Forensic Analysis using SNP and VNTR Signatures

Dr. Paul Keim

Northern Arizona University
And
Translational Genomics Research Institute (TGen)

Biothreat Agents

- *B. anthracis*
- *Y. pestis*

- *Francisella tularensis*
- *Burkholderia mallei, pseudomallei*
- *Brucella melitensis*, abortus, etc.

- *E. coli* O157:H7
- *Salmonella*

Molecular Microbial Forensics

- Molecular Assays
- Reference Databases
  - Population studies
- Real-life Validation Studies
- Models and Theory

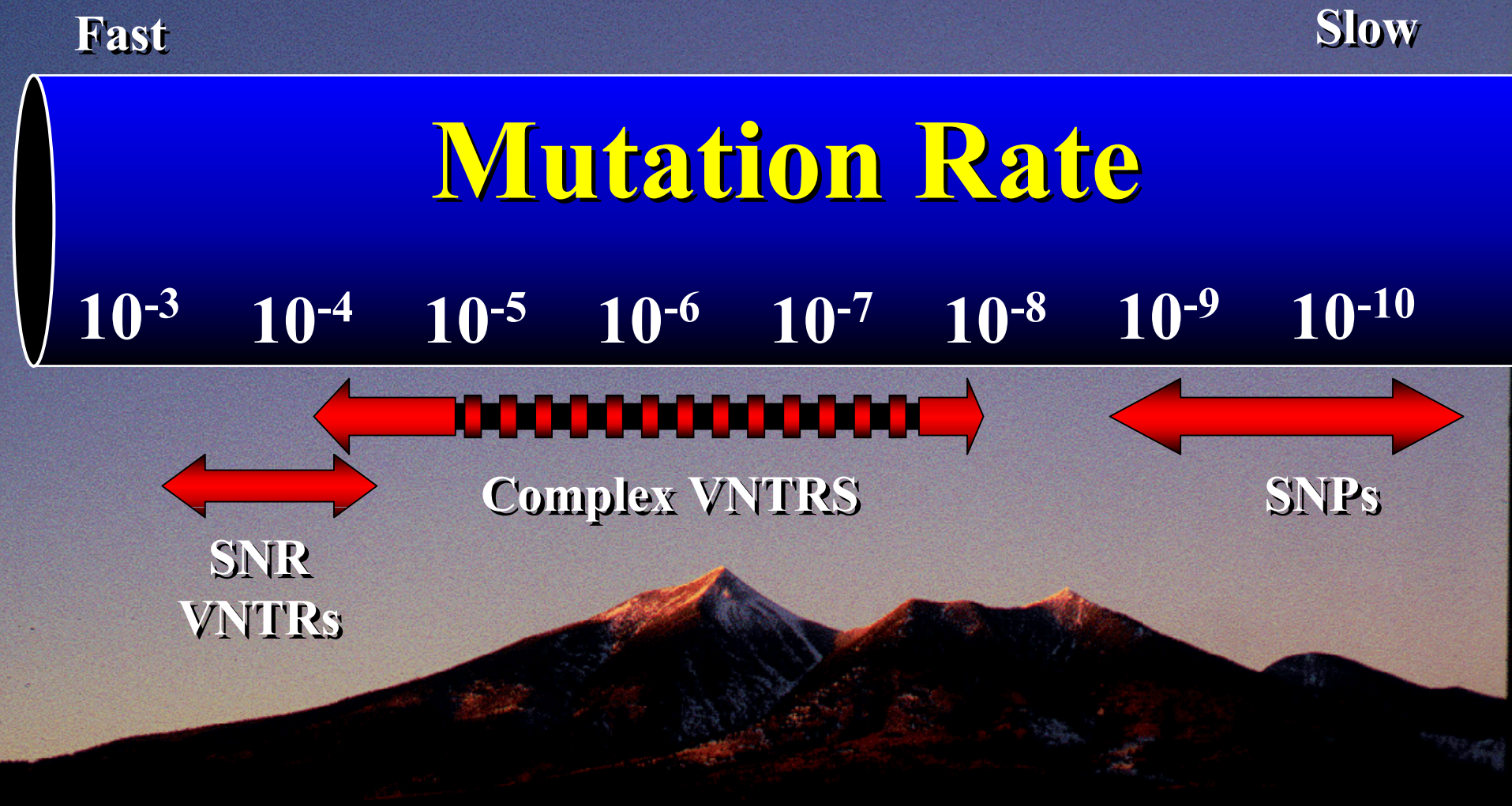
A genome is a complex and interesting place to work.

Not all loci are created equal for diagnostics.

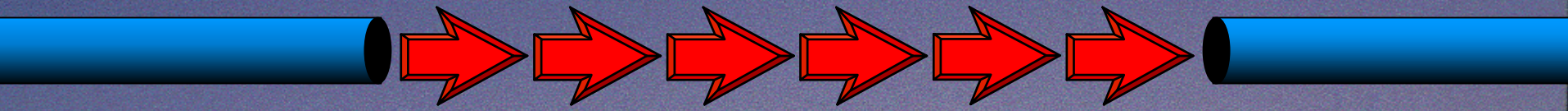
High Resolution "DNA Fingerprinting"

- Multiple-Locus VNTR Analysis (MLVA)
  - High Discrimination markers – rapidly mutating.
- Single Nucleotide Polymorphisms (SNPs)
  - Low Discrimination markers, but very stable!

Evolutionary Scale and Mutation Rate



Tandemly Repeated Sequences

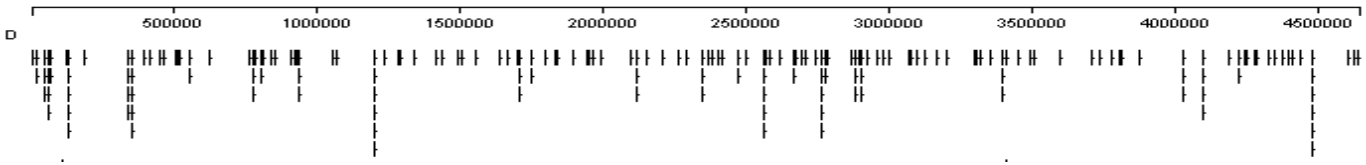


Genomic regions that vary greatly among worldwide isolates.

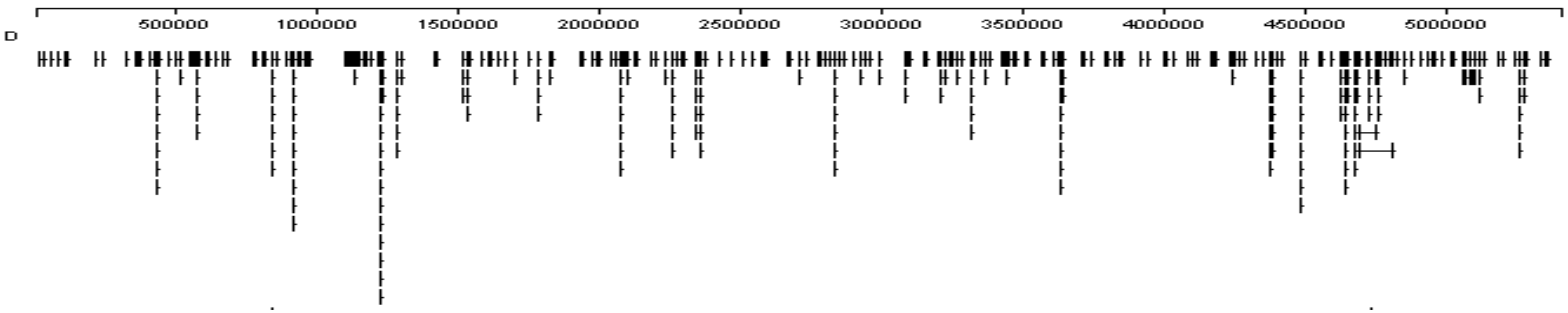
VNTR

Genomic Tandem Arrays

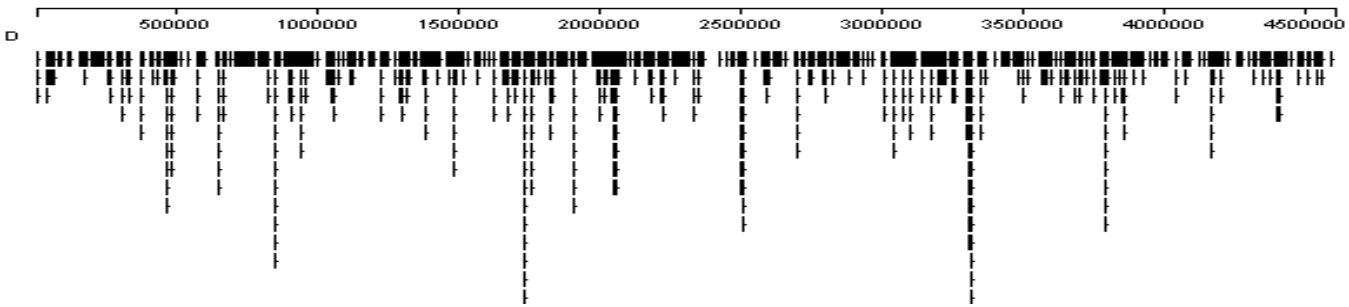
E. coli K12



B. anthracis Ames

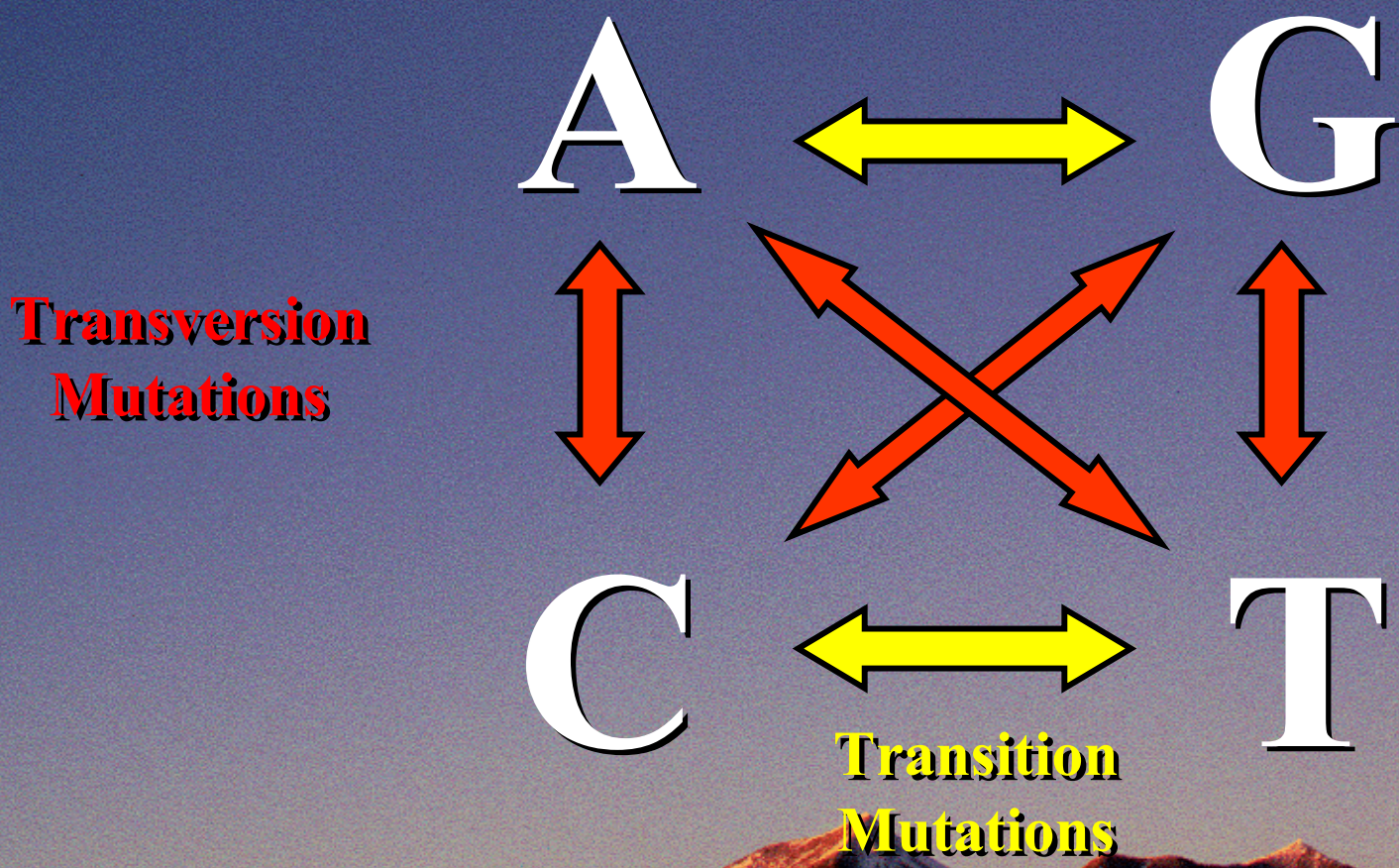


Y. pestis CO92



DNA Sequence Evolution

Kimura 2-parameter model



Two Approaches to SNP Discovery:

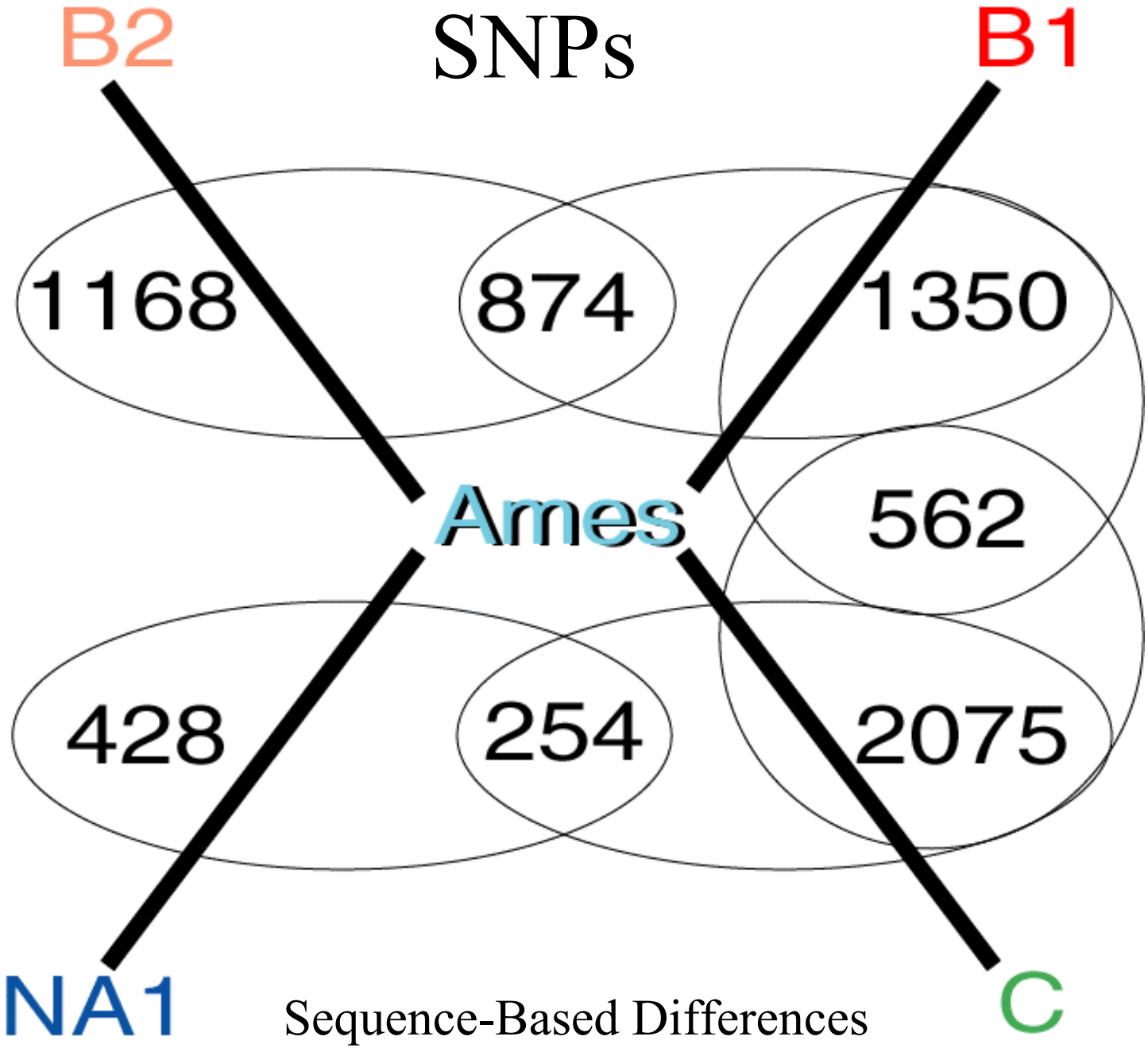
- **Limited Discovery:**
  - Multiple-Locus Sequence Typing

  Appropriate when diversity is high.

- **Exhaustive Discovery:**
  - Whole Genome Sequencing

  Appropriate when diversity is low.

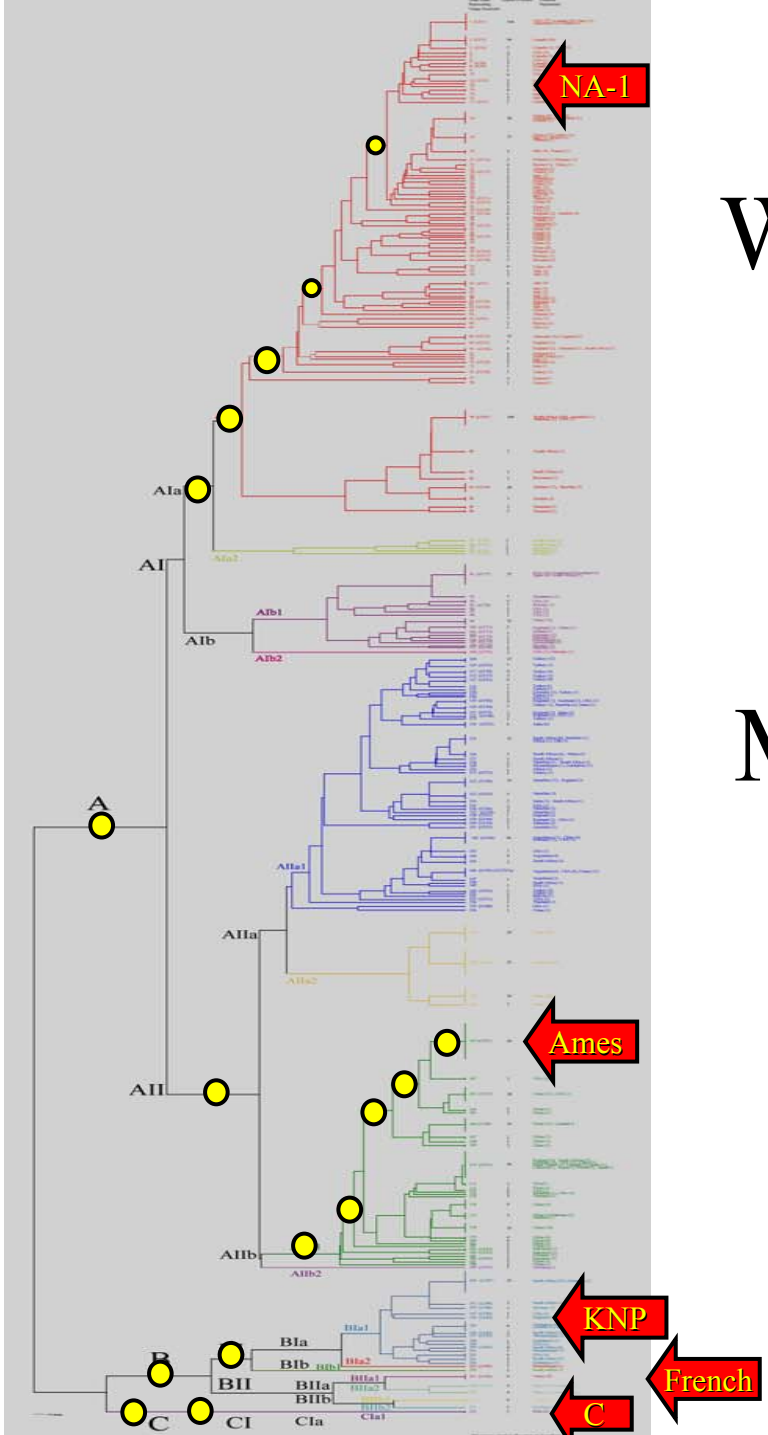
SNPs



Sequence-Based Differences

Canonical Characters

- A phylogenetic character (such as a SNP) that defines an evolutionary point.
- In practice, SNPs are chosen to represent a whole groups of SNPs that mark a common phylogenetic branch.

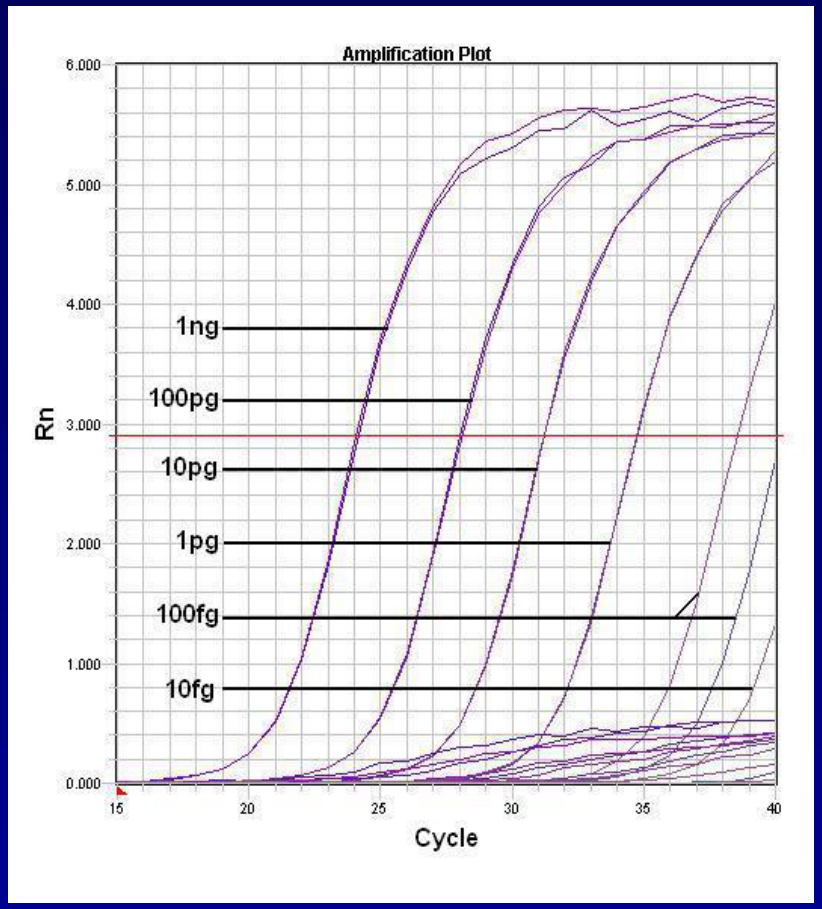


World Wide *B. anthracis* Canonical SNPs

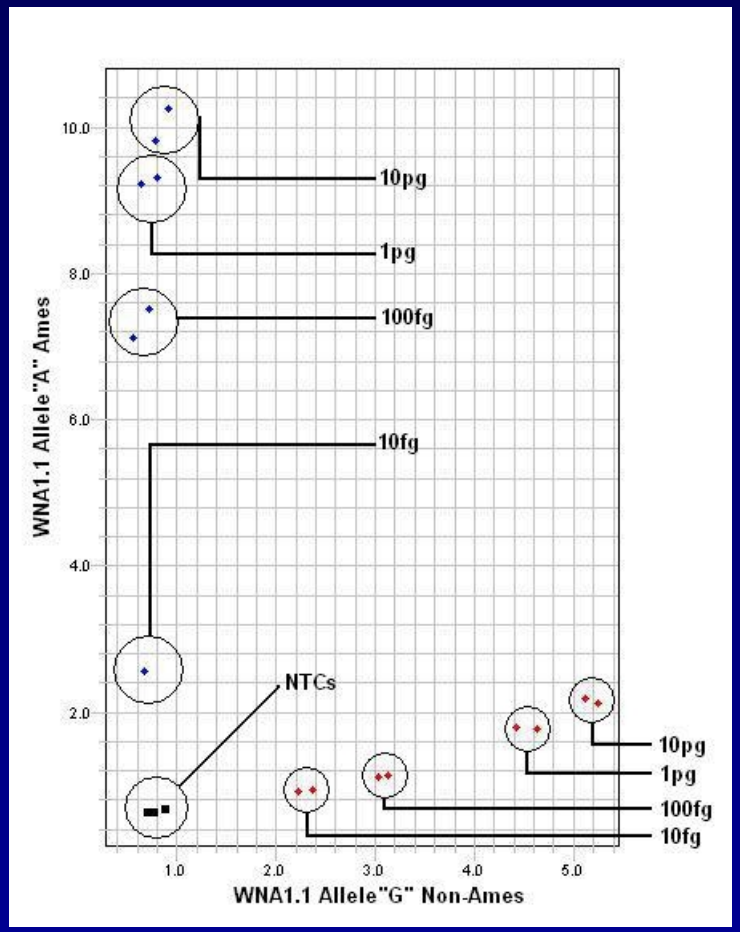
Marking major branches and specific strains

Ames-Specific RT-PCR

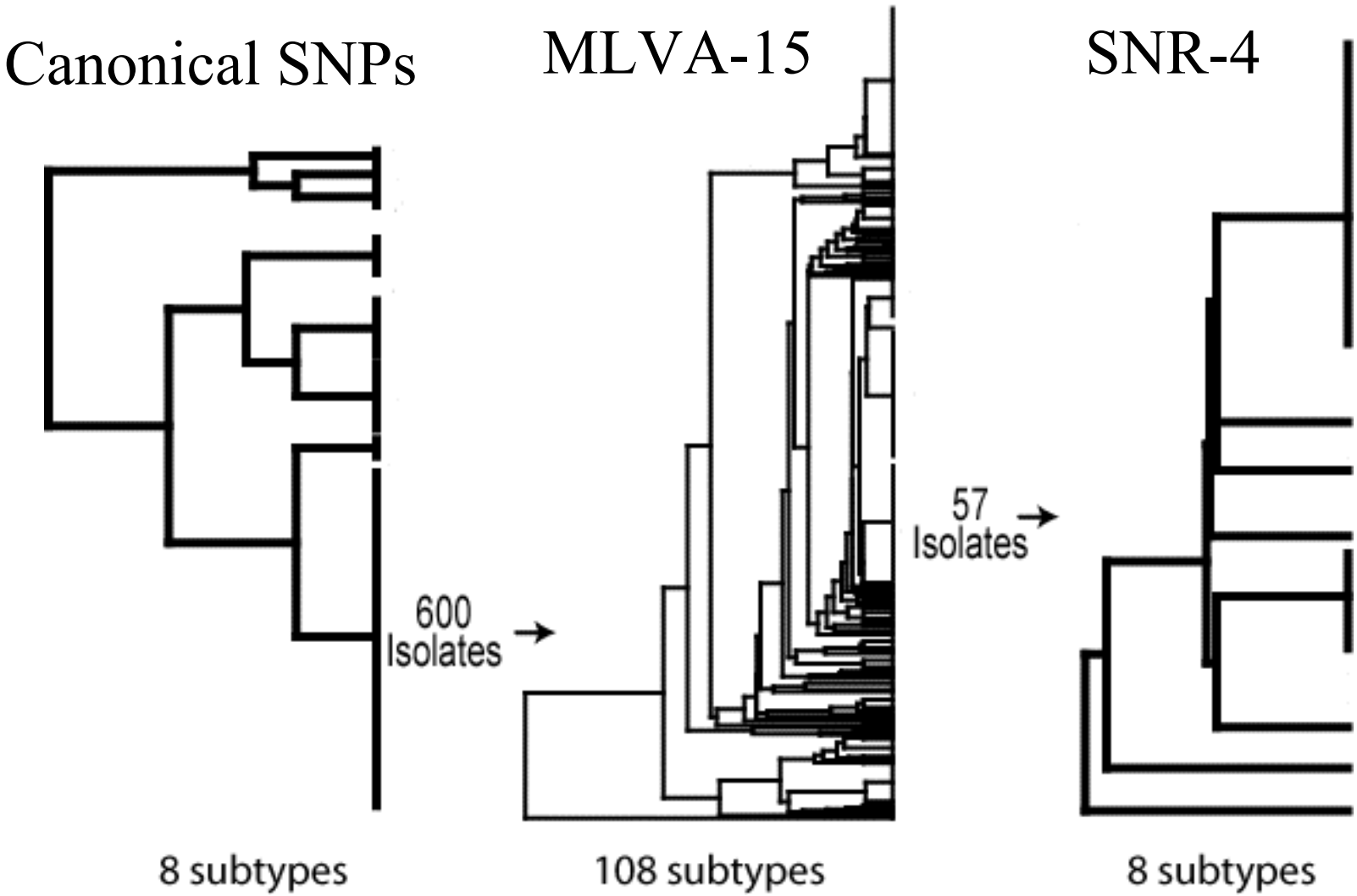
a)



b)



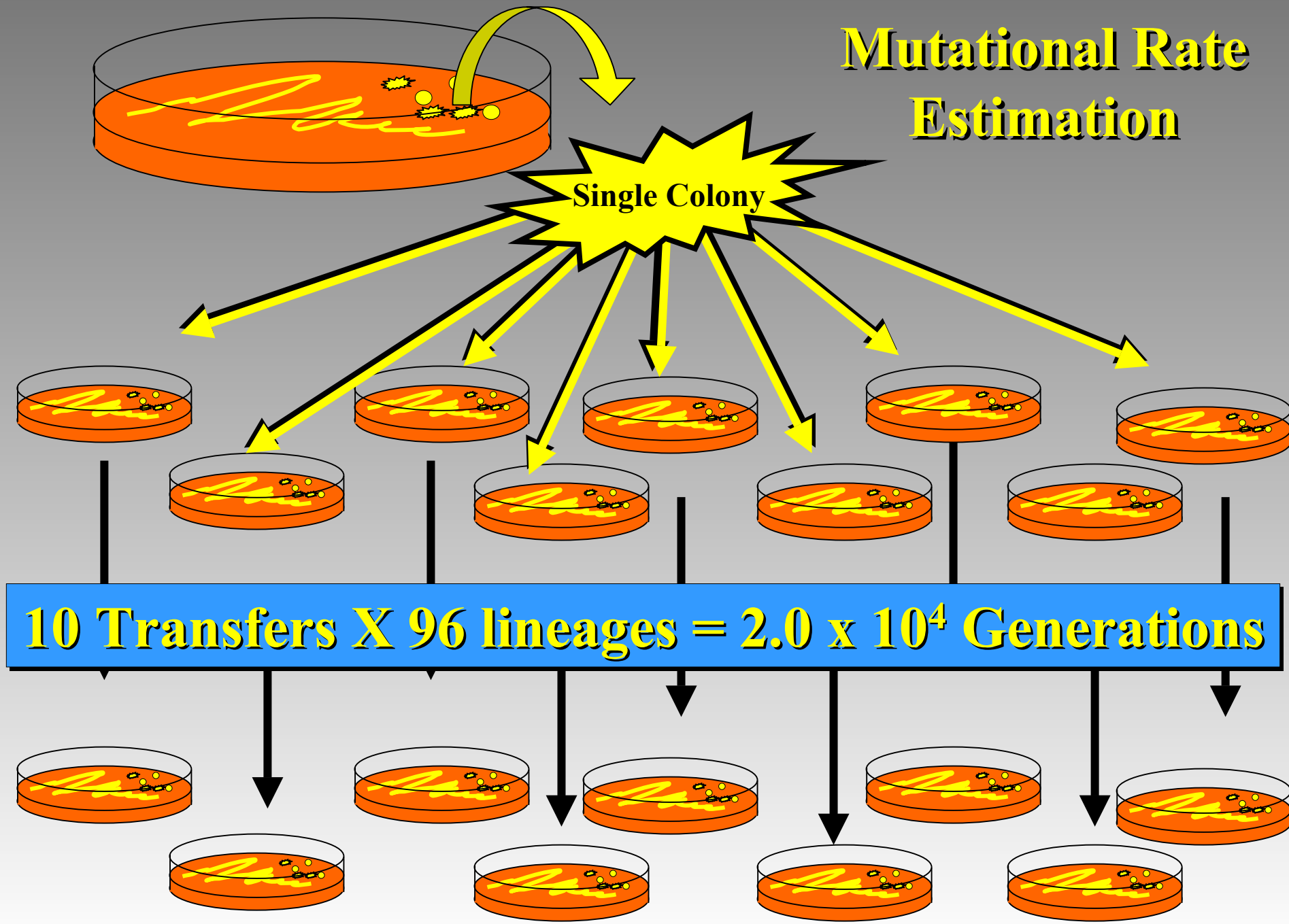
B. anthracis PHRANA (1067 isolates)



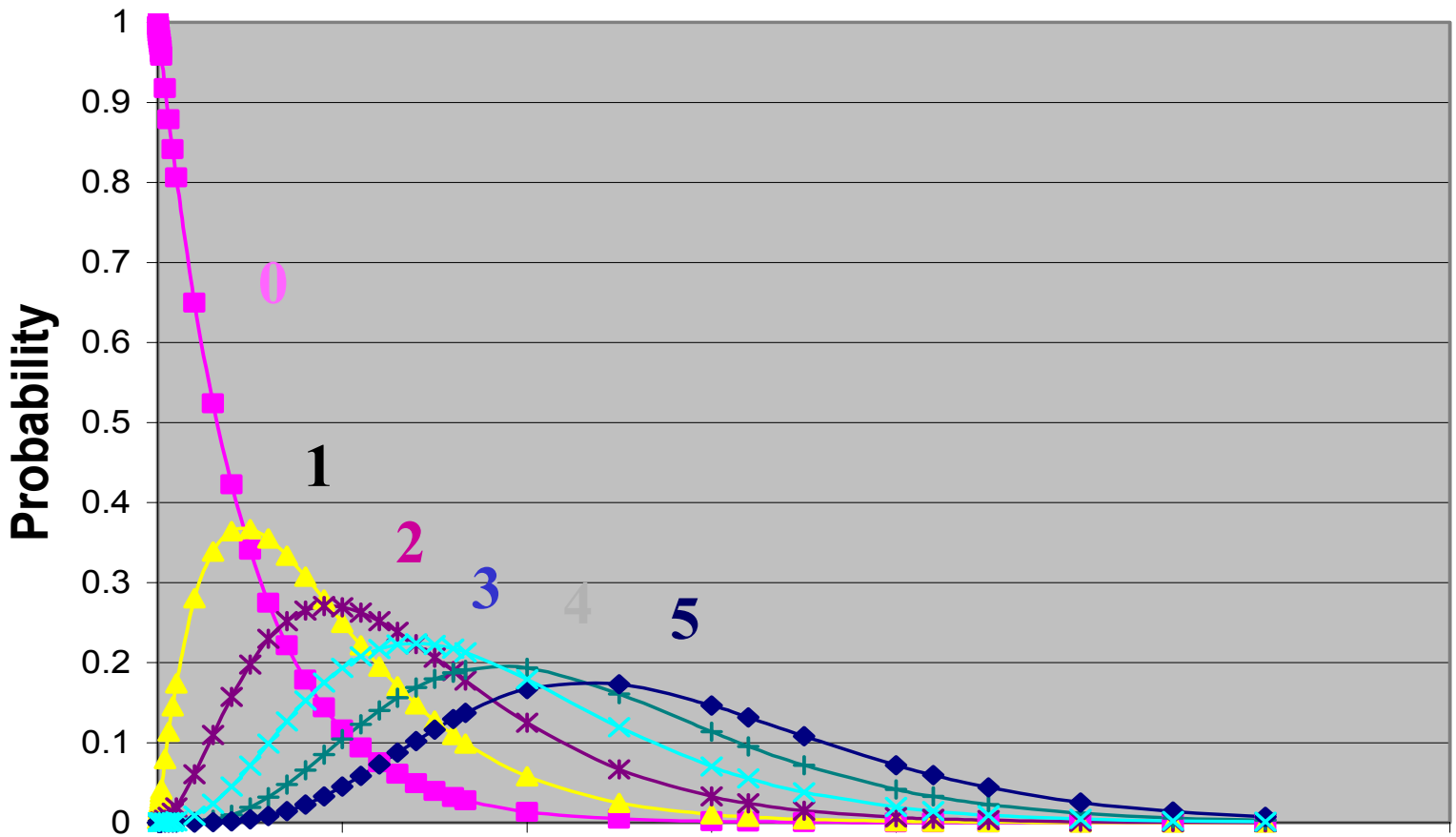
(208 GTs)

(476 GTs)

Mutational Rate Estimation

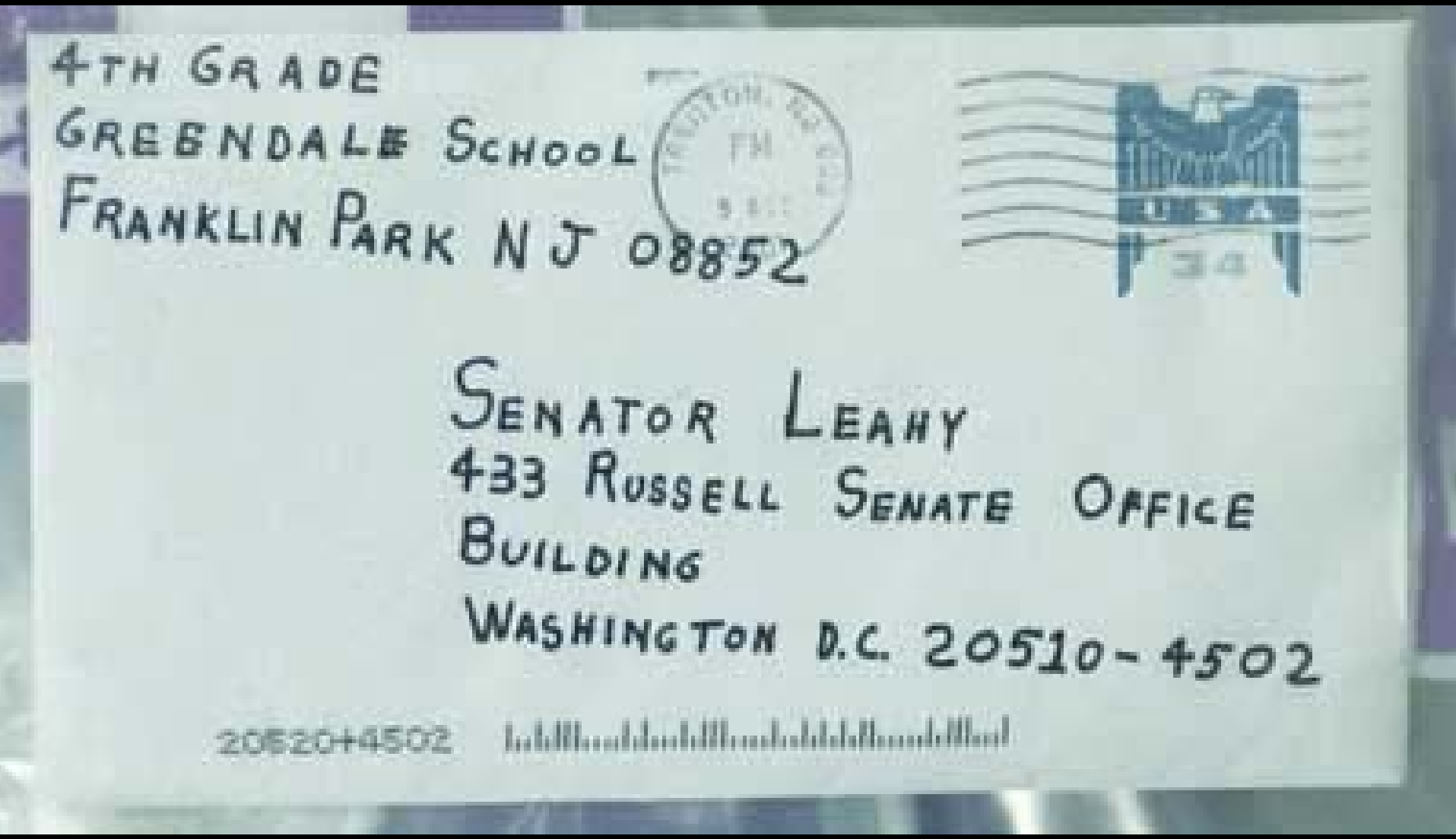


Mutation Rate Models



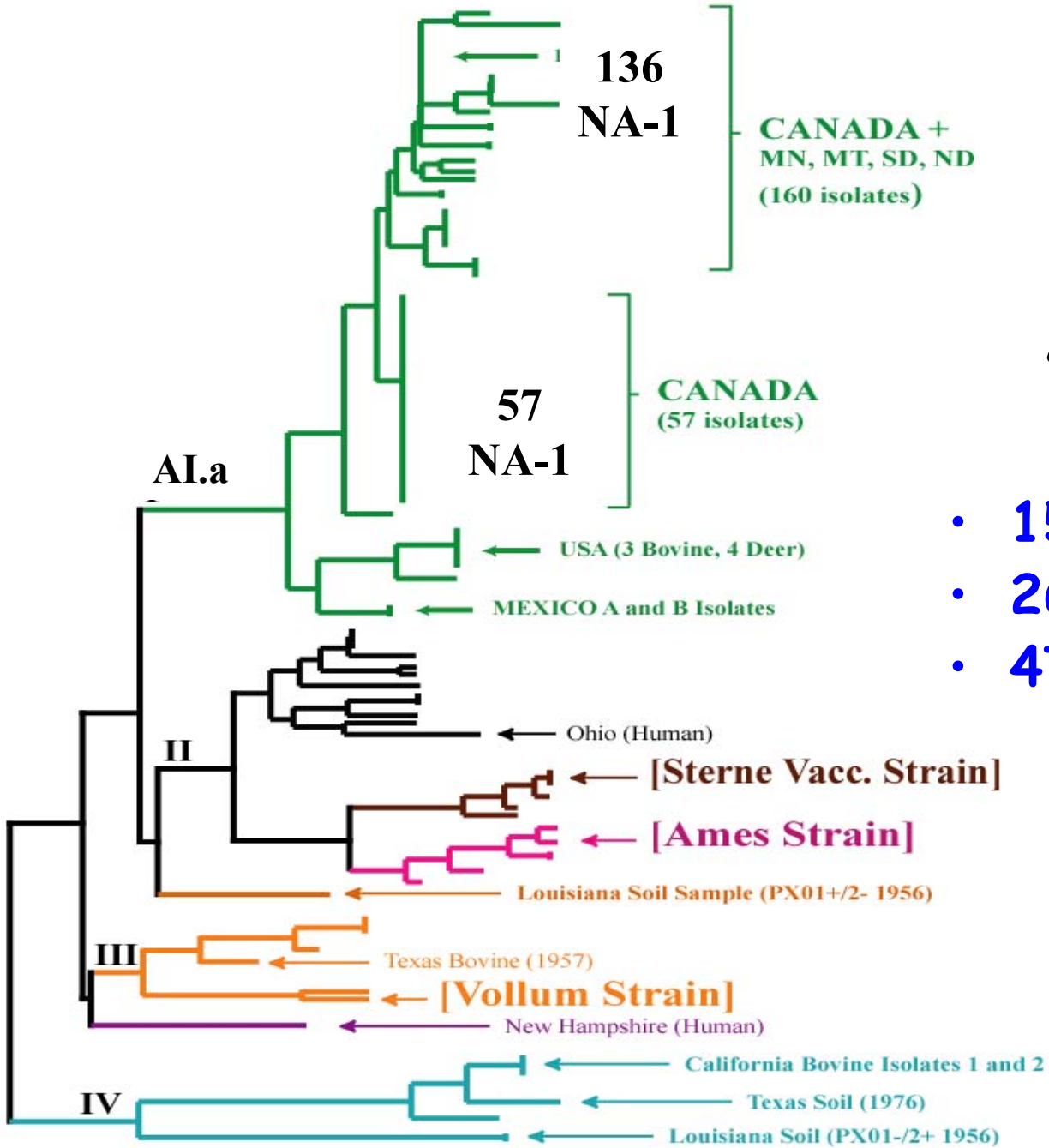
Generations Separating Two Isolates

Anthrax Letter Attacks



BACILLUS
BRRAT002
BT000676-01
4 OCT 2001
4363

North America Anthrax



- 15 VNTR Analysis
- 269 Isolates
- 47 genotypes

Matt Van Ert
unpublished

Molecular Forensics Analysis

Testimony in Court

- "A match – Inclusion"
  - Failure to exclude, could be the source.
- "Exclusion"
  - Very different and unlikely the same.
- "Kind of Close?"
  - Maybe yes? Maybe no?

Collaborators

- TIGR
  - Jacques Ravell, Tim Read
- CDC Atlanta
  - Alex Hoffmaster, Tanja Popovic
- CDC Ft. Collins
  - May Chu, Ken Gage, Jennifer Lowell
- Louisiana State University
  - Martin Hugh-Jones. Pamala Coker
- Institut Pasteur
  - Agnes Fouet, Michele Mock, Elisabeth Carniel
- Umeå Sweden Group
  - Anders Johansson, Mats Forsman, Anders Sjöstedt
- Livermore National Lab
  - Gary Andersen, Emilio Garcia, Patrick Chain, Peter Agron
- Los Alamos
  - Paul Jackson, Rich Okinaka, Karen Hill
- Keim Lab Characters
  - Talima Pearson, Joseph Busch,
  - Tatum Simonson, Shane Rhoton
  - Jana U'Ren, Ryan Easterday,
  - Matt Van Ert, Meghan Dukerich,
  - Lynn Huynh, Shaylan Zanecki,
  - Zack Jay, Michelle Maldonado,
  - Amy Vogler, Jessica Girard,
  - Kimothy Smith, Leo Kenefic, Ryan Easterday, Chris Allender
  - Jason Farlow, Dave Wagner,
  - Christine Keys, Lance Price, Dan Solomon, Sasha Klevytska, Jim Schupp

Research Funding

Dept. of Homeland Security
Bio-Chem Countermeasures

Department of Justice
Federal Bureau of Investigation

National Institutes of Health
Pathogen Evolution -NIGMS