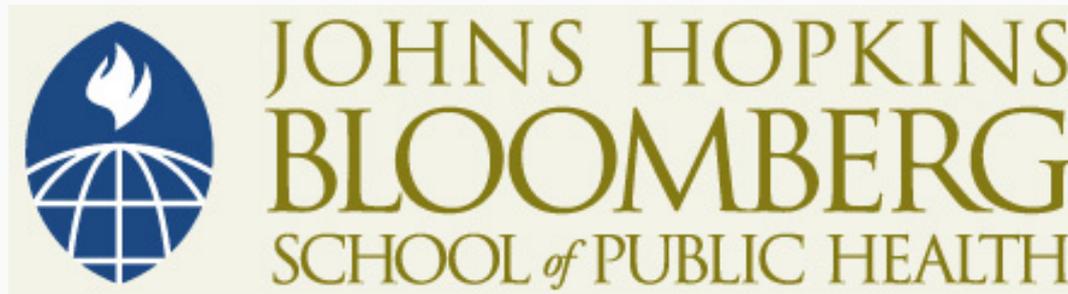


This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2007, The Johns Hopkins University and David Bishai. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL *of* PUBLIC HEALTH

# *The Molecular Epidemiology of Infectious Disease*

---

William R. Bishai, MD, PhD

The Johns Hopkins University

- Professor, Johns Hopkins University School of Medicine
- Co-Director, Center for Tuberculosis Research
- Interests
  - Molecular pathogenesis of TB
  - Cell division in mycobacteria
  - Pathogenesis of caseation and cavitation



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL *of* PUBLIC HEALTH

## *Section A*

---

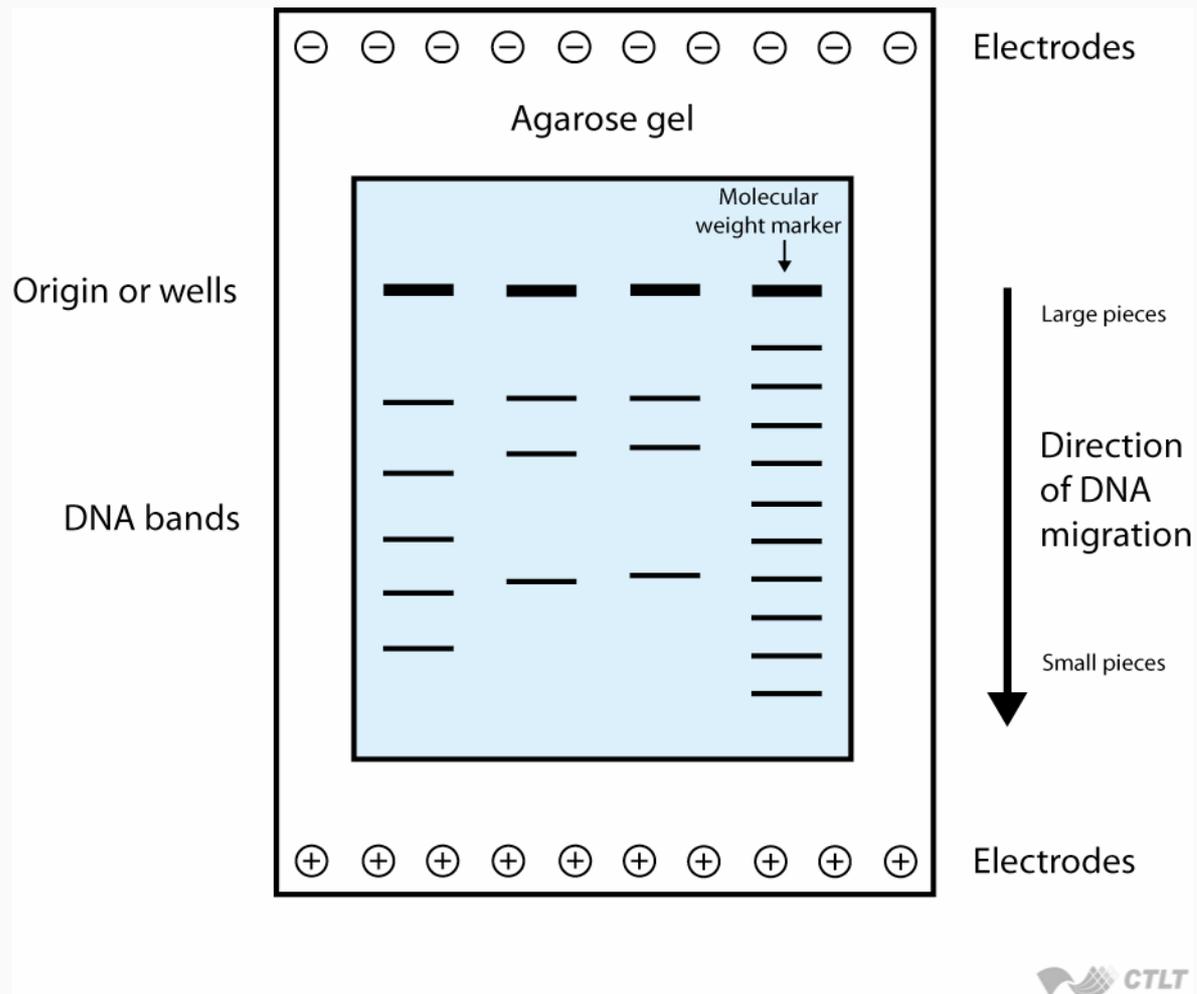
Laboratory Tools for Molecular Epidemiology of Bacterial Pathogens

# Conventional, Non-molecular Techniques

- **Biotyping**—specific biochemical reactions, colony morphology, environmental tolerances; rarely has enough discriminatory power for use as the sole epidemiologic marker
- **Antimicrobial susceptibility typing**—compares profiles of isolates to a panel of antibiotics; problem is that closely related species can acquire new resistances readily
- **Serotyping**—antibodies may be raised against a number of bacterial surface structures (e.g., LPS, capsule, outer membrane proteins, flagellae); historically, serotyping has been poorly reproducible, but it has improved as monoclonal antibodies are being used
- **Bacteriophage typing**—lytic bacteriophages (viruses of bacteria) infect certain strains, but not others; it was the standard means of typing *Salmonella* and *Staph. aureus* strains until recently; drawbacks: technically tricky, need to maintain a valid stock of phages, many strains non-typable
- **Bacteriocin typing**—bacterial toxins uniquely active against some members of a species; has been used for *Pseudomonas aeruginosa*; same drawbacks as bacteriophage methods bacteriocin

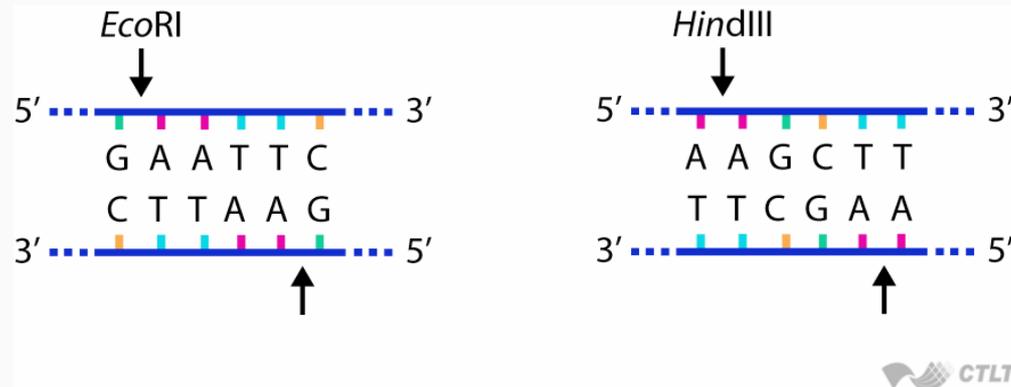
- Laboratory tools for molecular epidemiology of bacterial pathogens: a primer on DNA restriction and DNA separation techniques

# Conventional Gel Electrophoresis



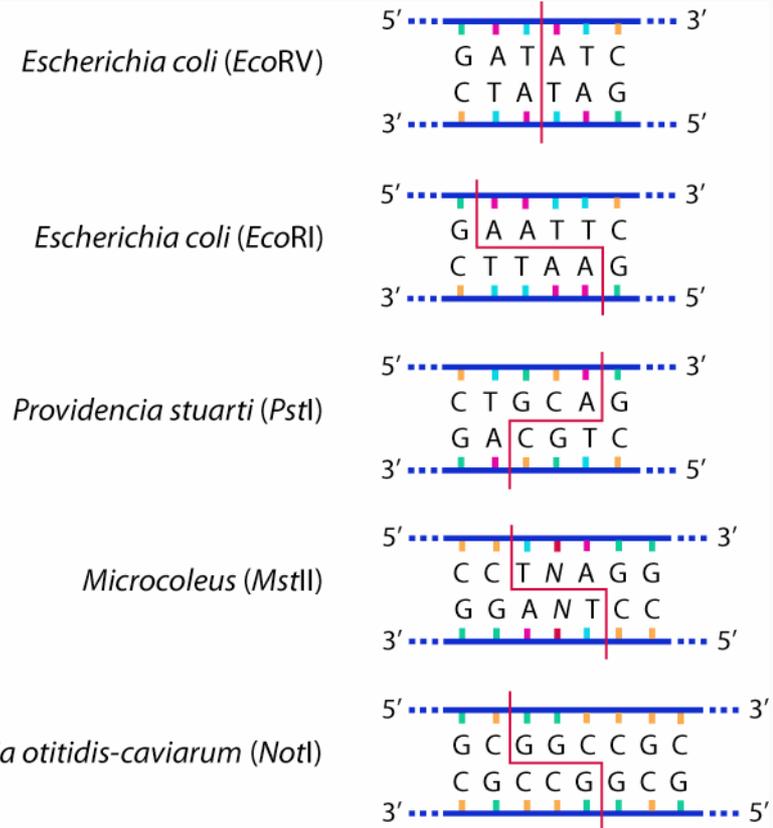
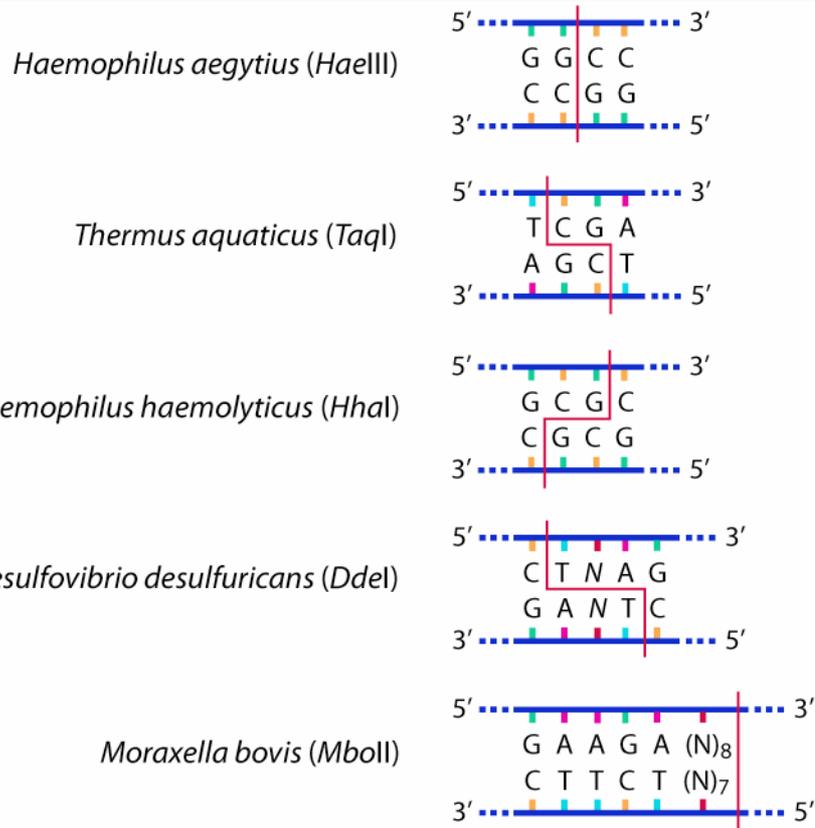
# Nucleic Acid (DNA-, RNA-) Based Methods

- Restriction analysis: restriction enzymes (REs) cut DNA at sequence-specific sites, e.g.,—

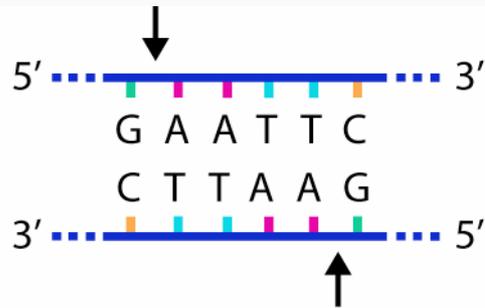


- Due to genetic drift or mutations, different isolates will have variations in the number and spacing of sites for an RE. This is called **restriction-fragment length polymorphism (RFLP)**.
  - ▶ REs with 4 base recognition sites cut every  $4^4$  (256) bp
  - ▶ REs with 6 base recognition sites cut every  $4^6$  (4,096) bp
  - ▶ REs with 8 base recognition sites cut every  $4^8$  (65,536) bp
  - ▶ Frequency of cuts also depends on the GC% of the bug, e.g., *M. tb.* (67% GC) has 0 *PacI* sites (TTAATTAA) and only 10 *DraI* sites (TTTAAA)

# Some Restriction Enzymes and Cleavage Sequences



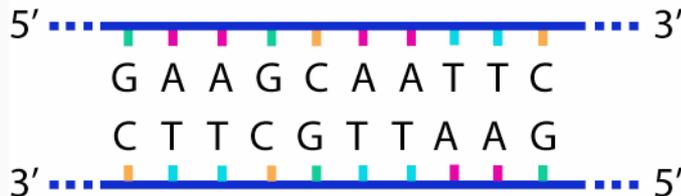
# Restriction Analysis and RFLP



Recognition Sequence for *Eco*RI.  
Arrows indicate site of enzyme cleavage.



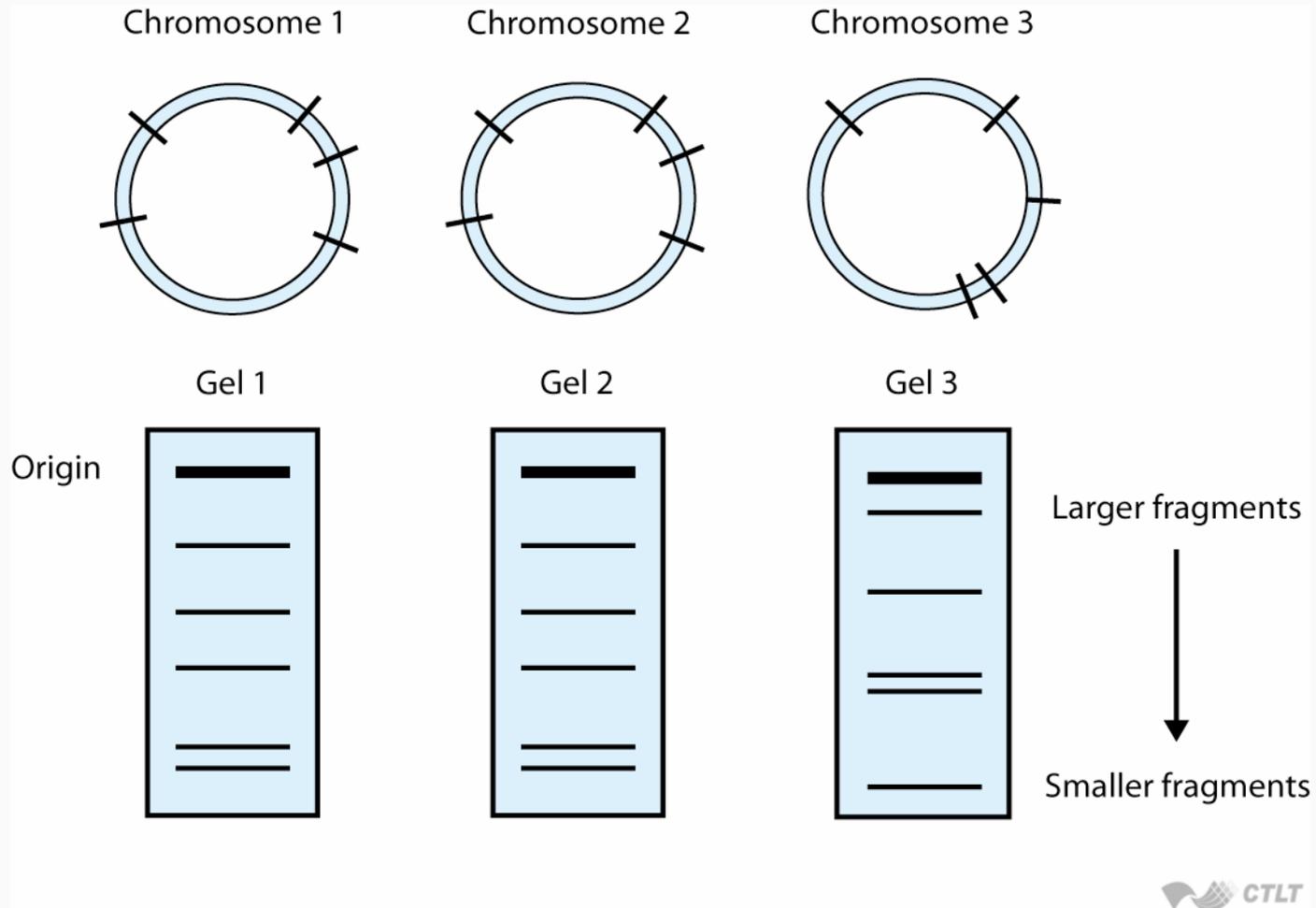
A point mutation occurs. AT base pair changed to CG. Restriction site specificity is lost.



A DNA insertion occurs. Addition of four base pairs. Restriction site specificity is lost.

Sequence recognized by restriction enzyme *Eco*RI. Point mutations or DNA insertions/deletions cause loss of restriction specificity as shown.

# Restriction Fragment Length Polymorphism (RFLP)





JOHNS HOPKINS  
BLOOMBERG  
SCHOOL *of* PUBLIC HEALTH

## *Section B*

---

Four Key Techniques

- Plasmid DNA restriction analysis
  - Same as above except small satellite pieces of DNA called plasmids (2,000–20,000 bp) are isolated and cut with REs
  - Here conventional electrophoresis may be used
  - Drawbacks: nonspecific, not all isolates will have plasmids; plasmids are shared readily in nature, hence isolates may be identical except one has acquired a plasmid

# *Nucleic Acid (DNA-, RNA-) Based Methods*

## ■ Hybridization techniques

— The strategy here is to cut chromosomal DNA with a frequent-cutting enzyme like a 6-base cutter (e.g., 4,000,000 bp genome would be cut into 977 fragments averaging 4,096 bp), but then to highlight only certain polymorphic bands within the many fragments by hybridization to a labeled probe by Southern blotting

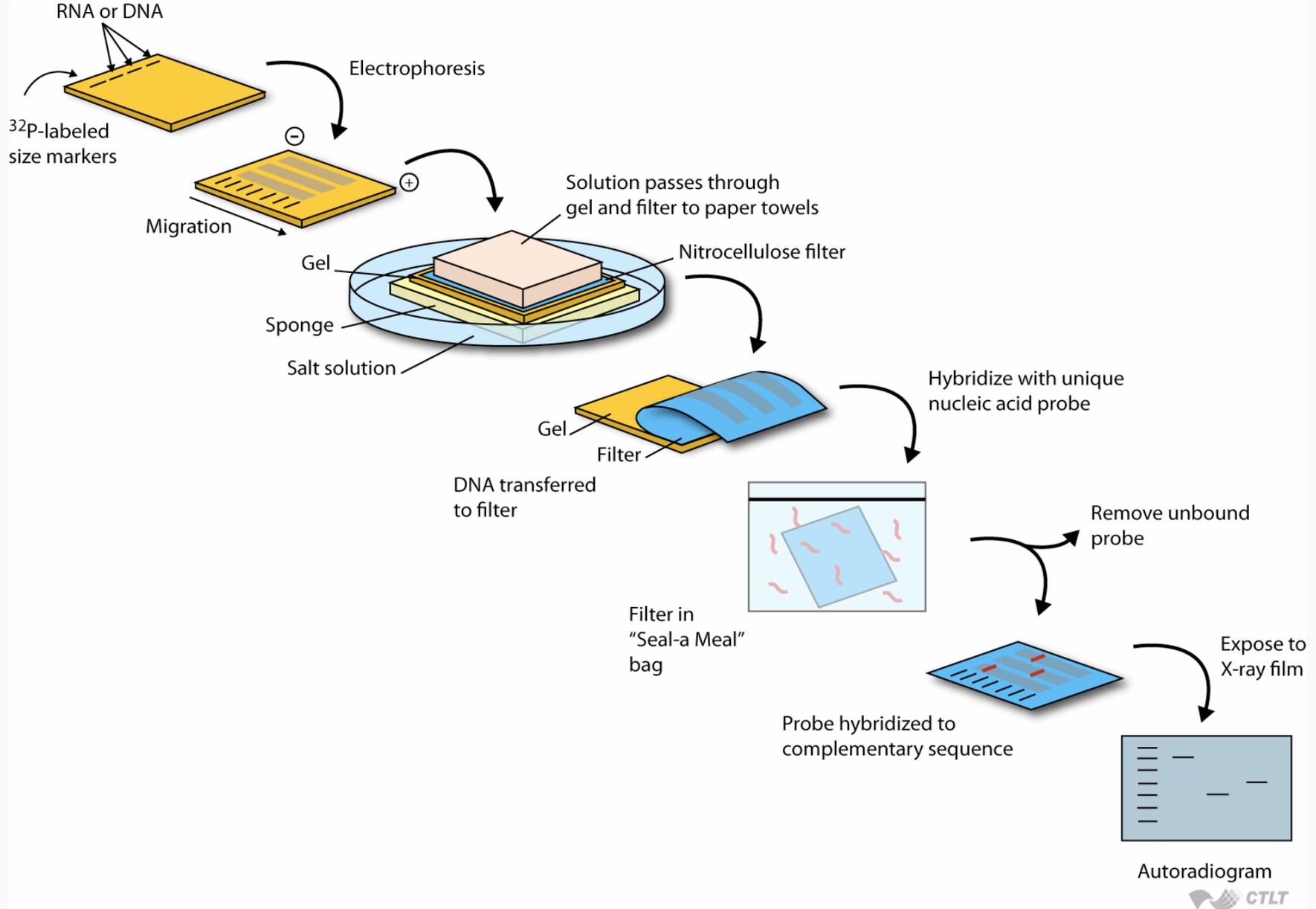
i) Randomly cloned probes

ii) Probes based upon specific virulence factors

KEY → iii) Probes derived from insertions sequences

KEY → iv) Probes directed against ribosomal DNA (rDNA):  
ribotyping

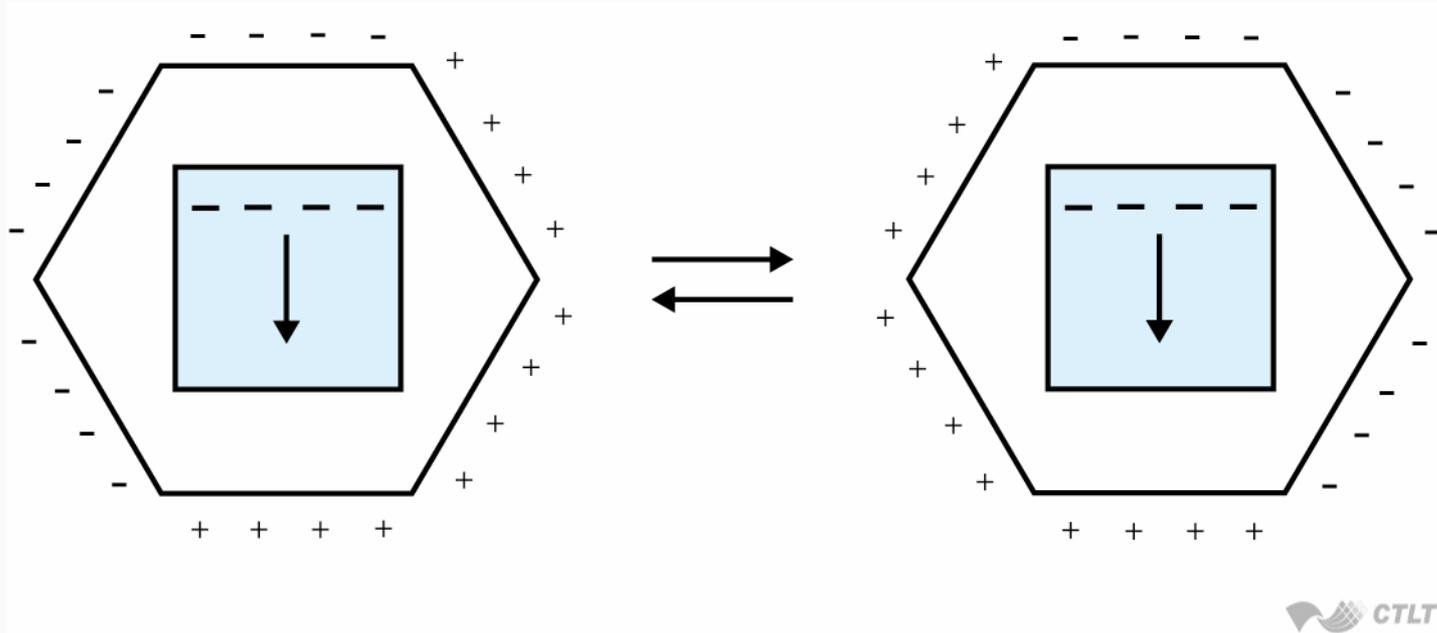
# Southern Blotting 2 of 2



# *Nucleic Acid (DNA-, RNA-) Based Methods*

- Chromosomal DNA restriction analysis; pulsed field gel electrophoresis (PFGE)
  - Detects the loss or gain of a restriction site due to genetic variation
  - An 8-base cutting RE used on DNA from a 4,000,000 bp chromosome will give about 60 fragments with an average size of 65,536 bp
  - 65,000 bp fragments are too large to be separated by standard DNA gels, but special electrophoresis equipment (PFGE) works
  - Hence using PFGE, chromosomal DNA fingerprinting is possible without the need for hybridization techniques to visualize only certain bands
  - Quite sensitive and specific, looks at entire genome
  - Drawbacks: technically demanding, expensive

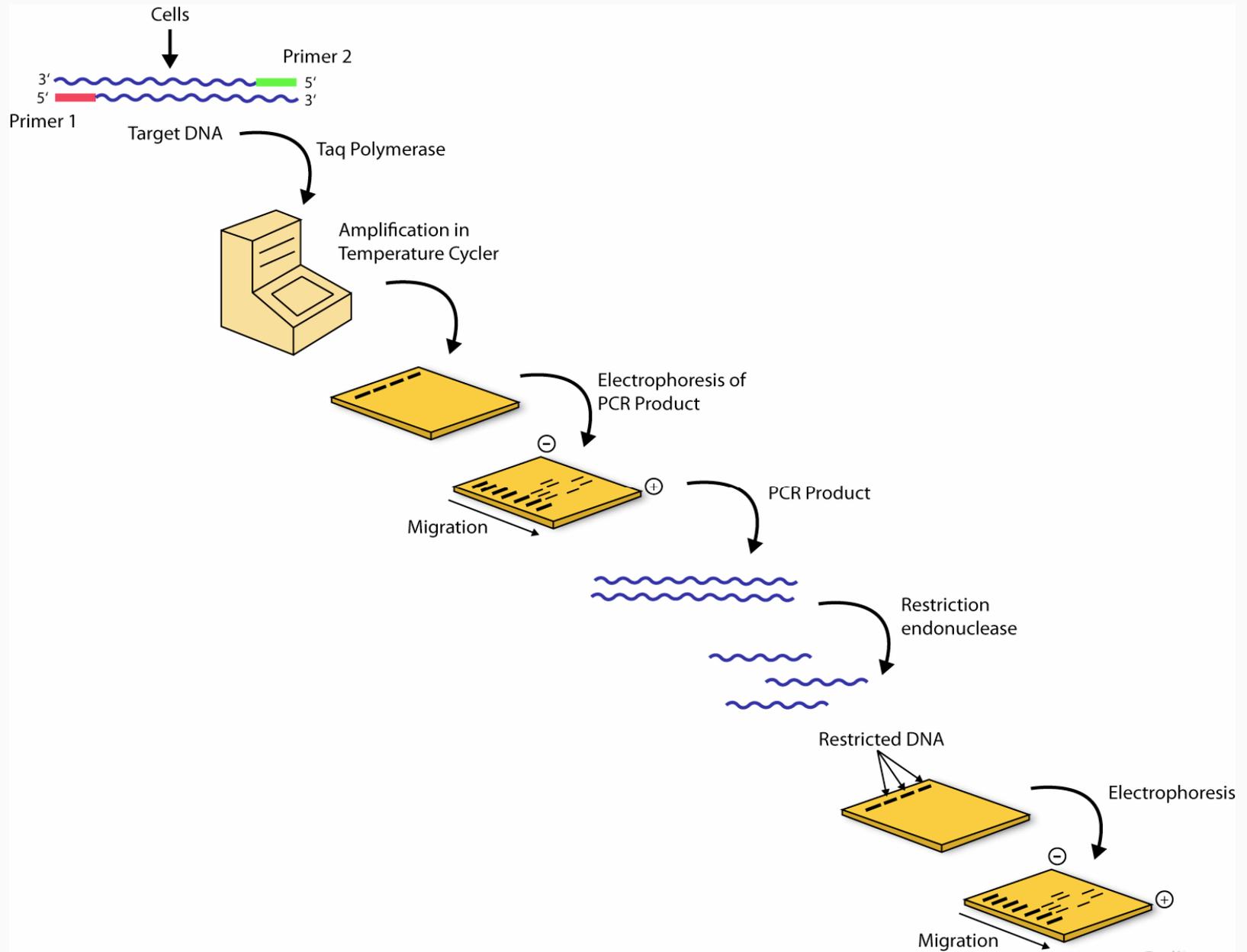
# PFGE Analysis



Schematic diagram of pulsed-field gel electrophoresis (PFGE) by the CHEF technique. Alternation of current is shown. The figure on the left indicates current from left to right. The figure on the right shows the current from right to left. Vertical arrows indicate net migration of DNA fragments in the gel.

- PCR-based subtyping methods
  - Polymerase chain reaction allows one to amplify a known sequence of DNA rapidly
  - Relatively easy, cost coming down
  - Drawbacks (1) cross-contamination is a big problem, (2) often poorly reproducible from machine to machine
    - A. PCR-RFLP
    - B. Arbitrarily primed PCR (AP-PCR) or random amplification of polymorphic DNA (RAPD)

# PCR Approaches



# *Cost: Material vs. Labor*

<b>Component</b>	<b>Cost/sample (\$) for:</b>			
	<b>Restriction of genomic DNA</b>	<b>Ribotyping</b>	<b>PFGE</b>	<b>PCR-RFLP</b>
Material	1.75	2.75	12.00	4.31
Labor	14.05	24.00	10.00	8.88
Total	15.80	26.75	22.00	13.19



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL *of* PUBLIC HEALTH

## *Section C*

---

Using the Laboratory Tools: Three examples from the literature

# *Antimicrobial Resistance Traced to Dairy Farms*

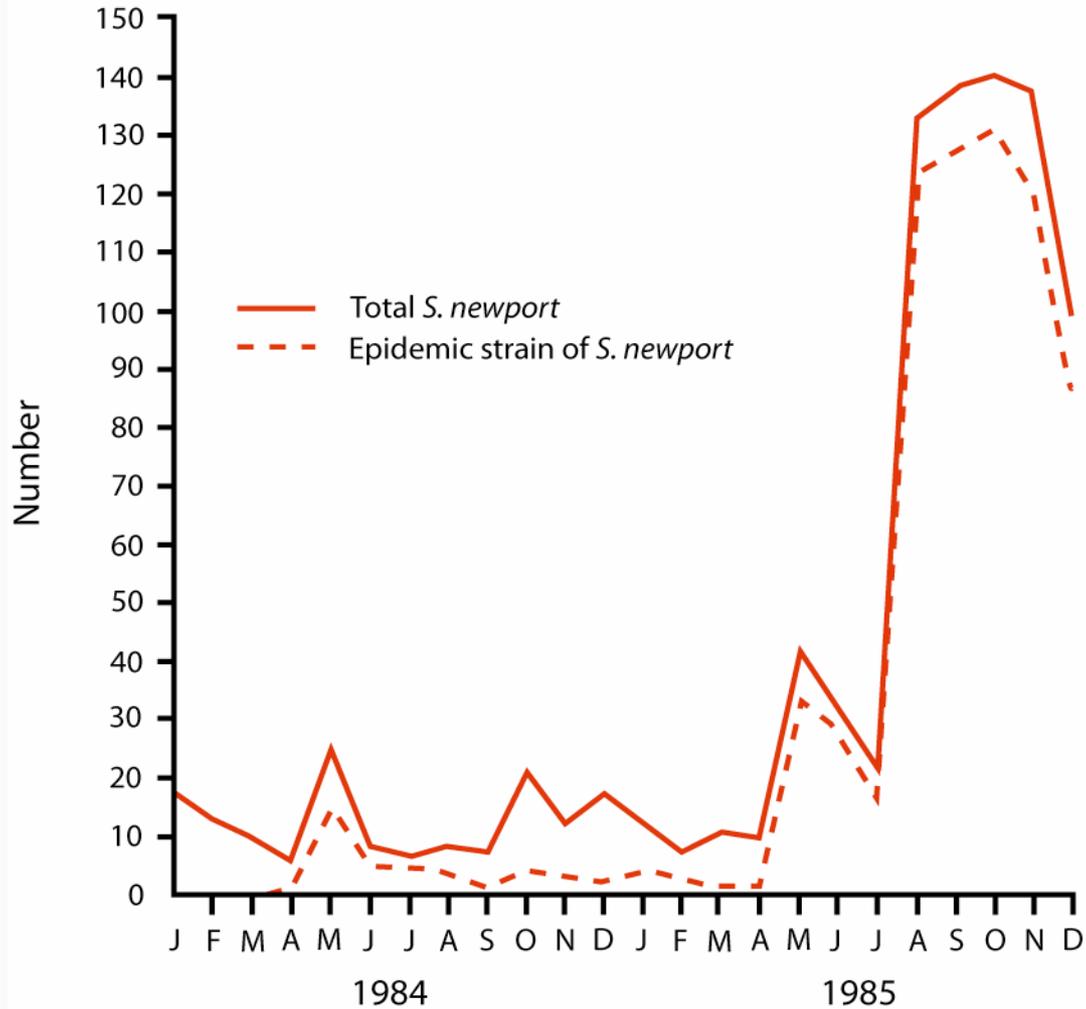
## **Chloramphenicol-resistant *Salmonella newport* traced through hamburger to dairy farms. A major persisting source of human salmonellosis in California**

*JS Spika, SH Waterman, GW Hoo, ME St Louis, RE Pacer, SM James, ML Bissett, LW Mayer, JY Chiu, B Hall, and et al.*

**Abstract:** Animal-to-human transmission of drug-resistant salmonella and the role of antimicrobial use in food animals in the emergence of these bacteria are controversial subjects. Investigation of a 4.9-fold increase in *Salmonella newport* isolations from Californians in 1985 showed that 87 percent of the isolates had an unusual antimicrobial-resistance pattern (including chloramphenicol resistance) and a single, identical plasmid. Interviews of 45 patients and 89 matched controls in Los Angeles County showed that illness was associated with penicillin or tetracycline use during the month before onset ( $P$  less than 0.001) and with eating ground beef during the week before onset ( $P = 0.052$ ). The epidemic strain was isolated from hamburger products eaten by cases, abattoirs where the animals from which the meat came were slaughtered, dairies that sent cows for slaughter on days when culture-positive products were processed, and ill dairy cows. Isolation of salmonella from beef carcasses in abattoirs correlated with the proportion of dead or moribund animals received for slaughter ( $r = 0.60$ ,  $P$  less than 0.05). Isolation of chloramphenicol-resistant salmonella from dairy farms was associated with the use of chloramphenicol at those dairies. We conclude that food animals are a major source of antimicrobial-resistant salmonella infections in humans and that these infections are associated with antimicrobial use on farms. *NEJM* 1987;316:565-570

# Salmonella newport: California, 1985

*S. newport* Isolates Serotyped at the California State Laboratory, by Month, in 1984 and 1985



# Antimicrobial Susceptibility Typing for Tracing

- Antimicrobial susceptibility typing to trace a California *Salmonella* outbreak
  - 4.9-fold rise in *Salmonella newport* in California in 1985
  - 85% of isolates were chloramphenicol-resistant and carried an identical plasmid
  - Epidemic strain isolated from:
    - ▶ Hamburger consumed by case patients
    - ▶ Abattoirs where cattle which produced the infected meat were slaughtered
    - ▶ Dairies which sent cattle for slaughter on days when tainted meat was produced
  - Chloramphenicol-resistant *Salmonella* statistically associated with dairies using chloramphenicol as a fattener as a fattener ( $P = 0.023$ )

	Fraction using chloramphenicol
Dairies with resistant <i>Salmonella</i>	3/8 (38%)
Dairies without resistant <i>Salmonella</i>	4/67 (6%)

## Example 2

### The Epidemiology of Tuberculosis in San Francisco -- A Population-Based Study Using Conventional and Molecular Methods

*Peter M. Small, Philip C. Hopewell, Samir P. Singh, Antonio Paz, Julie Parsonnet, Delaney C. Ruston, Gisela F. Schechter, Charles L. Daley, and Gary K. Schoolnik*

**Background:** The epidemiology of tuberculosis in urban populations is changing. Combining conventional epidemiologic techniques with DNA fingerprinting of *Mycobacterium tuberculosis* can improve the understanding of how tuberculosis is transmitted. **Methods:** We used restriction-fragment-length polymorphism (RFLP) analysis to study *M. tuberculosis* isolates from all patients reported to the tuberculosis registry in San Francisco during 1991 and 1992. These results were interpreted along with clinical, demographic, and epidemiologic data. Patients infected with the same strains were identified according to their RFLP patterns, and patients with identical patterns were grouped in clusters. Risk factors for being in a cluster were analyzed. **Results:** Of 473 patients studied, 191 appeared to have active tuberculosis as a result of recent infection. Tracing of patients' contacts with the use of conventional methods identified links among only 10 percent of these patients. DNA fingerprinting, however, identified 44 clusters, 20 of which consisted of only 2 persons and the largest of which consisted of 30 persons. In patients under 60 years of age, Hispanic ethnicity (odds ratio, 3.3;  $P = 0.02$ ), black race (odds ratio, 2.3;  $P = 0.02$ ), birth in the United States (odds ratio, 5.8;  $P < 0.001$ ), and a diagnosis of the acquired immunodeficiency syndrome (odds ratio, 1.8;  $P = 0.04$ ) were independently associated with being in a cluster. Further study of patients in clusters confirmed that poorly compliant patients with infectious tuberculosis have a substantial adverse effect on the control of this disease. **Conclusions:** Despite an efficient tuberculosis-control program, nearly a third of new cases of tuberculosis in San Francisco are the result of recent infection. Few of these instances of transmission are identified by conventional contact tracing. *NEJM* 1994;330:1703-1709

# Two Hypothetical Patterns of TB transmission

Experiment: Fingerprint all the active cases in Baltimore in 1995

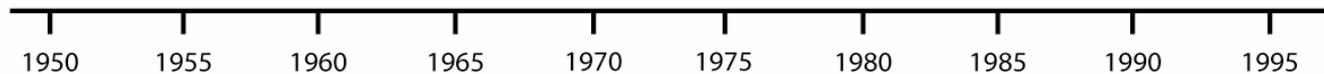
## Model 1: Active TB cases are due to reactivation only

1995 active cases have different fingerprints because they were acquired at different times and places



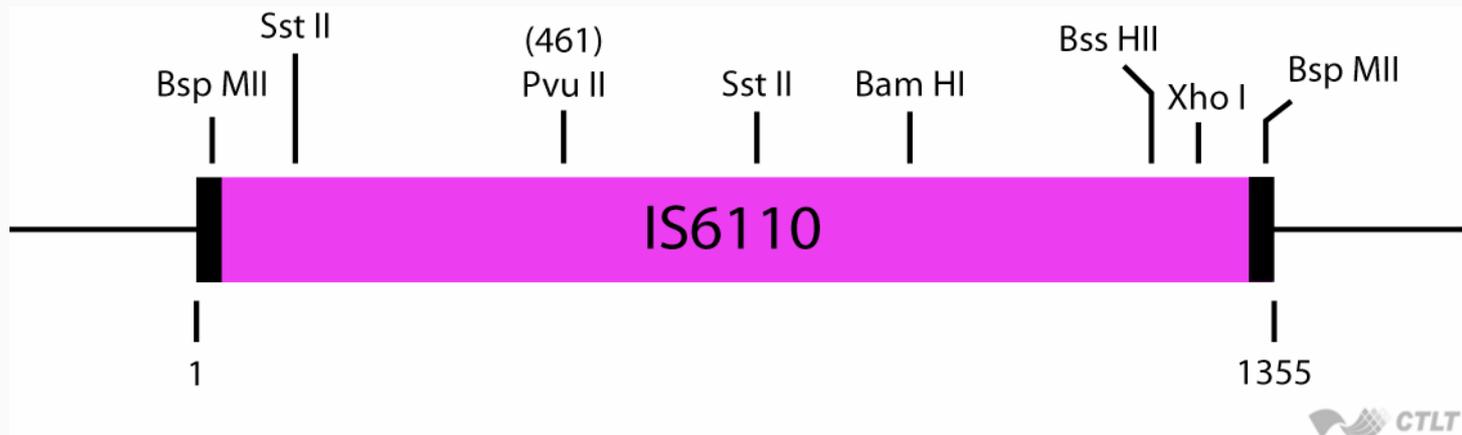
## Model 2: Active TB cases are due to primary active disease only

some 1995 active cases will have the same fingerprint because they were acquired at the same time and place



# *M. Tuberculosis* IS6110

- *M. tuberculosis* IS6110: an insertion sequence useful for Southern blot RFLP analysis



# *Implications for Public Health Officials*

- Old dogma
  - 90% of active TB due to reactivation
  - 10% due to active primary disease
- New data
  - San Francisco and New York: 30–40% of isolates show DNA fingerprint clustering
- Primary active TB may occur in 30–40% of urban cases rather than 10%
- Responses
  - Reactivation TB: large-scale tuberculin screening
  - Primary active TB: aggressive case-finding and contact investigation