

HIV Phylogenetics

**American Society for Microbiology
Conference for University Educators
The Johns Hopkins University
June 3, 2011**



HIV Drug Resistance Program
NCI-Frederick



Genetic Diversity

No one knows the diversity in the world, not even to the nearest order of magnitude. ... We don't know for sure how many species there are, where they can be found or how fast they're disappearing. It's like having astronomy without knowing where the stars are. E. O. Wilson

Genetic Diversity

Ye that are of good understanding note the doctrine that is hidden under the veil of the strange verses.

Lasciate ogni speranza voi ch'entrate
Dante Canto III

Genetic Diversity-Methodology

Cladistics

- Arranging organisms into groups that include a common ancestor
- Focus on
 - Shared derived characteristics
 - Parsimony
 - ORIGINS
- Hypothesis generated

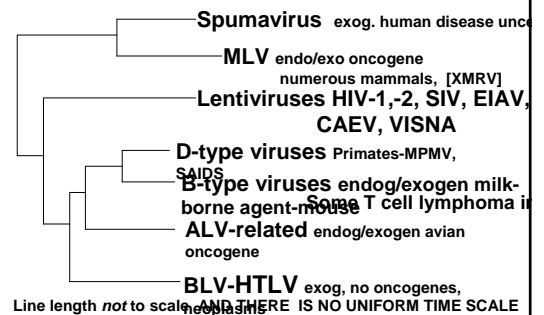
Genetic Diversity-Methodology

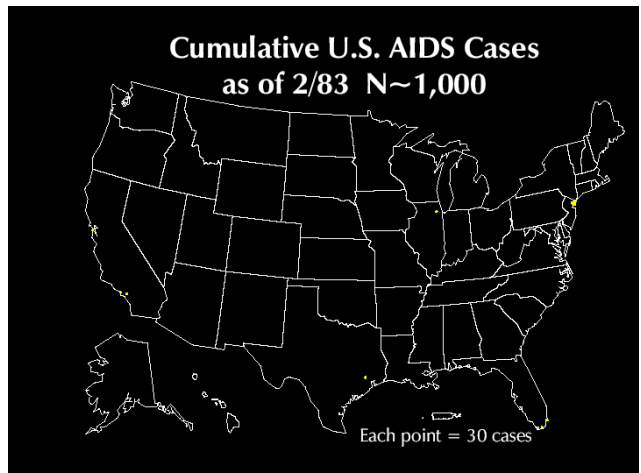
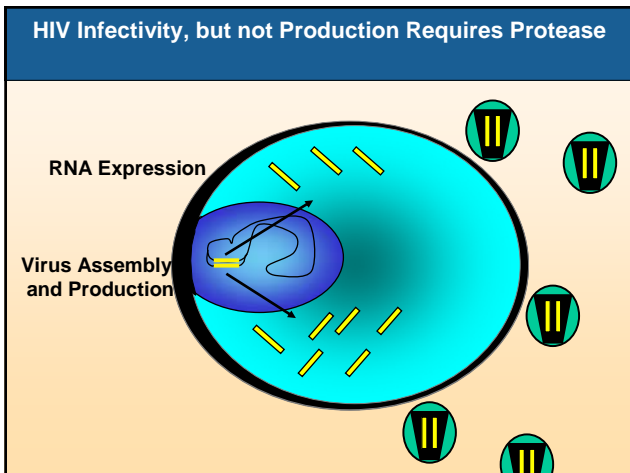
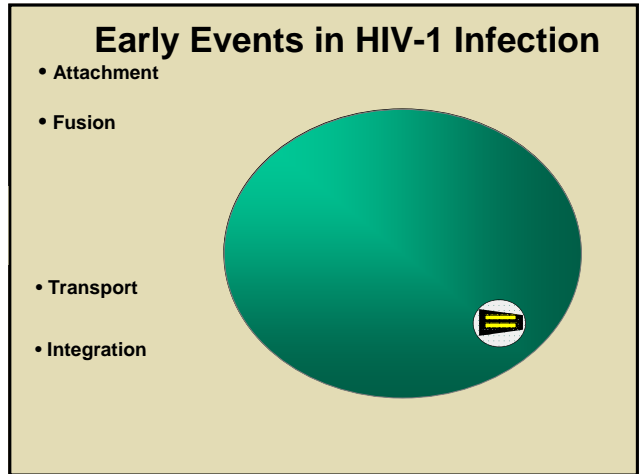
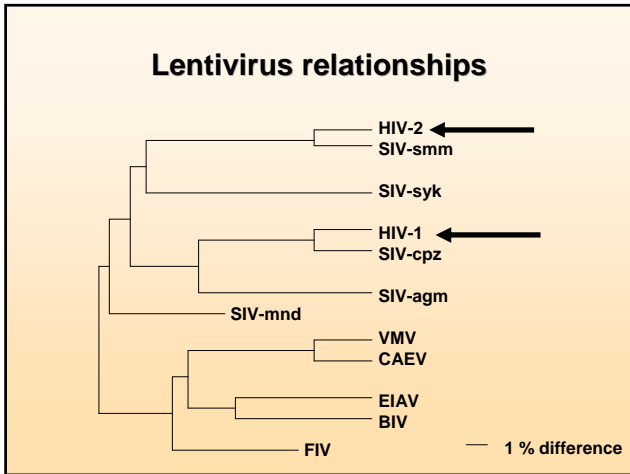
Cladistics

- Arranging organisms into groups
 - Focus on origins
 - Groups are related by descent
 - Divergence occurs by bifurcation
 - Shared derived characteristics
 - Parsimony
 - Change occurs in lineages over time
- Hypothesis generated



Retroviruses Classification by RT Sequence into Seven Families

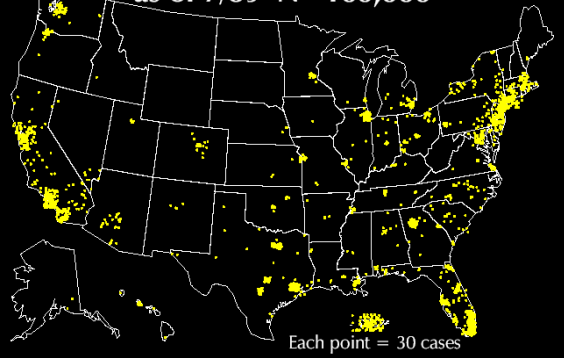




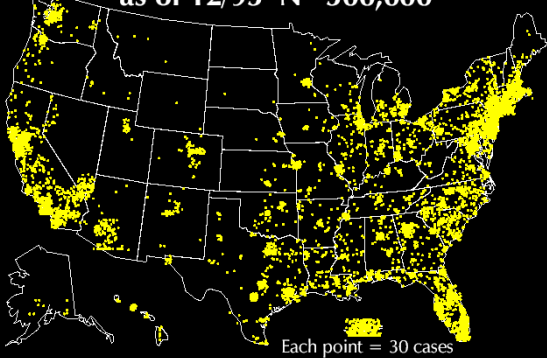
Cumulative U.S. AIDS Cases
as of 5/85 N~10,000



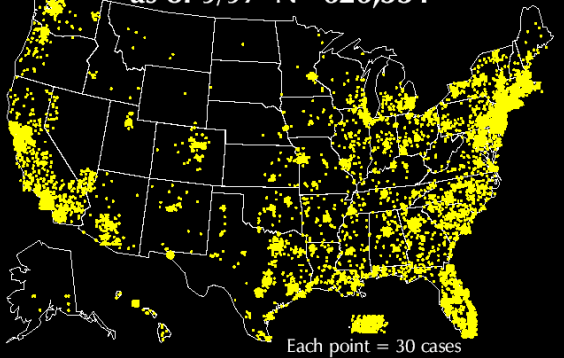
Cumulative U.S. AIDS Cases
as of 7/89 N~100,000

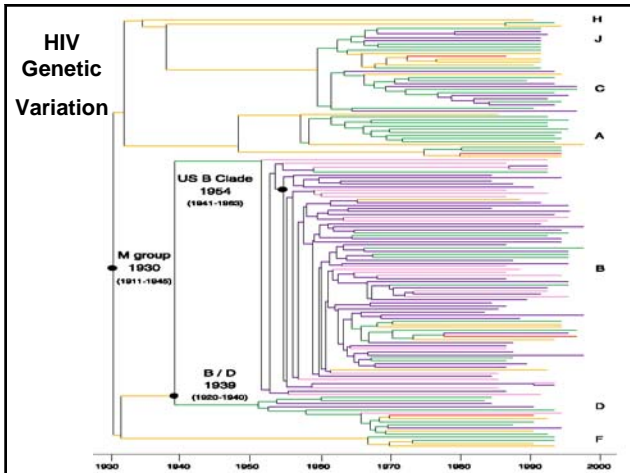
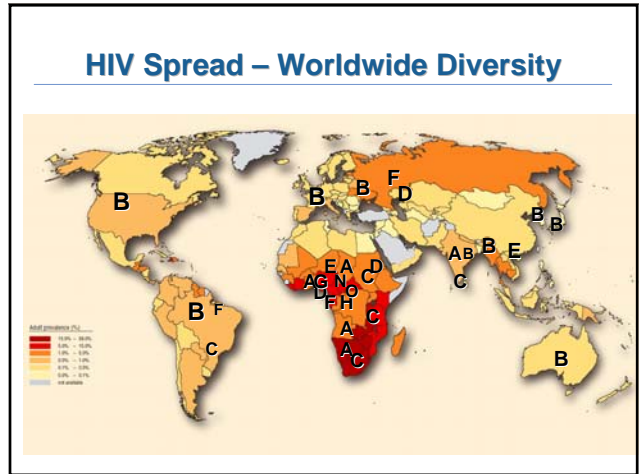
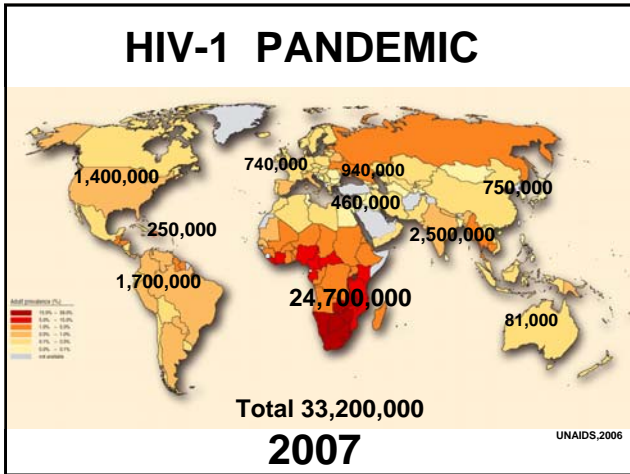


Cumulative U.S. AIDS Cases
as of 12/95 N~500,000



Cumulative U.S. AIDS Cases
as of 9/97 N~626,334





Bushmeat Trade in Central and West Africa



Poacher



Chophouse



Bushmeat market



Home



Restaurant

www.karlammann.com

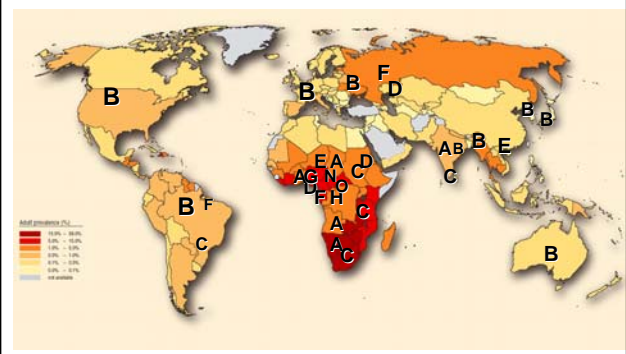
Chimpanzee Origins of HIV-1



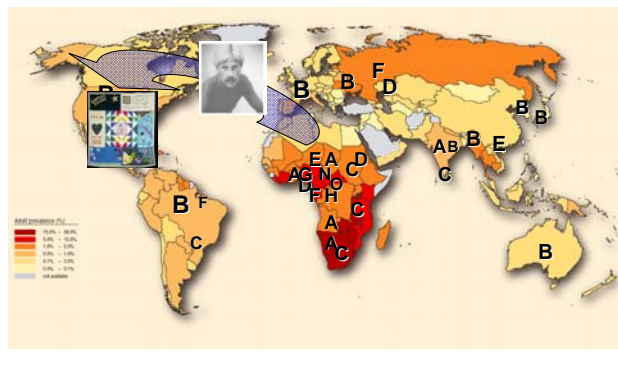
FIG. 2. Evolutionary relationships of SIVcpz and HIV-1 inferred from maximum likelihood phylogenetic analyses of full-length envelope protein sequences compiled from 105 SIVcpz strains from *P. troglodytes* and *P. schweinfurthii* as highlighted in red and blue, respectively. Representative strains of SIV-1 groups M, N, and O from humans are highlighted. Bootstrap values are shown at nodes with a minimum posterior probability of 95% or higher. The scale bar denotes 0.1% nucleotide divergence.

Hahn and coworkers JV 2005

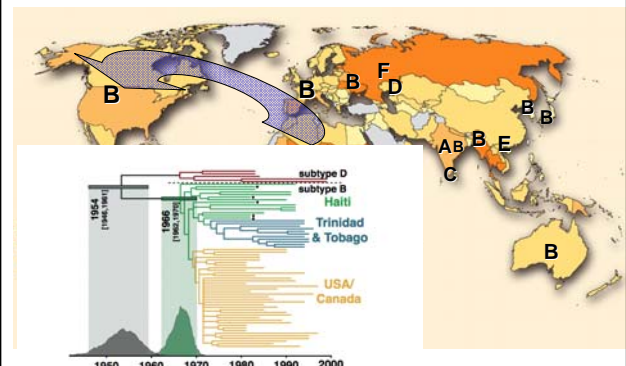
HIV Spread – Worldwide Diversity

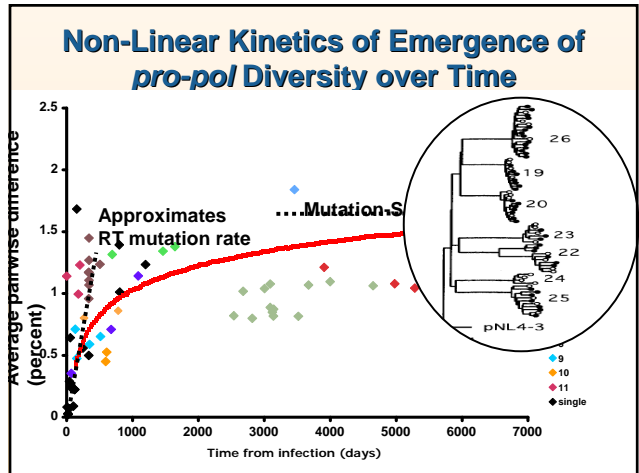
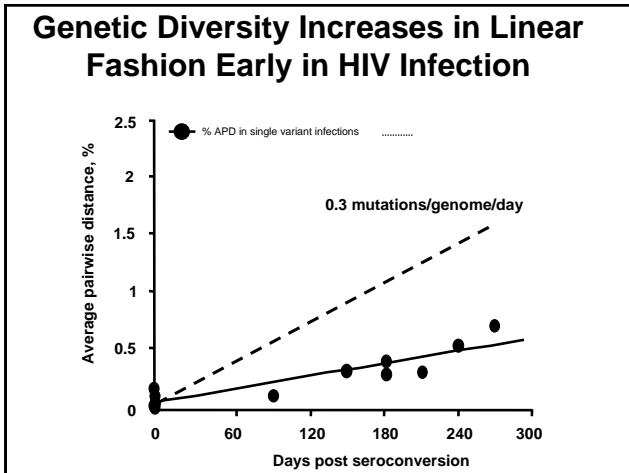
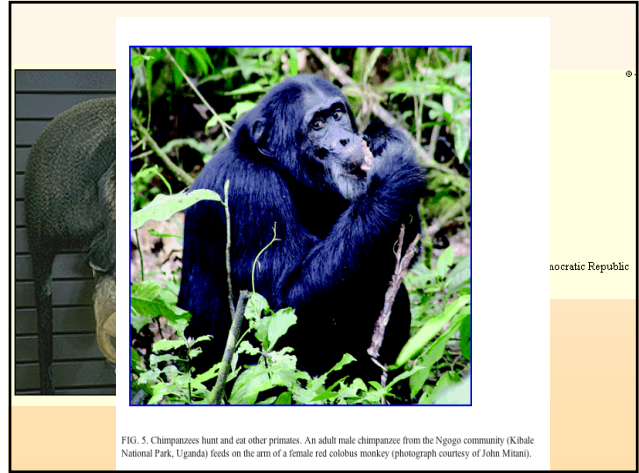
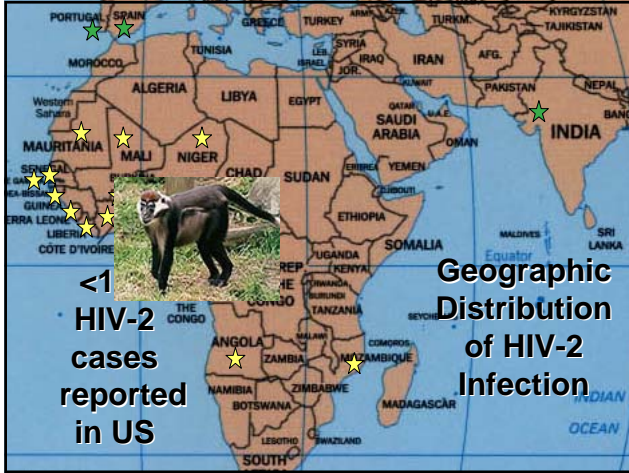


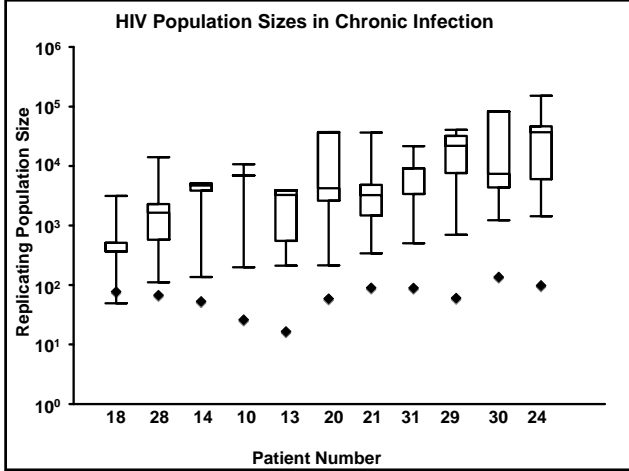
HIV Spread- Founder effects



HIV Spread- Founder effects

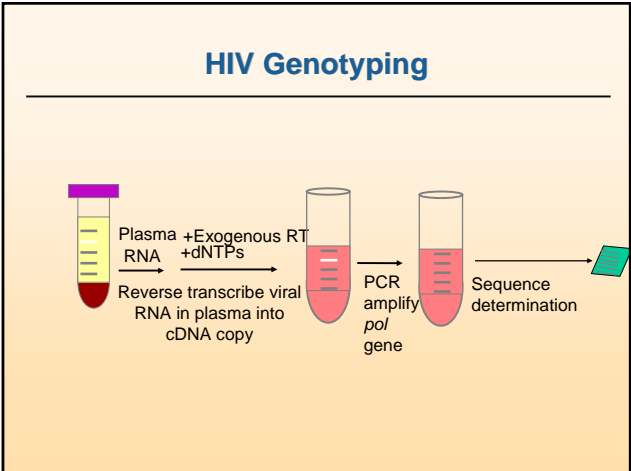
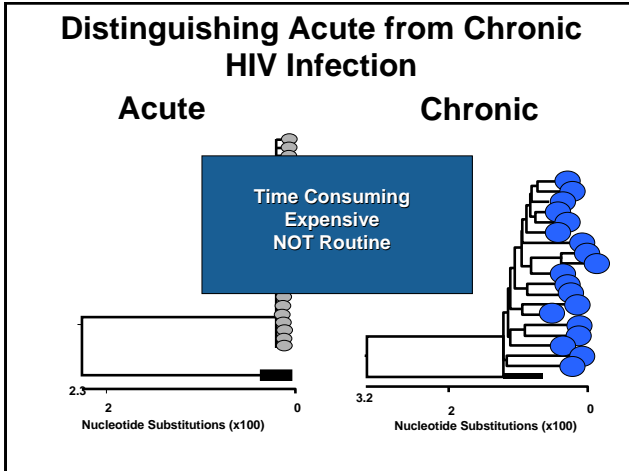




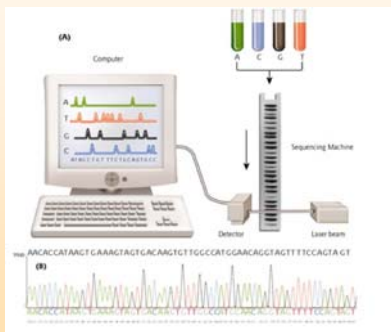


HIV Population Genetics

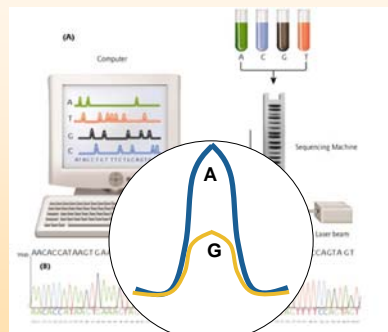
- HIV populations PRIOR to therapy
 - Diverse
 - Large
 - Change slowly



Sequence Processing



Sequence Processing



AMBIGUOUS BASE CALL

Applied Population Genetics with Population Based Sequencing

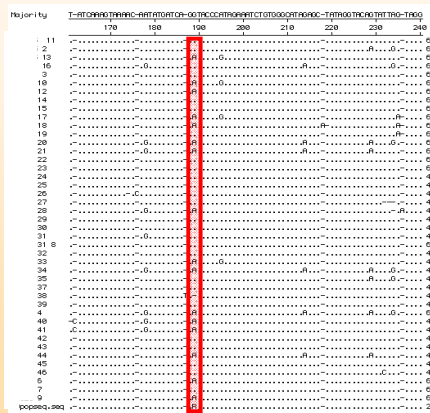
A. Acute Infection: No ambiguities

TGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTTGGGCTGAAAAATCCATACAATACTCCAGTATTGGCC
 ATAAAGAAAAAGCAGTACTAAATGCGAAATTTAGTAGATTTTCAGAGAACTTAATAAGCAGACTCAAGACTTGT
 GGGAGTCCAATTTAGGAATACCACATCCCGCAGGGTTAAAAAGAAAAAATCAGTAAACAGTACTGGATGGGT
 GATGCATATTTTCAGTCCCTTAGATAAGGACTTCAGGAAGTATACGCATTTACCATACCTAGTATAAAACAATG
 AGACACCAGGGATAG
 ATACCAGTACAATGTGGTTCCACAAGGATGAAAGGATCACCAGCAATATTTCAAAATAGCATGACAAAAATCTT
 AGACCTTTGAGAAAGAAAATCCAGACATAGTTATCTCAATACATGGATGTTGTATGTAGGATCTGACITTA
 GAAATAGGCAGCATAGAATAAAGTAGGAACTGAGAGAATCTGTGGAGTGGGGTTTACCACACCGA
 CAAAAAGCATCAGAAAGAACTCCATCTTTGGATGGGTTATGAACCTCATCTGATAAATGGACAGTACAGCC
 TATAGAGCTGCCA

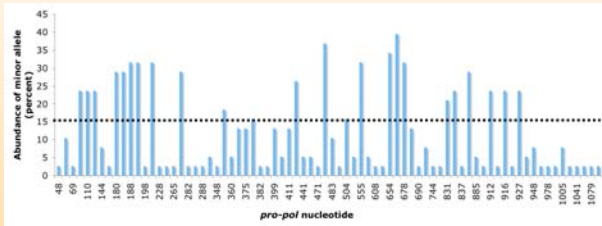
B. Chronic Infection: 0.7% of base calls are ambiguous

TGTRCAGAACTAGAAAAGGAAAGGAAATTTCAAAAATTTGGGCTGAAAAATCCATATAATACTCCAGTATTGGC
 CATAAAGAAAAAGACAGTACCAAAATGGAGA
 AAAGTAGTAGATTTCCAGAGAACTGAATAAAGAACTCAAGACTTCGGGAAGTCCAATTTAGGAATACCACATCC
 CGCAGGGATAAAAAAGAAACAAGTCAGTAACA
 GTATGGATGGGGTATGCATATTTCTCAGTTCCTTAGATGAAGATTTCCAGGAAGTATACTGCATTCACCATAC
 CCAATATAACAATGAGACACCGGATGATATCAGTACAATGTGCTCCACAGGGATGGAAGGGGTCAACA
 GCAATATCCAAGTAGCATGACAAAATCTAGAGCCCTTTAGAAAACAGAATCCAGACATAGTTATCTATCA
 TACGTRGATGATTTGTATAGGATCTGACTAGAAAATAGGCAACATAGAGCAAAATAGAGGAACCTGAGA
 MAACATCTGTTGAAATGGGGTTTTGCACACAGATGAAAAGCATCAGAAAAGAACTCCATCTCTGGATGG
 GTATGAACTCCATCTGATAAATGGACAGTACAGCCTATAGTCTGCCA

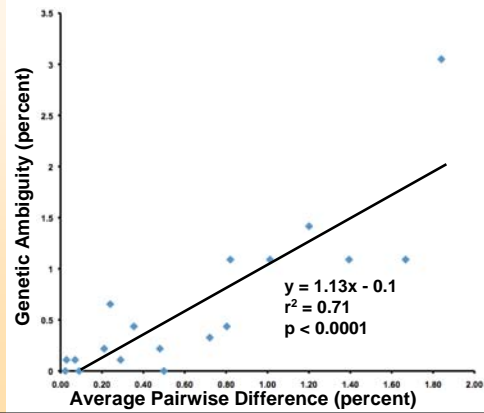
Ambiguity Captures A Portion of Sequence Diversity



Polymorphisms that Accumulate in HIV-1 *pro-pol* During Chronic Infection Are Frequent And Abundant



Sequence Ambiguity Reflects Genetic Diversity



Sequence Ambiguity Distinguishes Chronic From Recent HIV Infection

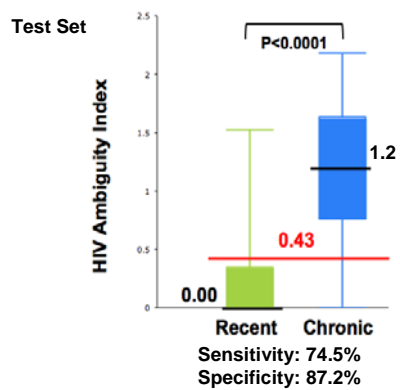
Patients

- **Recent:**
 - N=100
 - NIAID/CCMD
 - Public Health Agency of Canada
 - Documented seroconversion
- **Chronic:**
 - N=97 NIAID/CCMD
 - Western blot positive > 1 year
 - No Antiretroviral therapy

Methods

- Commercial Sequencing
- Randomized sample training set
 - Maximize Sensitivity/Specificity
- Test Set

Sequence Ambiguity Distinguishes Chronic From Recent HIV Infection



Identifying Recent HIV-1 Infection

- HIV-1 *pro-pol* diversity increases in bimodal fashion following infection
- During chronic infection, polymorphisms are frequent and abundant
- Polymorphisms are detectable by population based sequencing as ambiguous bases
- Bioinformatic methods can distinguish recent from chronic infections
 - Estimate HIV incidence
 - Simultaneous determination of transmitted resistance
 - Phylogenetic analyses
 - Identify transmission networks
 - Determine relative contributions of acute and chronic infections to transmission.

NCI HIV Drug Resistance Program NIAID/CCMD/VRC

- J. Coffin
 - M. Keiser
 - V. Bolz
 - H. Mens
 - W. Shao
 - J. Spindler
 - A Wiegand
 - F. Cossarini
 - E. Anderson
 - E. Wilson
 - T. Tanzosh
 - S. Hill
 - Study Coordinator
- Patient Volunteers**
- M Sneller
 - S. Kottitil
 - I. Sereti
 - R. Davey
 - H. Masur
 - H. C. Lane
 - D. Douek
 - N. Sandler

NNMC

- Anu Ganesan

Karolinska Institute

- S. Palmer
- L. Josefsson

•UCLA

- E. Daar

•Johns Hopkins

- J. Margolick

•UCSD

- E. Reid

NCI HAMB

- R. Yarchoan
- T. Uldrick

University of Pittsburgh

- J. Mellors
- D. McMahon