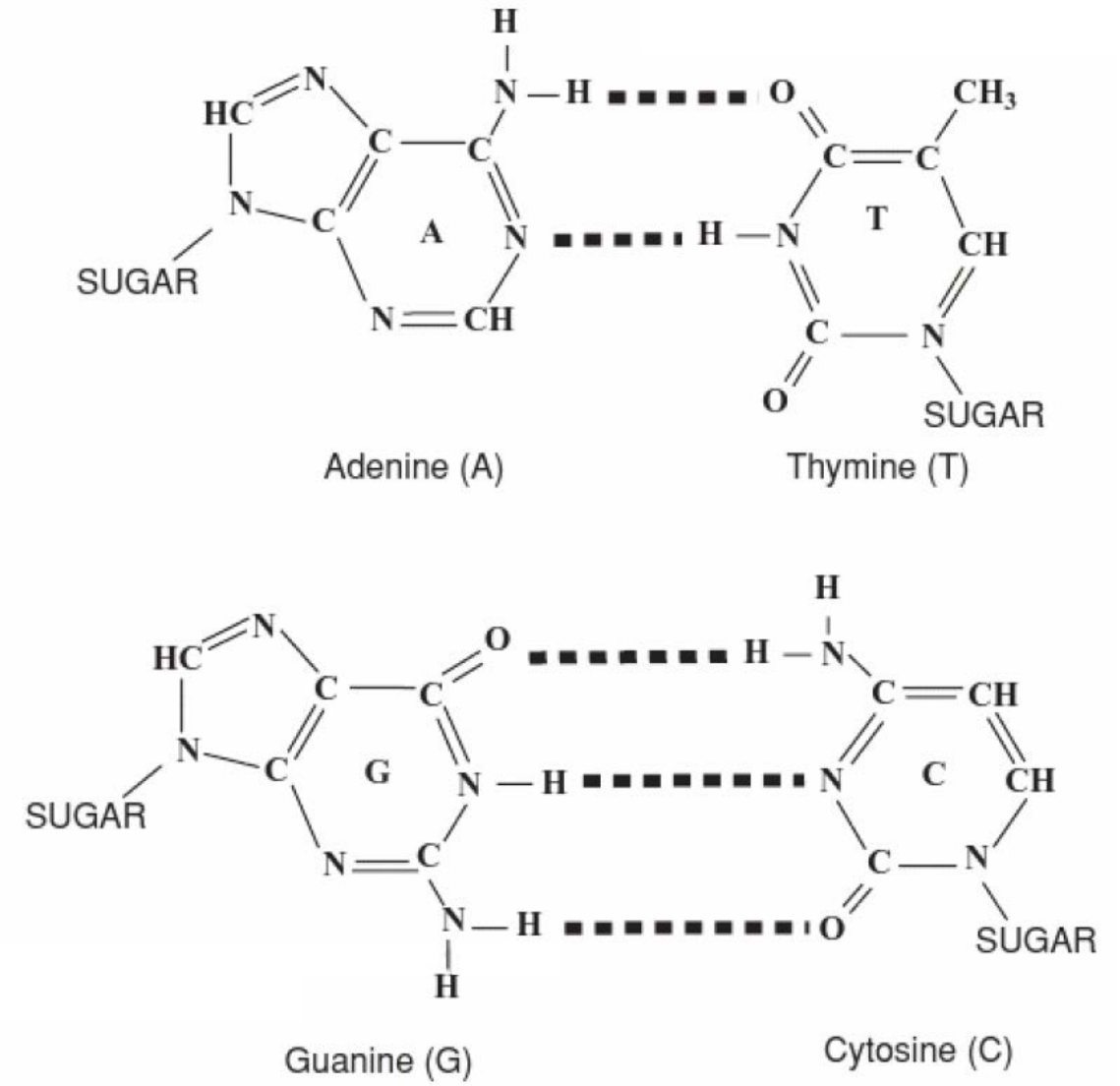
DNA/RNA Structure

- DNA are linear polymers of nucleotides
- A gene is a segment of neucleotides that codes for one polypeptide chain.

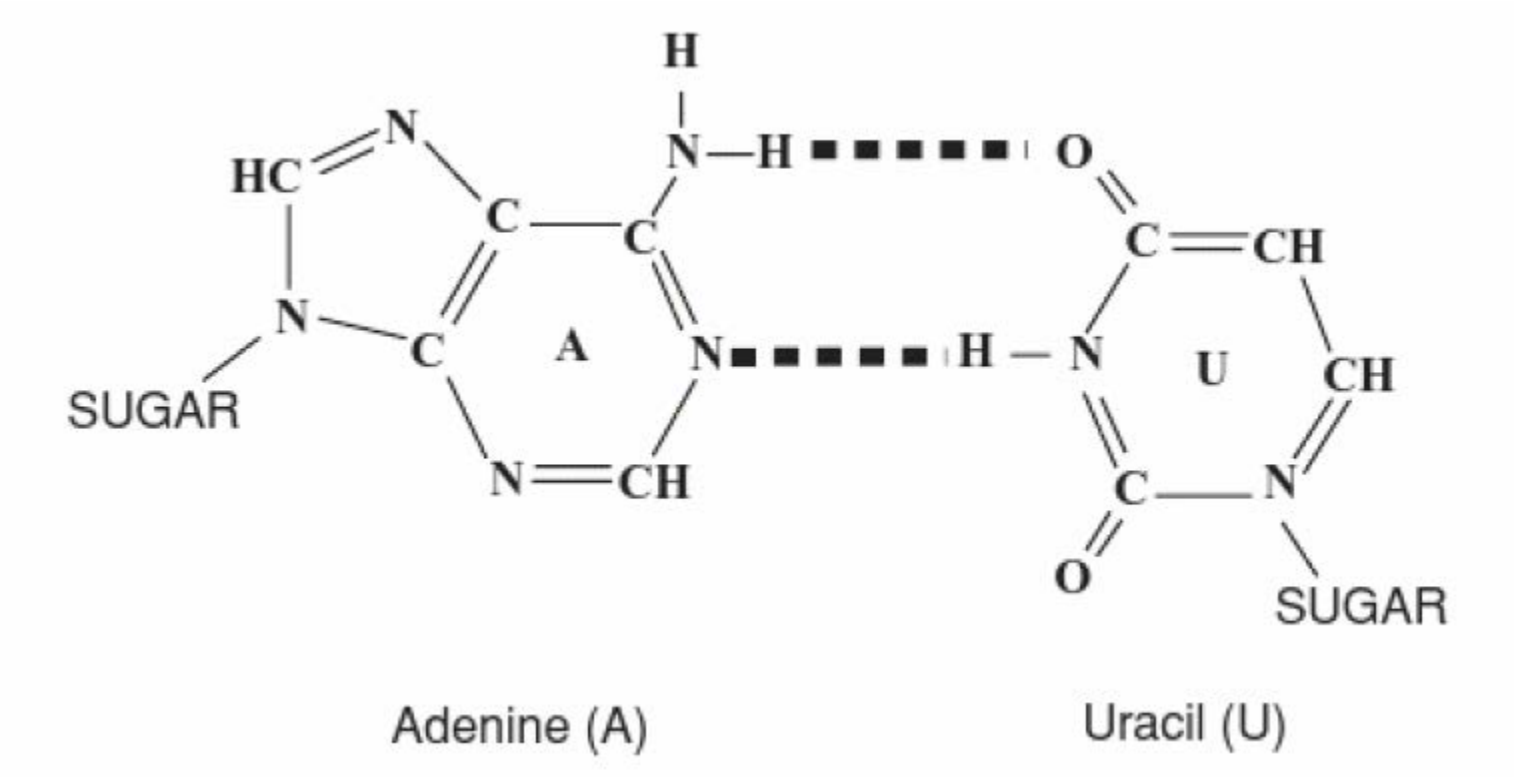




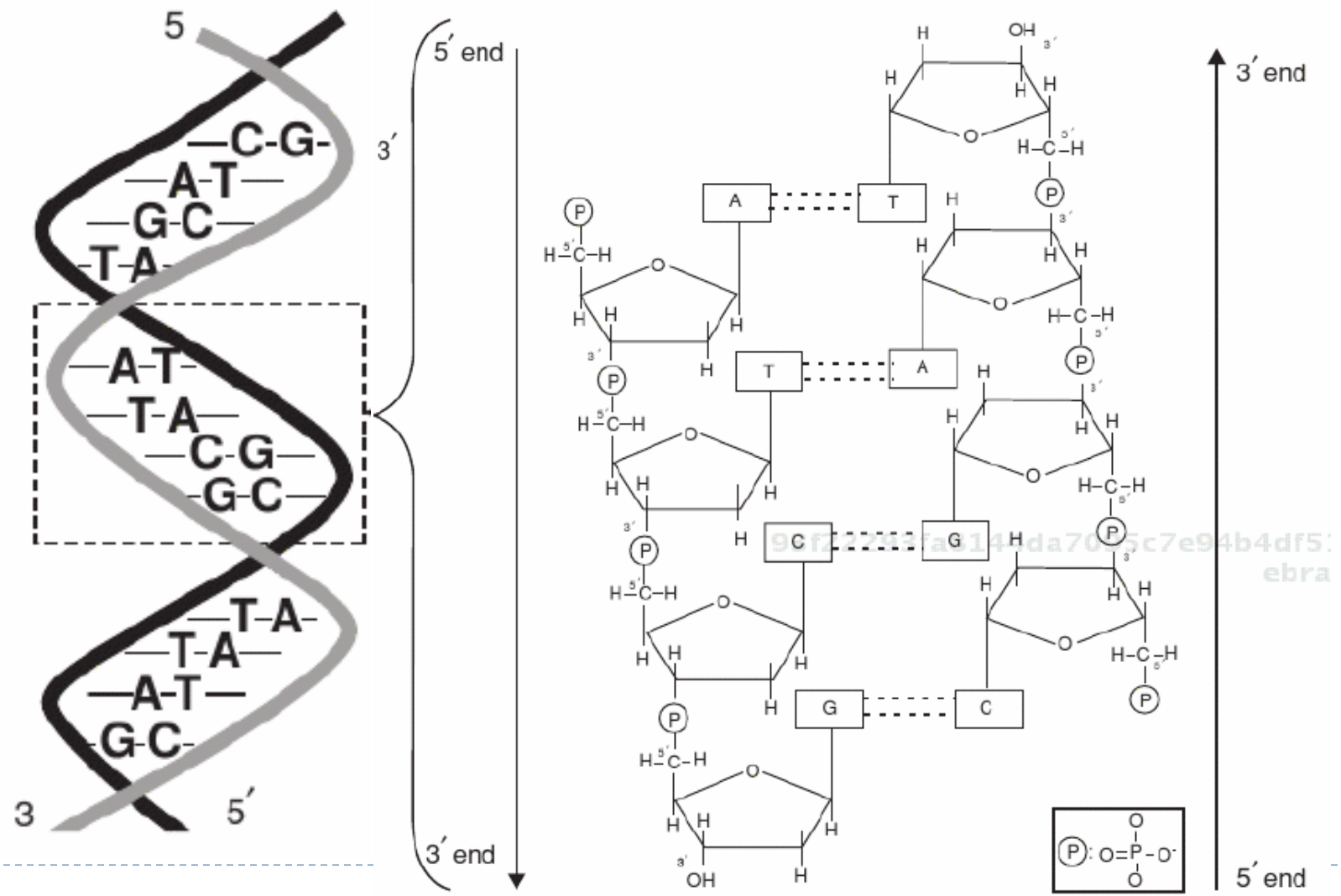
DNA Structure



RNA Structure

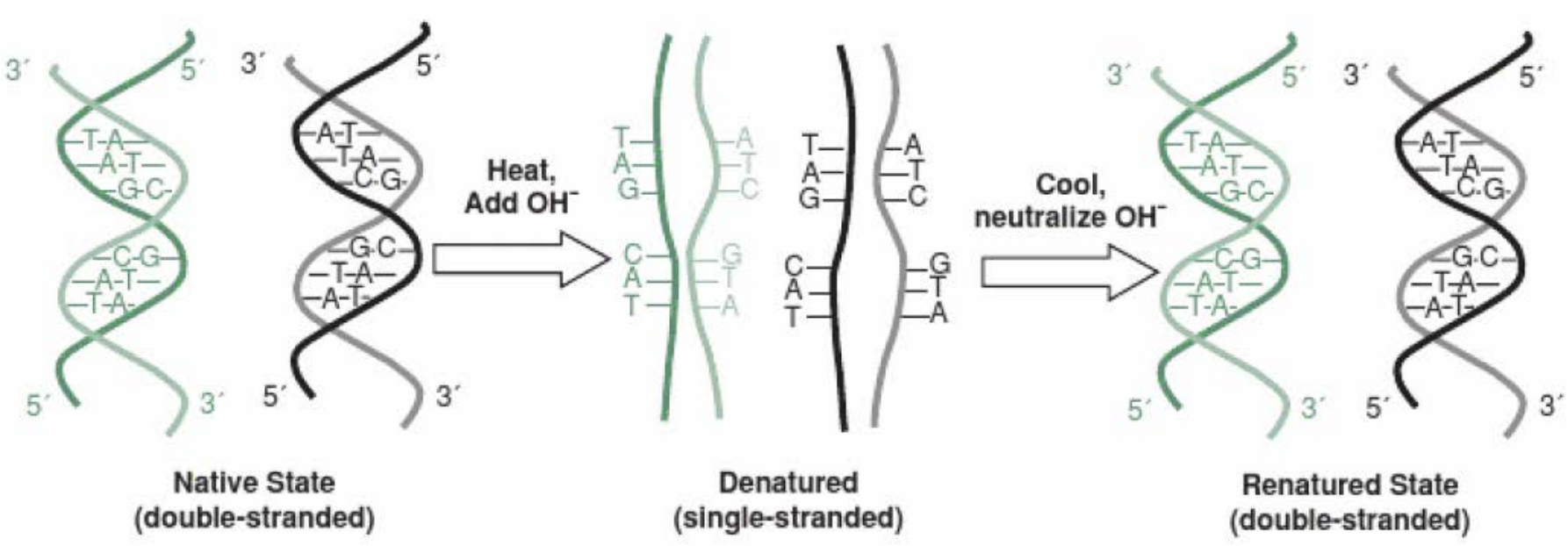


DNA Structure



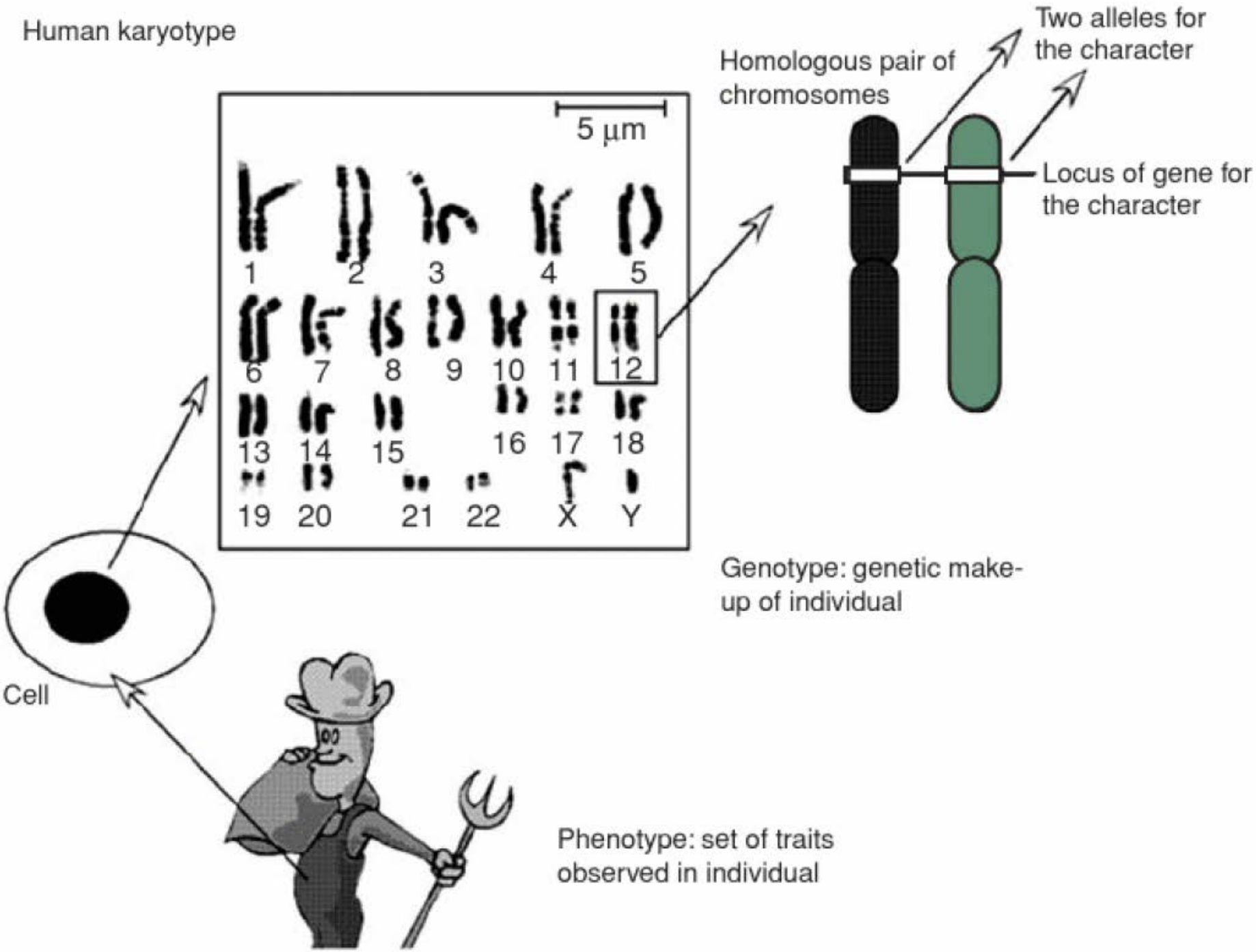
DNA Structure

- The two strand of DNA in a double helix are held together by Hydrogen bonds between base pairs.
- The hydrogen bonds can be disrupted (denatured) to separate the two strands.



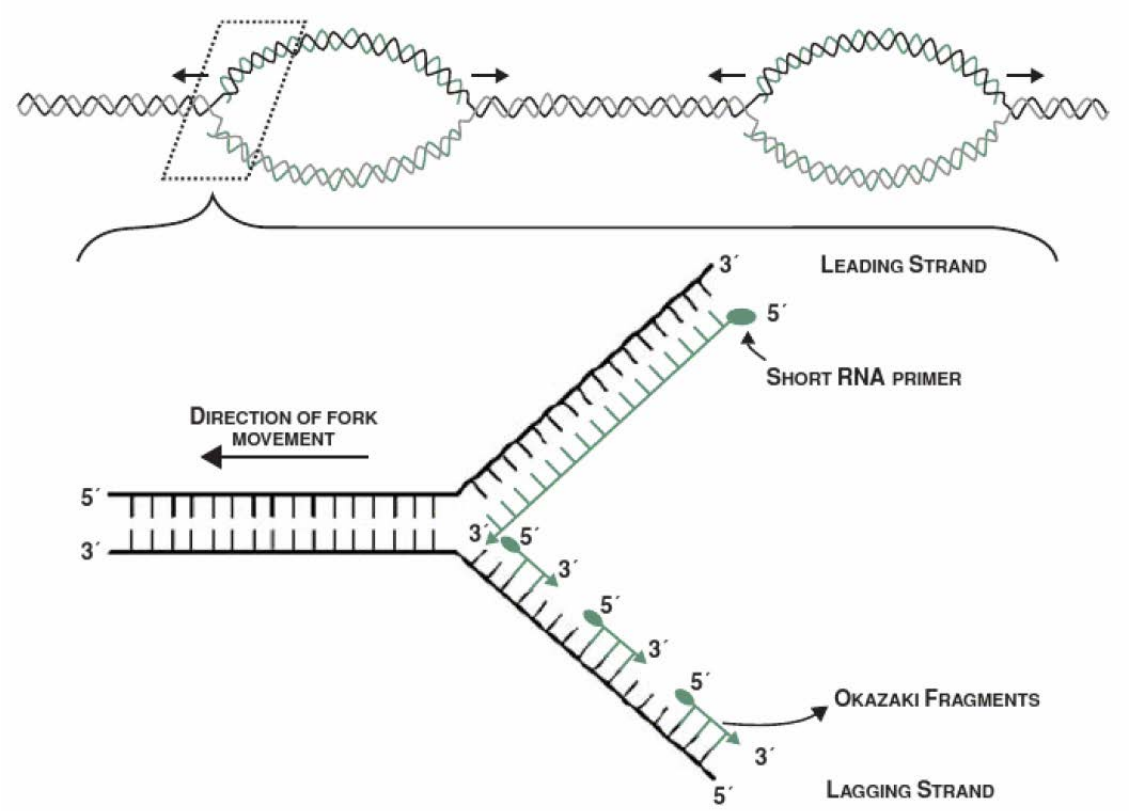
Chromsomes

- Chromosomes are strands of DNA packed with the nucleus of most cells.
  - Nearly 2m of DNA is compacted into a nucleus that is 5-10 μm in diameter.
  - The paired chromosomes are called homologous chromosomes
  - Genes are a segment of DNA sequence on the chromosome that encode specific proteins.


Human karyotype
5 µm
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
Homologous pair of chromosomes
Two alleles for the character
Locus of gene for the character
Genotype: genetic make-up of individual
Cell
Phenotype: set of traits observed in individual

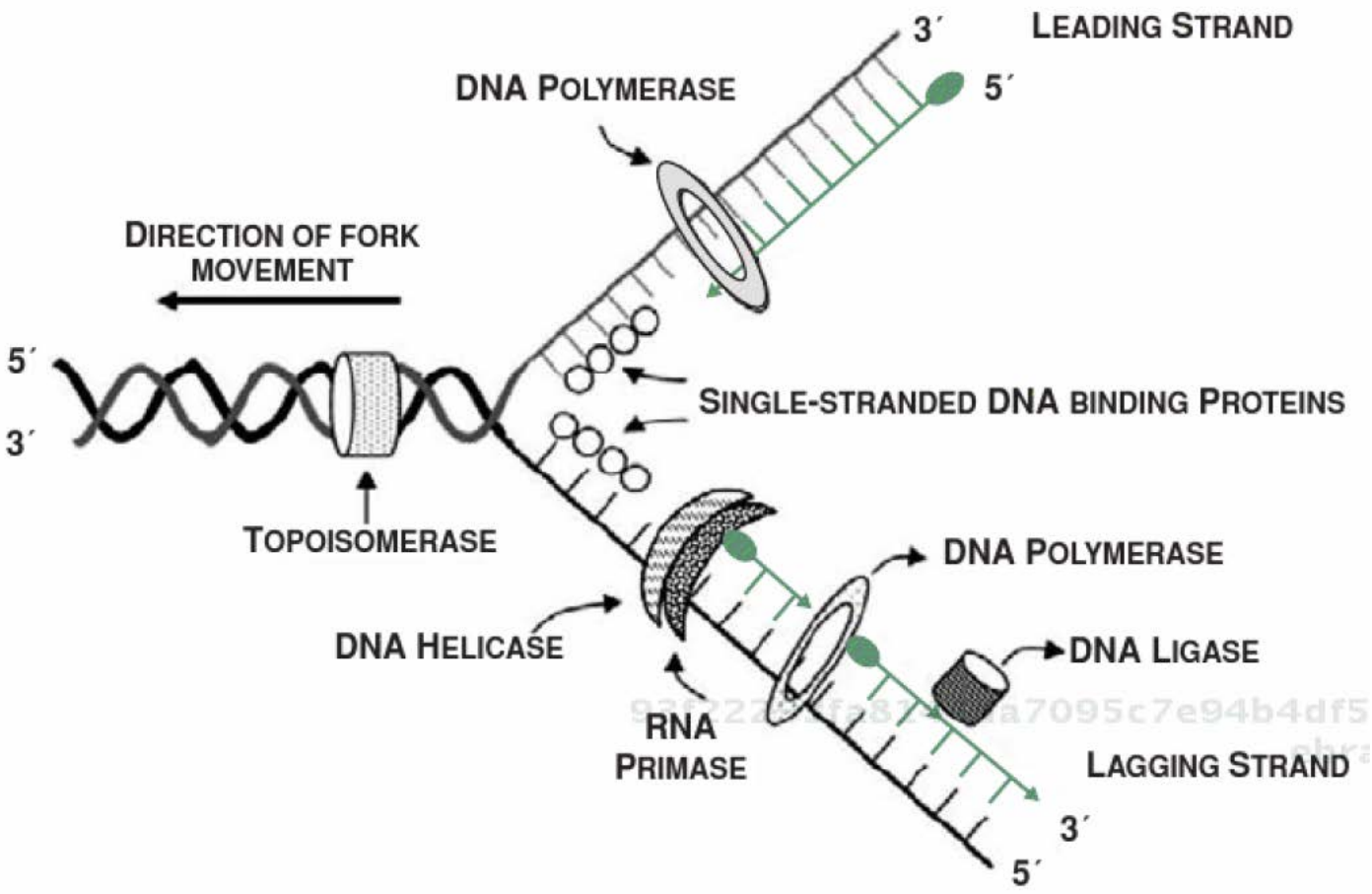
DNA Replication

- How is double-stranded DNA copied to ensure that daughter cells receive an exact version of the parent DNA?

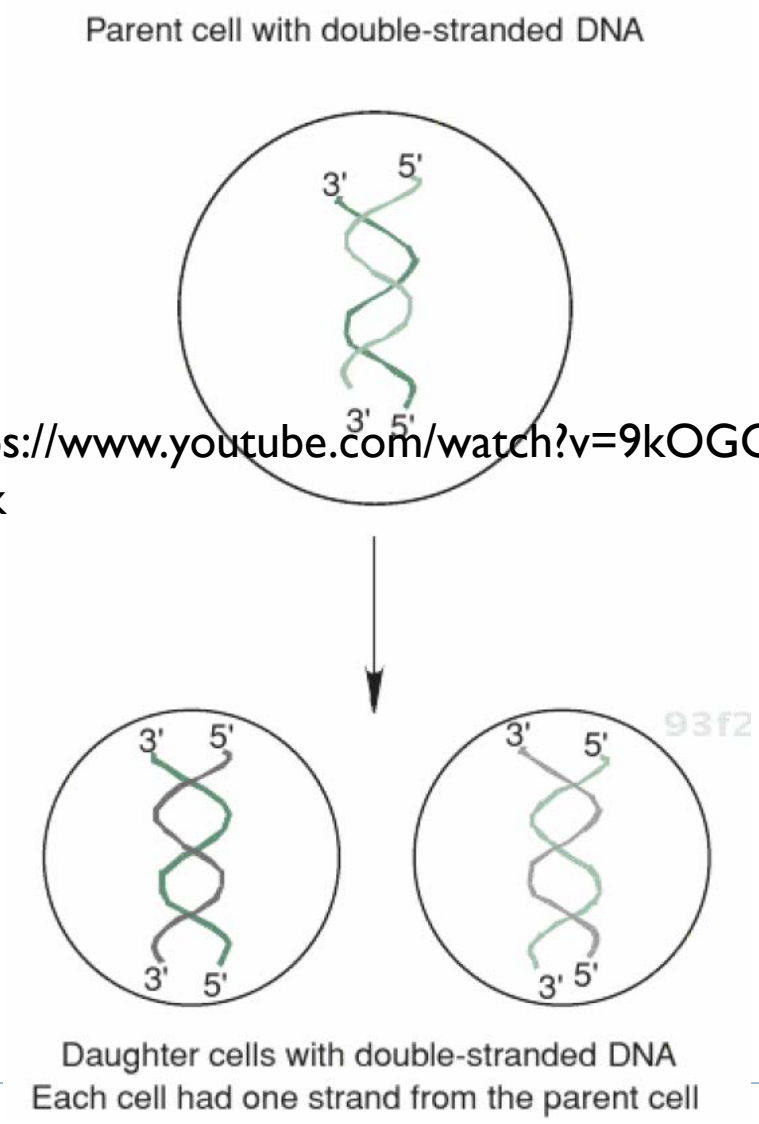


DNA Replication

- How is double-stranded DNA copied to ensure that daughter cells receive an exact version of the parent DNA?

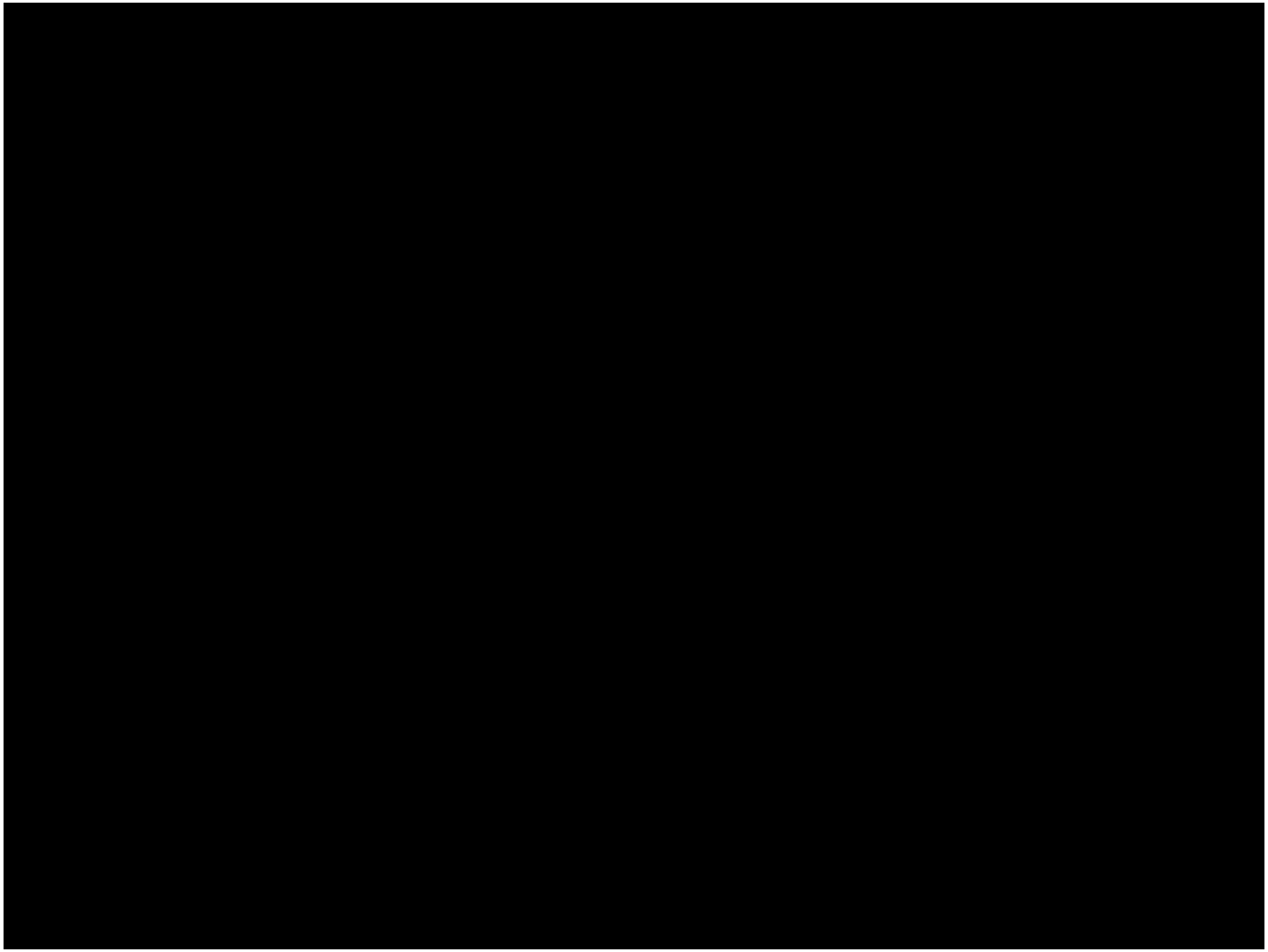


DNA Replication

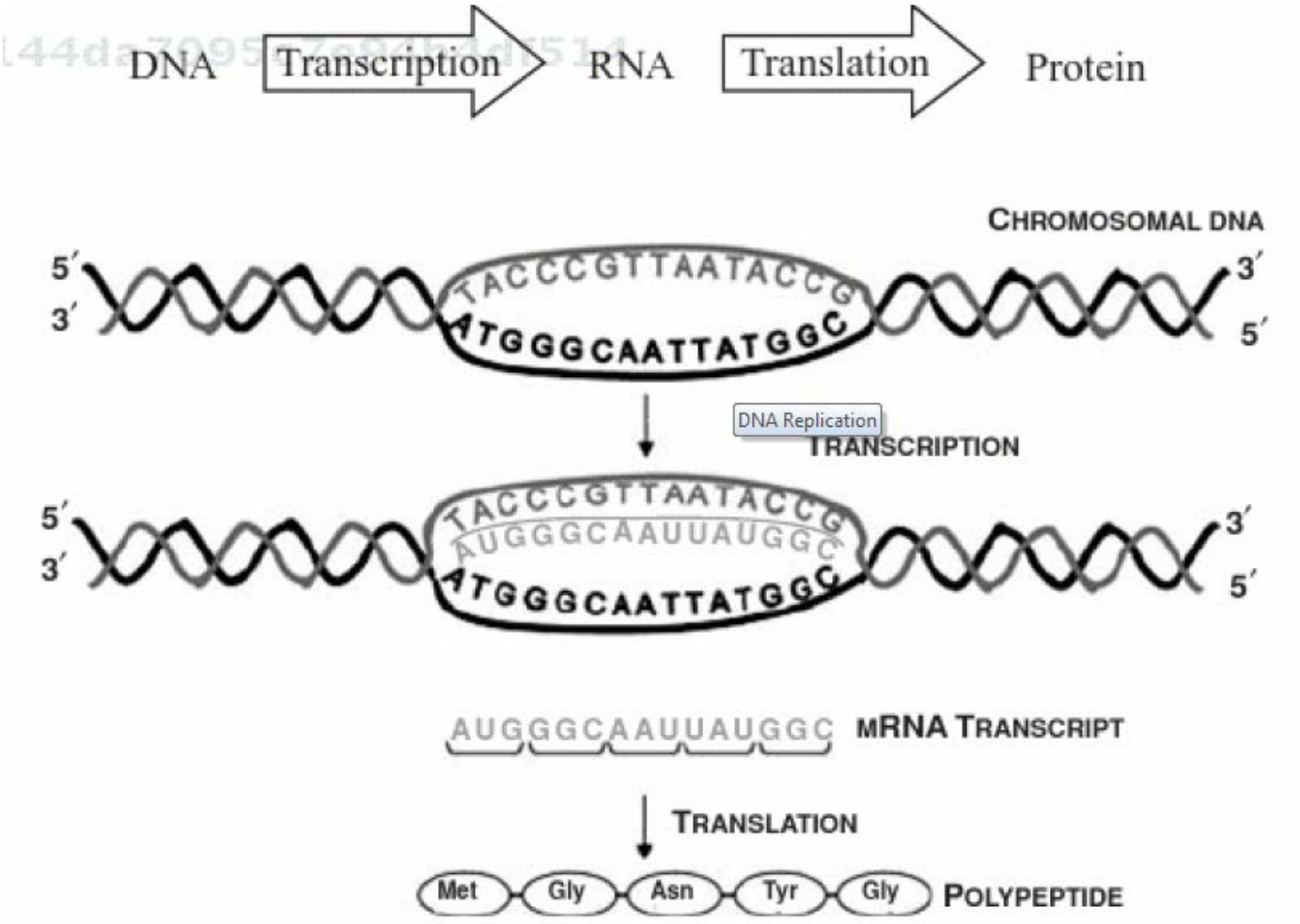


https://www.youtube.com/watch?v=9kOGOY7vthk

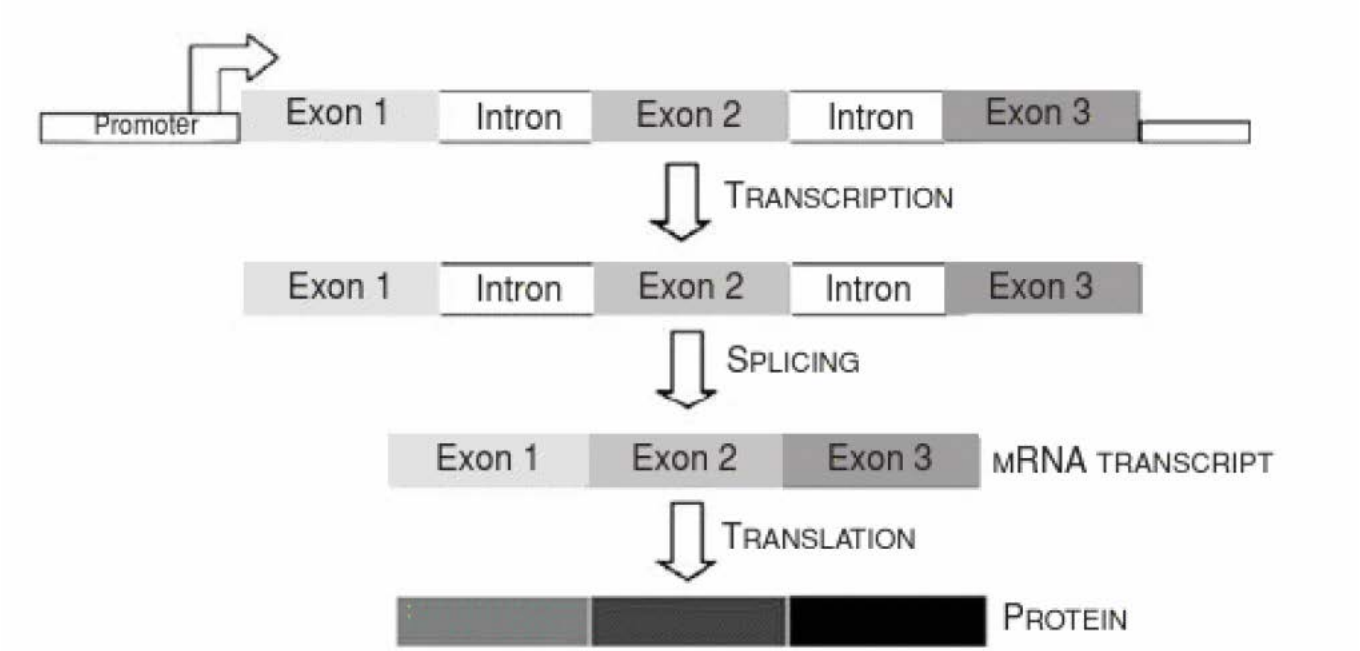
The Central Dogma: Transcription and Translation



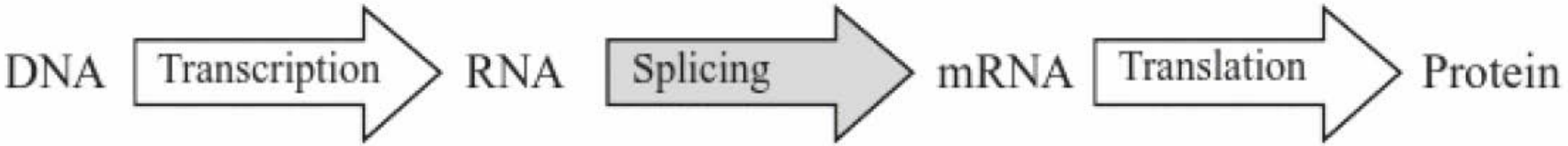
The Central Dogma: Transcription and Translation



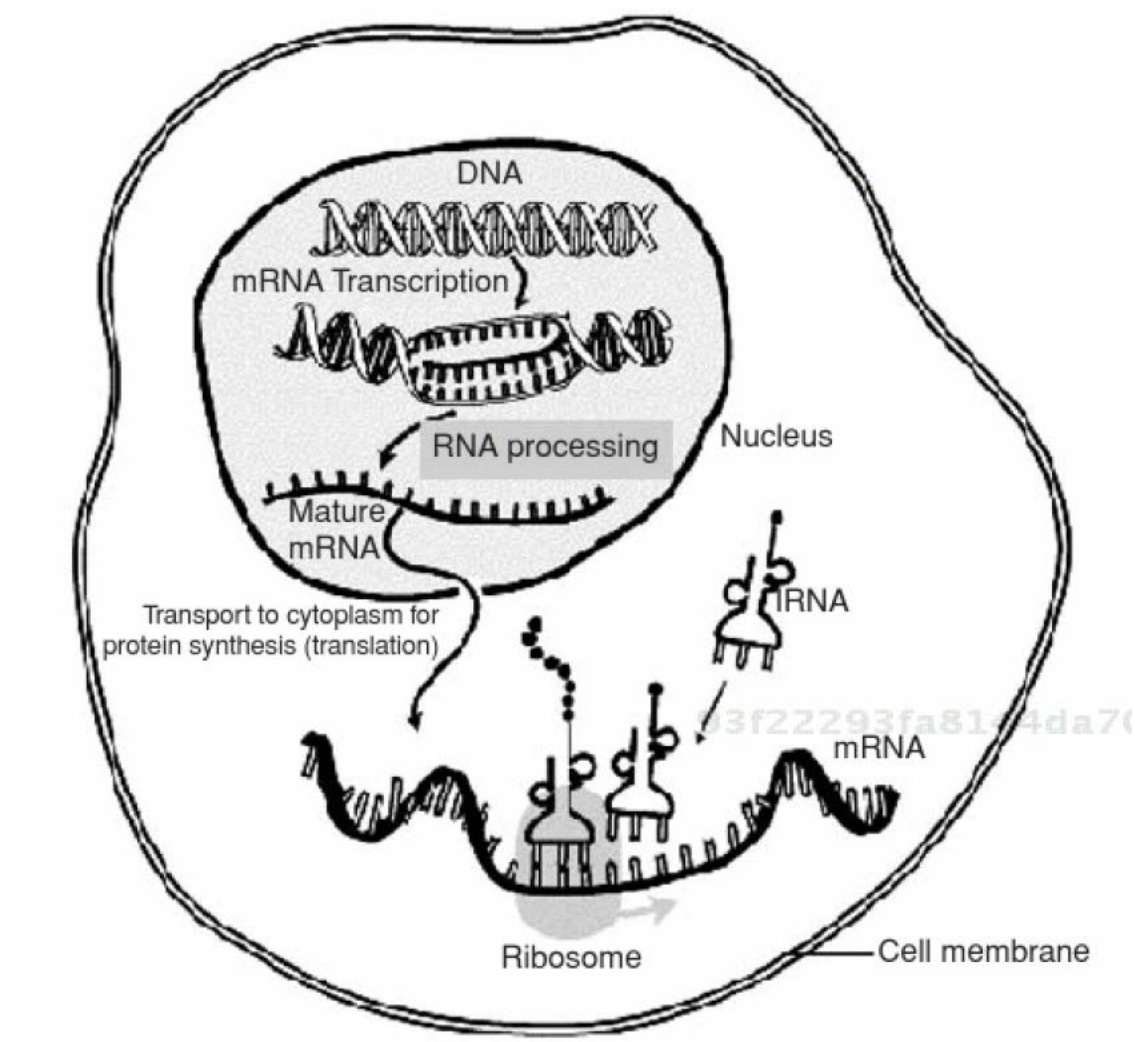
RNA Splicing



Revised dogma



mRNA translation: Protein synthesis



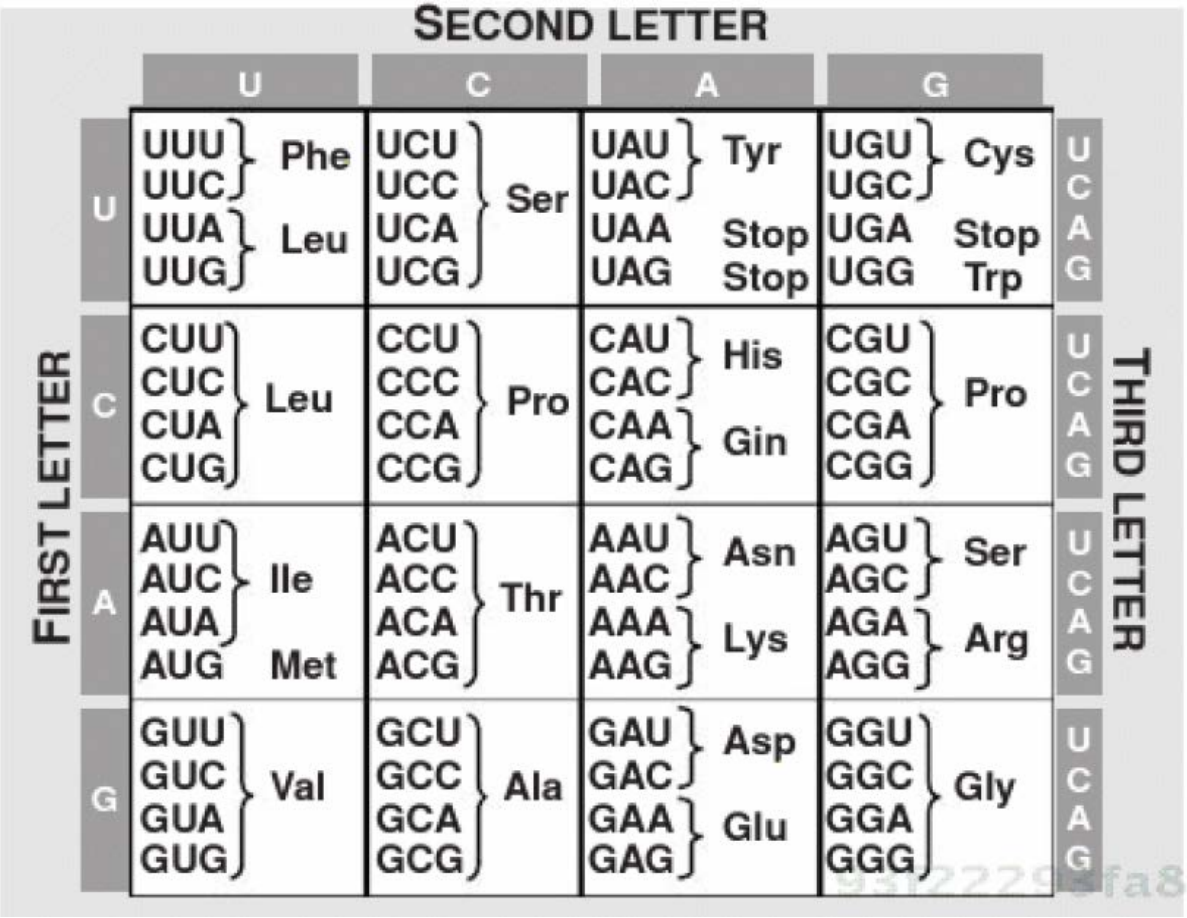
mRNA translation: Protein synthesis

- Triplet rule: three-base word in the mRNA transcript is called a **codon.**
- why not 3???

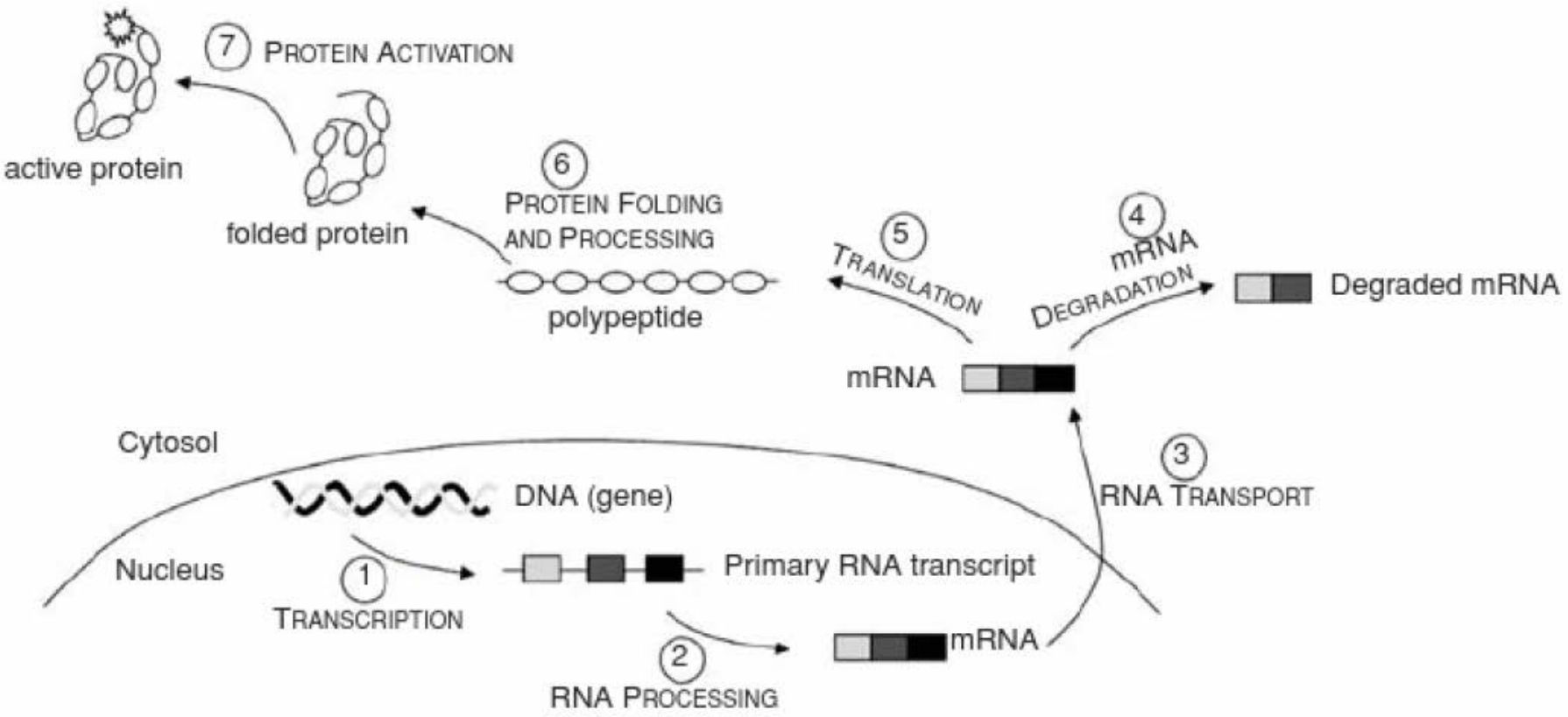
- Combination: 4x4x4 =64 possible codes
- Multiple codon for each amino acid

mRNA translation: Protein synthesis

- Triplet rule: three-base word in the mRNA transcript is called a **codon.**



Control of gene expression



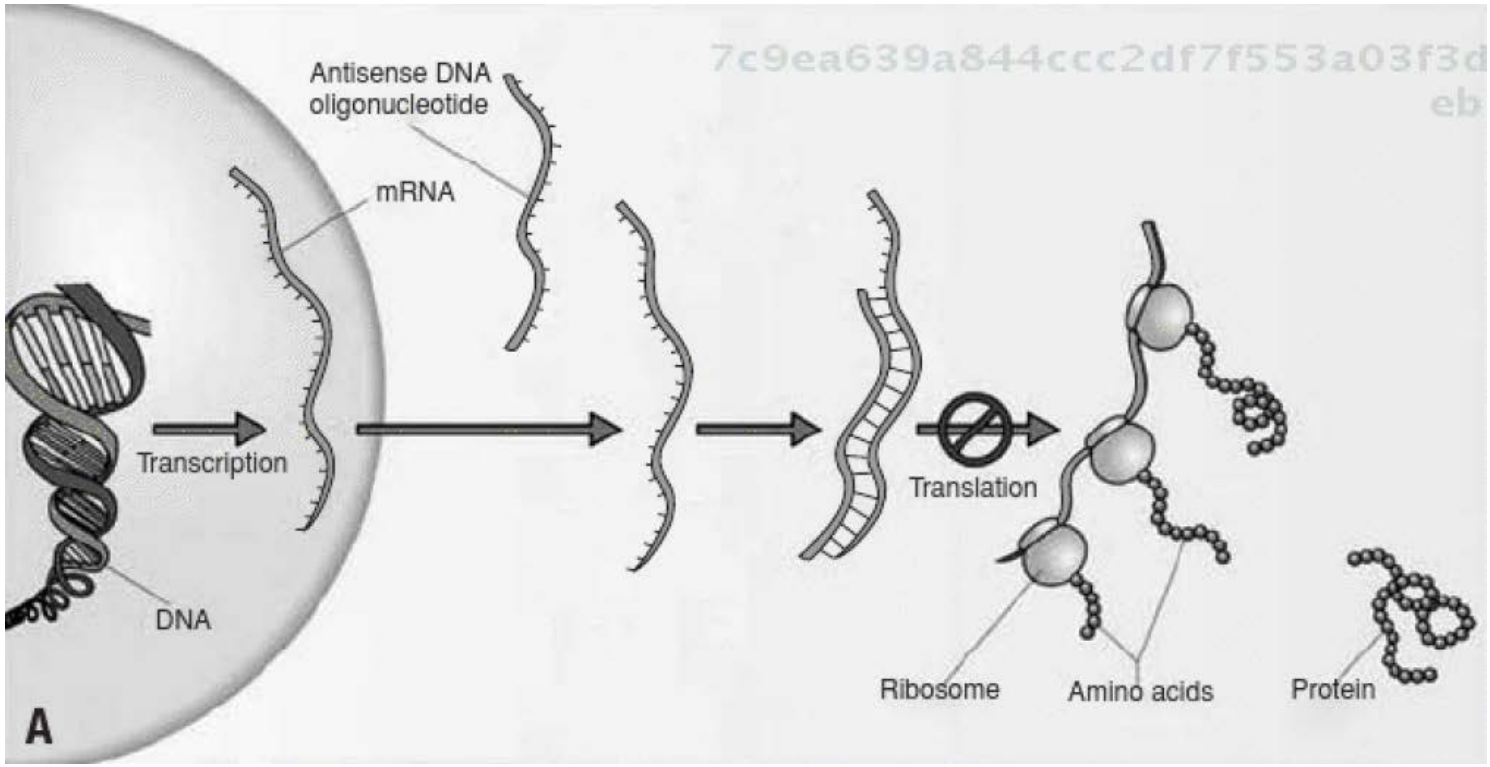
Control of gene expression

- Question:
  - All of the cells of the human body contain the same genetic information, but they are functionally different, why?
- Gene expression can be regulated at each step in the pathway of converting DNA into protein.
  - Transcription factors
  - Alternative splicing
  - Transport of mRNA
  - Stability of mRNA transcript
  - Proteins can be inactivated

Control of gene expression

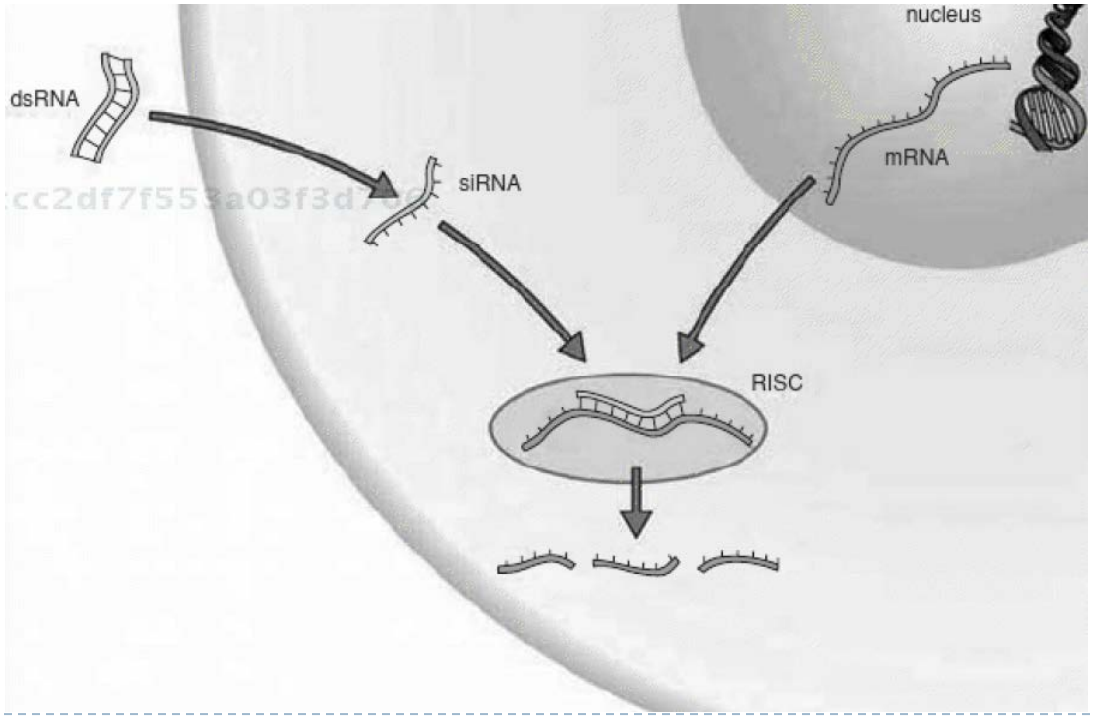
Control of gene expression (Engineering)

- Antisense nucleic acid



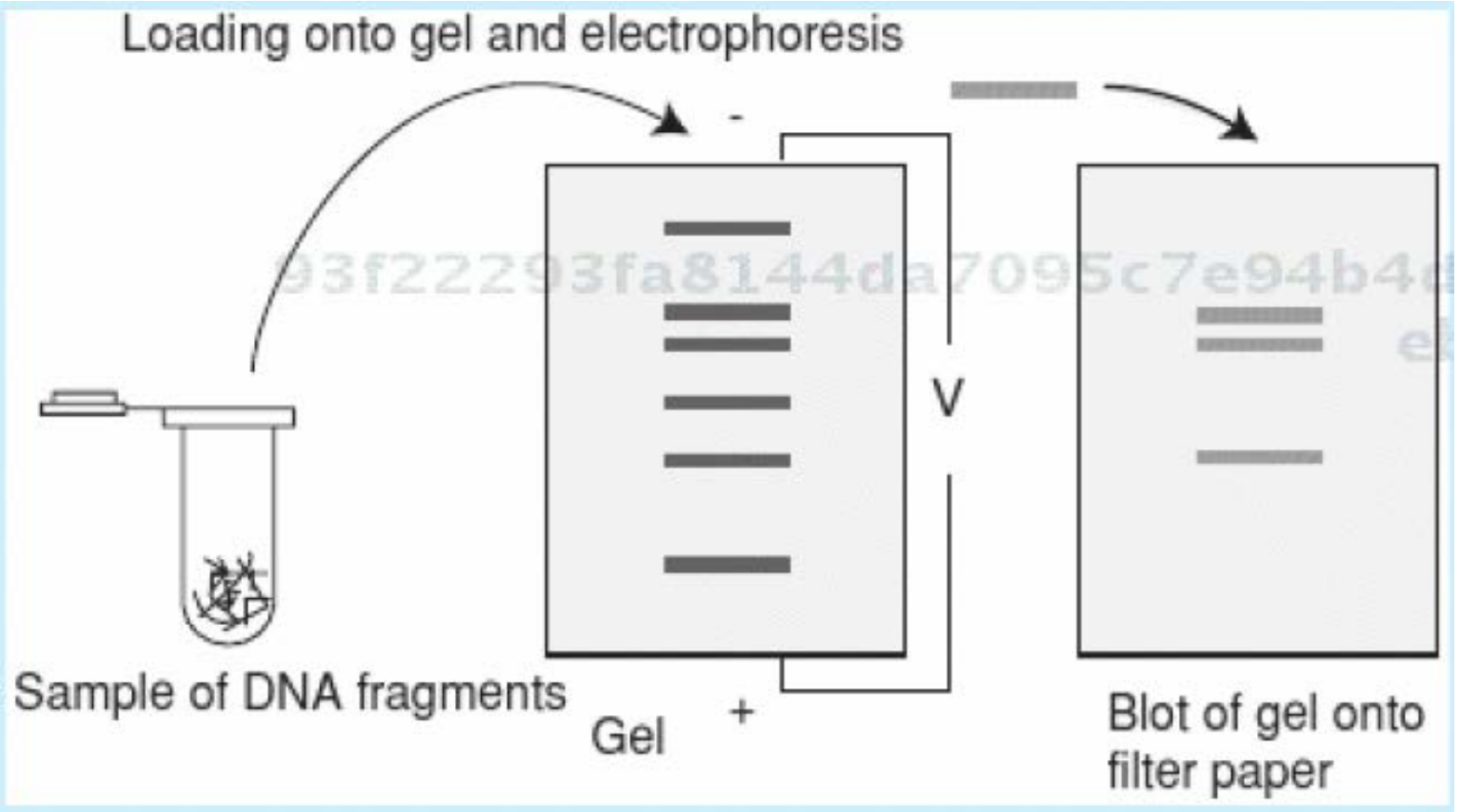
Control of gene expression (Engineering)

- Small interference RNA, siRNA
  - Duplex strands of less than 30 bp
  - Binds to the complementary bases in the target mRNA transcript
  - Cause the degradation of the mRNA (persisted effect)

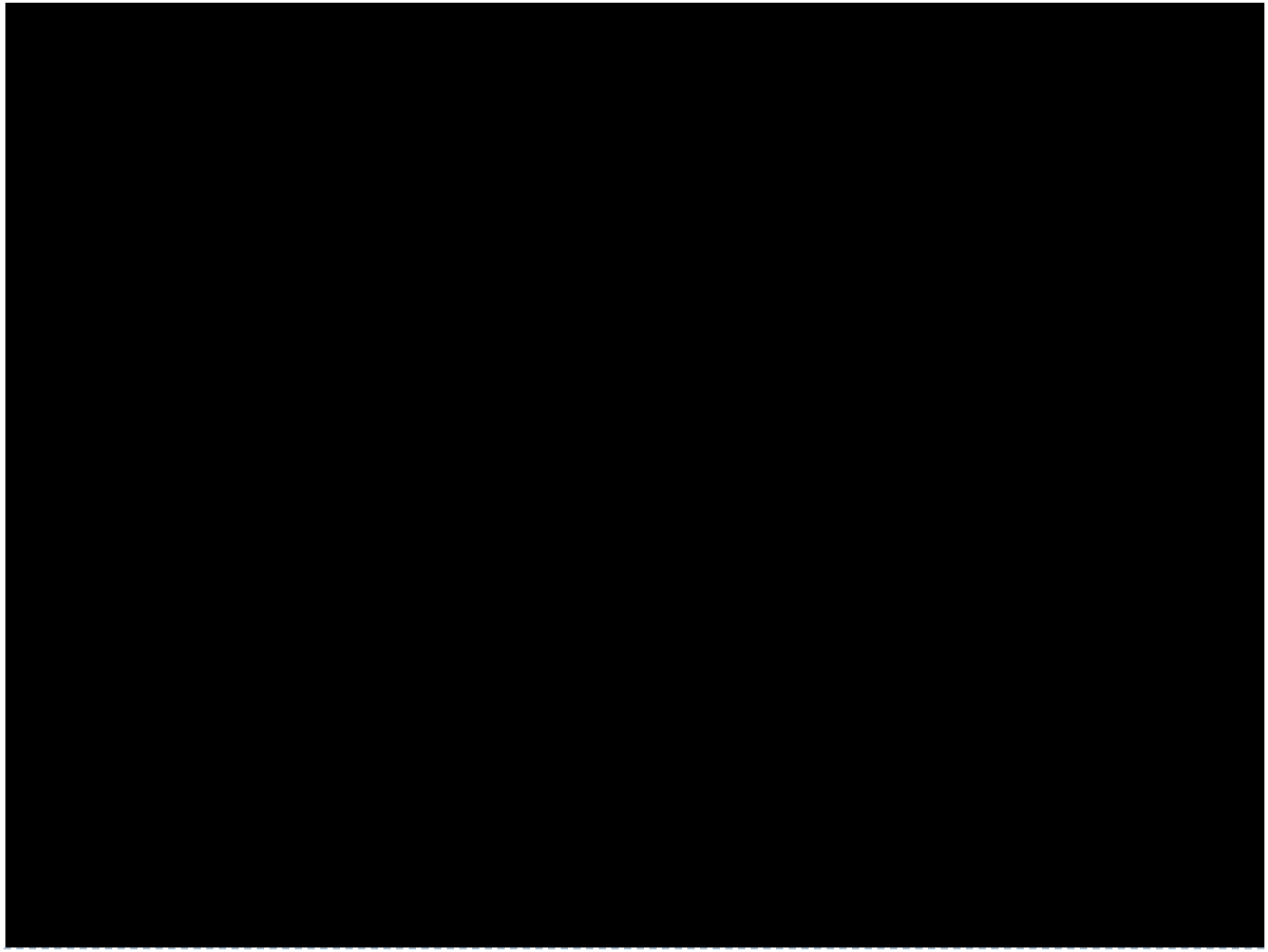


DNA Fingerprinting

- Unique evidence of an individual: within saliva, in hair,



DNA Fingerprinting



Genetic Engineering

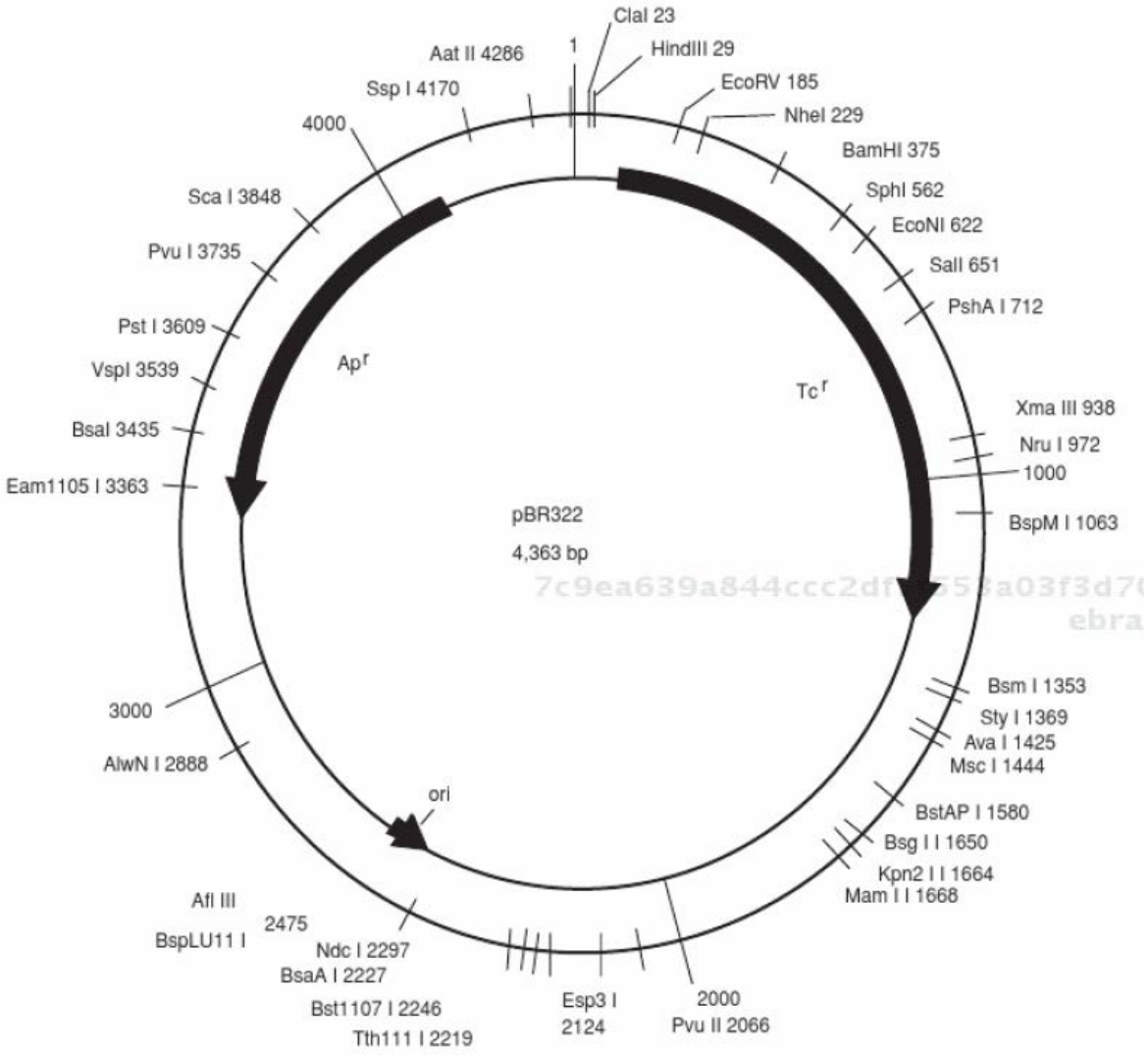
- Cloning

Cloning

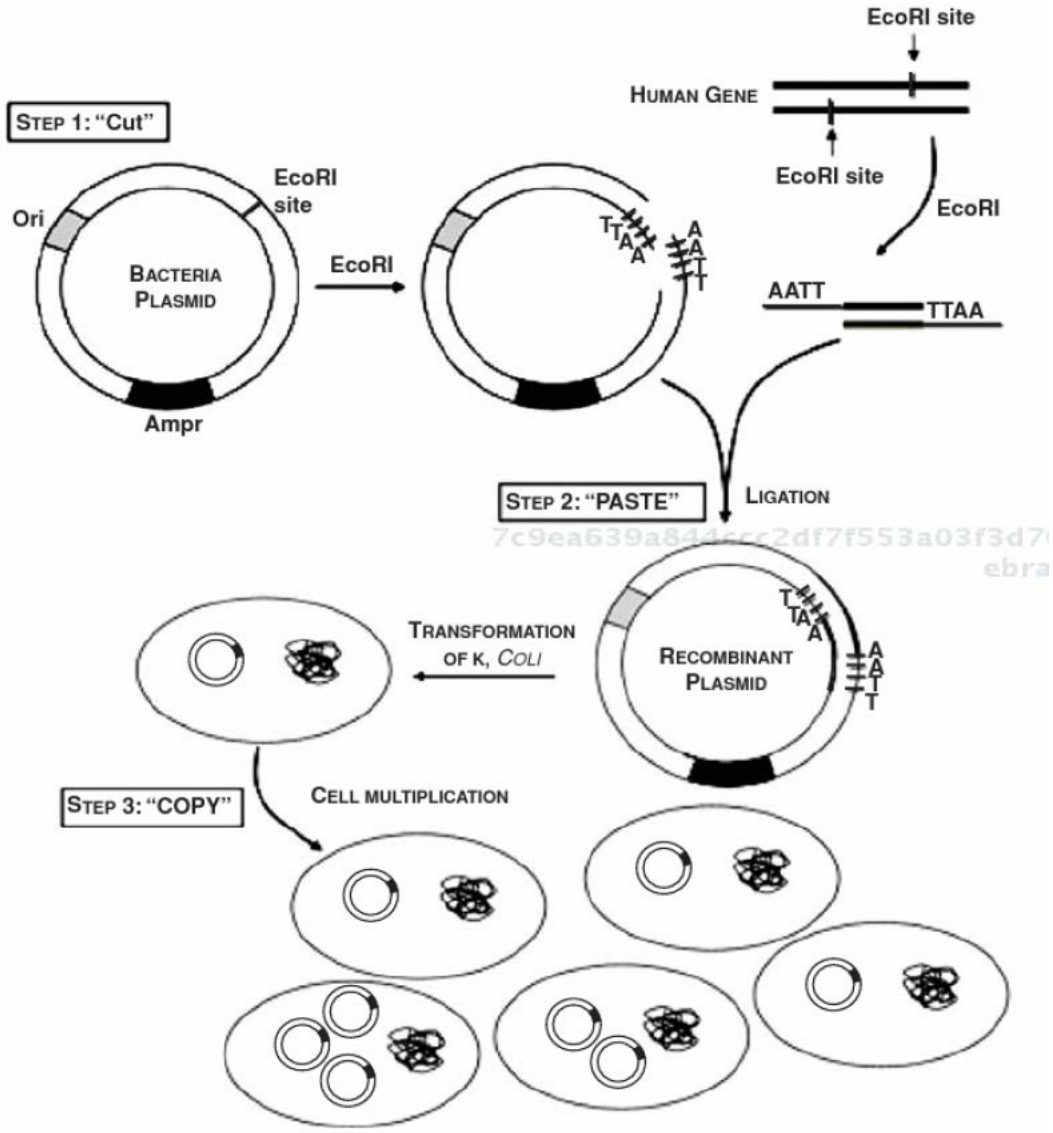
DNA Cloning

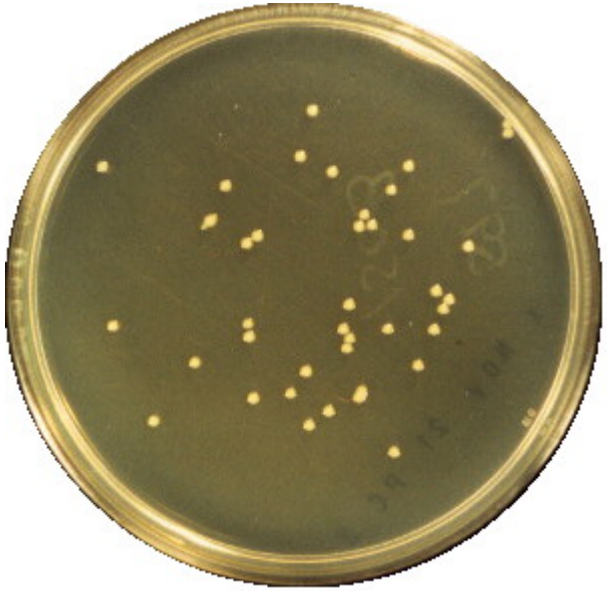
- Making identical copies of DNA
  - **Plasmids,** small circular DNA that replicate independently of chromosome in bacteria.
  - Occurred naturally in microorganisms to move genes between individual microorganisms.
  - **Restriction enzymes**: microorganism's defense system, allowing them to cut up and destroy genes of invaders, recognize specific short DNA sequences
  - These two tools have been used extensively in molecular biology

Plasmid


ClaI 23
HindIII 29
EcoRV 185
NheI 229
BamHI 375
SphI 562
EcoNI 622
SalI 651
PshA I 712
Xma III 938
Nru I 972
1000
BspM I 1063
Bsm I 1353
Sty I 1369
Ava I 1425
Msc I 1444
BstAP I 1580
Bsg I I 1650
Kpn2 I I 1664
Mam I I 1668
2000
Pvu II 2066
Esp3 I
2124
Bst1107 I 2246
Tth111 I 2219
BsaA I 2227
Ndc I 2297
2475
AfI III
BspLU11 I
AlwN I 2888
3000
BsaI 3435
Eam1105 I 3363
Vspl 3539
Pst I 3609
Pvu I 3735
Sca I 3848
4000
Ssp I 4170
Aat II 4286
1
ori
Ap^r
Tc^r
pBR322
4,363 bp

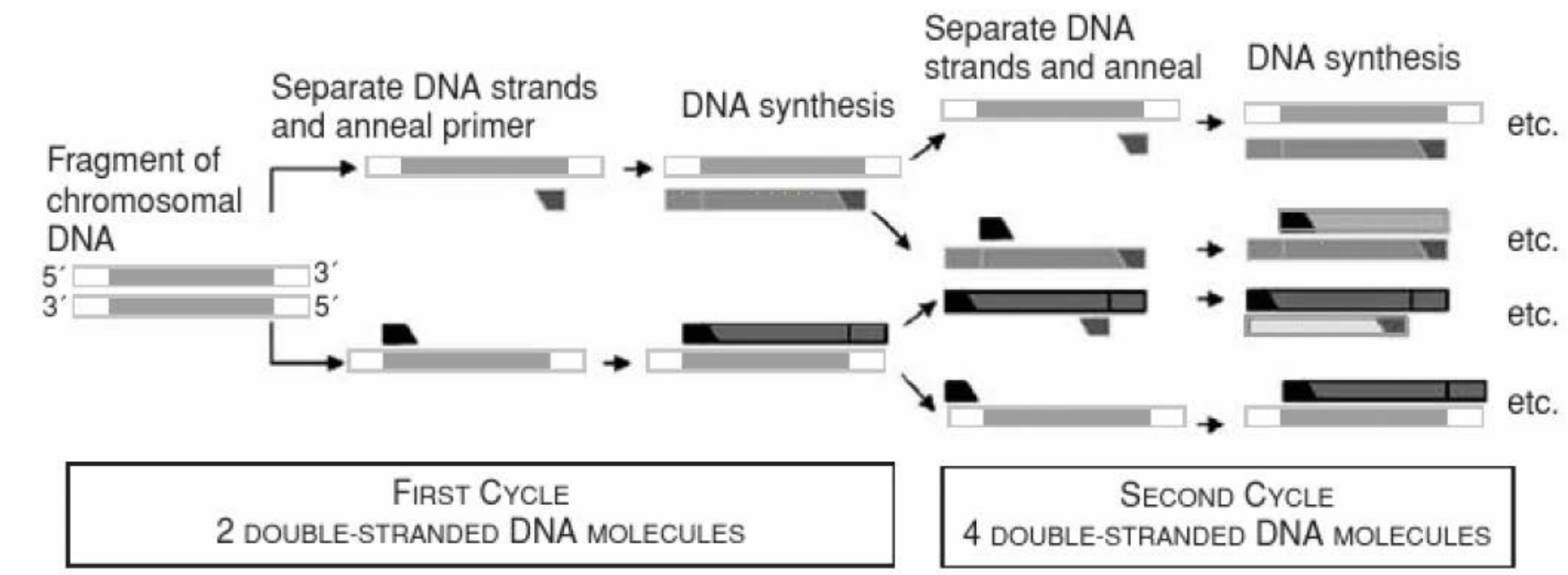
Transform


EcoRI site
HUMAN GENE
EcoRI site
STEP 1: "Cut"
Ori
EcoRI site
BACTERIA PLASMID
Ampr
EcoRI
EcoRI
AATT
TTAA
STEP 2: "PASTE"
LIGATION
TRANSFORMATION OF K, COLI
RECOMBINANT PLASMID
STEP 3: "COPY"
CELL MULTIPLICATION



Polymerase Chain Reaction (PCR)

- An engineering method to clone DNA
- Use DNA polymerase (Tag)



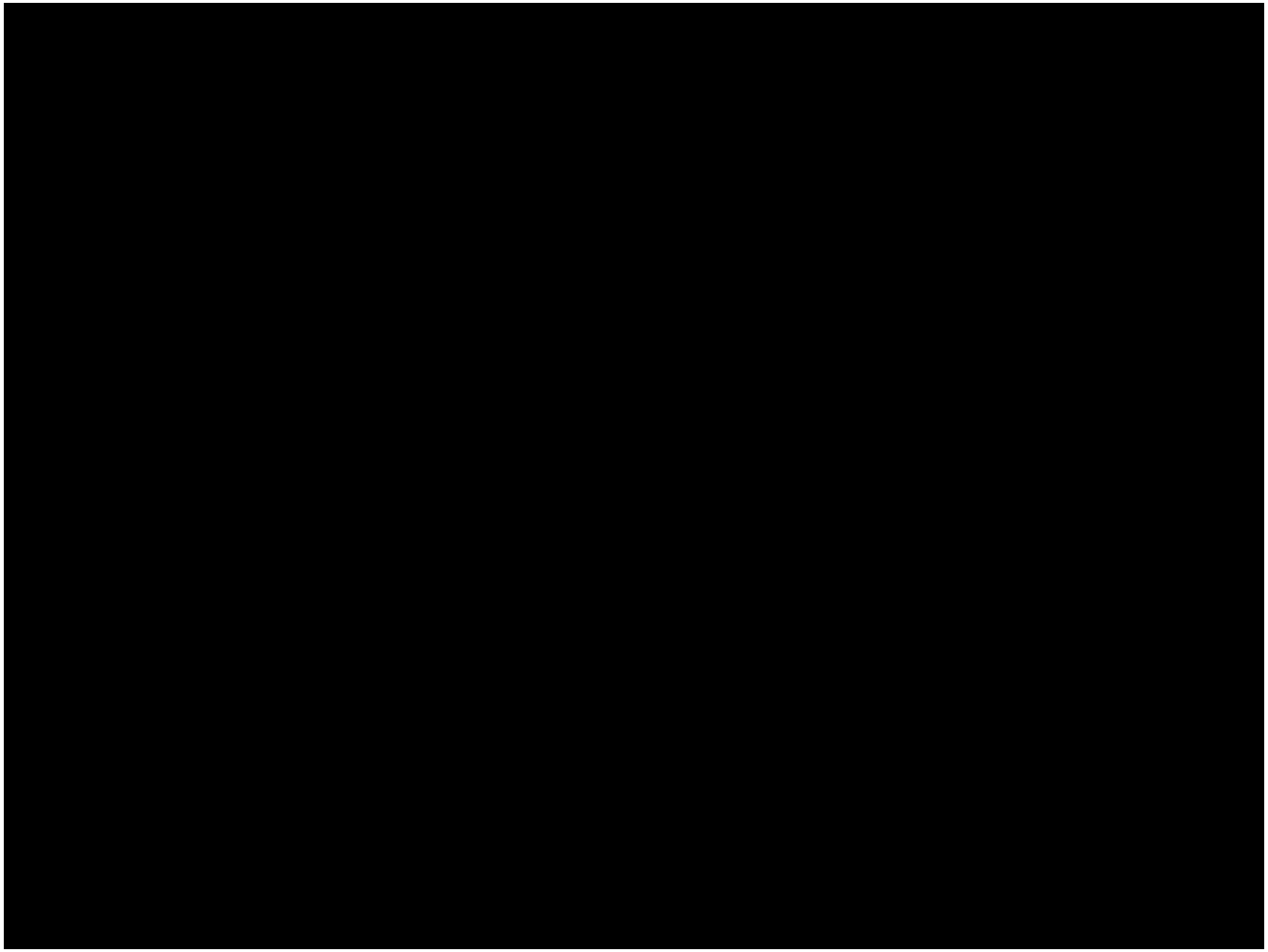
Biomedical Application of Genetic Engineering

Making human protein in bacteria

- Prior to DNA technology, insulin was extract from animals, similar, but not identical
- Make plasmid containing human gene encoding insulin
- Insert the plasmid to bacteria, which will produce human insulin as they translate the plasmid DNA
- Genentech developed the technology for recombinant human insulin, and approved for use in treating human diabetes in 1983.

Make human protein in bacteria

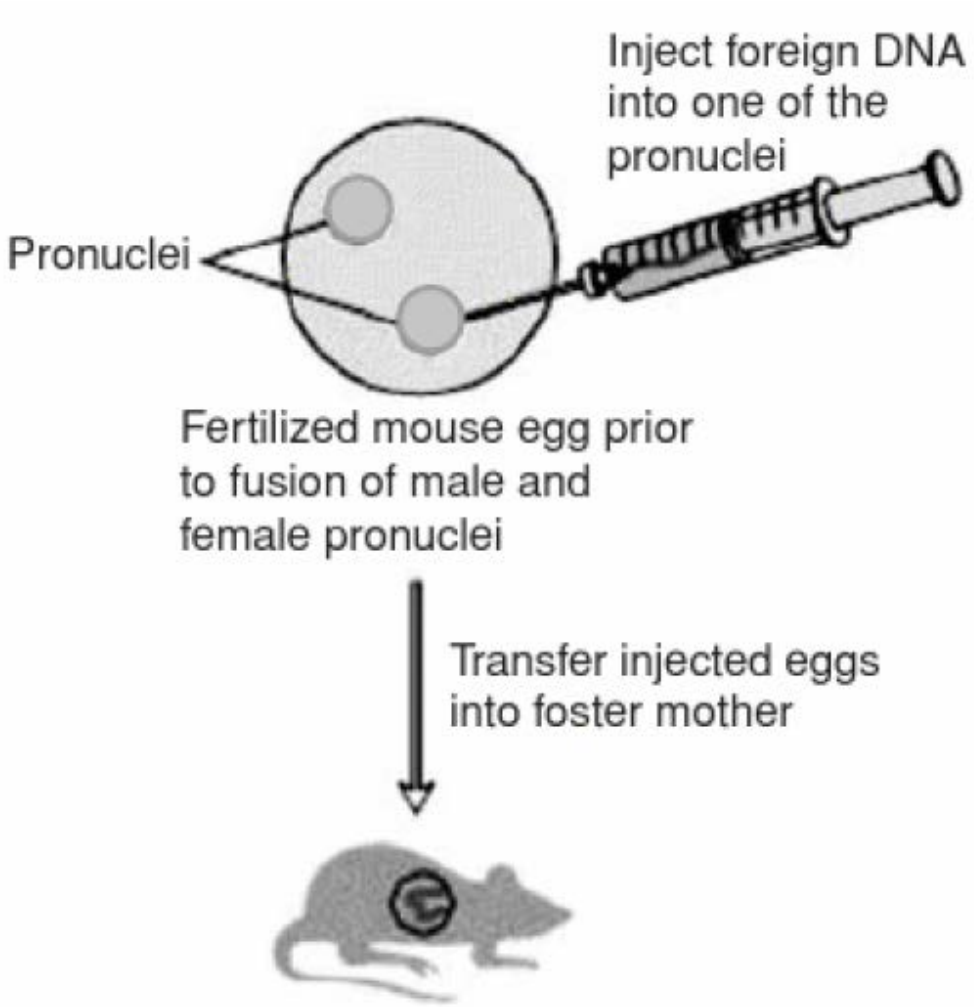
Insulin from Bacterial



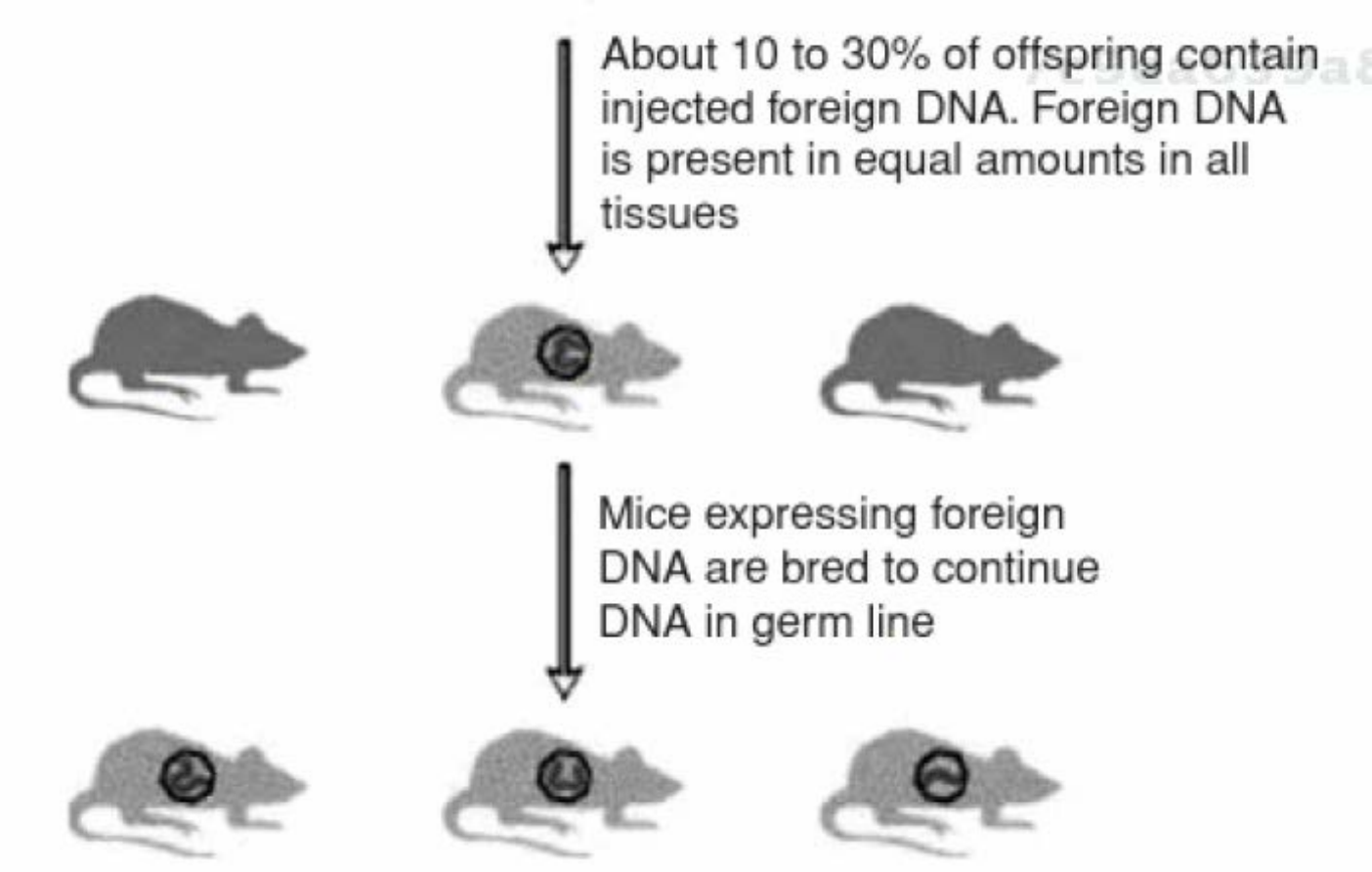
Animal models of human diseases

- The genome of rodents and other animals can be altered by recombinant DNA technique to produce phenotypes that mimic human diseases: **transgenic animals**

Animal models of human diseases



Animal models of human diseases



Gene Therapy

Disease	Location of Defect	Incidence	Treatment
	Enzyme involved		
Lesch–Nyhan syndrome	Hypoxanthine–guanine phosphoribosyltransferase	1 in 100,000–380,000 (males only)	
Adenosine deaminase deficiency	Adenosine deaminase		Enzyme infusion; gene therapy
Gaucher's disease	Lysosomal glucocerebrosidase		Enzyme infusion
Phenylketonuria	Phenylalanine hydroxylase		Diet
Sickle cell anemia	Hemoglobin	0.4% of African-American males	
Hemophilia	Factor VIII	1 in 10,000 males	Blood transfusion; protein infusion
	Membrane protein involved		
Cystic fibrosis	Cystic fibrosis transmembrane conductance receptor (CFTR)	1 in 2000 live births in Caucasians	Gene therapy not yet successful
Disaccharidase deficiency (lactose intolerance)	Lactase	Frequent	Diet
Duchenne's muscular dystrophy	Dystrophin gene product	1 in 3500 males	None
	Protein involved		
Huntington's disease	Huntington		None

Goal of gene therapy:

Management and correction of human diseases

a. Inherited and acquired disorders

b. cancer

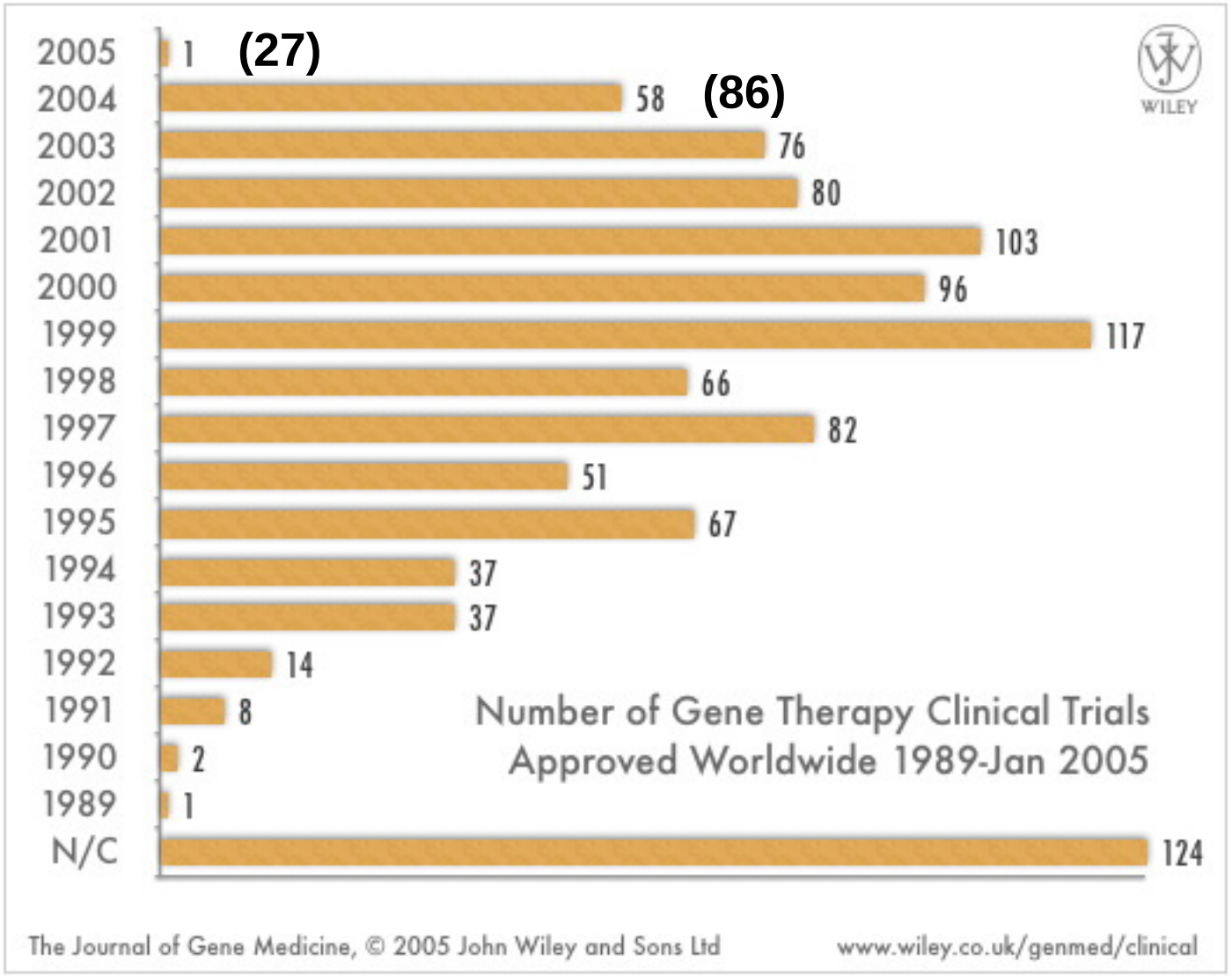
c. AIDS/HIV

<u>Good news</u>: Promising advances during the last two decades in recombinant DNA technology.

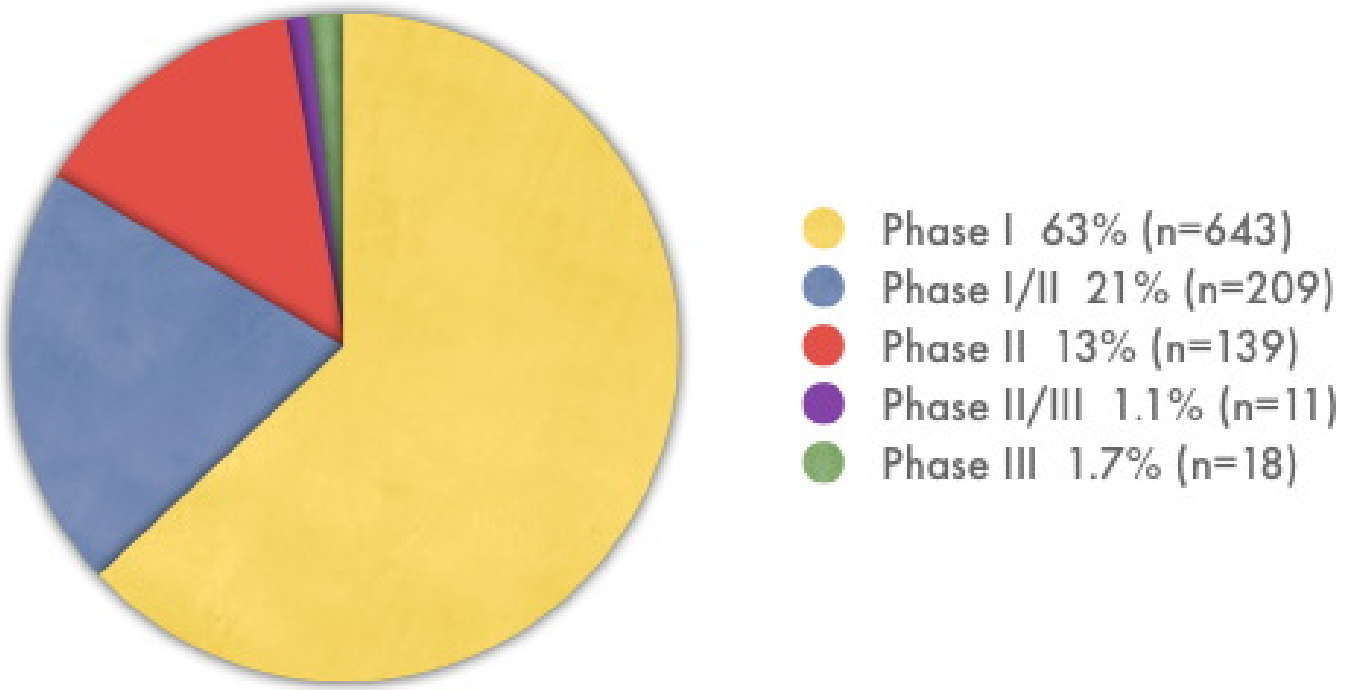
<u>Bad news</u>: (Until recently?) Efficacy in any gene therapy protocol is not definitive.

1. Shortcomings in gene transfer vectors.

2. Inadequate understanding of biological interactions of vector and host.


2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
N/C
1
58
76
80
103
96
117
66
82
51
67
37
37
14
8
2
1
124
(27)
(86)
WILEY
Number of Gene Therapy Clinical Trials
Approved Worldwide 1989-Jan 2005
The Journal of Gene Medicine, © 2005 John Wiley and Sons Ltd
www.wiley.co.uk/genmed/clinical

Phases of Gene Therapy Clinical Trials



www.wiley.co.uk/genmed/clinical

Genetic diseases:

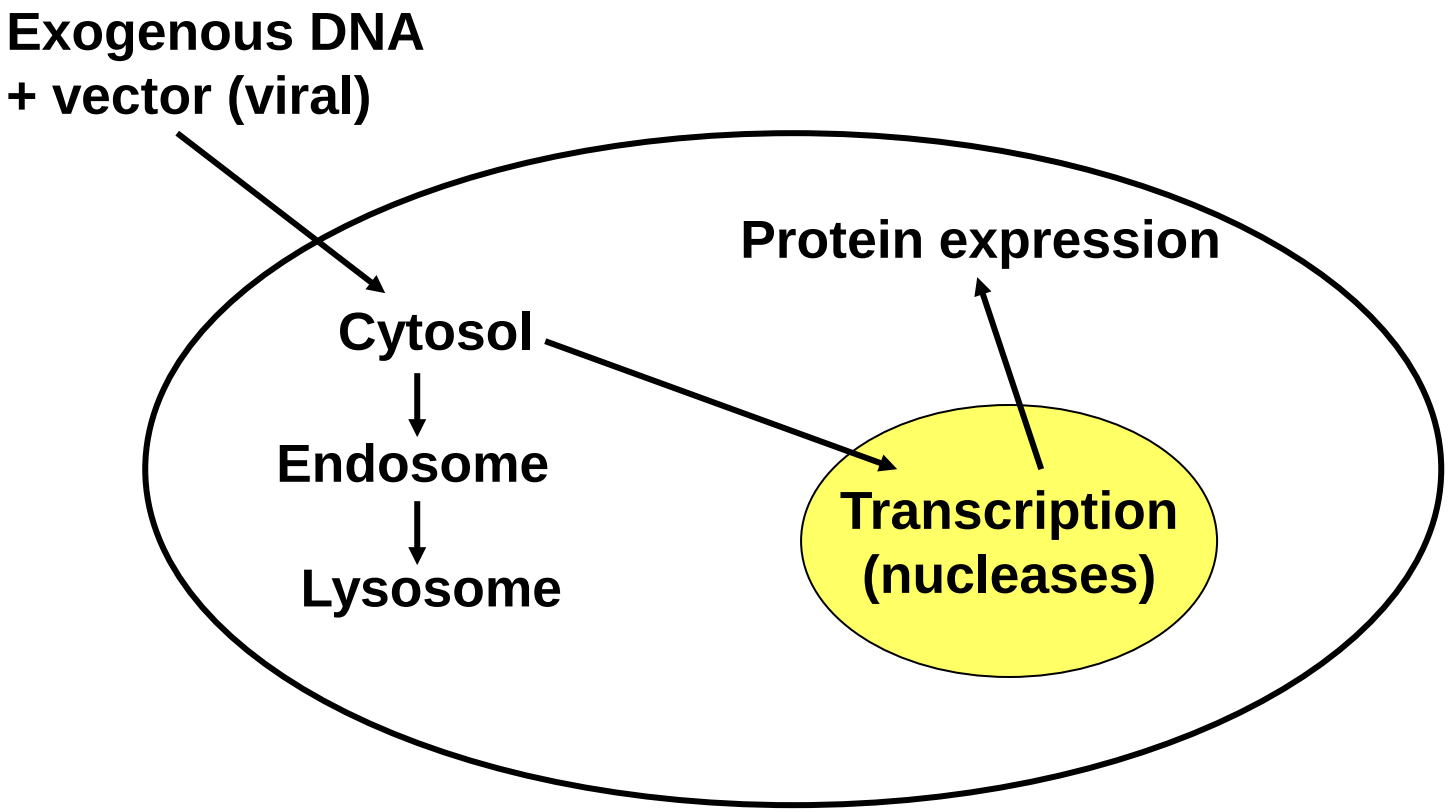
Type 1: Single locus (gene) is defective and responsible for the disease, 100% heritable.

examples: **Sickle cell anemia,**
Hypercholesterolemia
Cystic fibrosis

Type 2: Polygenic traits, <100% heritable, may be dependent on environmental factors and lifestyle.

examples: **Heart disease**
Cancer
Diabetes
Alcoholism
Schizophrenia
Criminal behavior?
etc….?

Gene Transfer



Barriers that prevent transfer of exogenous DNA

Vectors for gene transfer:

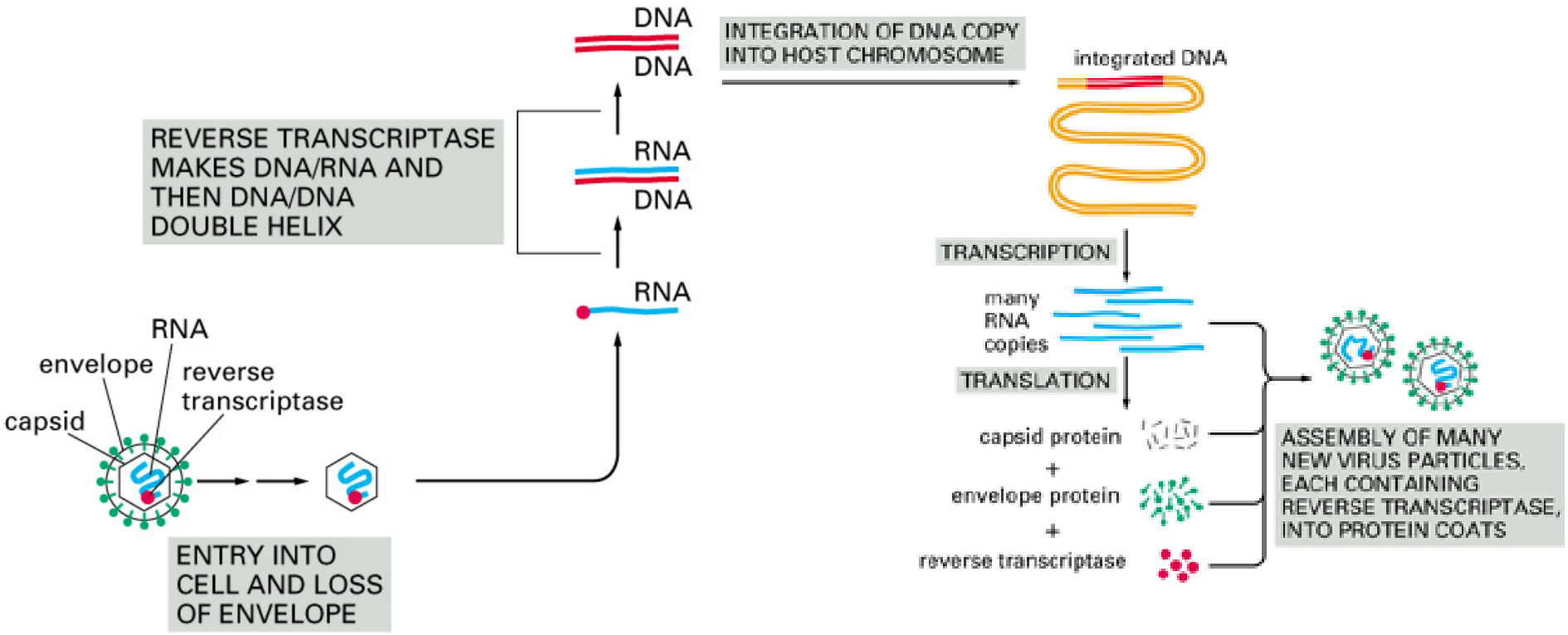
<u>Retroviruses</u>: promising candidates and widely used.

- Insert genetic material into host DNA.
- Insertion may disrupt a host gene.
- Insertion may be in a region that doesn't producevery much of the desired protein
- Can trigger immune response.

Ideal vector characteristics:

- Insert size: one or more genes.
- Targeted: limited to a cell type.
- No immune response.
- Stable: not mutated.
- Production: easy to produce high concentrations (titer).
- Effective: produce enough protein to cause an effect.

Retrovirus life cycle



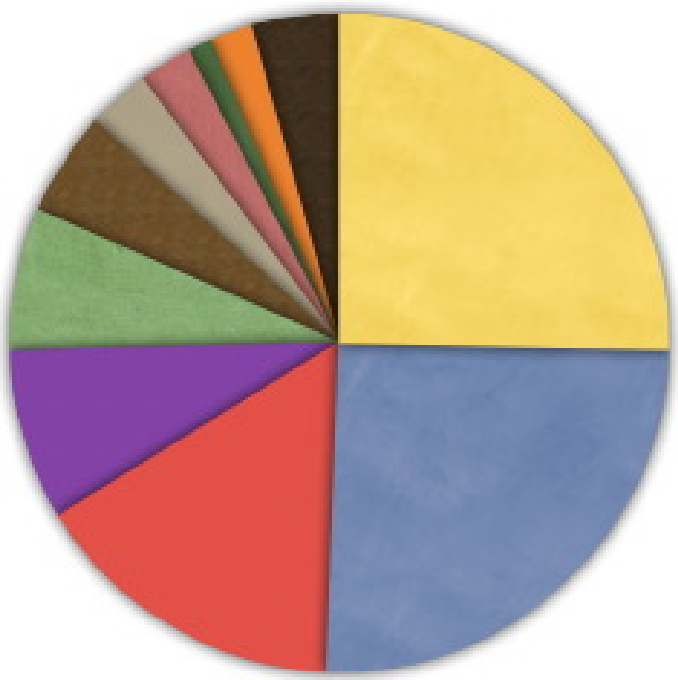
Retrovirus life cycle

Retrovirus as delivery vectors

- Retrovirus with replication defective can be used as gene delivery vectors.

- Genome integration of inserted gene
  - Stable expression
  - Random insertion into genome, may disrupt normal genome function
  - Only transduce dividing cells
  - Less than 7kb gene

Vectors Used in Gene Therapy Clinical Trials



- Retrovirus 27% (n=263)
- Adenovirus 26% (n=262)
- Naked/Plasmid DNA 15% (n=162)
- Lipofection 8.6% (n=87)
- Pox virus 7.2% (n=73)
- Vaccinia virus 4.7% (n=53)
- Herpes simplex virus 3% (n=32)
- Adeno-associated virus 2.5% (n=27)
- RNA transfer 1.1 (n=12)
- Others 2.2% (n=22)
- N/C 4.2% (n=45)

www.wiley.co.uk/genmed/clinical

Non-viral DNA carriers:

Cationic liposomes: Positively charged lipids interact with negatively charged DNA. (lipid-DNA complex).

-Transverses cell membranes

<u>Advantages:</u>

a. Stable complex
b. Can carry large sized DNA
c. Can target to specific cells
d. Does not induce immunological reactions.

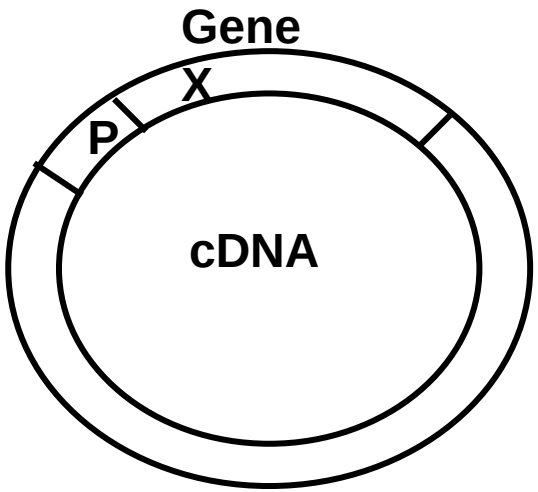
<u>Disadvantages:</u>

a. Low transfection efficiency
b. Transient expression
c. Inhibited by serum
d. Some cell toxicity

Non-viral DNA carriers:

2. Naked plasmid DNA injection

Promoter + gene of interest
(P) (gene X)



Expression observed in thymus, skeletal and cardiac muscle, skin.

Gene therapy to turn off genes:

<u>Antisense approach</u>:

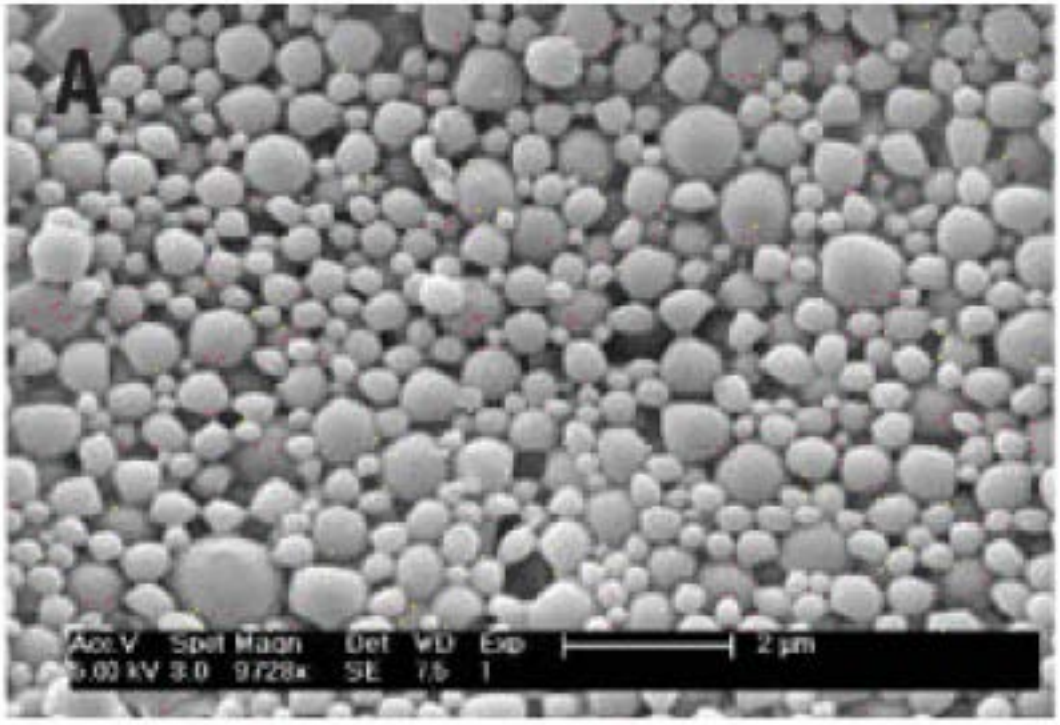
***DNA makes mRNA, mRNA makes protein.**
Antisense complements mRNA (sense) and prevents protein expression.

***Small interfering RNA (siRNA) molecules.**

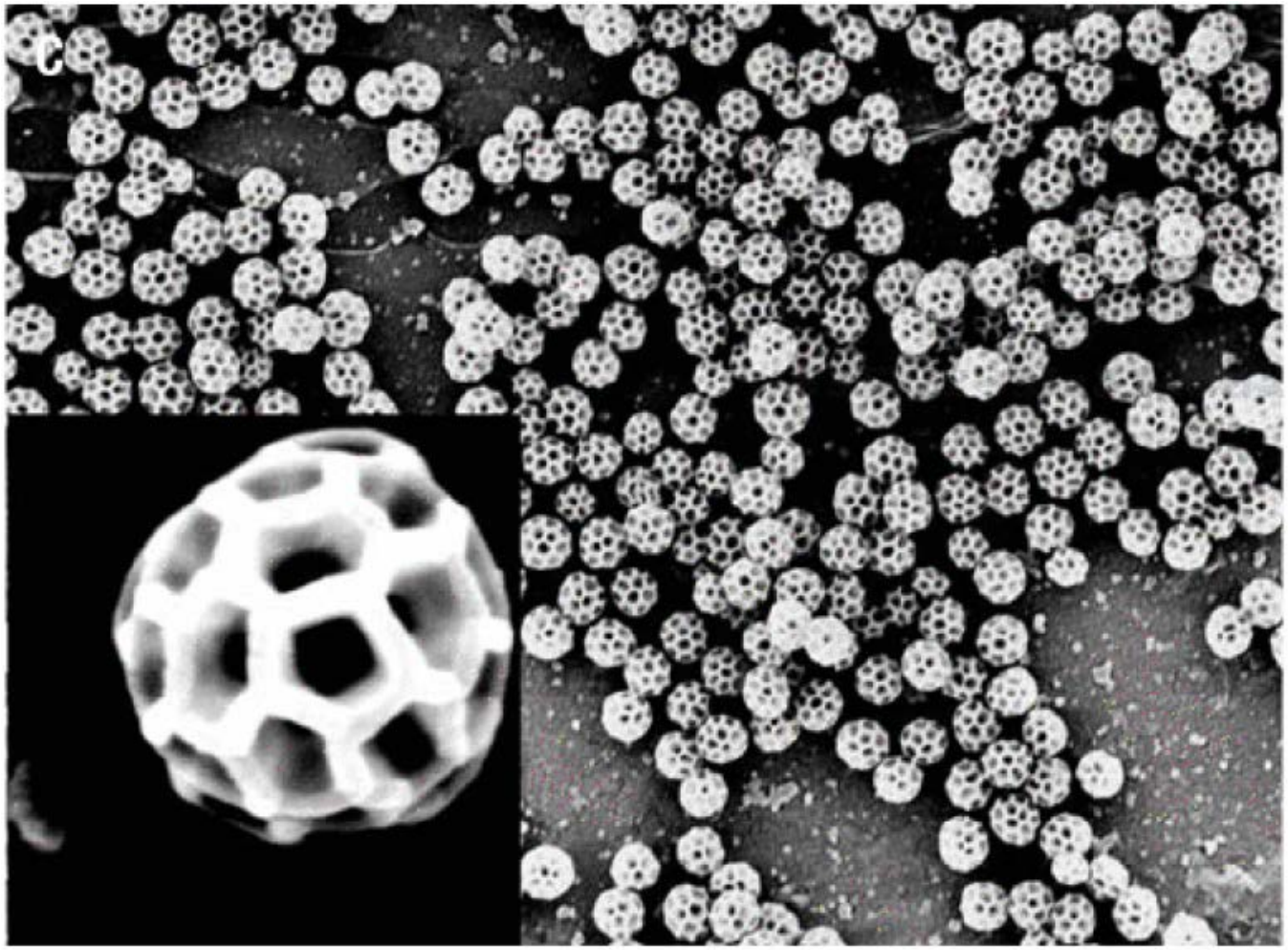
The silent treatment: siRNAs as small molecule drugs.

***Ribozymes**

Non-viral Gene Delivery Methods:



Non-viral Gene Delivery Methods:



Case study: Jesse Gelsinger

***First documented patient to die from gene therapy treatment. (may have been others).**

<u>Disease</u>: liver enzyme deficiency (ornithine transcarbamylase, OTC) – controls ammonia metabolism

- **Vector used to deliver OTC – modified adenovirus**

<u>Goal</u>: deliver vector to liver cells and express OTC.

<u>Problem</u>: Very low transfer efficiency (1%), difficult to get enough functioning OTC expressed to do any good.

<u>Solution</u>: Infect with higher dose of viral particles. (38 trillion)

Outcome:

-Vector not only delivered gene to liver but to other tissue.
-Triggered systemic inflammatory response.
-Patient acquired fever, coma, death.

Why?

-Animal studies suggested dose was OK (?).
-Adenoviral vectors known to induce inflammatory response.
-Patient already compromised:
Patient had higher than allowed ammonia levels.

Results of follow-up investigation:

-**<u>3 month investigation by FDA concluded.</u>**

- **patient enrollment in study despite ineligibility.**
- **participants misled on safety and toxicity issues.**
- **loosening of criteria for accepting volunteers.**
- **informed consent document did not reveal results of animal studies.**

*** Other investigators may not have disclosed important information on patient deaths in gene therapy trials.**

- **Adenovirus safety: Engineered to prevent viral replication.**
- **Mutation from replication incompetent to competent?**

- **Shut down of Univ. of Penn. Institute for Human Gene Therapy**
- **Lawsuits**

Some successes:

Treatment of Severe Combined Immuno Deficiency (SCID)

- **Genetic defects cause decreased T and B cells and NK cells.**
- **Affects 1-75,000 births.**
- **Mostly males (most common form is X-linked)**
- **Types: ADA (adenine deaminase) or Gamma chain (γc).**
- **Success in treating children observed in Italy, Israel, England, France, and USA.**
- **Phase 1 trial: collect bone marrow, isolate $CD34^+$ stem cells, and infect with retroviral vector containing the gene encoding the γ-common chain. Inject two infants with 14-26 million $CD34^+$ cells/kg (5- 9 million contained the introduced gene).**

Recent successes continued:

Phase I clinical trials results:

Detectable levels of NK and T cells containing the introduced gene were found in the blood within 30 and 60 days, respectively, and their numbers increased progressively until normal levels were reached. After 3 months, the two patients were also able to make antibodies in response to vaccination against diphtheria, tetanus, and pertussis.

10/3/02: France and US (FDA) halted SCID gene therapy due to leukemia-like side effects in one child. Not clear whether this is related to the gene therapy itself.

1/14/03: FDA suspended 30 gene therapy trials using retrovirus vectors due to another case of leukemia.

Human Genome Project

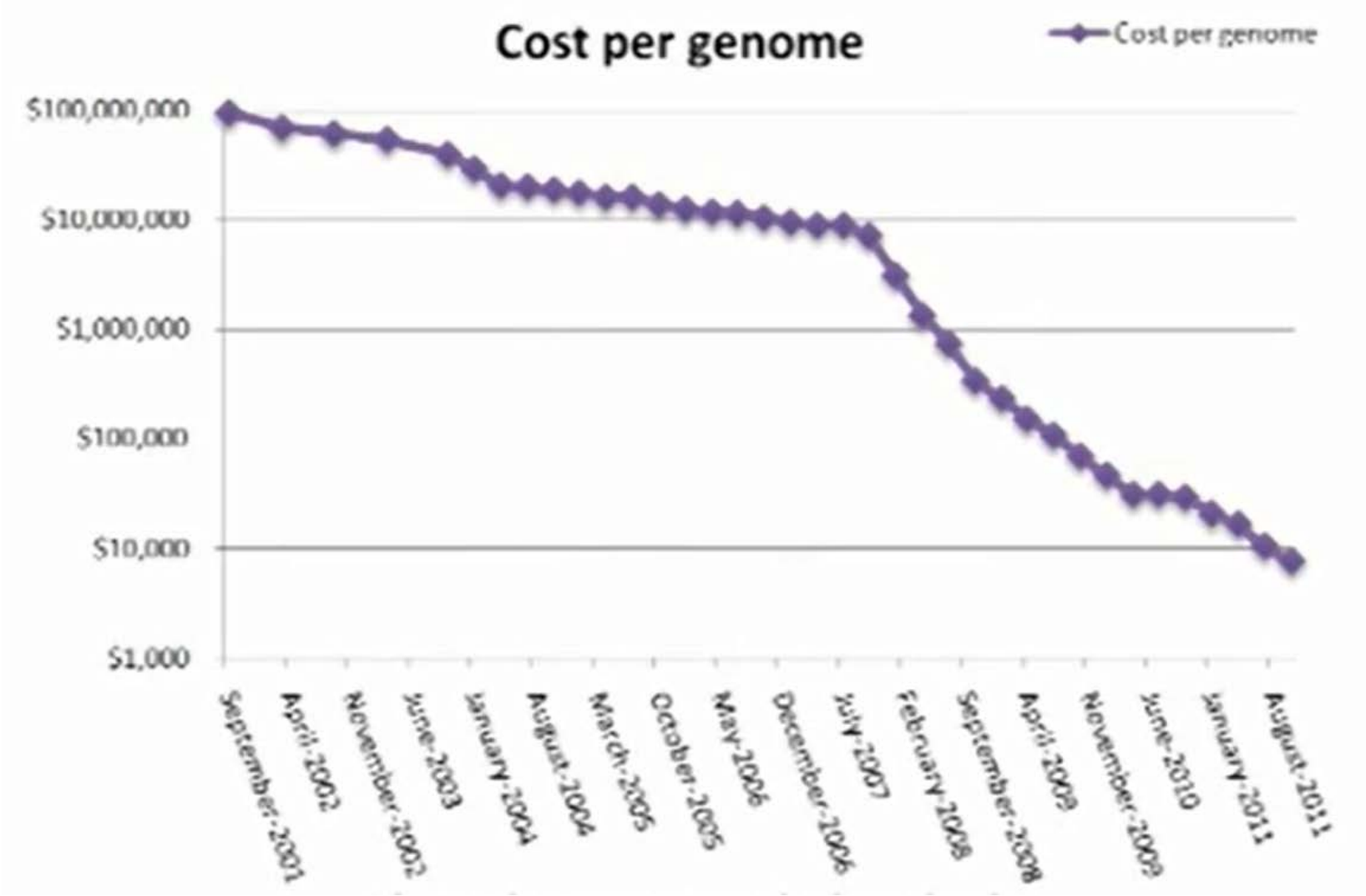
Impacting many disciplines

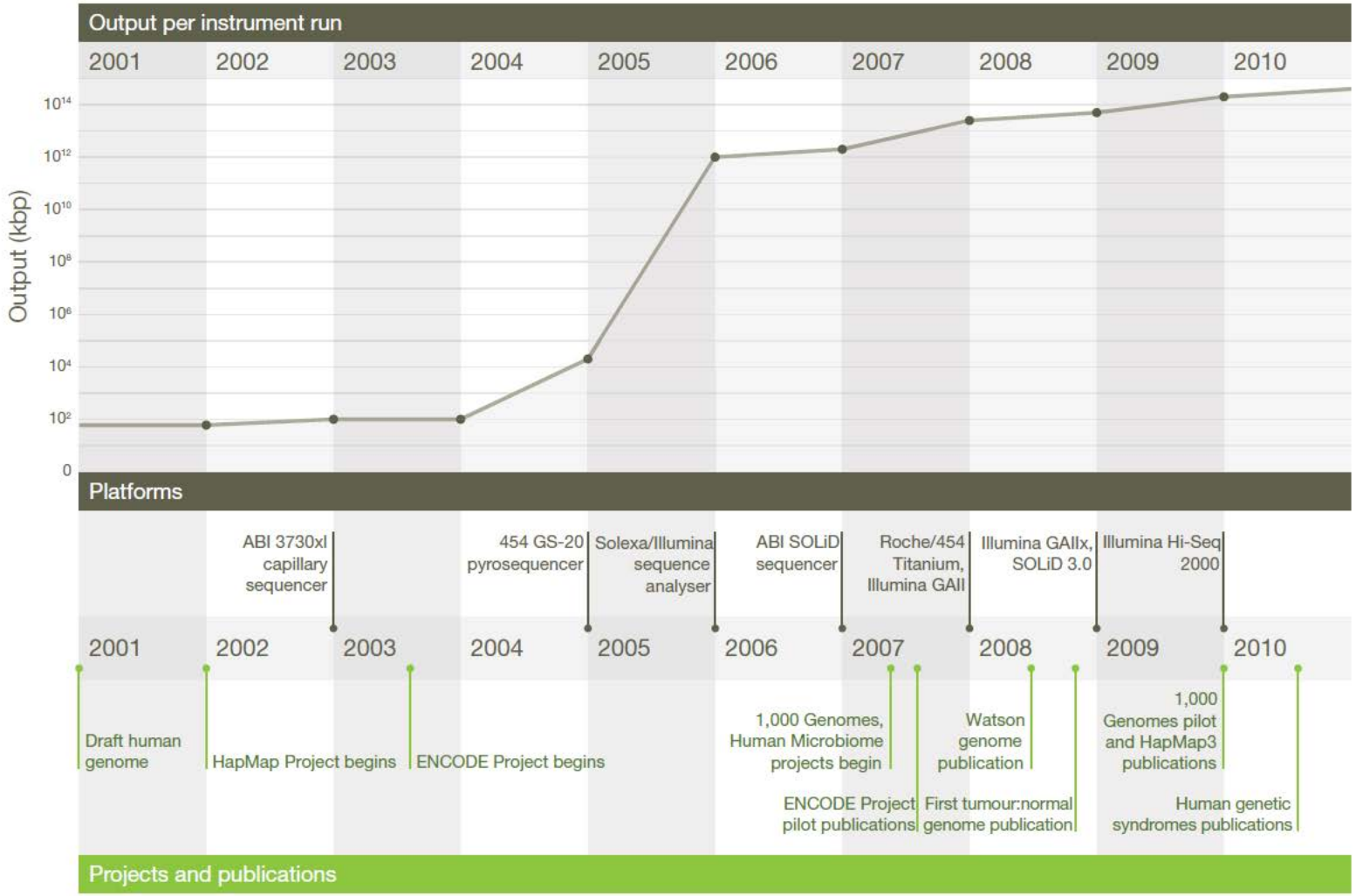
Courtesy
U.S. Department of Energy
Human Genome Program

Global Carbon Cycles
Industrial Resources • Bioremediation
Evolutionary Biology • Biofuels • Agriculture • Forensics
Molecular and Nuclear Medicine • Health Risks

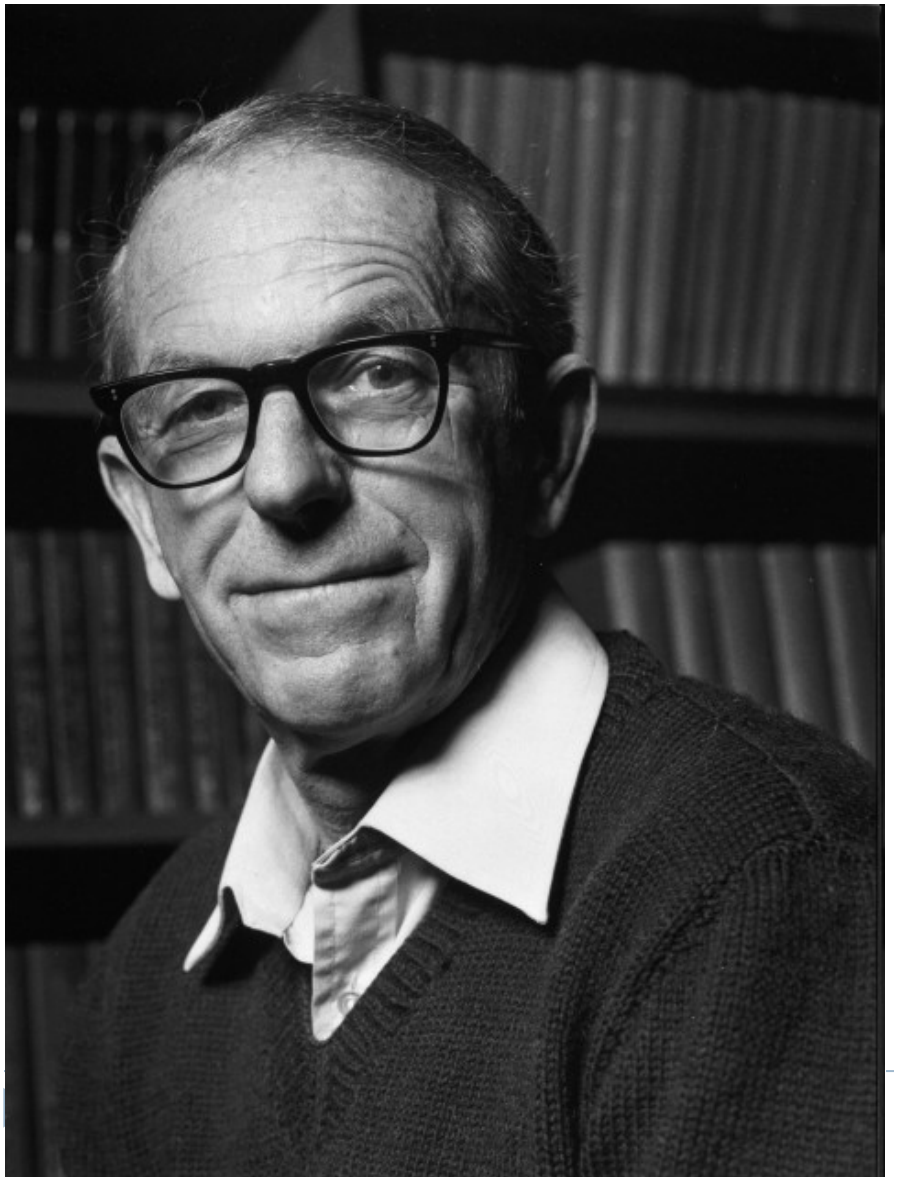
YGA 99-1133R

DNA sequencing technology




Output per instrument run
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Output (kbp)
10^14
10^12
10^10
10^8
10^6
10^4
10^2
0
Platforms
ABI 3730xl capillary sequencer
454 GS-20 pyrosequencer
Solexa/Illumina sequence analyser
ABI SOLiD sequencer
Roche/454 Titanium, Illumina GAII
Illumina GAIIx, SOLiD 3.0
Illumina Hi-Seq 2000
Draft human genome
HapMap Project begins
ENCODE Project begins
1,000 Genomes, Human Microbiome projects begin
ENCODE Project pilot publications
First tumour:normal genome publication
Watson genome publication
1,000 Genomes pilot and HapMap3 publications
Human genetic syndromes publications
Projects and publications

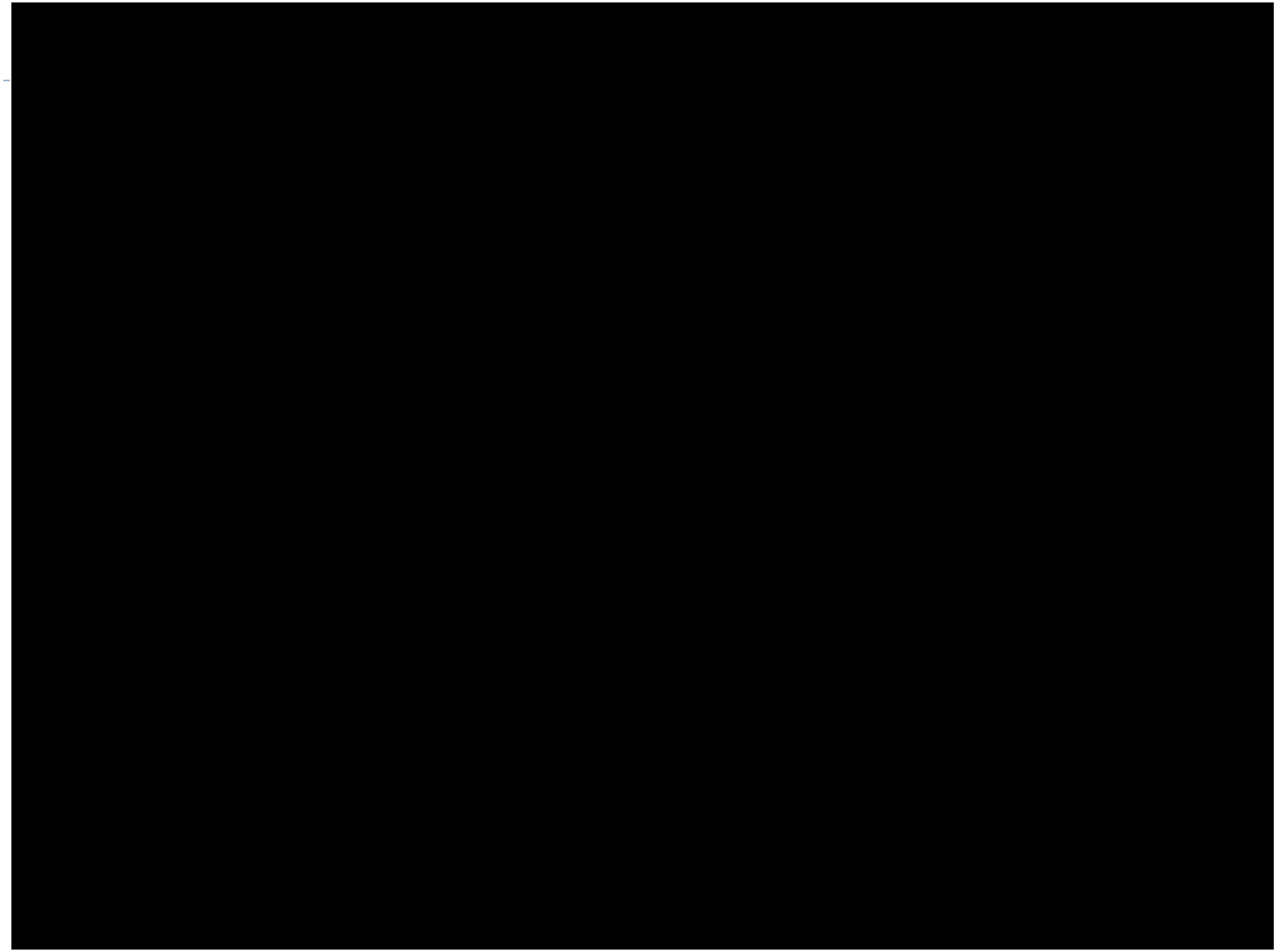
Frederick Sanger
1918-



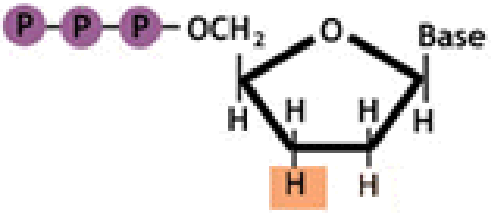
Sanger is the only chemist to have received two Nobel Prizes in Chemistry, the first as the sole recipient in 1958 for his work as the first to sequence a protein, the sequencing of insulin; and the second in 1980, shared with Paul Berg and Walter Gilbert, for the sequencing of nucleic acids.

Dideoxy (Sanger) Method

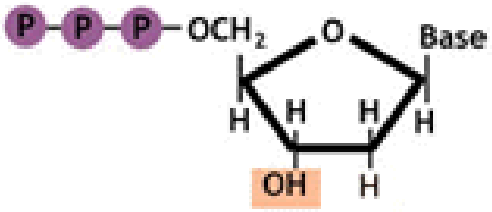
- 4 Steps:
  1. Denaturation
  2. Primer attachment and extension of bases
  3. Termination
  4. Gel electrophoresis



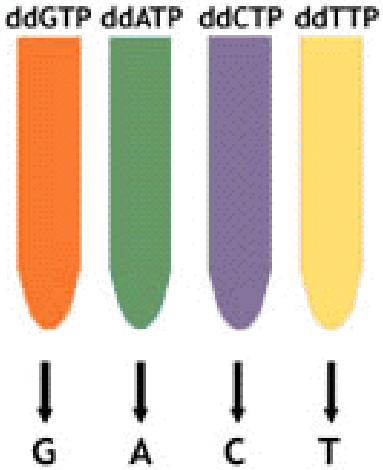
Sanger Method: Dideoxy Chain Termination



dideoxynucleotide (ddNTP)

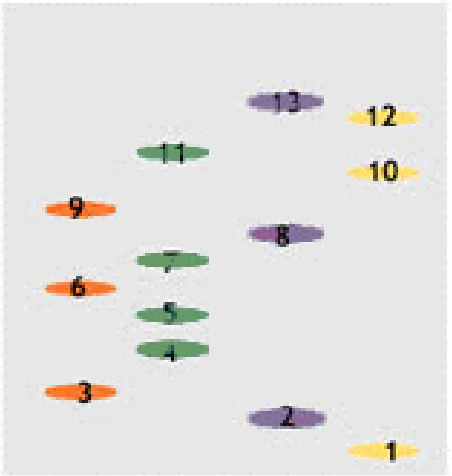


deoxynucleotide (dNTP)



Add ddGTP, ddATP, ddCTP, ddTTP, one to each of four tubes containing target DNA. Load each onto a separate lane on a gel.

300-500 bases

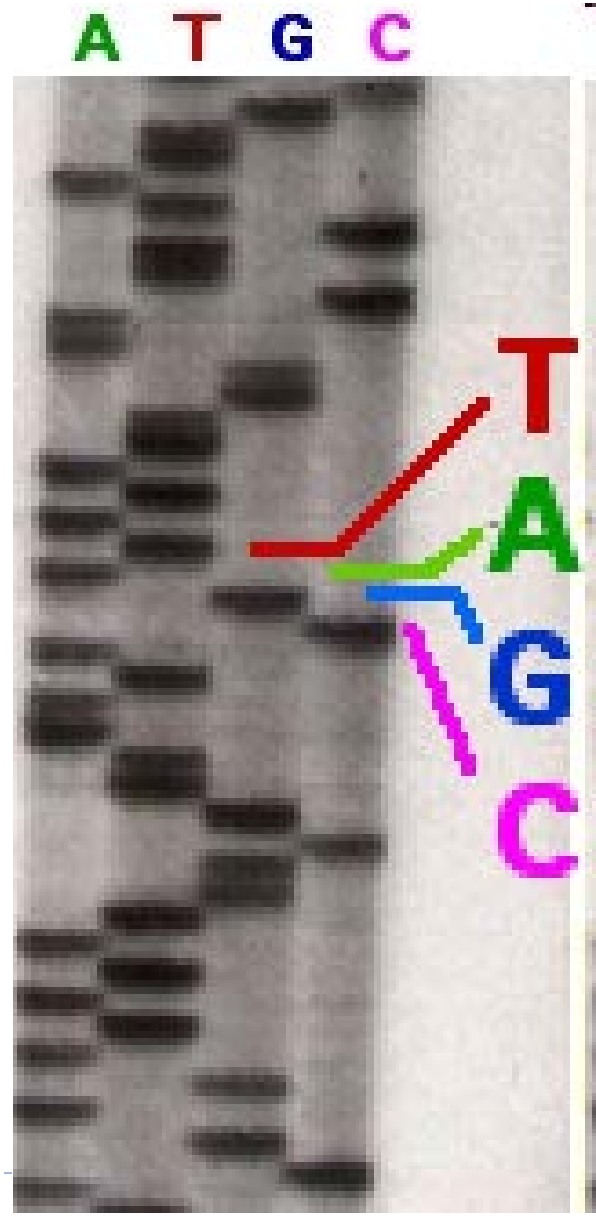


Largest

To sequence, read the order of bases from the smallest to the largest.

TCGAAGACGTATC

Smallest



Deep Sequencing
or
2nd Generation Sequencing
Technology

454 SEQUENCING

Roche

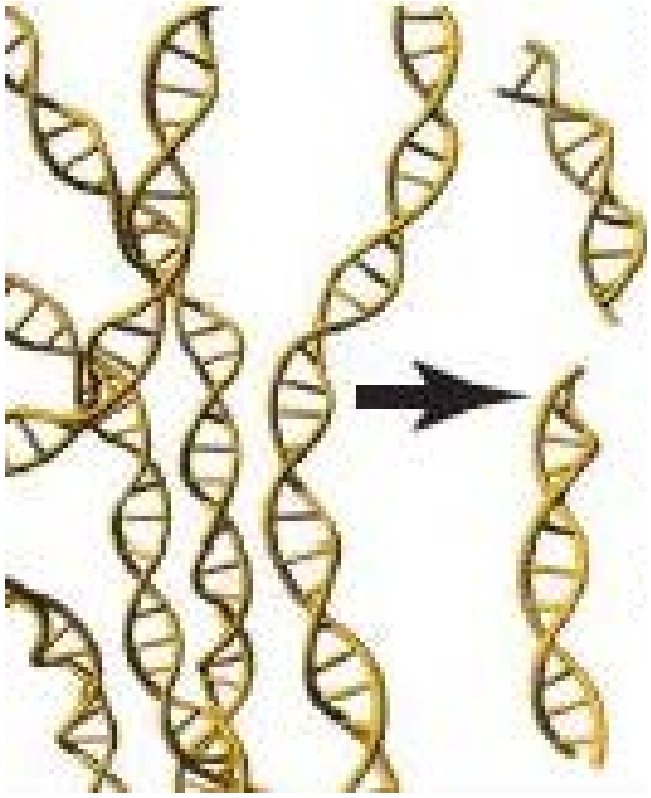
A "Pyrosequencing" method.

System Workflow:

- One Fragment = One Bead = One Read
- Four main steps:
  - Generation of a single-stranded template DNA library
  - Emulsion-based clonal amplification of the library
  - Data generation via sequencing-by-synthesis
  - Data analysis using different bioinformatics tools

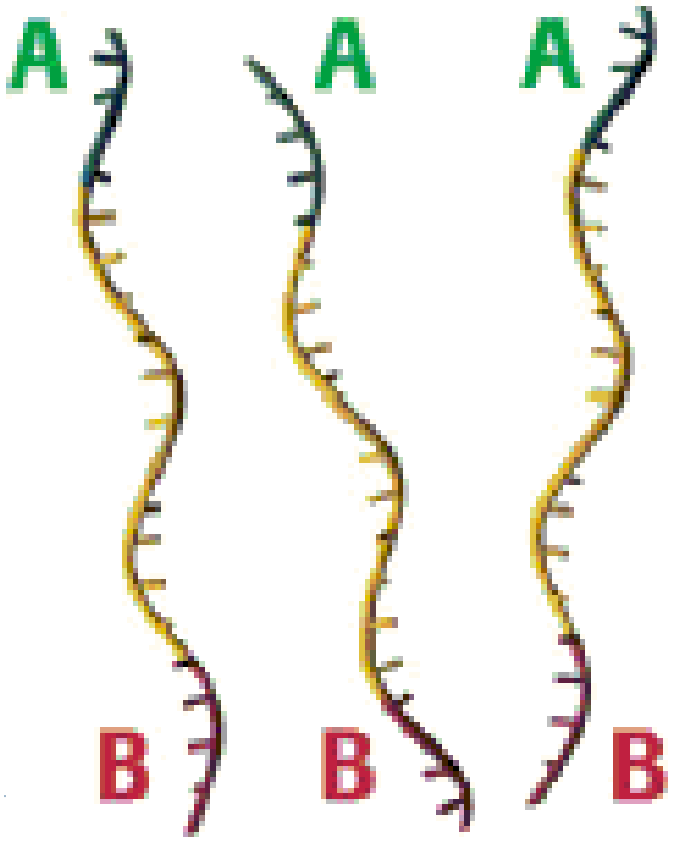
The Roche 454/GS FLX Sequencing Technology

1. Samples consisting of longer sequences are first sheared into a random library of 300-800 base-pair long fragments.



2. Adaptors essential for purification, amplification and sequencing are added to both ends of the fragments.

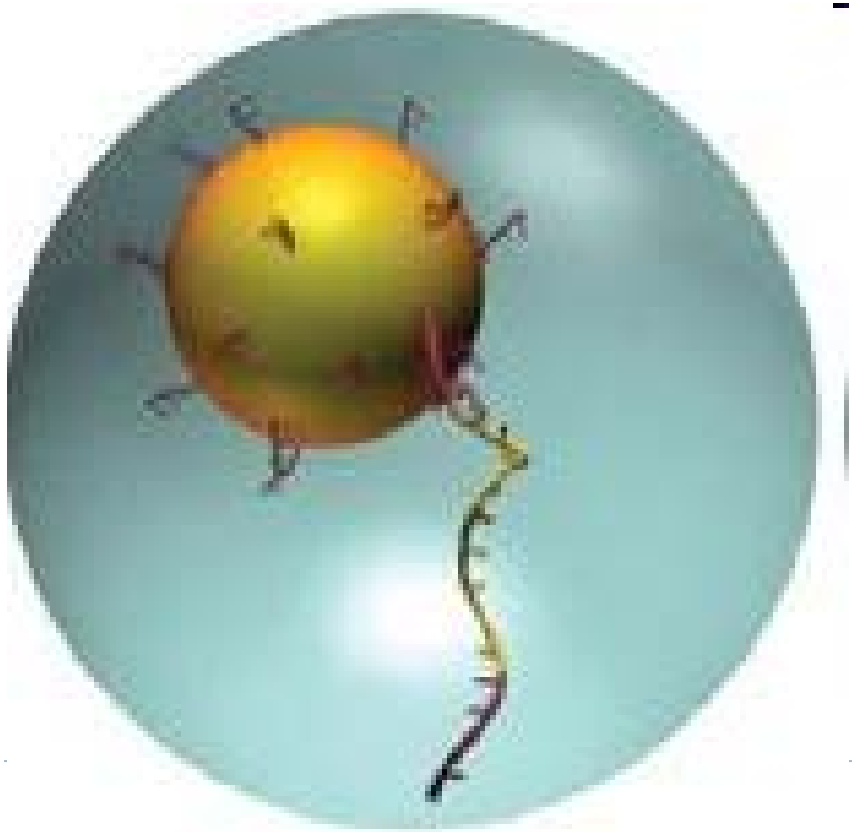
If the sample is double stranded one strand is removed and the remaining single strands are used in the following steps.



3. Using the adaptors, individual fragments are captured on own unique beads.

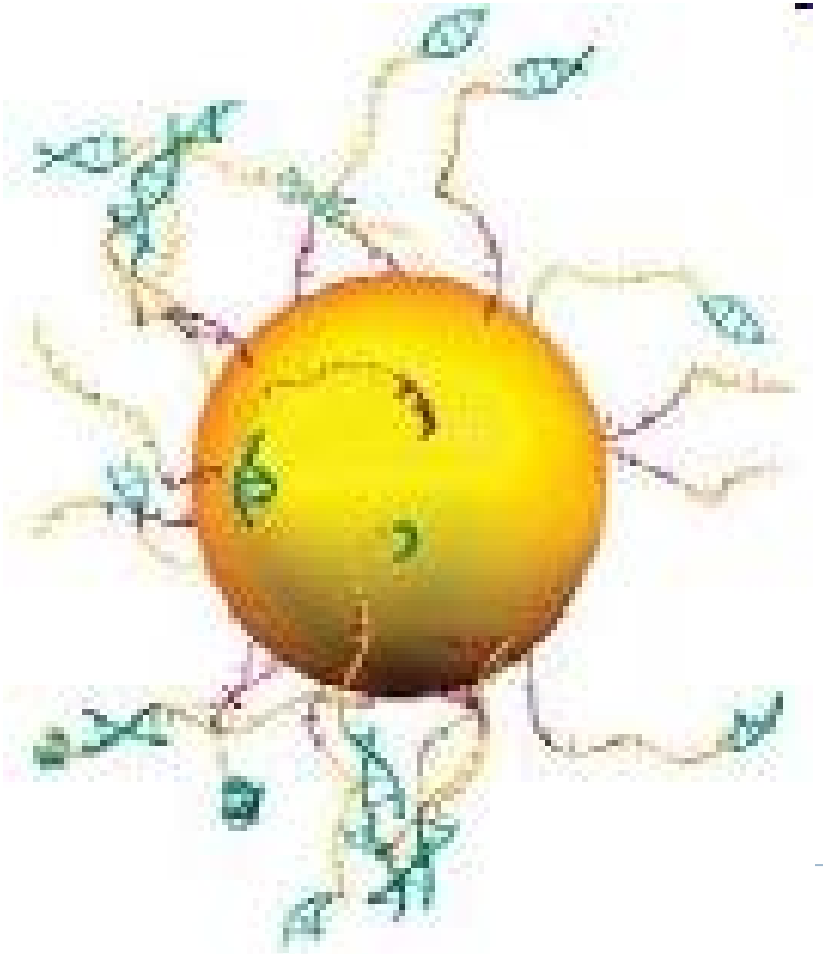
A bead and the bound fragment together with a water-in-oil emulsion form a microreactor so that each fragment can be amplified without contamination via the so called emulsion PCR (emPCR).

The entire fragment collection is amplified in parallel.

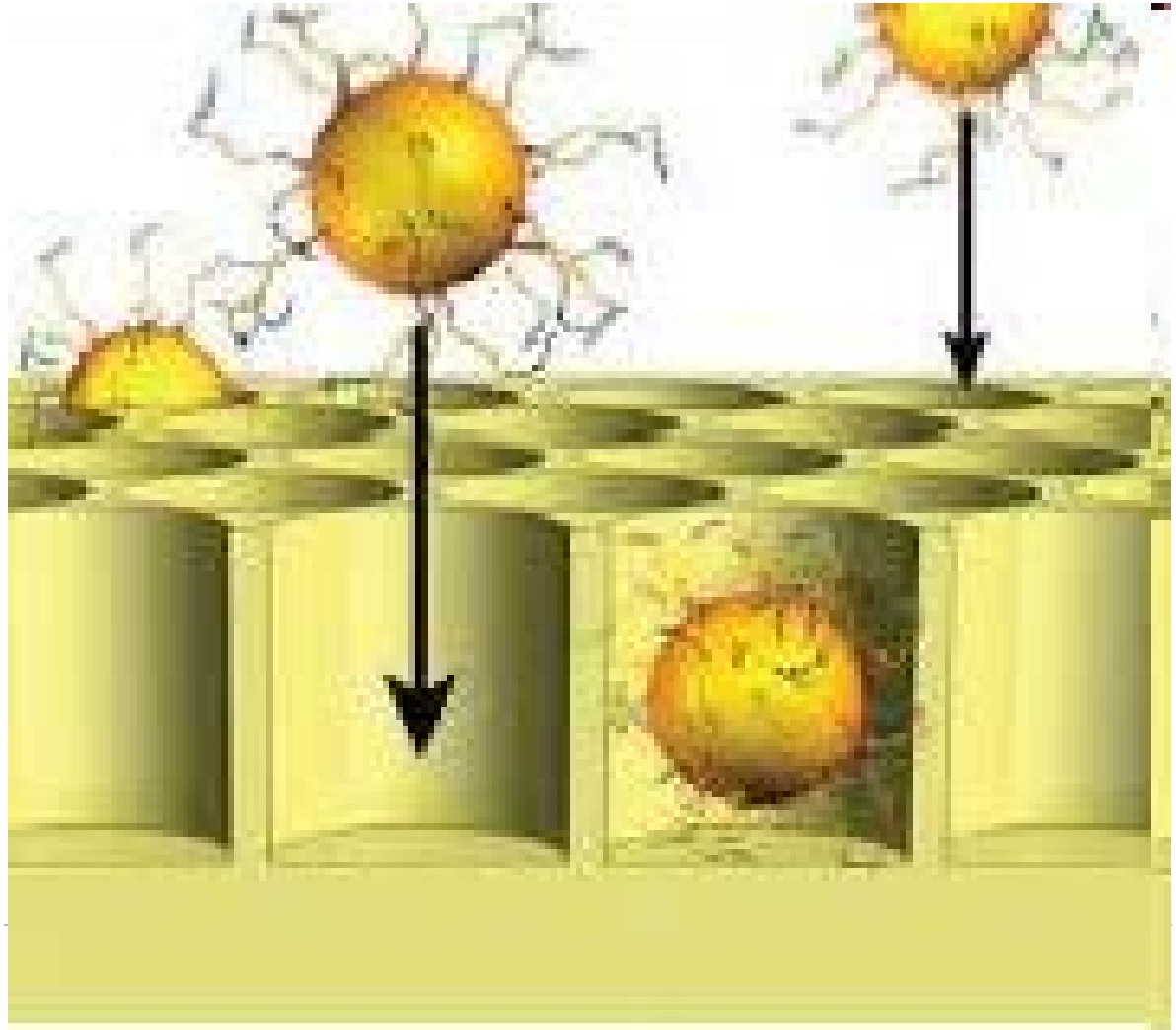


4. The emPCR amplifies each fragment several million times.

After amplification, the emulsion shell is broken and the clonally amplified beads are ready for loading onto the fibre-optic PicoTiterDevice for sequencing.



5. The PicoTiterPlate is loaded with one fragment carrying bead per well and smaller beads with the enzymes necessary for sequencing.



6. Sequencing is accomplished by synthesizing the complementary strands of the bead attached templates.

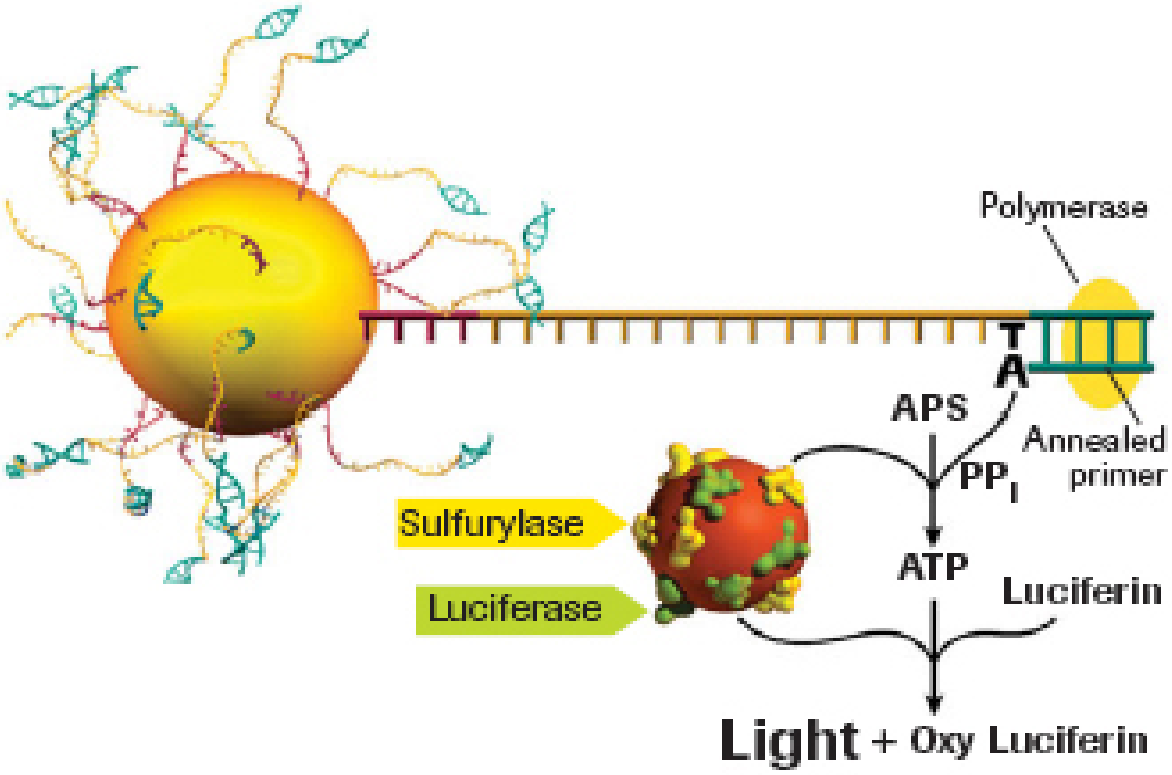
In a number of cycles the four bases (ATGC) are sequentially washed over the PicoTiterPlate.

The incorporation of a new base is associated with the release of inorganic pyrophosphate starting a chemical cascade.

This results in the generation of a light signal which is captured by a CCD camera.

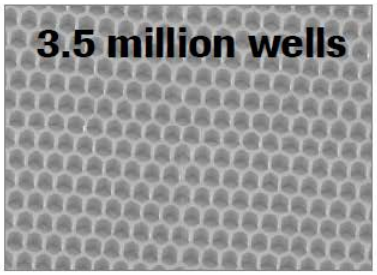
The released PPi is converted to ATP with adenosine 5′-phosphosulfate (APS) and ATP sulfurylase.

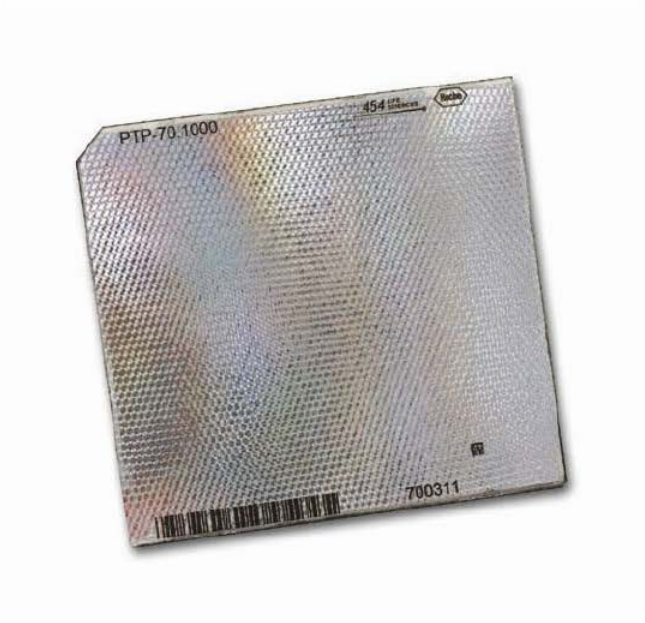
The ATP is used by luciferase in the metabolism of luciferin, which emits photons.



GS FLX Titanium Series

PicoTiterPlate Device





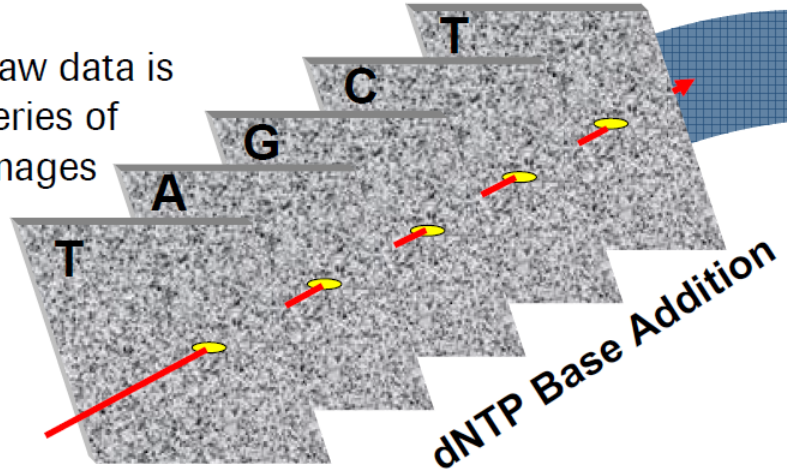
www.roche-applied-science.com

454 SEQUENCING

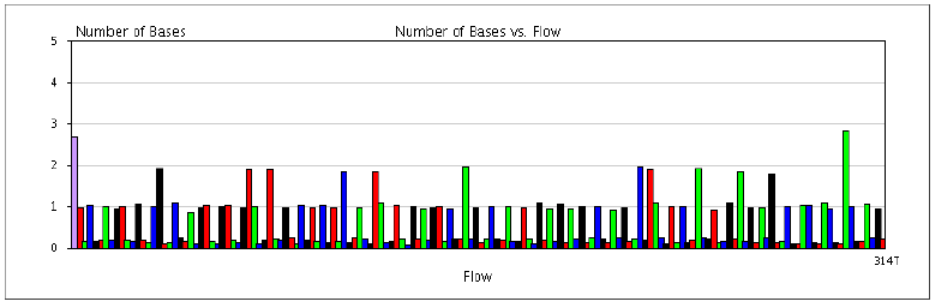
Genome Sequencer FLX Instrument Data

Image Processing Overview

1. Raw data is series of images



2. Each well's data extracted, quantified and normalized



3. Read data converted into "flowgrams"

www.roche-applied-science.com

454 SEQUENCING

3rd Generation