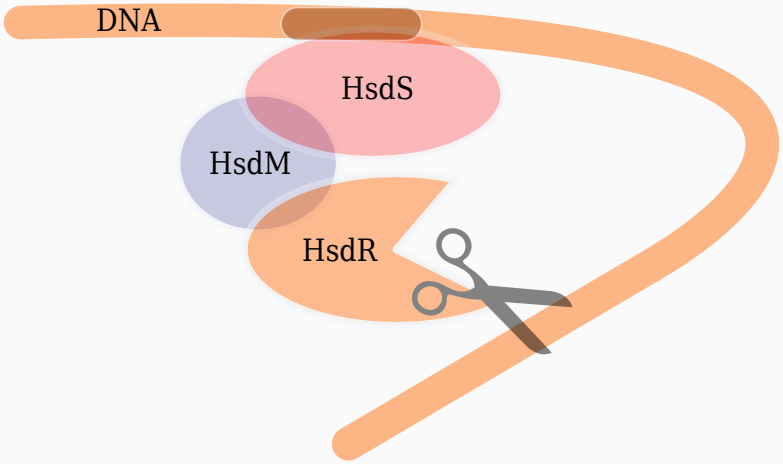
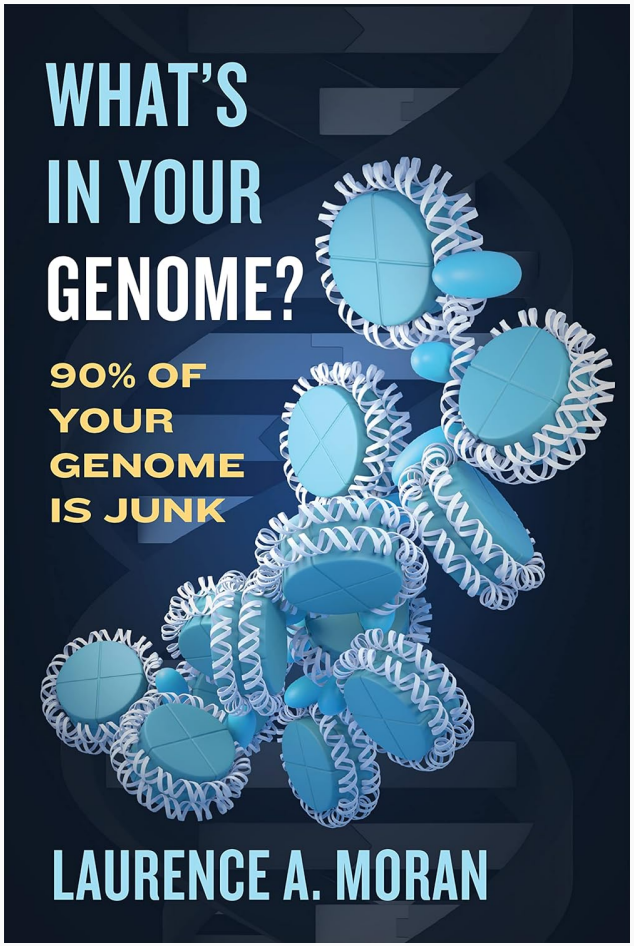


Restriction Enzymes

ICB 5751 - The Origins of Molecular Biology

May 16, 2024

On the shoulders of giants



A third problem may explain why so many scientists don't appreciate the arguments for junk DNA, and that's an ignorance of history. We live in a time where everything seems so new and exciting and baby boomers are supposedly out of touch with reality. According to this view, looking back at the work of boomers and their parents in the 1960s and 1970s is a waste of time because there can't be anything of value in that stuffy old literature. After all, they didn't even know about epigenetics and CRISPR, so how could they have understood genomes

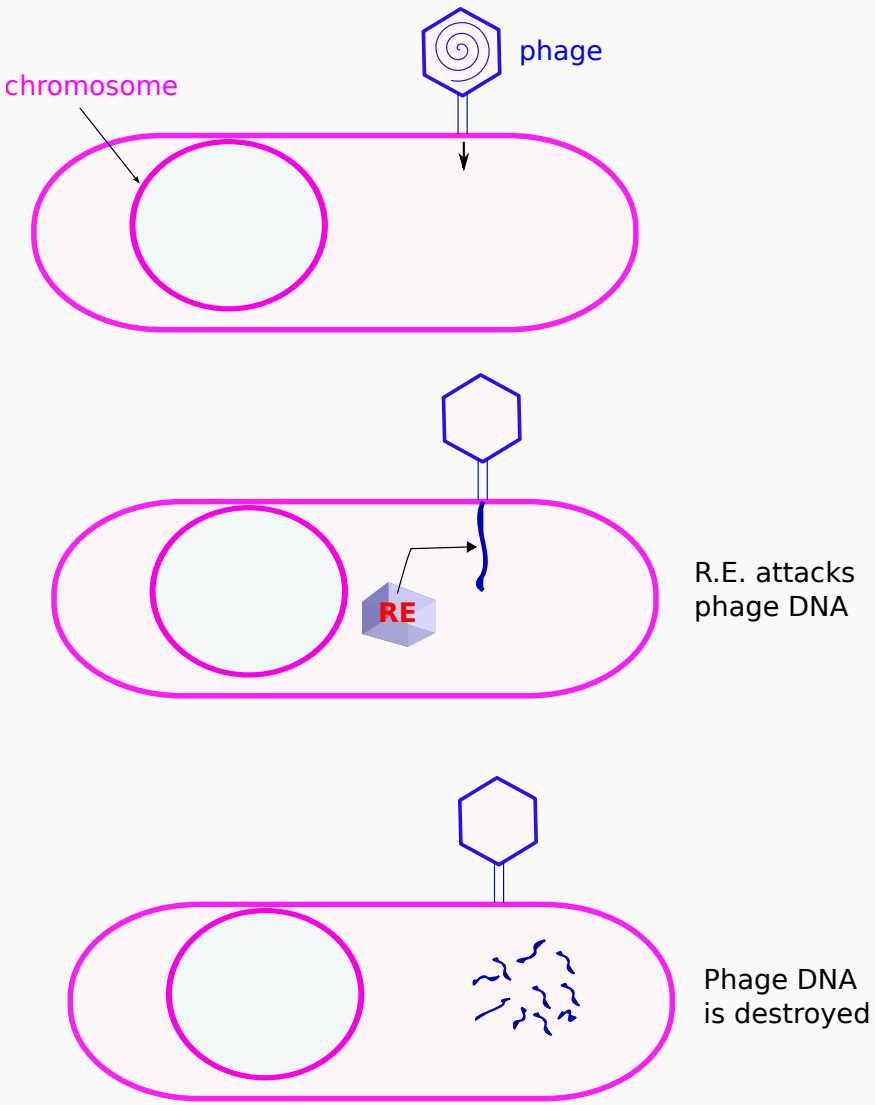
Today, it is difficult to imagine a time when our laboratory freezers were not well stocked with restriction enzymes, when DNA sequencing was not possible, or when genes were only accessible to the geneticists and could not be simply cloned out by recombinant DNA technology.

Richard J. Roberts, New England Biolabs, 2005

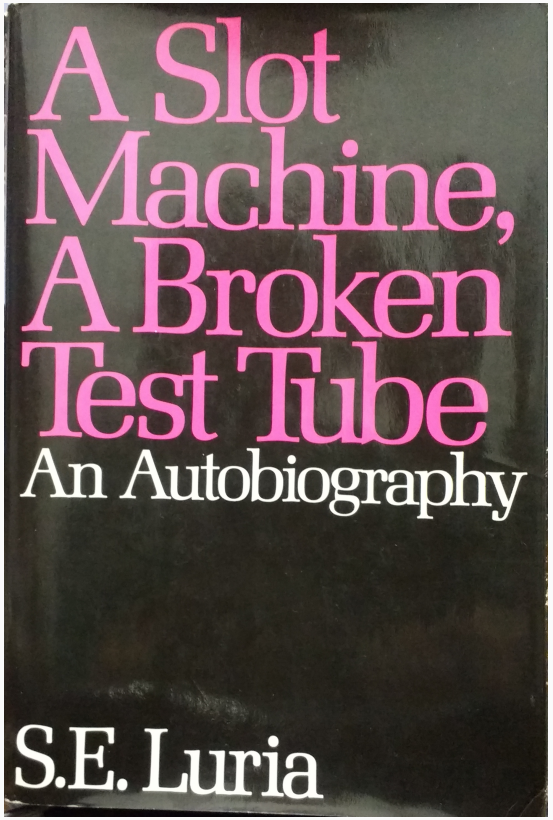
Endonucleases a.k.a. Restriction Enzymes

- Molecular scissors that cut double-strand DNA
- Produced by bacteria as a protection against phages
- As of May 2024, 5054 enzymes characterised
- 194,992 putative REs; 625 commercially available REs
- Nearly all bacteria carry at least one RE system
- Critical thinking questions:
  - Why REs are so abundant?
  - Why intracellular bacterial parasites do not usually carry RE genes?

Restriction enzyme: a deadly weapon against phages

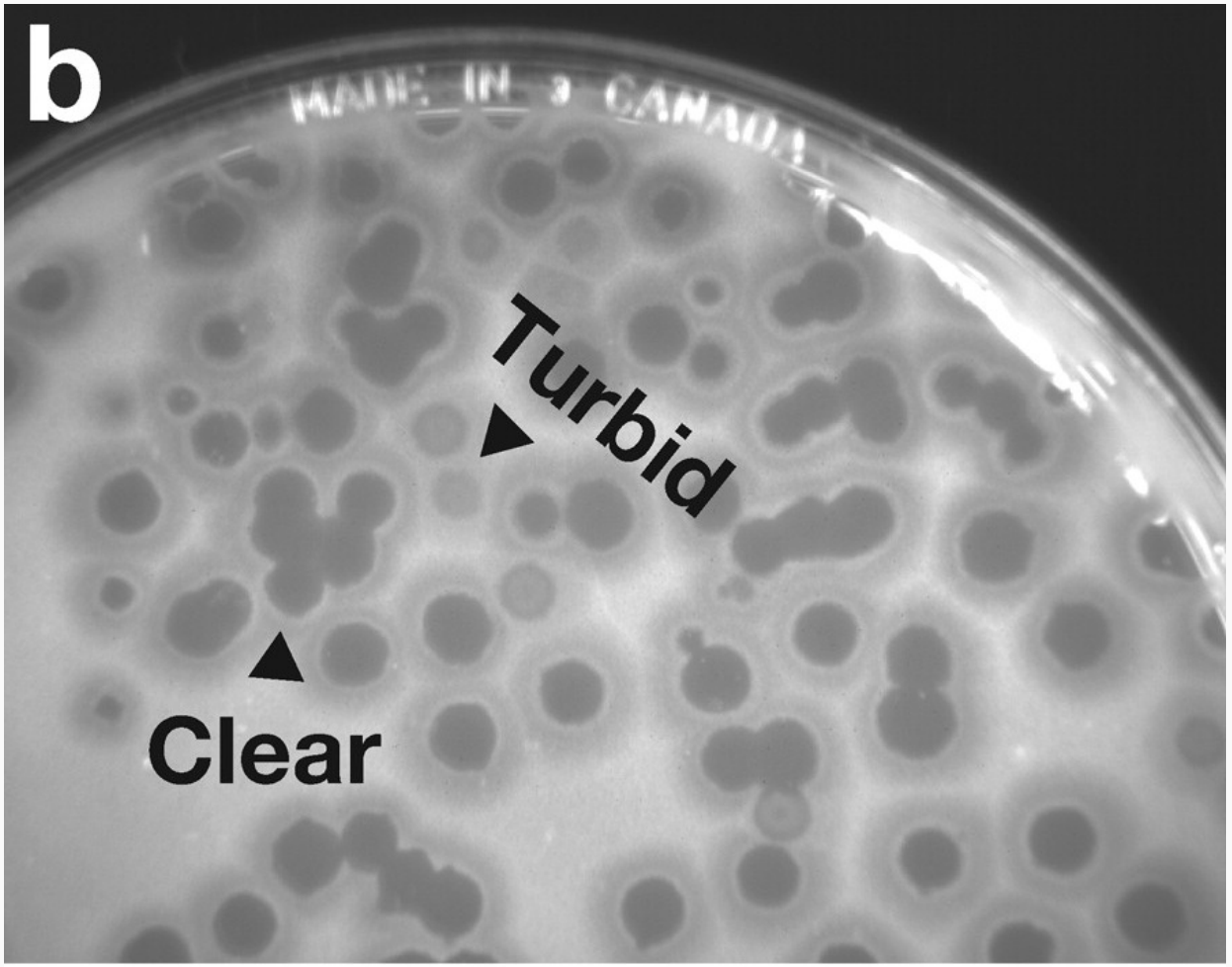


How it all began

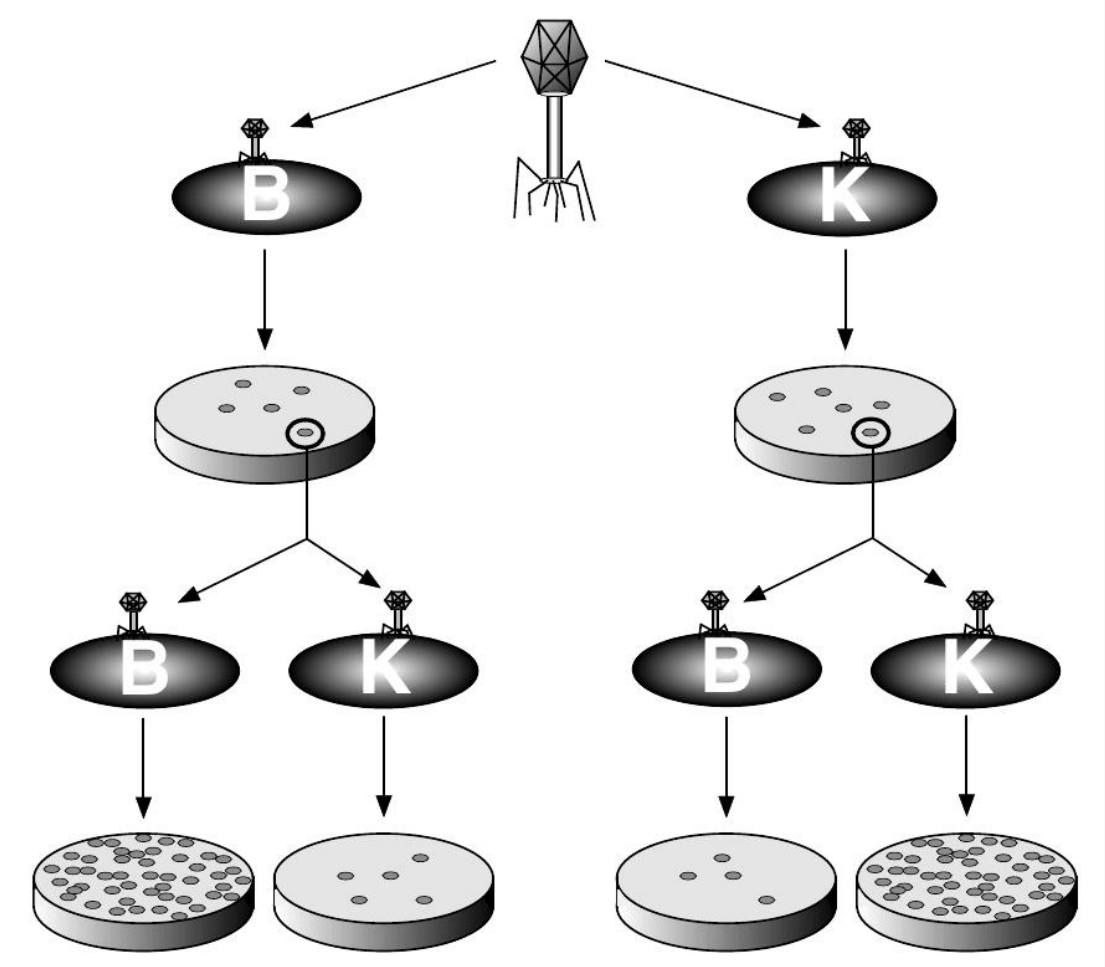


- *S. E. Luria and Mary L. Human.* A nonhreditary, host-induced variation of bacterial viruses. *J. Bacteriol.* **64**, 1952

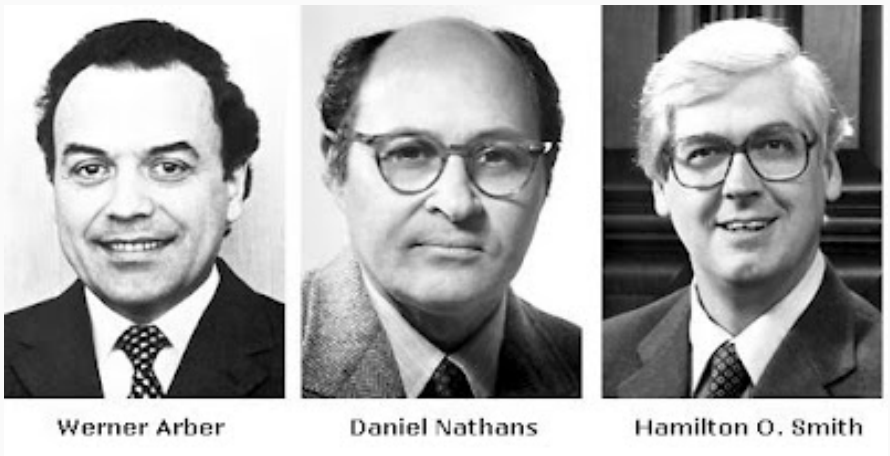
How does a phage plaque looks like?



The principle of phage restriction



The biotechnological revolution

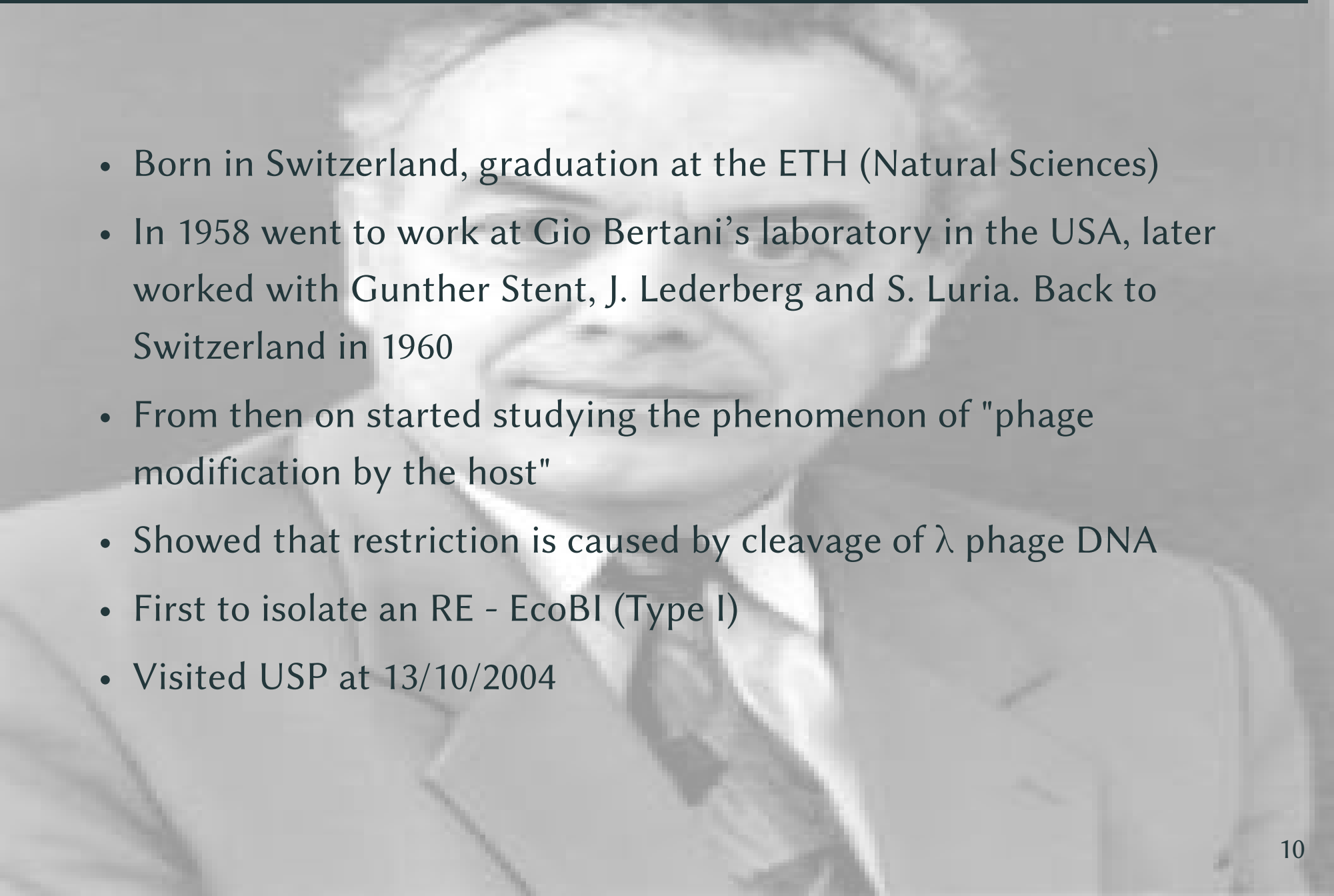


1978 Nobel Prize

Some examples:

- Insulin and other hormones
- Recombinant vaccines
- Plants resistant to pests, herbicides, drought
- Gene therapy
- …

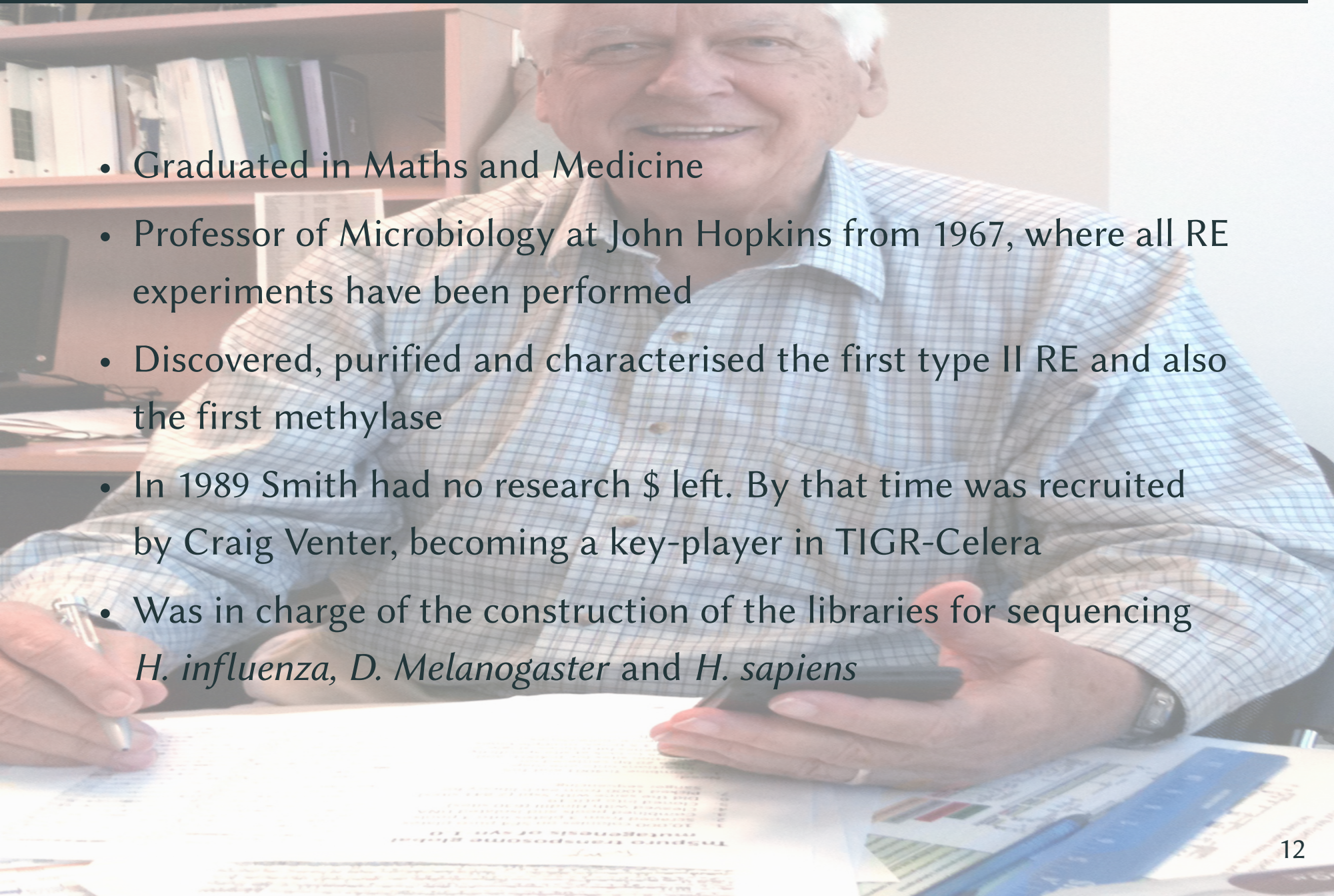
Werner Arber (1929-)



- Born in Switzerland, graduation at the ETH (Natural Sciences)
- In 1958 went to work at Gio Bertani's laboratory in the USA, later worked with Gunther Stent, J. Lederberg and S. Luria. Back to Switzerland in 1960
- From then on started studying the phenomenon of "phage modification by the host"
- Showed that restriction is caused by cleavage of λ phage DNA
- First to isolate an RE - EcoBI (Type I)
- Visited USP at 13/10/2004

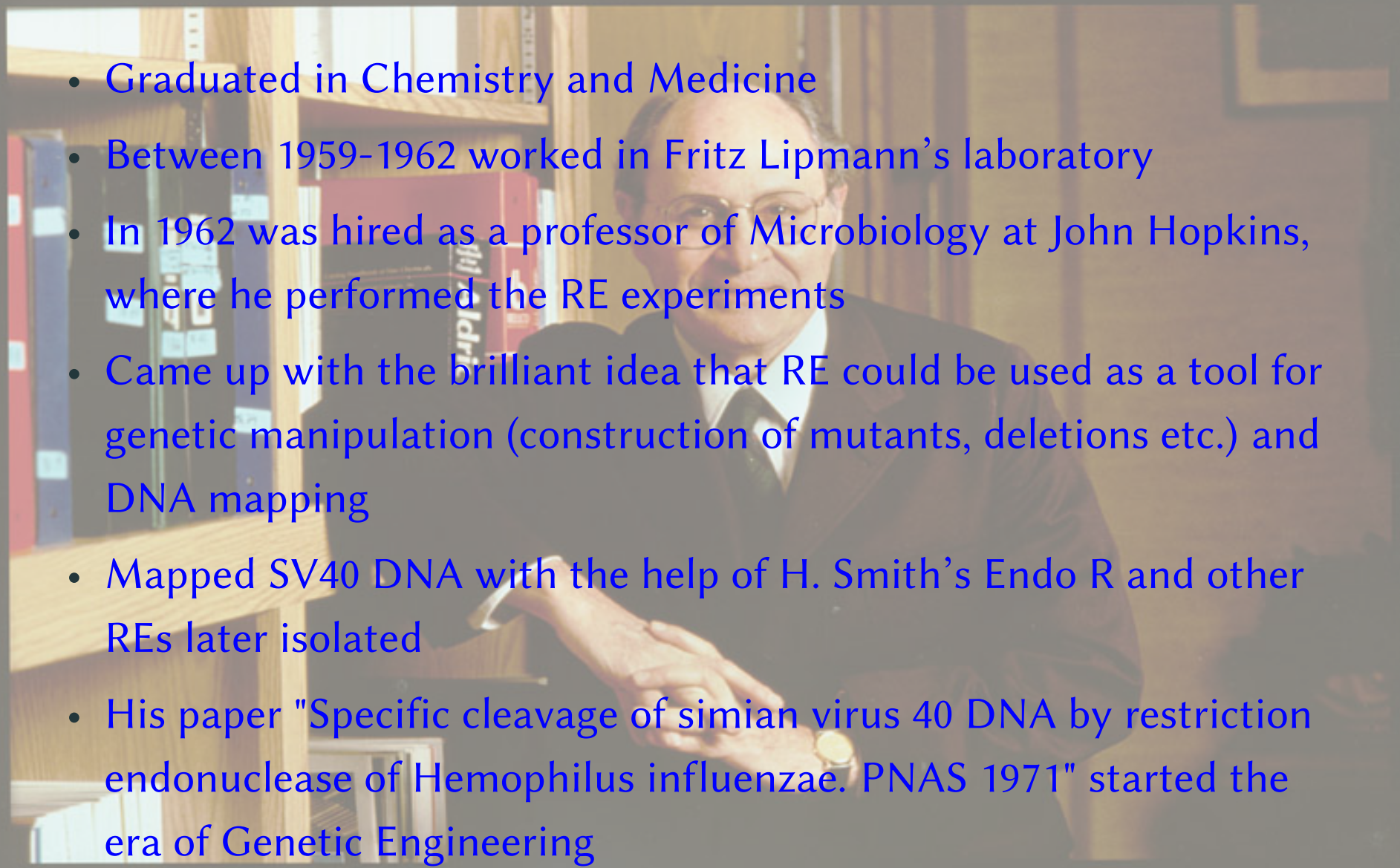
Werner Arber at USP

Hamilton Smith (1931-)



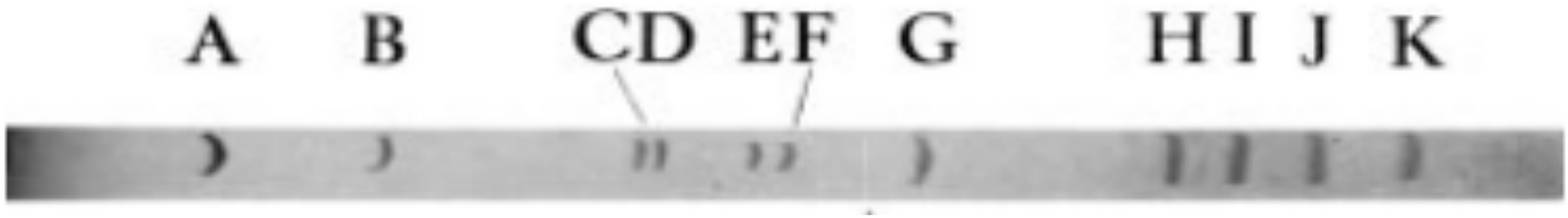
- Graduated in Maths and Medicine
- Professor of Microbiology at John Hopkins from 1967, where all RE experiments have been performed
- Discovered, purified and characterised the first type II RE and also the first methylase
- In 1989 Smith had no research $ left. By that time was recruited by Craig Venter, becoming a key-player in TIGR-Celera
- Was in charge of the construction of the libraries for sequencing *H. influenza, D. Melanogaster* and *H. sapiens*

Daniel Nathans (1928-1999)



- Graduated in Chemistry and Medicine
- Between 1959-1962 worked in Fritz Lipmann's laboratory
- In 1962 was hired as a professor of Microbiology at John Hopkins, where he performed the RE experiments
- Came up with the brilliant idea that RE could be used as a tool for genetic manipulation (construction of mutants, deletions etc.) and DNA mapping
- Mapped SV40 DNA with the help of H. Smith's Endo R and other REs later isolated
- His paper "Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae. PNAS 1971" started the era of Genetic Engineering

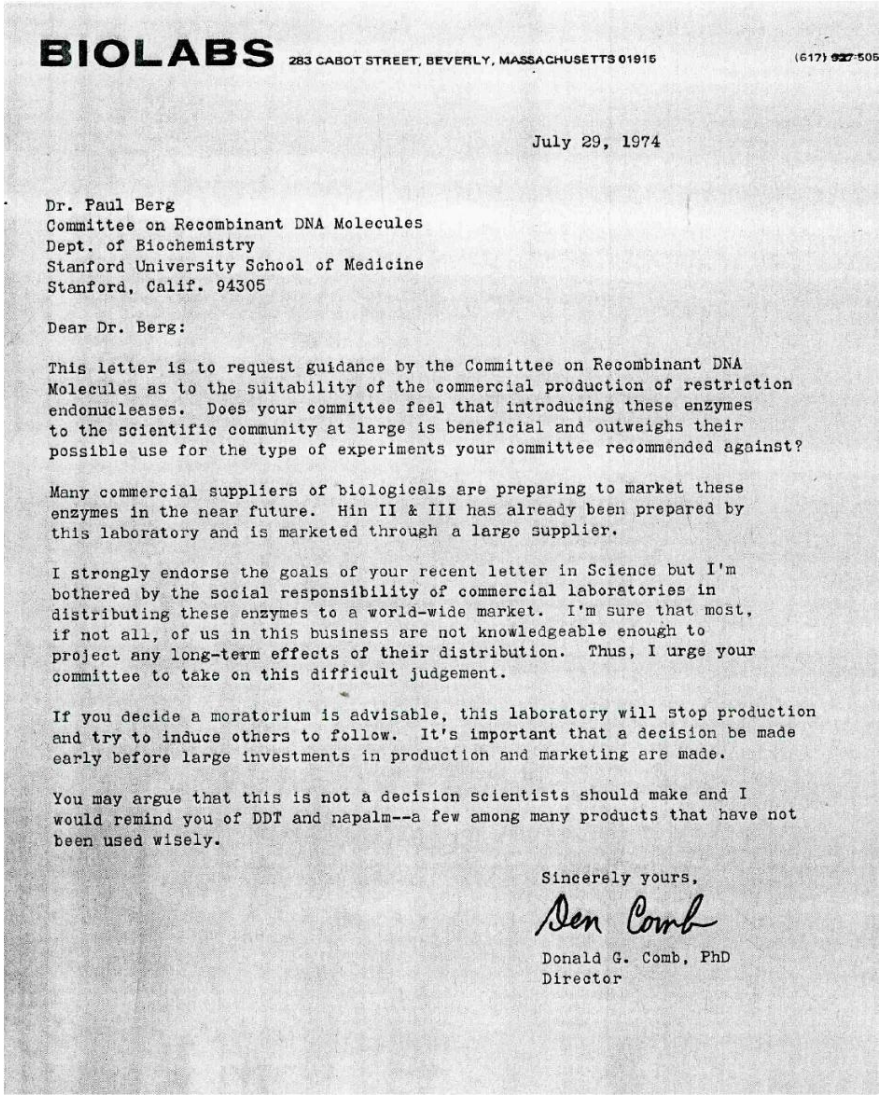
- Nathans was the first to use polyacrylamide electrophoresis (in glass tubes) to resolve DNA fragments



Radioautogram of ^{14}C-labeled SV40 DNA cleaved with endonuclease R (Danna & Nathans 1971)

- He later sugested using REs for genomic mapping. It was the precursor of RFLP

The "danger" of Restriction Enzimes



BIOLABS 283 CABOT STREET, BEVERLY, MASSACHUSETTS 01915 (617) 927-505

July 29, 1974

Dr. Paul Berg
Committee on Recombinant DNA Molecules
Dept. of Biochemistry
Stanford University School of Medicine
Stanford, Calif. 94305

Dear Dr. Berg:

This letter is to request guidance by the Committee on Recombinant DNA Molecules as to the suitability of the commercial production of restriction endonucleases. Does your committee feel that introducing these enzymes to the scientific community at large is beneficial and outweighs their possible use for the type of experiments your committee recommended against?

Many commercial suppliers of biologicals are preparing to market these enzymes in the near future. Hin II & III has already been prepared by this laboratory and is marketed through a large supplier.

I strongly endorse the goals of your recent letter in Science but I'm bothered by the social responsibility of commercial laboratories in distributing these enzymes to a world-wide market. I'm sure that most, if not all, of us in this business are not knowledgeable enough to project any long-term effects of their distribution. Thus, I urge your committee to take on this difficult judgement.

If you decide a moratorium is advisable, this laboratory will stop production and try to induce others to follow. It's important that a decision be made early before large investments in production and marketing are made.

You may argue that this is not a decision scientists should make and I would remind you of DDT and napalm--a few among many products that have not been used wisely.

Sincerely yours,

Don Comb

Donald G. Comb, PhD
Director

Why bacterial DNA is not affected by its own RE?

- Werner Arber and Hamilton Smith showed that REs are always accompanied by DNA modification activity

Restriction (phage DNA)

5'- G A A T T C - 3'
3'- C T T A A G - 5'

Digestion with EcoRI

5'- G A A T T C - 3'
3'- C T T A A G - 5'

Modification (bacterial DNA)

5'- G A A T T C - 3'
3'- C T T A A G - 5'

Modification with EcoRI

CH_3

5'- G A A T T C - 3'
3'- C T T A A G - 5'

CH_3

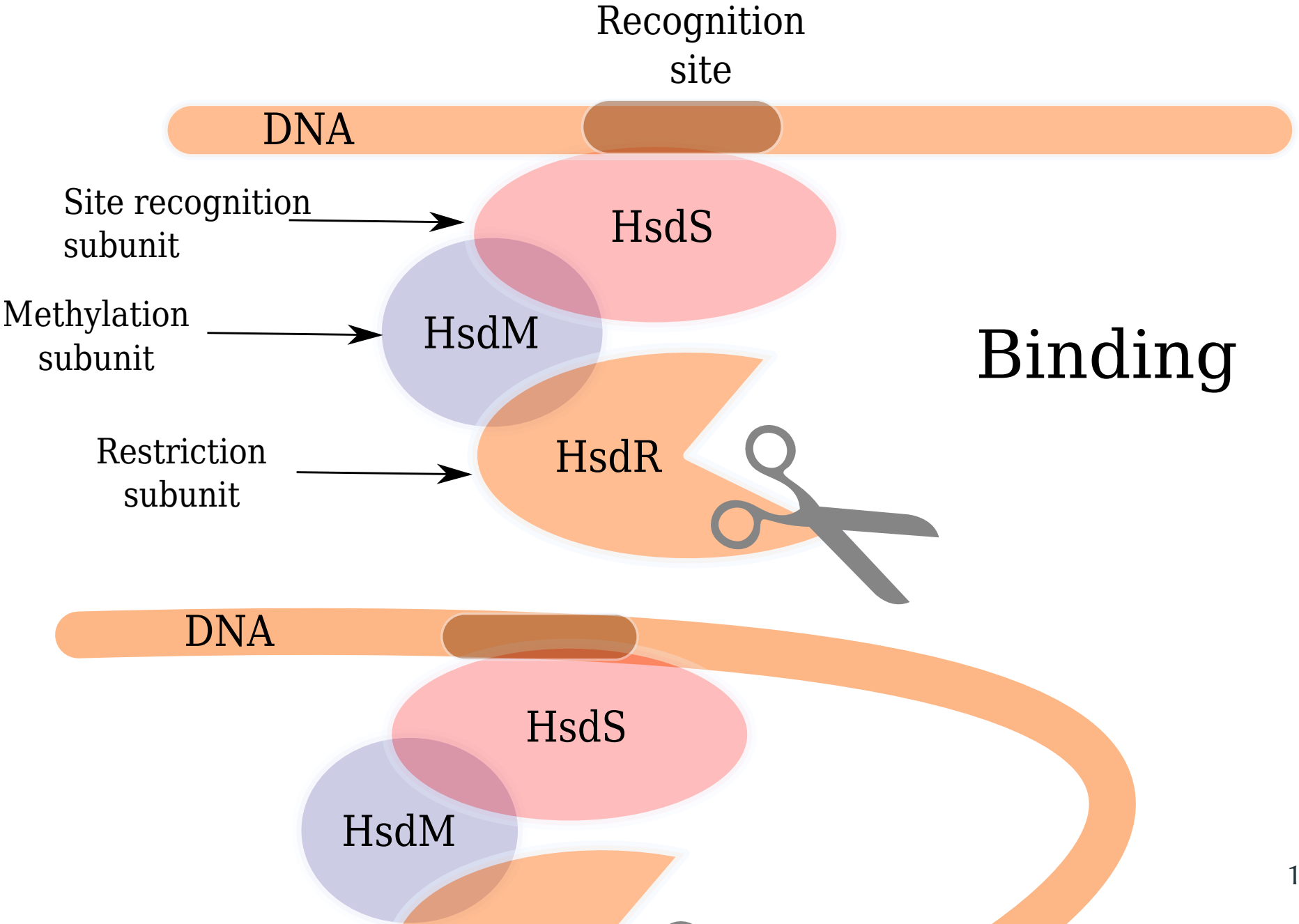
Critical thinking questions

- It is possible to isolate R^- M^- and R^-M^+ bacteria, but not R^+M^-. Why mutants of this last category cannot be isolated?
- Upon chromosome replication, only one DNA strand of each new chromosomes is methylated. How do R-M systems operate?
- Why bacteria do not methylate an infecting DNA virus?
- Is R-M methylation a case of epigenetic inheritance?

Restriction Enzyme types: I, II, III e IV

- Type I: Recognizes a specific sequence, but unspecifically cuts DNA 1000 bp ahead
- Built-up by 3 subunits: DNA-recognition, methylation and restriction (HsdS, HsdM e HsdR)
- Example: EcoK ($AACN_6GTGC$)
- Suicidal enzyme

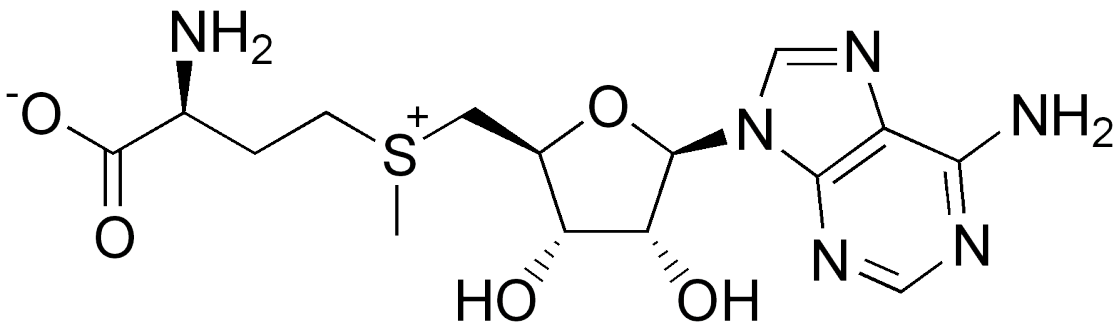
Type I Restriction Enzymes



Type III Restriction Enzymes

Features:

- A single protein with catalytic subunits for restriction and methylation
- Cuts at 24-26 bp away from the recognition site
- Example: EcoP15I (CAGCAGN_{25}^)
- RE type I and III require ATP and S-adenosyl-methionine



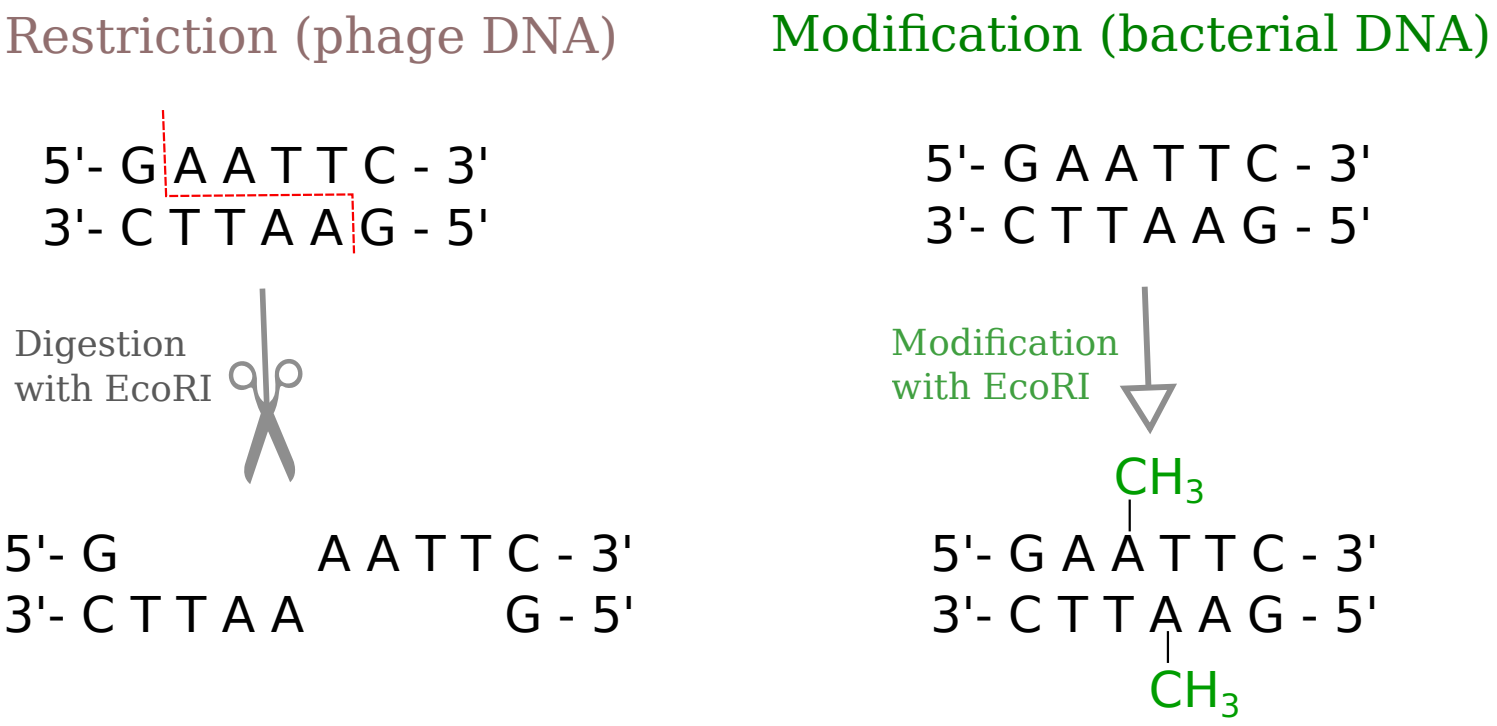
Type II Restriction Enzymes

Features

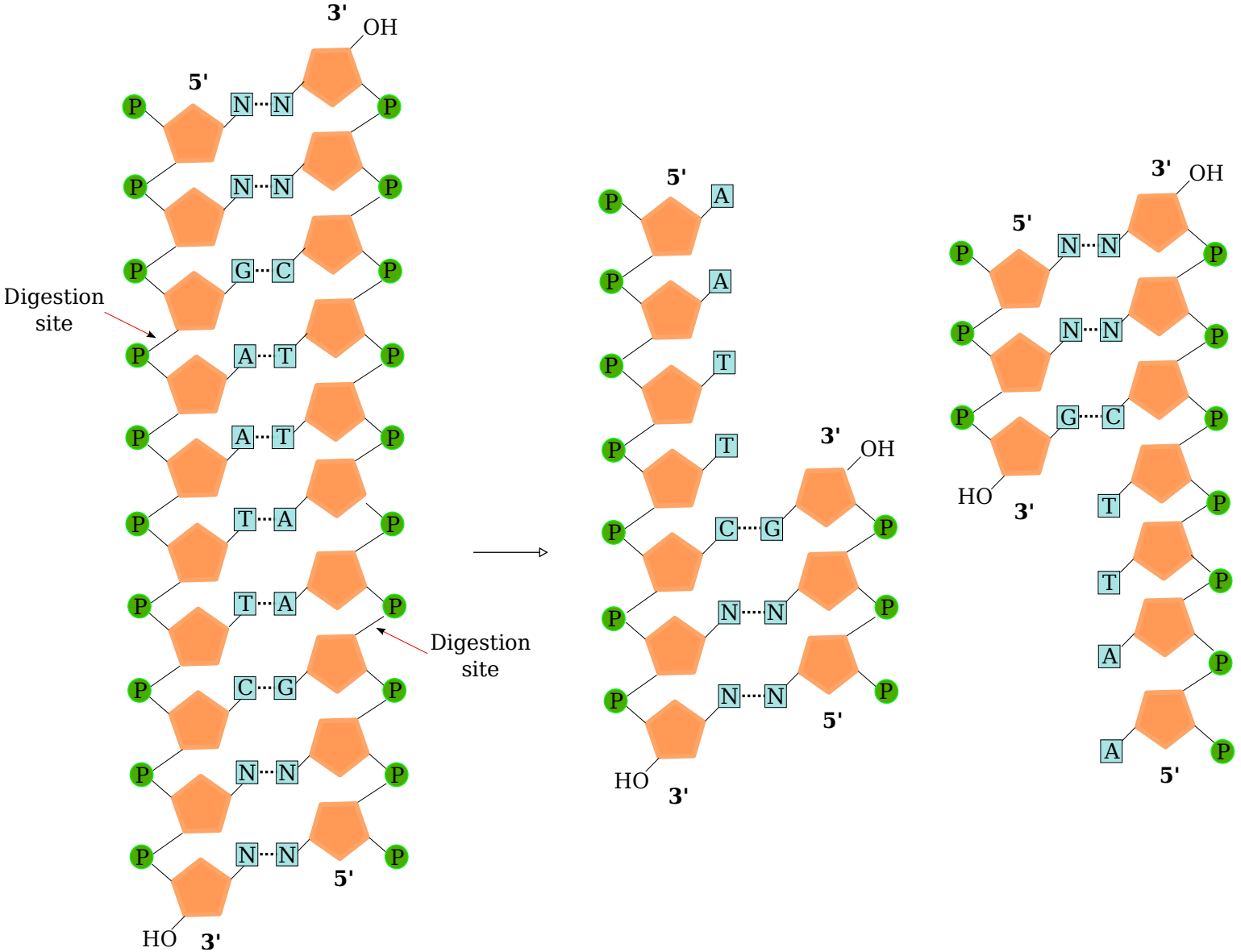
- Recognise palindromes, cleave at the same site
- Widely used in genetic engeneering
- Most abundant REs (4884 out of 5054 REs on May 2024)
- Two proteins: an endonuclease and a methylase
- RE does not require ATP or SAM, only Mg^{2+}

Restriction and Methylation by type II enzymes

- The RE breaks the phosphodiester backbone and the MTase methylates both strands
- DNA digestion generates 5'-cohesive ends, 3'-cohesive ends or blunt ends



EcoRI digestion



Examples of type II REs

Enzyme	Site	Ends	Enzyme source
EcoRI	G^AATTC	5' sticky	*Escherichia coli* RY13
BamHI	G^GATCC	5' sticky	*Bacillus amyloliquefaciens* H
BglII	A^GATCT	5' sticky	*Bacillus globigii*
PstI	CTGCA^G	3' sticky	*Providencia stuartii*
XmaI	C^CCGGG	5' sticky	*Xanthomonas malvacearum*
SmaI	CCC^GGG	blunt	*Serratia marcescens*
Sau3A	^GATC	5' sticky	*Staphylococccus aureus* 3A
AluI	AG^CT	blunt	*Arthrobacter luteus*
NotI	GC^GGCCGC	5' sticky	*Nocardia otitidis-caviarum*
PacI	TTAAT^TAA	3' sticky	*Pseudomonas alcaligenes*

Exercise

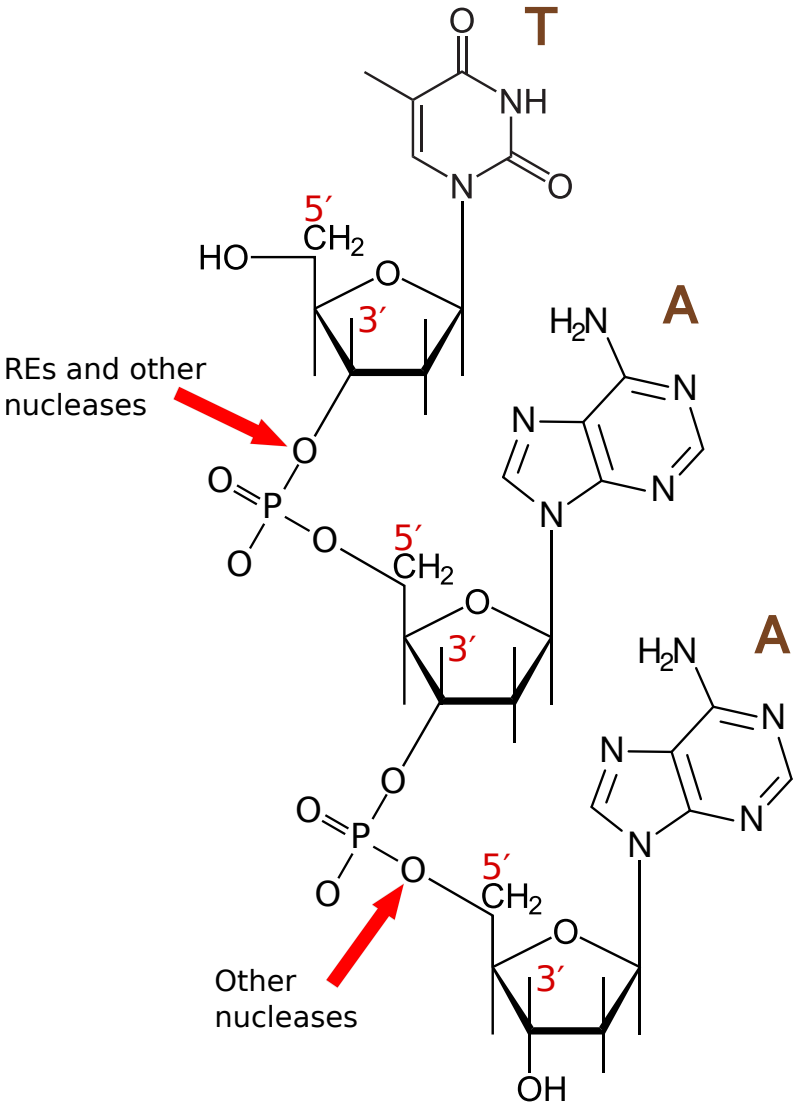
- Draw the ends of both DNA strands obtained after digestion with BamHI (G^GATCC), PstI (CTGCA^G), Sau3A (^GATC), SmaI (CCC^GGG) or XmaI (C^CCGGG)

Example: EcoRI

5'GAATTC → G-OH P-AATTC

3'CTTAAG CTTAA-P HO-G

Other endonucleases

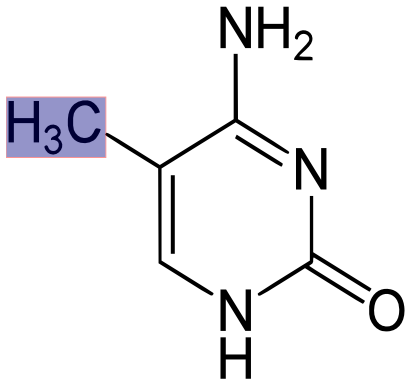


Examples of palindromes

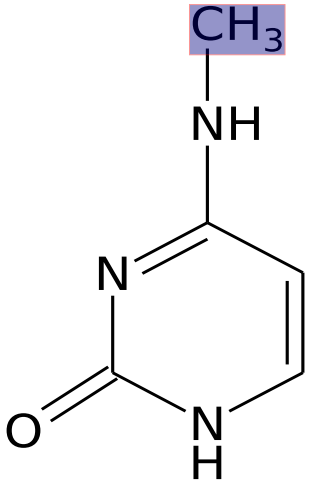
- noon, civic, radar, level, rotor, kayak, racecar, madam, refer
- Mr. Owl ate my metal worm
- Never odd or even
- 5’-GAATTC CTGCAG GATC
  3’-CTTAAG GACGTC CTAG

Types of Methylation

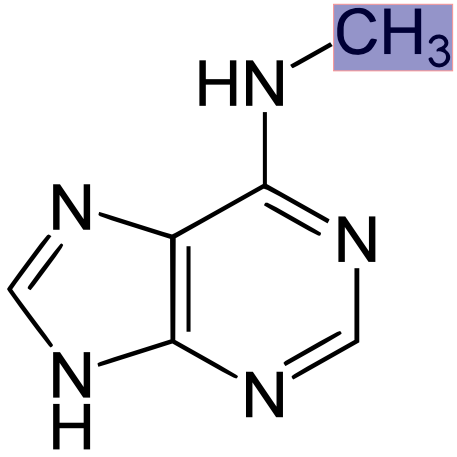
- The CH_3 donor is always S-adenosyl-methionine (SAM)
- Type II REs do not depend on SAM (unlike RE types I and III)
- The methylated atom varies as follows:



5-methylcitosine



N4-methylcitosine



N6-methyladenine

Nomenclature

- Suggested by Hamilton Smith and Daniel Nathans
- Four letters:
  - first letter = **G**enus
  - Second and third letters = **s**pecies
  - Fourth letter = **S**train name or another characteristic
  - Numeral = indicates the order in which the enzyme was isolated from that particular organism
- Examples:
  - BamHI = *Bacillus amyloliquefaciens*, strain H, Enzyme n° 1
  - EcoRV = *Escherichia coli*, strain RY13, Enzyme n° 5

Frequency of RE sites in the genome

- What is the expected frequency of 4 bp, 5 bp, 6 bp, 8 bp and 10 bp site in *E. coli* (50% GC)?
- What is the expected frequency of HincII (GTPy^PuAC)?
- What is the expected frequency for the CCC^GGG (SmaI) site in an organism with 33% GC?
- Real frequency depends on the % of G+C in the genome and other factors

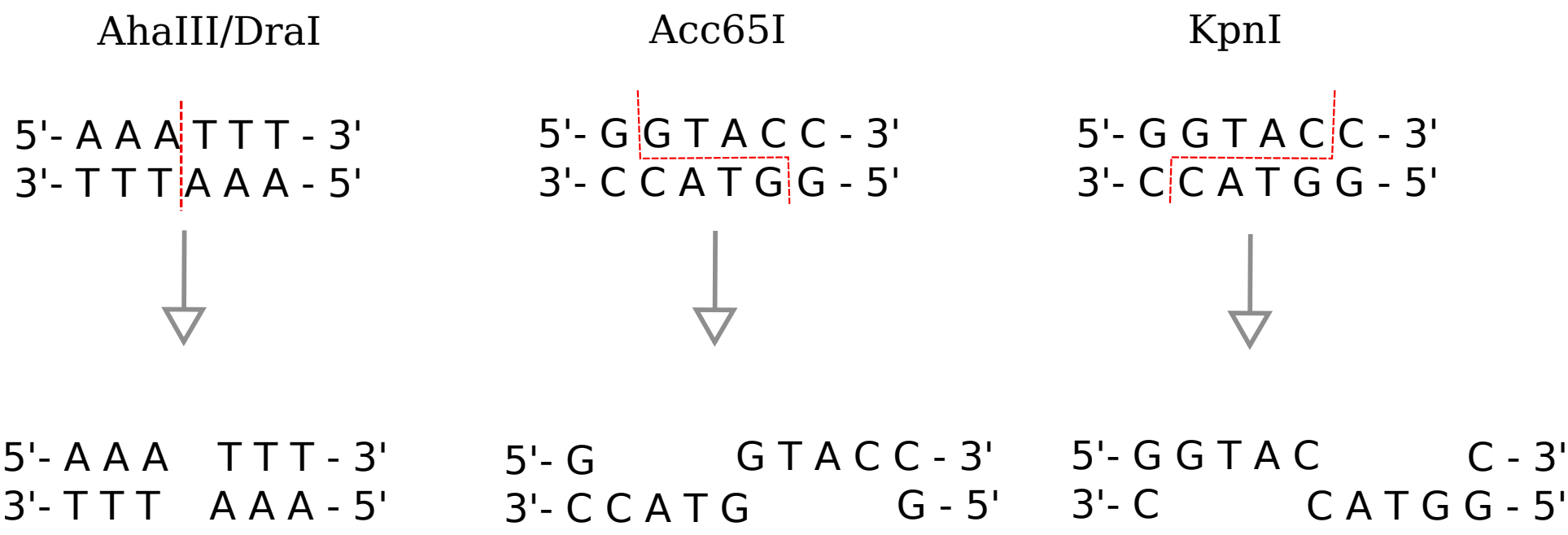
In *E. coli* (50% GC):

Enzyme	Site	N° Sites	Theor. Freq.	Real Freq.
BamHI	G^GATCC	494	1/4096	1/9392
DraI	TTT^AAA	1679	1/4096	1/2763
EcoRI	GAA^TTC	645	1/4096	1/7193
HindIII	A^AGCTT	556	1/4096	1/8344
SmaI	CCC^GGG	426	1/4096	1/10891
NotI	GC^GGCCGC	23	1/65536	1/201725
HaeIII	GG^CC	12571	1/256	1/369

Isoschizomers

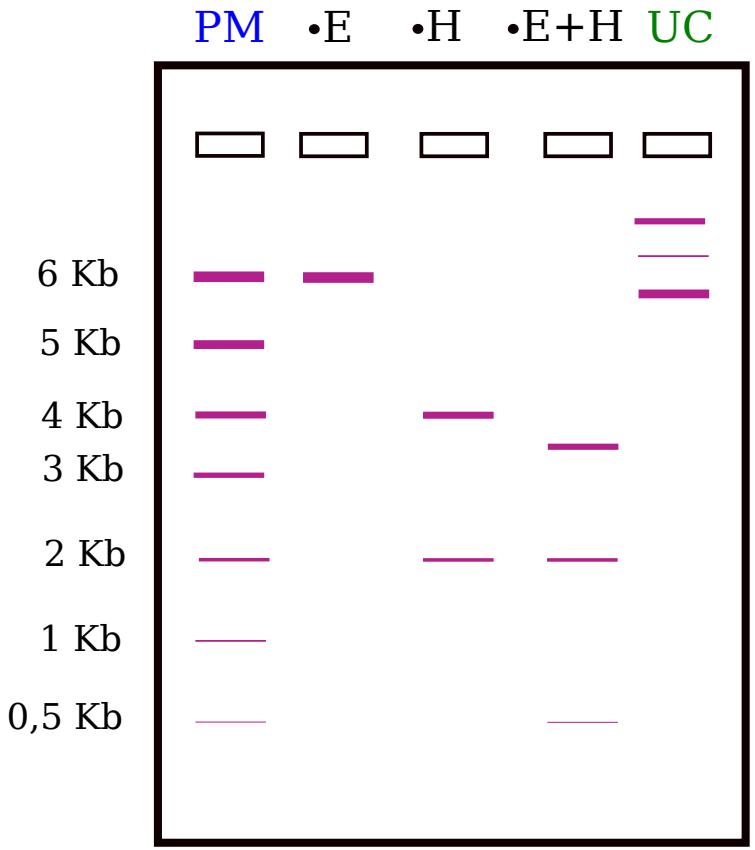
Definition
Different enzymes that act on the same site, generating (or not) identical ends

- From greek: *iso* = same + *skhizo* = cut
- Examples:



Restriction Enzyme Applications

Draw based on the gel below a restriction map. What is the size of the DNA molecule? Is the DNA linear or circular?



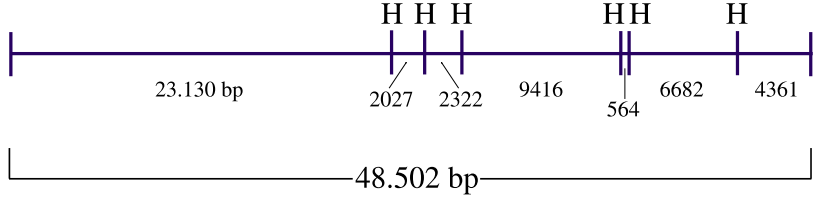
DNA Mapping

A DNA fragment of 15 kb was digested with BamHI, EcoRI and PstI, as shown in the table below, resulting in fragments with the following sizes:

BamHI	EcoRI	PstI	BamHI +EcoRI	BamHI +PstI	EcoRI +PstI	BamHI +EcoRI +PstI
14	12	8	11	8	7	6
1	3	7	3	6	5	5
			1	1	3	3
						1

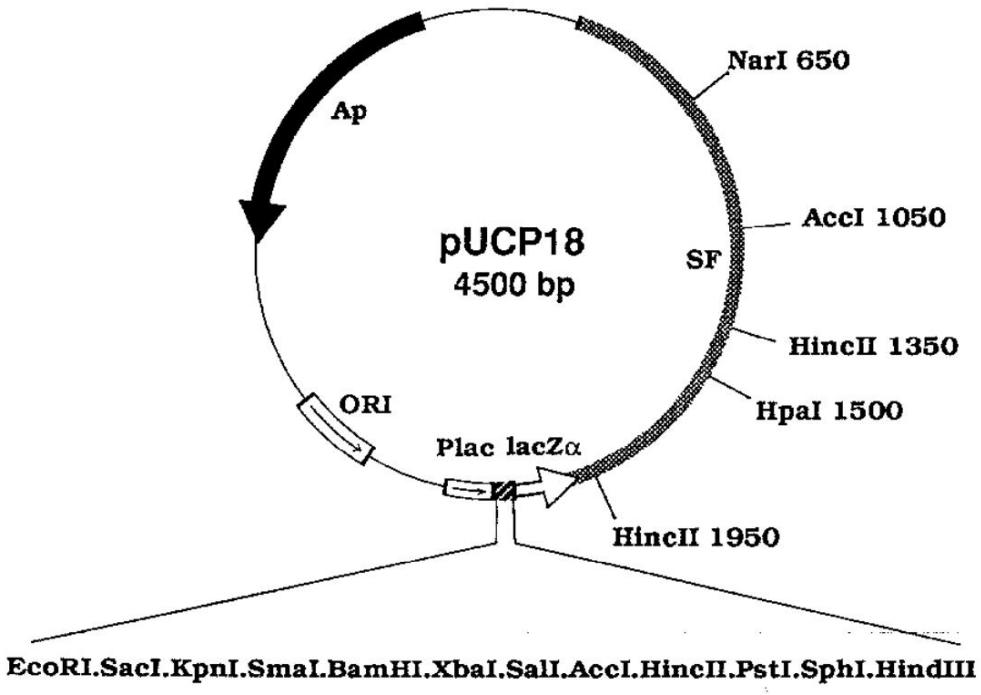
Draw a 15 Kb DNA fragment map naming the restriction enzymes and the distances between the RE sites.

Example (λ phage restriction map):



Exercise

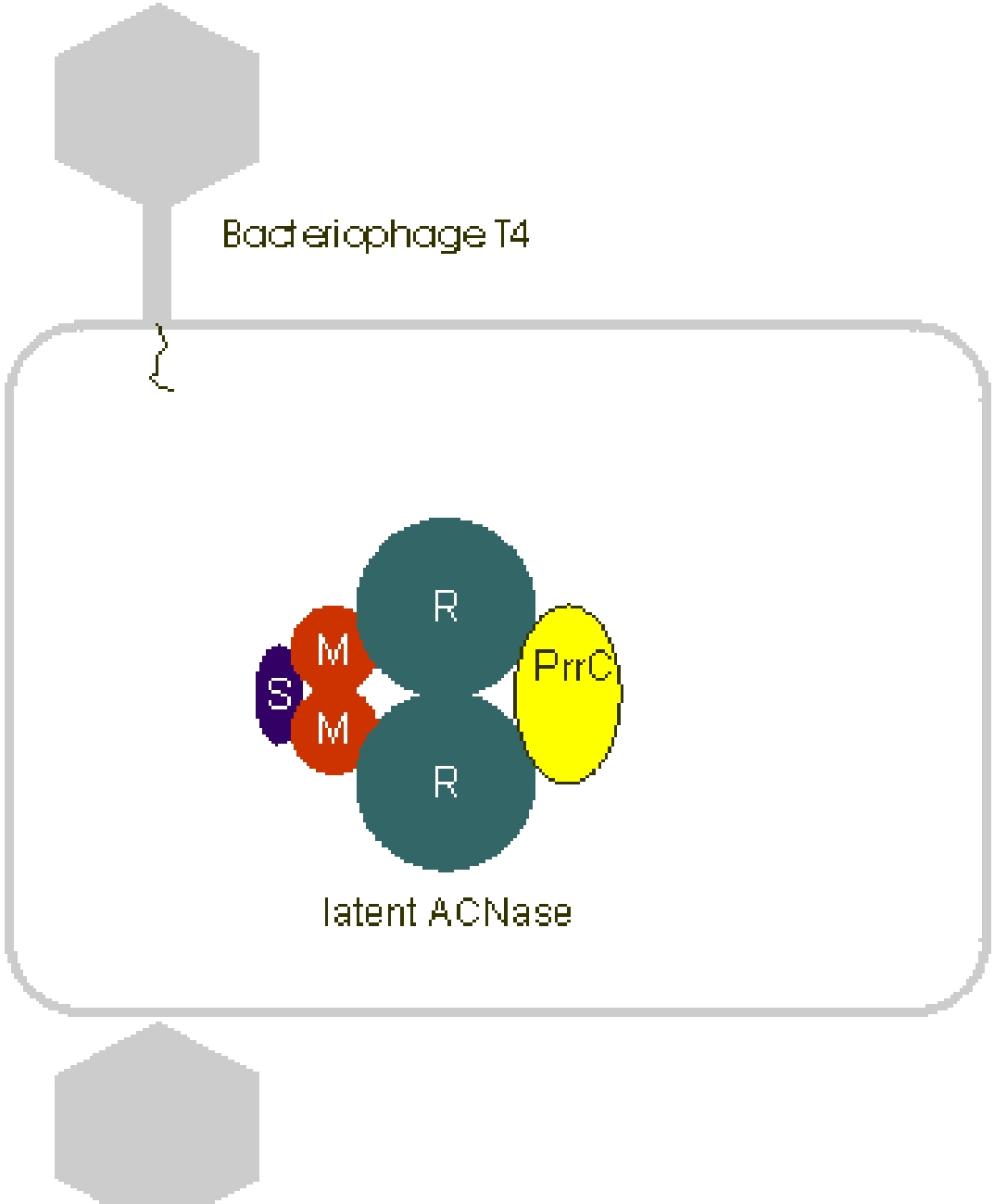
Congrats! You have cloned your favourite gene (*yfg*) of 2.7 Kb in the BamHI site of pUCp18. The *polylinker* (MCS) of pUCp18 and the restriction map of *yfg* are shown below:



Start	End	Enzyme	Restriction_site
1	6	BamHI	GGATCC
68	73	BclI	TGATCA
200	205	PvuII	CAGCTG
429	434	BclI	TGATCA
478	483	SspI	AATATT
562	567	AhaIII*	TTTAAA
675	680	HindII	GTYRAC
915	920	ClaI	ATCGAT
928	933	PvuII	CAGCTG
1077	1082	EcoRI	GAATTC
1173	1178	NcoI	CCATGG
1183	1188	ClaI	ATCGAT
1324	1329	ClaI	ATCGAT
1408	1413	EcoRI	GAATTC
1452	1457	BclI	TGATCA
2701	2706	BamHI	GGATCC

- What are the expected fragments if *yfg* (not cloned yet) would be digested with BamHI, BclI, ClaI ou EcoRI?
- How would you determine the *yfg* gene orientation relative to the P*lac* promoter?

The empire counter-attacks

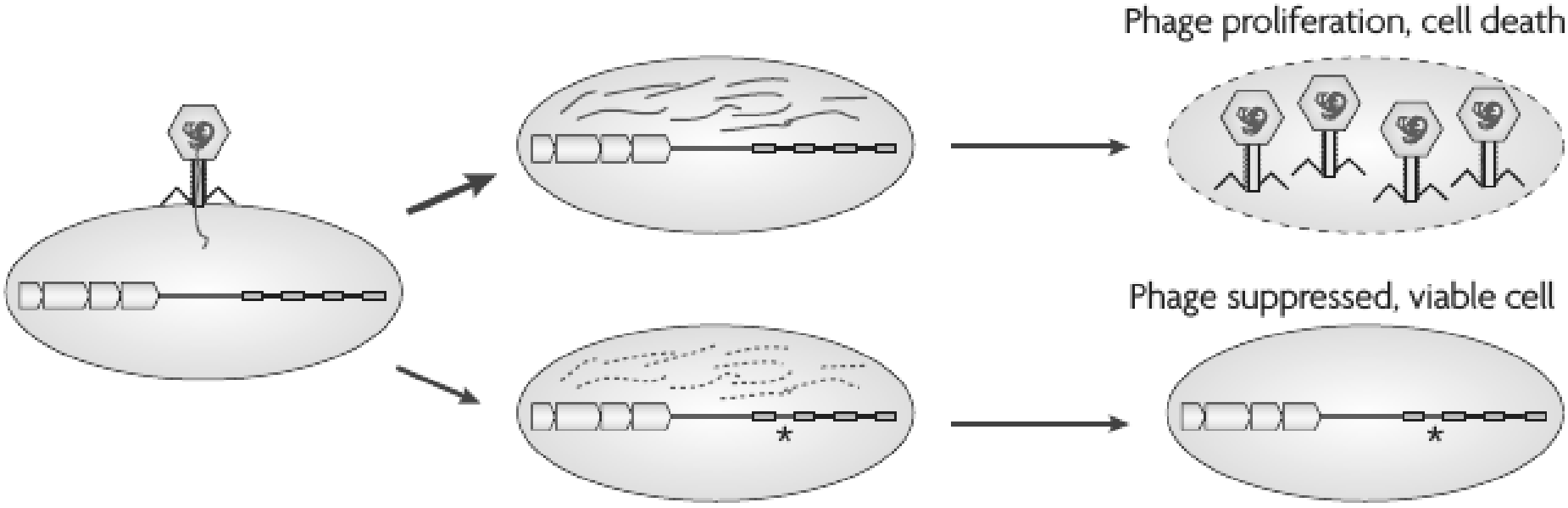


CRISPR-Cas system

- Clustered, Regularly Interspaced Short Palindromic Repeats (CRISPR) – 40% of Bacteria and 90% of Archae genomes
- Direct repeats (24-47 bp) separated by “spacers” (26-72 bp)
- Most repeats contain palindromic sequences (5-7 bp)
- Spacers contain phage DNA sequences
- A 550 bp leader sequence (rich in AT) at CRISPR 5’
- CAS genes encode nucleases, helicases, regulatory proteins

CRISPR-Cas function in bacteria

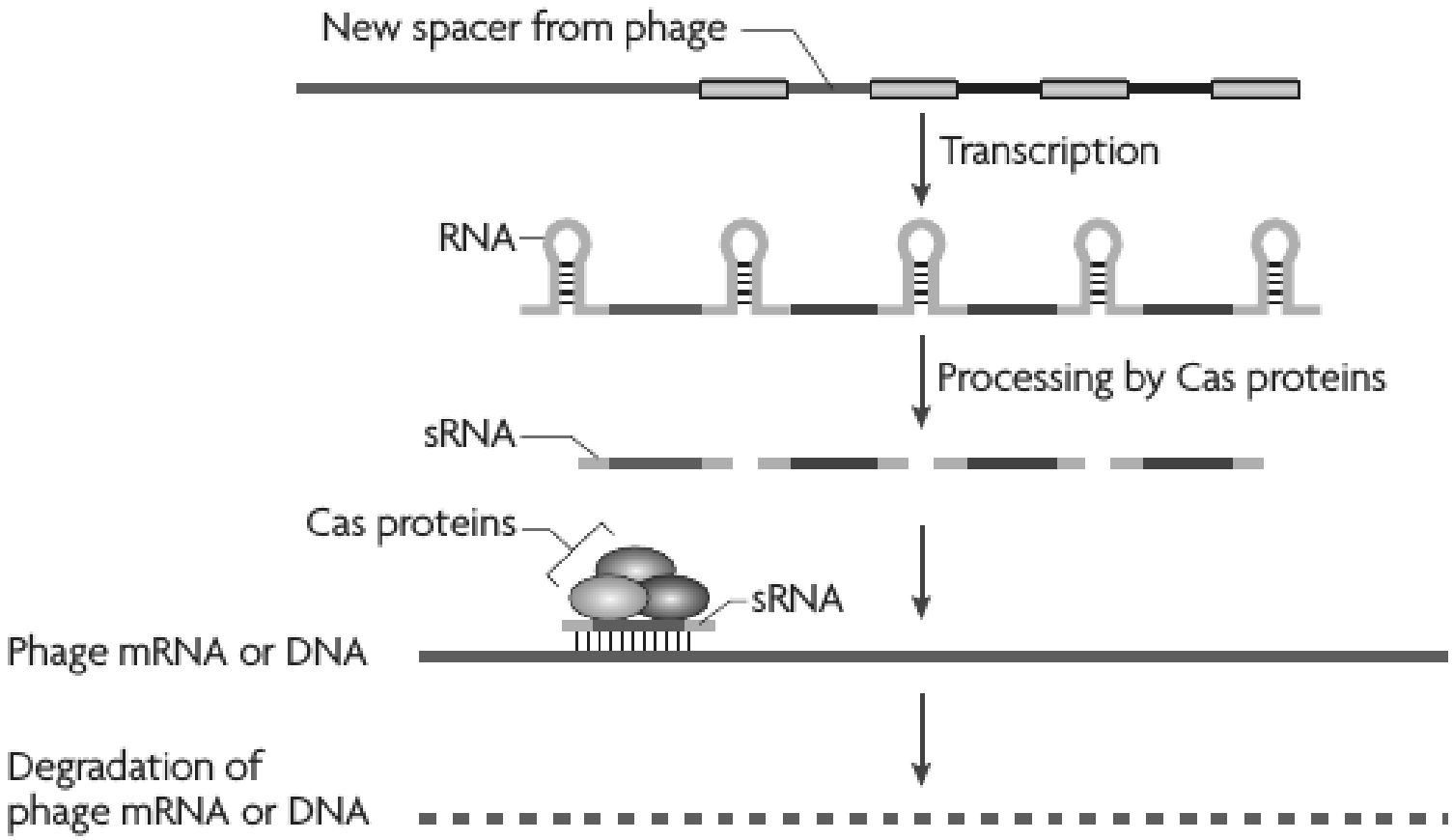
- Acquiring DNA sequences from invading phages (spacers)



Sorek et al (2008). Nature Reviews. vol. 6: 181-186

Mechanism of action of CRISPR-Cas

- cas3 (helicase)
- cas4 (nuclease)



Sorek et al (2008). Nature Reviews. vol. 6: 181-186