

Sequencing Applications in HIV Forensics



Steve Scherer

Human Genome Sequencing Center
Department of Molecular
and Human Genetics
Baylor College of Medicine

ABRF 2013

Application of Next Generation
Sequencing Technologies for Whole
Transcriptome and Genome Analysis
Workshop

Phylogenetic Linkage Cases

Establishing Linkage:

- The Florida Dentist – 6 patient samples were more closely related to each other than unrelated controls
- The French Surgeon
- The French Nurse
- Criminal cases of HIV transmission in Sweden, Australia, Belgium, Denmark, Germany and Scotland

Established Non-Linkage:

- Baltimore Surgeon
- UK obstetrician/gynecologist

Molecular Forensics

- National: 2001 Anthrax in the Mail Case – Amerithrax
 - Lengthy FBI/DOJ investigation – Bruce Ivins
 - Highly controversial
 - National Academy of Sciences Review: “impossible to reach any definitive conclusion about the origins of the anthrax in the letters, based solely on the available scientific evidence”
- International: The Poison Umbrella Case
 - Georgi Markov: Bulgarian activist, playwright & satirist was working for the BBC World Service in London

SEARCH

The Web CNN.com

Search

Home Page

World

U.S.

Weather

Business at CNNMoney

Sports at SI.com

Politics

Law

Technology

Science & Space

Health

Entertainment

Travel

Education

Special Reports

SERVICES

Video

E-mail Newsletters

Ricin and the umbrella murder

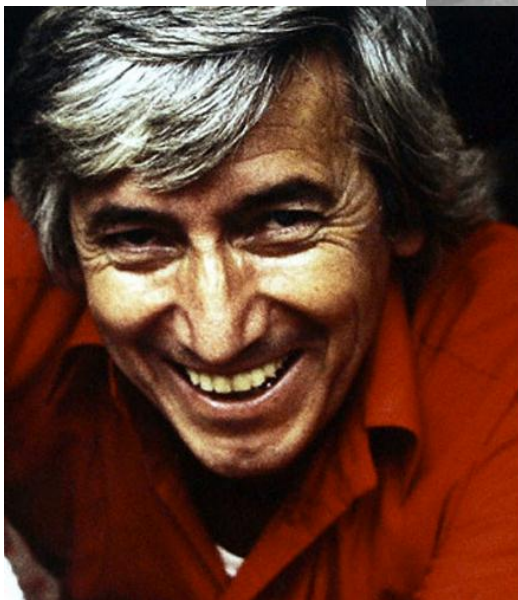
Thursday, October 23, 2003 Posted: 2:27 AM EDT (0627 GMT)

LONDON, England -- It was one of the most notorious acts of assassination carried out during the Cold War.

Bulgarian dissident Georgi Markov was killed by poison dart filled with ricin and fired from an umbrella in London in 1978.

Markov, a communist defector working for the BBC World Service, left his office at Bush House in the UK capital on September 11 and walked across Waterloo Bridge to take the train home to Clapham in south-west London.

advertisement



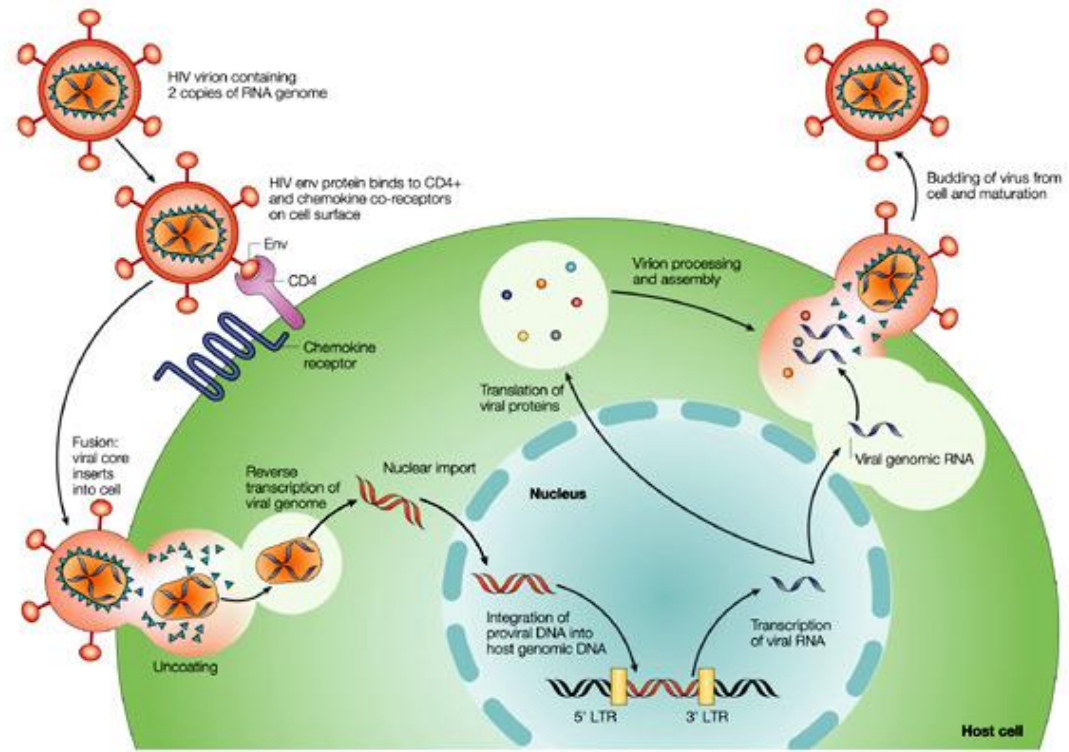
DNA Testing in the Judicial System

- In use since the mid '80's
- Usually used to link perpetrators to violent crime scenes
- Generally stable allowing the use of polymorphic marker DNA fingerprinting - error correcting machinery

Dynamic evolution of HIV

Individuals infected with HIV-1 contain a dynamically evolving population of related genomes

- Viral expansion
 - Mutation rate
 - Recombination rate
 - Production of 10^8 to 10^{10} virions per day
- Lineage extinction
 - Non-replicating virions
 - Immune system
 - Drug therapy



Phylogenetics in Forensics

Viral dynamics limit use the common practice of matching DNA profiles

Phylogenetic methods are ideally suited for determining the HIV pattern of descent in suspected transmission cases

In support of the *a priori* hypothesis, HIV forensics can identify case samples to be “more closely related” than to unrelated samples

In opposition of the *a priori* hypothesis, HIV forensics can show case samples to be unrelated

HIV forensics

**State of Louisiana v.
Richard J. Schmidt**

Patient → Trahan



Metzker *et al.* (2002)
PNAS **99**: 14292-14297

**State of Washington v.
Anthony E. Whitfield**

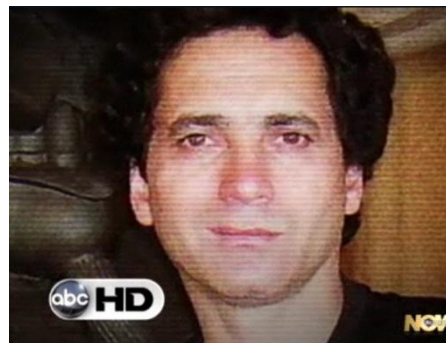
Whitfield → 5 partners



Scaduto *et al.* (2010) *PNAS*
107: 21242-21247

**State of Texas v.
Philippe Padieu**

Padieu → 6 partners



State of Louisiana v. Richard J. Schmidt

Patient → Victim

Proceedings of the National Academy of Sciences of the United States of America

PNAS

CURRENT ISSUE // ARCHIVE // NEWS & MULTIMEDIA // FOR AUTHORS // ABOUT PNAS COLLECTED ARTICLES / BROWSE BY TOPIC / EARLY EDITION

↑ > Current Issue > vol. 99 no. 22 > Michael L. Metzker, 14292–14297

Molecular evidence of HIV-1 transmission in a criminal case

Michael L. Metzker^{*†}, David P. Mindell[‡], Xiao-Mei Liu^{*§}, Roger G. Ptak[¶], Richard A. Gibbs^{*}, and David M. Hillis^{**}

Author Affiliations 

^{*}Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030; [†]Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor, MI 48109-1079; [‡]School of Dentistry, Biologic and Materials Sciences, University of Michigan, Ann Arbor, MI 48109; and [¶]Section of Integrative Biology and Center for Computational Biology and Bioinformatics, University of Texas, Austin, TX 78712

Edited by Walter M. Fitch, University of California, Irvine, CA, and approved September 4, 2002 (received for review May 2, 2002)



This Issue
October 29, 2002
vol. 99 no. 22
[Table of Contents](#)

[◀ PREV ARTICLE](#) [NEXT ARTICLE ▶](#)

Published online before print
October 18, 2002, doi:
[10.1073/pnas.222522599](https://doi.org/10.1073/pnas.222522599)
PNAS October 29, 2002 vol. 99 no.
22 14292-14297

Classifications 

Biological Sciences
Evolution

Prosecution argued successfully:

Methods used for:

- DNA isolation
- PCR
- DNA sequencing
- Phylogenetic analysis of HIV-1 positive samples

Met the standards of judicial evidence admissibility

Due to:

- Are subject to empirical testing
- Can be assessed for error
- Subject to peer review and publication
- Generally accepted in the scientific community

Case Facts



- August 4, 1994 – a Lafayette, LA gastroenterologist created a cocktail of two blood samples; one from a man infected with HIV-1 and a second from a patient infected with Hep-C, then infected his former girlfriend by intramuscular injection
- Victim had 7 sexual relationships between 1984 and 1995; all tested negative for HIV
- Victim was a nurse in the Lafayette area; no documented needle sticks; one HIV + saliva splash; tested negative
- Victim donated blood and tested negative 10/92, 5/93 & 4/94
- Victim tested positive in January 1995 – accused the physician
- Police raid discovered potential source blood draw record

Testing Protocol

1. Data was derived from two separate labs and two independent blood draws; the second lab was blinded.
2. Samples were tested one at a time – lab scrub between samples.
3. PCR amplified *env* and RT genes
 - exhibit different biological functions
 - subject to different selective pressures
 - known to undergo different rates of evolution.
4. 50 molecular clones from sample were sequenced to further delineate *env* genetic diversity.
5. Geographic HIV-1+ controls were recruited from the Lafayette metro area based on previous studies demonstrating geographic subtype stratification of HIV-1 sequences

Table 1. Summary of LA control group sample sources

Risk factors			
Homosexual	57%	(16/28)	
Heterosexual	18%	(5/28)	
Blood transfusion	11%	(3/28)	
Bi-sexual	7%	(2/28)	
IV drug user	4%	(1/28)	
Sharps	4%	(1/28)	
Date infected			
1983–1989	57%	(16/28)	
1990–1992	36%	(10/28)	
Unknown	7%	(2/28)	
CD4 ⁺ cell counts			AIDS
>500	36%	(10/28)	(2/10)
200–500	25%	(7/28)	(3/7)
<200	25%	(7/28)	(6/7)
ND	14%	(4/28)	(4/4)

Risk factors and dates infected were obtained by anonymous questionnaire. ND, not determined.

Phylogenetic Analyses

For the case: parsimony and minimum evolution
using maximum-likelihood distances

Subsequently: Markov-chain Monte Carlo Bayesian Analysis
based on a General-Time-Reversible model of
sequence evolution, with γ -distributed rate heterogeneity
among sites and a calculated proportion of invariable sites

Table 2. Means and 95% confidence intervals for parameters of the GTR + Γ + I model for gp120 sequences

Parameter	Mean	95% Confidence interval
C–T substitution rate	5.03	3.60–7.03
C–G substitution rate	0.97	0.57–1.54
A–T substitution rate	0.75	0.52–1.07
A–G substitution rate	3.87	2.91–5.10
A–C substitution rate	2.34	1.60–3.34
Frequency of A	0.40	0.37–0.43
Frequency of C	0.15	0.13–0.17
Frequency of G	0.23	0.21–0.25
Frequency of T	0.22	0.20–0.25
α (shape of Γ distribution)	0.53	0.43–0.68
Proportion of invariable sites	0.08	0.01–0.18

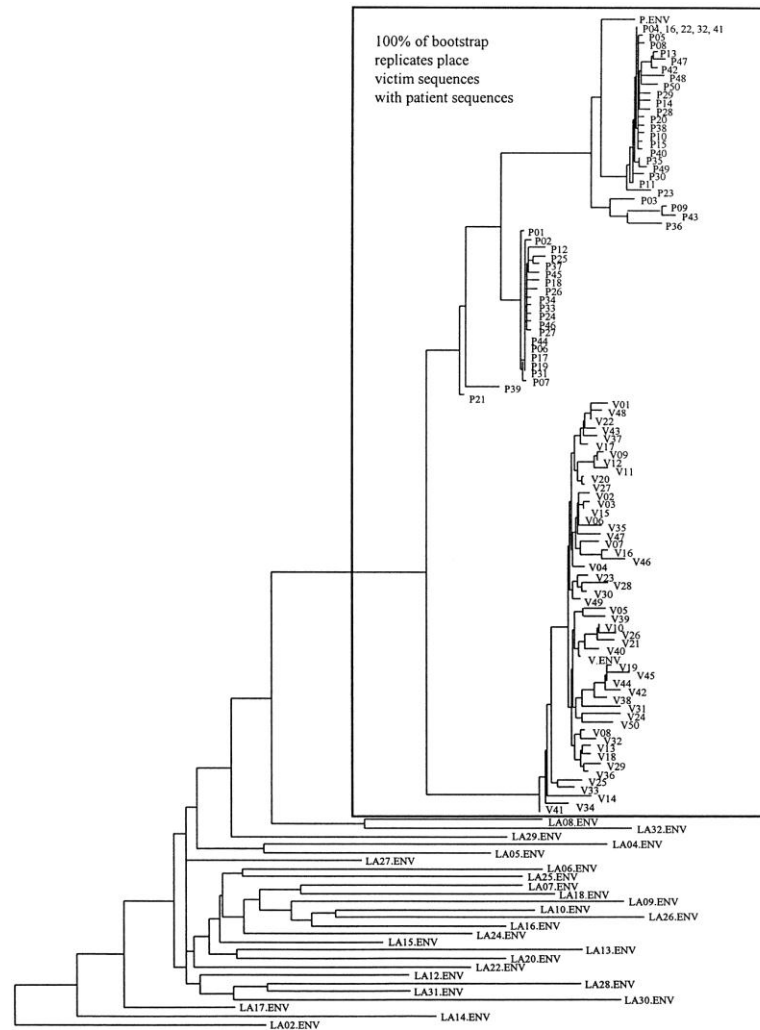
Data based on MCMC sampling (25). The rate of all substitution classes is shown relative to that of the G–T substitution class.

Table 3. Means and 95% confidence intervals for parameters of the GTR + Γ + I model for the RT sequences

Parameter	Mean	95% Confidence interval
C–T substitution rate	110.36	23.04–195.53
C–G substitution rate	17.59	2.82–42.02
A–T substitution rate	7.62	1.34–17.32
A–G substitution rate	83.01	16.29–171.17
A–C substitution rate	16.60	3.41–35.62
Frequency of A	0.40	0.36–0.43
Frequency of C	0.17	0.14–0.19
Frequency of G	0.20	0.17–0.23
Frequency of T	0.23	0.20–0.26
α (shape of Γ distribution)	0.94	0.38–1.94
Proportion of invariable sites	0.50	0.29–0.63

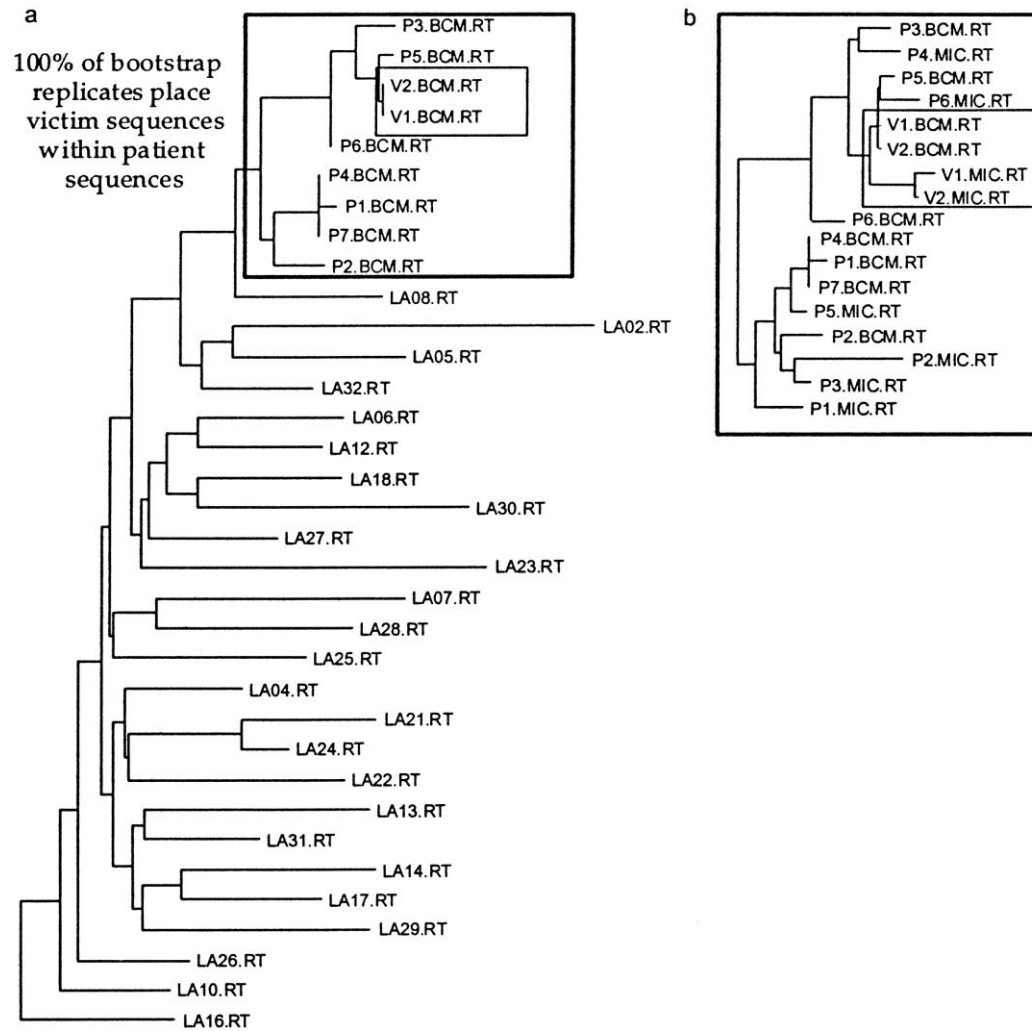
Data are based on MCMC sampling (25). The rate of all substitution classes is shown relative to the rate of the G–T substitution class.

Phylogenetic analysis of the gp120 region using a minimum evolution criterion and maximum likelihood distances assuming an HKY+ Γ model of evolution.



Metzker M L et al. PNAS 2002;99:14292-14297

Phylogenetic analysis of the RT region; details of the analysis are the same as for Fig. 1.



Results

For RT: Patient sequences were found to be paraphyletic with respect to victim sequences (victim sequences nested within patient sequences) by parsimony, minimum evolution, and Bayesian analyses.

For *env* gp120: Not as strong; supports a weak monophyletic grouping. Probably due faster evolution and strong immuno-selection.

Defendant was found guilty of attempted second degree murder; Appeal rejected by the Louisiana Supreme Court; U.S. Supreme Court declined to the case

The Second Set of Cases

Proceedings of the National Academy of Sciences of the United States of America

PNAS

[CURRENT ISSUE](#) // [ARCHIVE](#) // [NEWS & MULTIMEDIA](#) // [FOR AUTHORS](#) // [ABOUT PNAS](#) // [COLLECTED ARTICLES](#) / [BROWSE BY TOPIC](#) / [EARLY EDITION](#)

[↑](#) > [Current Issue](#) > [vol. 107 no. 50](#) > [Diane I. Scaduto, 21242–21247](#)

Source identification in two criminal cases using phylogenetic analysis of HIV-1 DNA sequences

Diane I. Scaduto^{a,b}, Jeremy M. Brown^{c,1}, Wade C. Haaland^{a,b}, Derrick J. Zwickl^{c,2}, David M. Hillis^{c,3}, and Michael L. Metzker^{a,b,d}

[Author Affiliations](#) [↗](#)

[Author Notes](#)

Contributed by David M. Hillis, October 20, 2010 (sent for review September 22, 2010)



This Issue
December 14, 2010
vol. 107 no. 50
[Masthead \(PDF\)](#)
[Table of Contents](#)

[◀ PREV ARTICLE](#) [NEXT ARTICLE ▶](#)

Published online before print
November 15, 2010, doi:
10.1073/pnas.1015673107
PNAS December 14, 2010 vol. 107
no. 50 21242–21247

[Classifications](#) [▼](#)

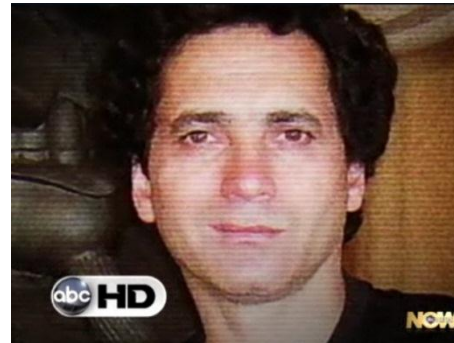
**State of Washington v.
Anthony E. Whitfield**

Whitfield → 5 partners



**State of Texas v.
Philippe Padieu**

Padieu → 6 partners



Direction of transmission

(source → recipient)

Providing evidence for the direction of transmission would further strengthen the *a priori* hypothesis.

Genetic bottleneck during transmission

- Paraphyly: Evidence for direction of transmission

Study design:

- identities of case subjects were blinded to investigators
- case sample handling were separated both temporally and spatially to eliminate the possibility of cross contamination
- case allegations were multiple transmissions from a single source

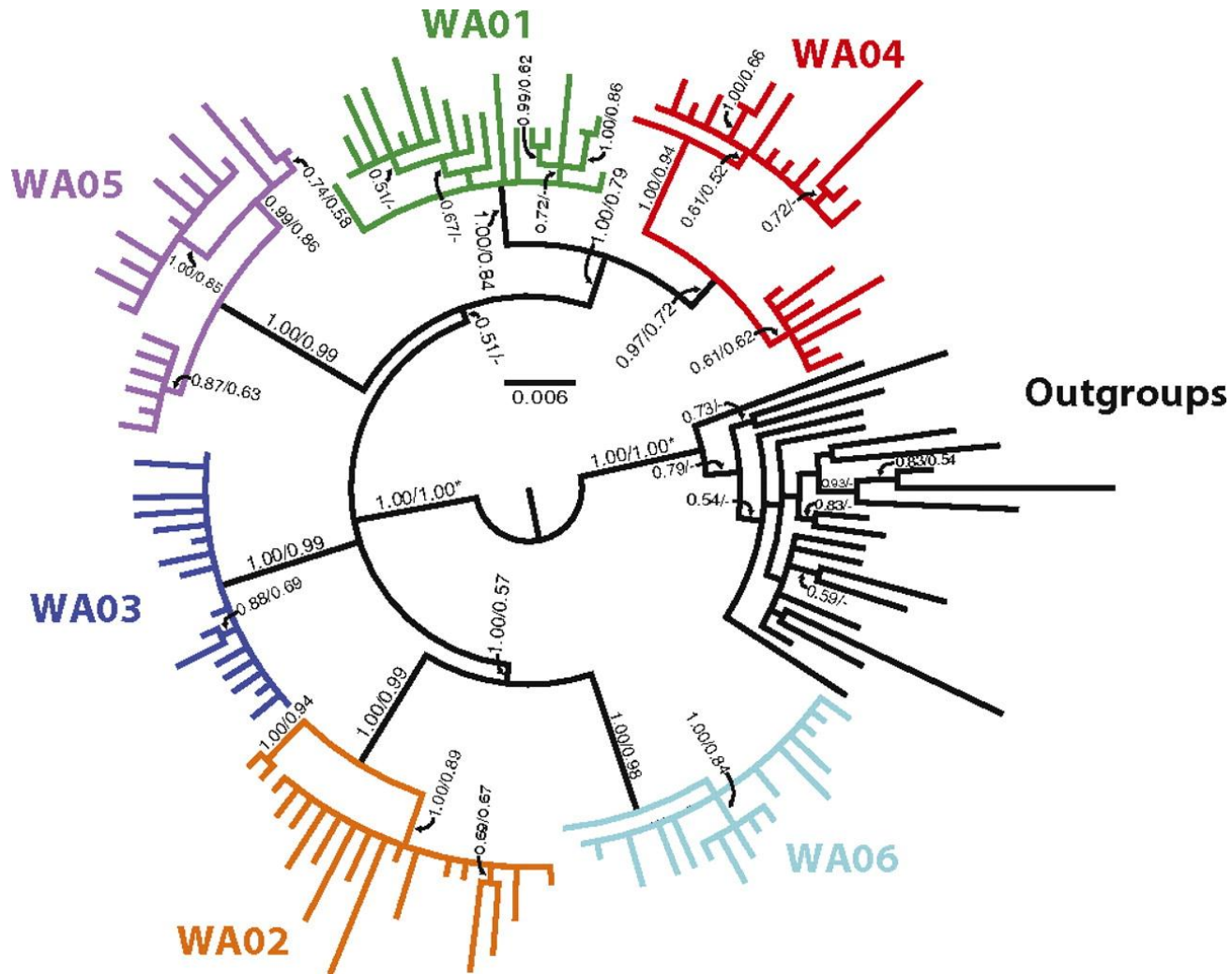
Washington Case

- Defendant allegedly learned of his HIV positive status in April 1992; 17 partners were exposed between 1999 and 2004; 5 tested positive between 2002 and 2004
- These 6 samples formed the basis of the *a priori* hypothesis of transmission from one source to multiple recipients

Texas Case

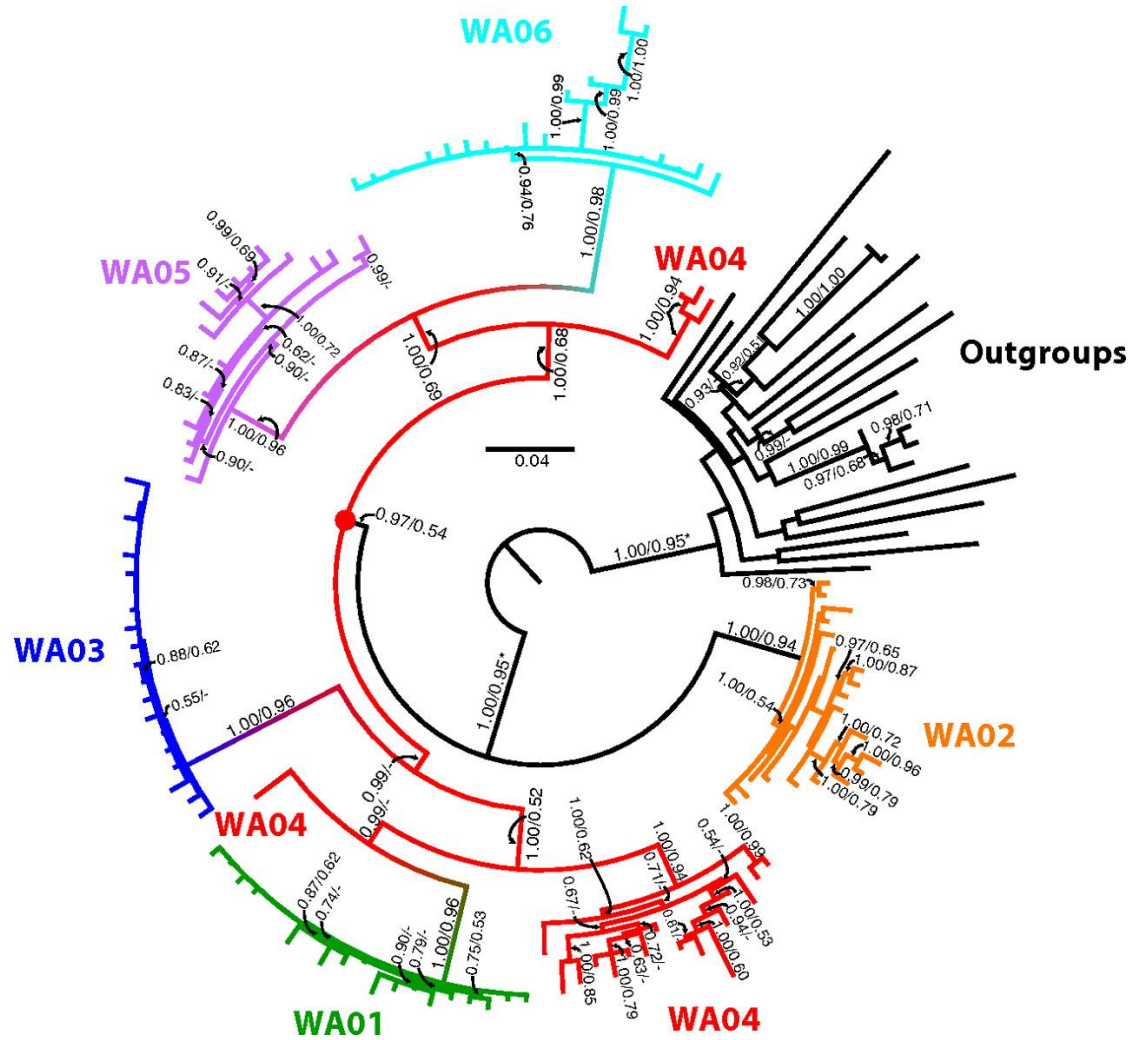
- Defendant allegedly learned of his HIV positive status in September 2005; 6 partners tested positive between April 2006 and March 2007
- These 7 samples formed the basis of the *a priori* hypothesis of transmission from one source to multiple recipients

Washington case: ML tree for the pol gene dataset using BLAST-selected GenBank controls.



Scaduto D I et al. PNAS 2010;107:21242-21247

Washington case: ML tree for the env gene dataset using BLAST-selected GenBank controls.



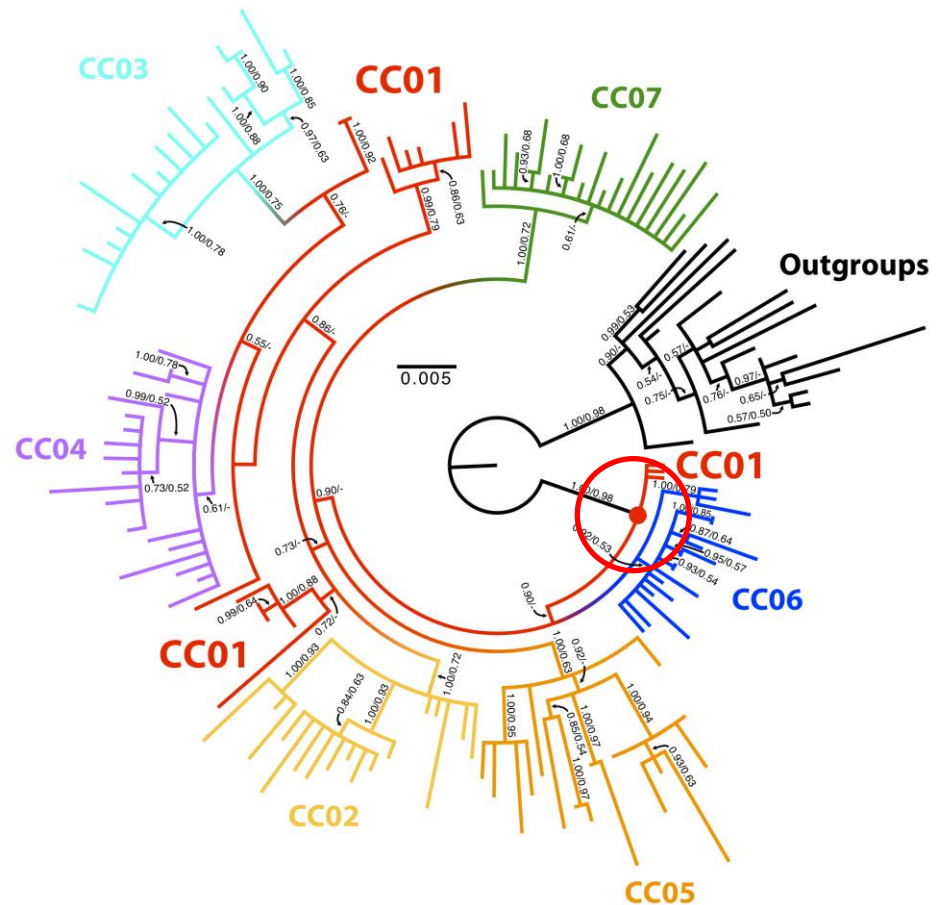
Scaduto D I et al. PNAS 2010;107:21242-21247

Texas case: pol tree

CC01 exhibited a paraphyletic relationship to all CC case sequences

- Bayesian posterior probabilities (1.00)
- ML bootstrapping proportions (0.98)

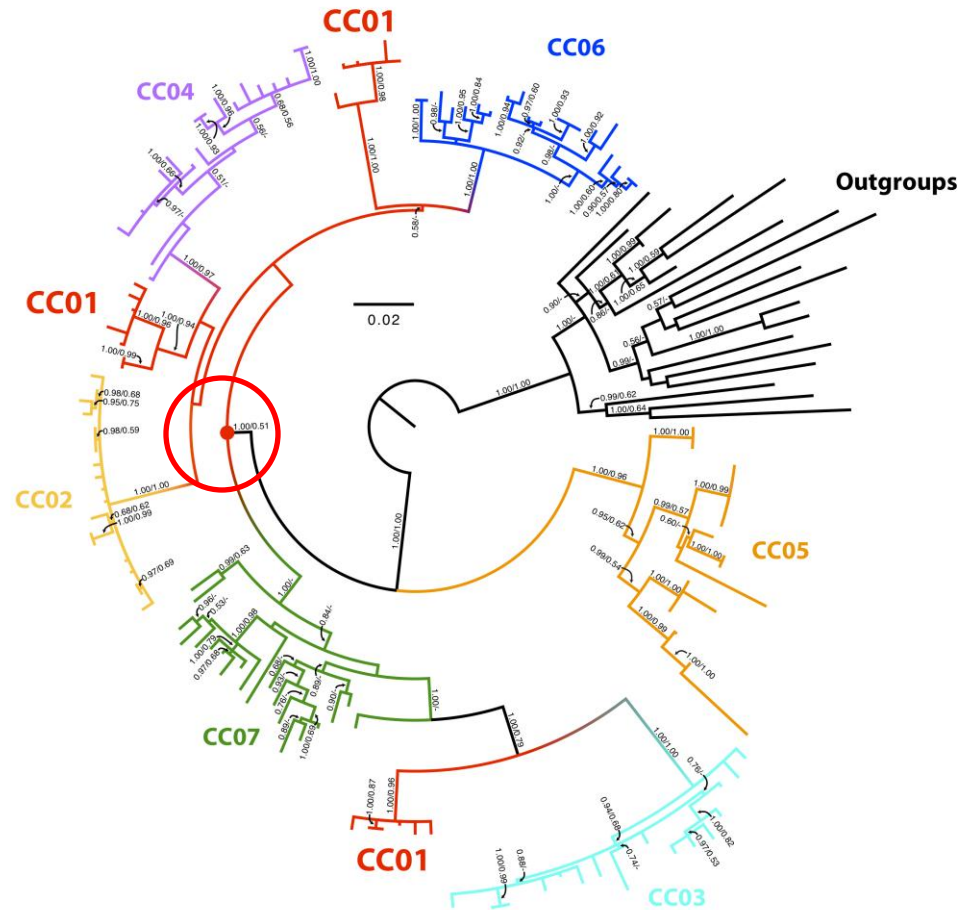
Red circle represents the most recent common ancestor of sequences from CC01



Texas case: *env* tree

CC01 exhibited a paraphyletic relationship to all CC case sequences but CC05

- Bayesian posterior probabilities (1.00)
- ML bootstrapping proportions (1.00)



Red circle represents the most recent common ancestor of sequences from CC01

Breaking the code

For the WA case, we inferred that sample WA04 was the source (i.e., index case)

At trial, the identity of sample WA04 was revealed to be that of Anthony E. Whitfield

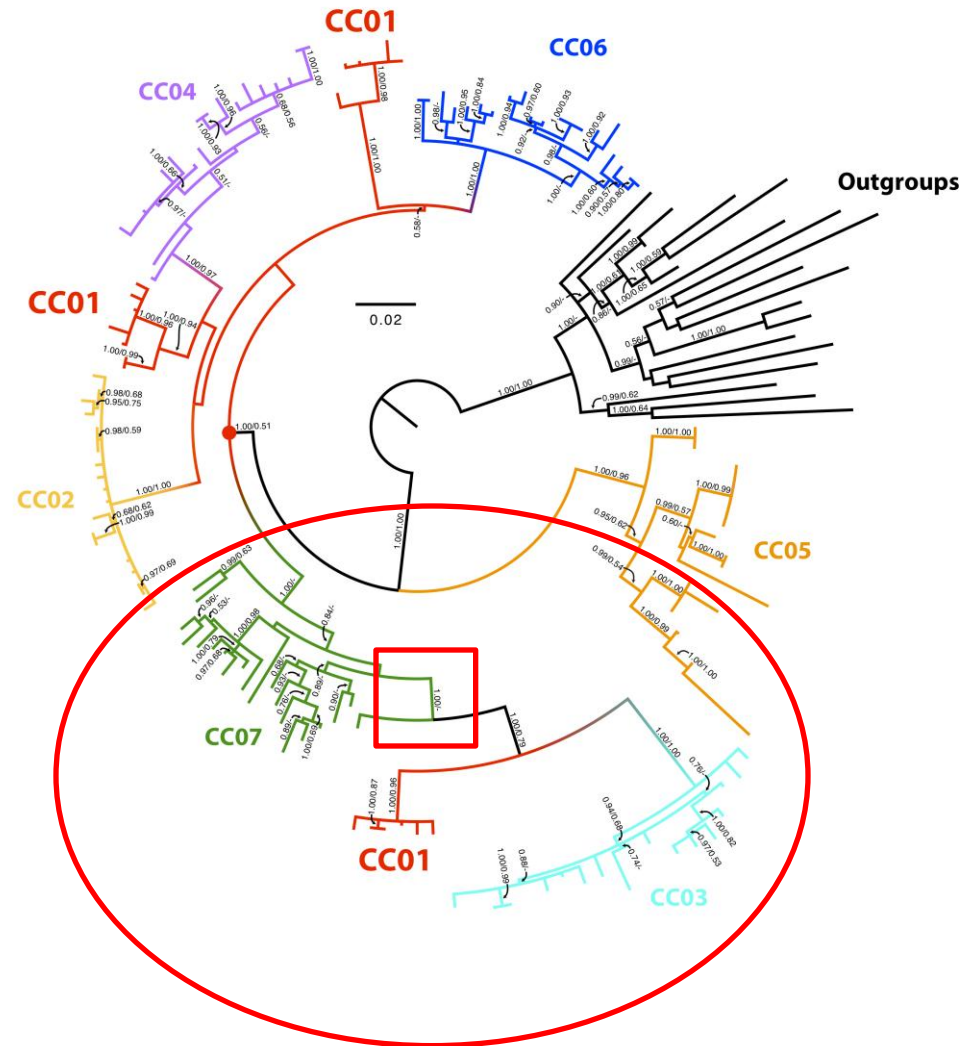
For the TX case, we inferred that sample CC01 was the index case

At trial, the identity of sample CC01 was revealed to be that of Philippe Padieu

Texas case: *env* tree

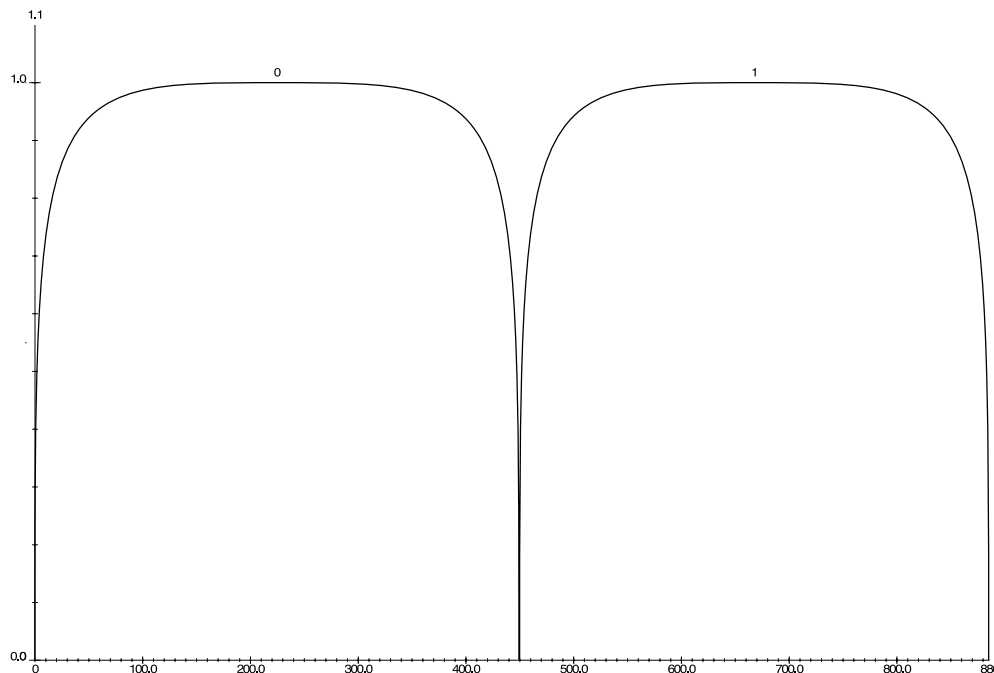
CC07 exhibited a paraphyletic relationship to several CC01 and all CC03 sequences

- Bayesian posterior probabilities (1.00)
- ML bootstrapping proportions (0.79)

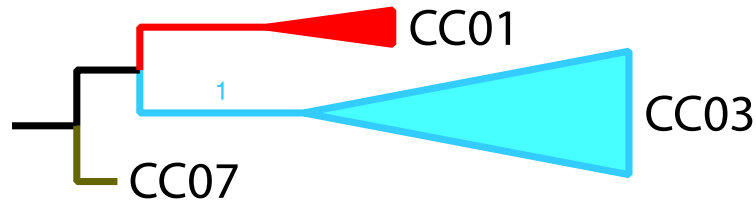


Testing for Recombination in env

Used a maximum likelihood method that detects recombination breakpoints using a hidden Markov model*
If two trees are allowed across the sequence, it estimated a recombination structure like this:



Divergent Signals Between Halves

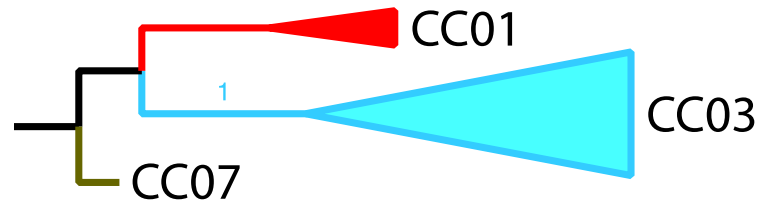


1st half - Posterior Probability: 0.00

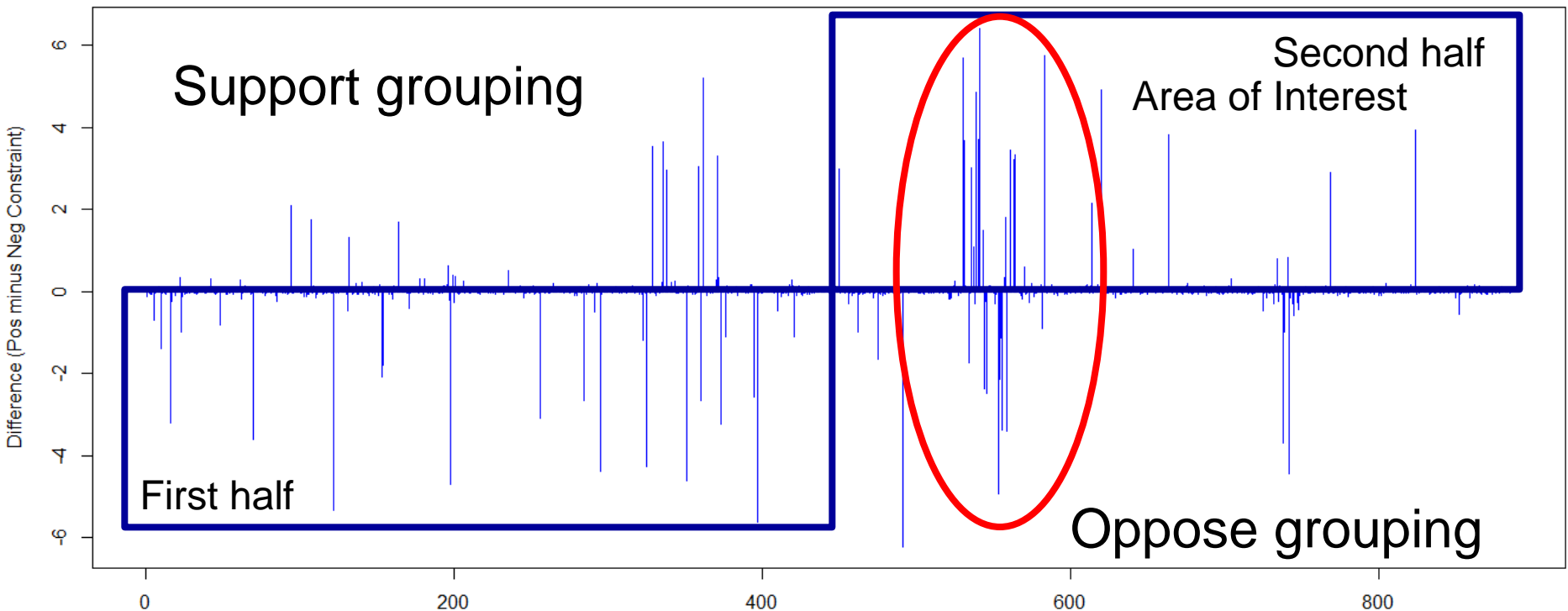
2nd – half - Posterior Probability: 1.00

Divergent Signals Between Halves

Site index



Maps within V4
loop of gp120



OPEN ACCESS

10,255

VIEWS

4

CITATIONS

13


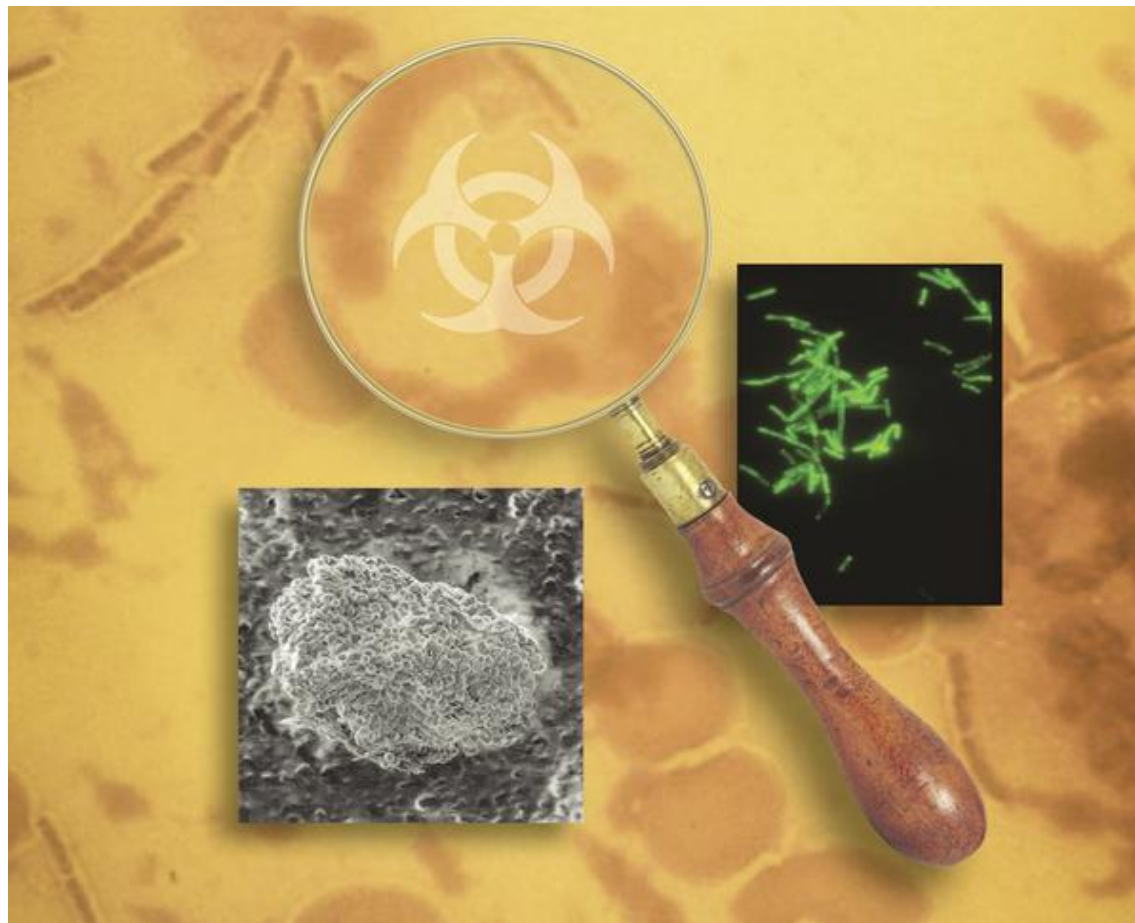
ACADEMIC
BOOKMARKS

2

SOCIAL
SHARES

POLICY FORUM

Biocrimes, Microbial Forensics, and the Physician

Steven E Schutzer , Bruce Budowle, Ronald M Atlas

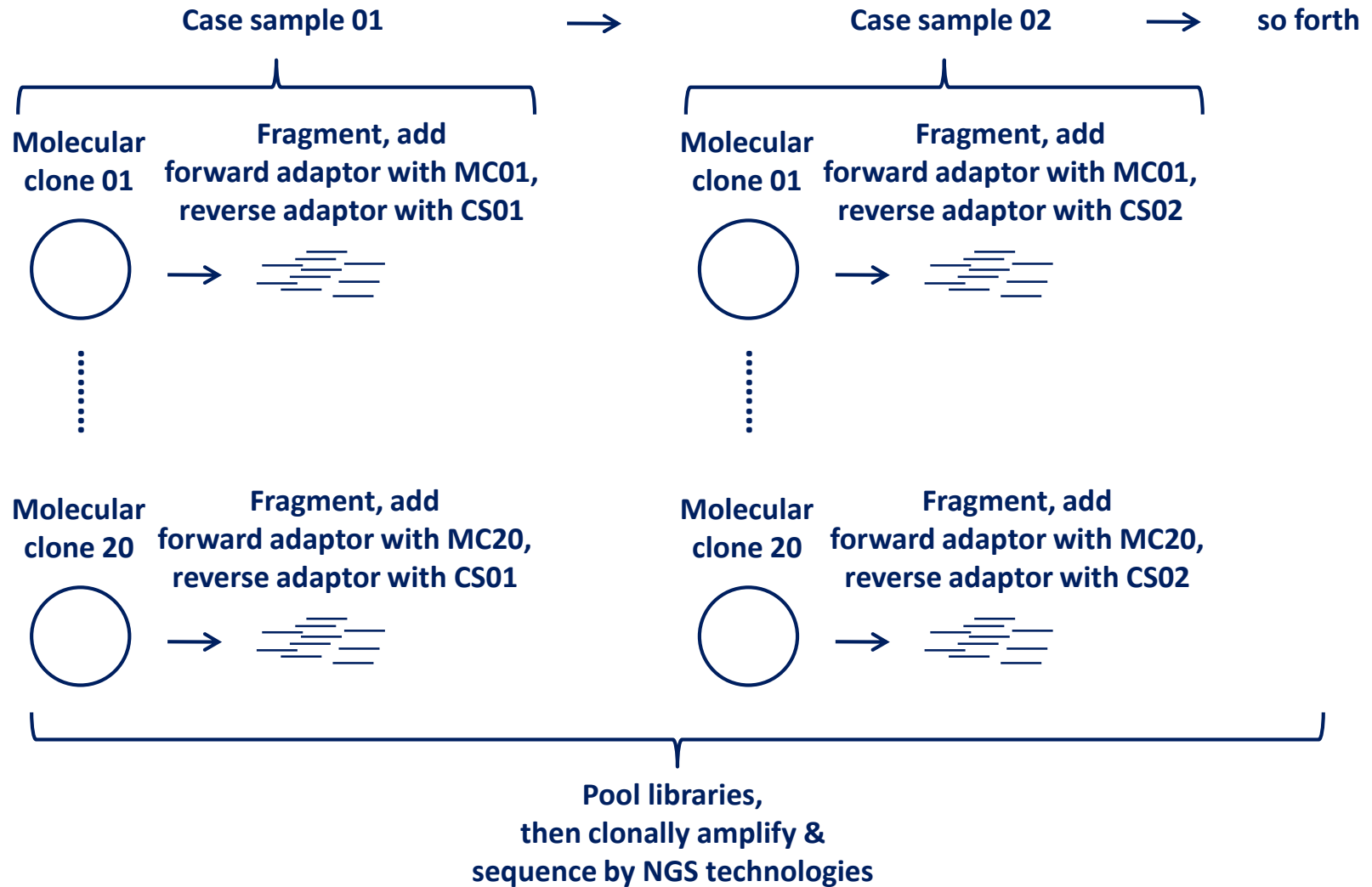
Schutzer SE, Budowle B, Atlas RM (2005) Biocrimes, Microbial Forensics, and the Physician. PLoS Med 2(12): e337.

doi:10.1371/journal.pmed.0020337

<http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020337>

NGS in HIV forensics

Development of the 'pathogen toolkit'



Further Cases and Uses

Virology Journal



Research

Open Access

A HIV-1 heterosexual transmission chain in Guangzhou, China: a molecular epidemiological study

Zhigang Han^{†1}, Tommy WC Leung^{†2}, Jinkou Zhao^{†3}, Ming Wang¹, Lirui Fan¹, Kai Li⁴, Xinli Pang⁴, Zhenbo Liang⁴, Wilina WL Lim² and Huifang Xu^{*1}

Address: ¹Guangzhou Center for Disease Control and Prevention, Guangdong, PR China, ²Virology Division, Public Health Laboratory Services Branch, Center for Health Protection, Hong Kong SAR, ³Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, PR China and ⁴Huadu Center for Disease Control and Prevention, Guangdong, PR China

Email: Zhigang Han - zhiganghan616@163.com; Tommy WC Leung - tommywcleung@yahoo.com; Jinkou Zhao - jinkouzhao@hotmail.com; Ming Wang - wangming@gzcdc.org.cn; Lirui Fan - fanlirui2010@hotmail.com; Kai Li - gzibwf@21cn.com; Xinli Pang - gzibwf@21cn.com; Zhenbo Liang - gzibwf@21cn.com; Wilina WL Lim - wllim@pacific.net.hk; Huifang Xu* - xuhuifang1027@21cn.com

* Corresponding author †Equal contributors

Published: 25 September 2009

Received: 25 June 2009

Virology Journal 2009, 6:148 doi:10.1186/1743-422X-6-148

Accepted: 25 September 2009

This article is available from: <http://www.virologyjournal.com/content/6/1/148>

© 2009 Han et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

OPEN ACCESS Freely available online



Using HIV Transmission Networks to Investigate Community Effects in HIV Prevention Trials

Joel O. Wertheim^{1*}, Sergei L. Kosakovsky Pond², Susan J. Little², Victor De Gruttola³

¹ Department of Pathology, University of California San Diego, San Diego, California, United States of America, ² Department of Medicine, University of California San Diego, San Diego, California, United States of America, ³ Department of Biostatistics, Harvard University, Cambridge, Massachusetts, United States of America

OPEN ACCESS Freely available online



Molecular Epidemiology of HIV-1 Transmission in a Cohort of HIV-1 Concordant Heterosexual Couples from Dakar, Senegal

Wim Jennes^{1*}, Jordan K. Kyongo¹, Evelyn Vanhommerig¹, Makhtar Camara², Sandra Coppens³, Moussa Seydi⁴, Souleymane Mboup², Leo Heyndrickx³, Luc Kestens¹

¹ Laboratory of Immunology, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium, ² Laboratory of Immunology, Department of Bacteriology-Virology, Centre Hospitalier Universitaire Le Dantec, Cheikh Anta Diop University, Dakar, Senegal, ³ Laboratory of Virology, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium, ⁴ Department of Infectious Diseases, Centre Hospitalier Universitaire Fann, Cheikh Anta Diop University, Dakar, Senegal

Haemophilia



Haemophilia (2012), 18, 291–299

DOI: 10.1111/j.1365-2516.2011.02620.x

ORIGINAL ARTICLE *Transfusion transmitted disease*

Molecular evidence of HIV-1 transmission in 20 Korean individuals with haemophilia: phylogenetic analysis of the *vif* gene

Y.-K. CHO,* Y. JUNG,* J.-S. LEE* and B. T. FOLEY†

*Department of Microbiology, University of Ulsan College of Medicine, Seoul, South Korea; and †HIV Databases, Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, USA

MRSA Transmission on a Neonatal Intensive Care Unit: Epidemiological and Genome-Based Phylogenetic Analyses

Ulrich Nübel^{1*}, Matthias Nachtnebel^{2,3,4}, Gerhard Falkenhorst², Justus Benzler², Jochen Hecht^{5,6}, Michael Kube^{5a}, Felix Bröcker⁵, Karin Moelling^{5,7}, Christoph Bühner⁸, Petra Gastmeier⁹, Brar Piening⁹, Michael Behnke⁹, Manuel Dehnert², Franziska Layer¹, Wolfgang Witte¹, Tim Eckmanns²

1 Department of Infectious Diseases, Unit of Nosocomial Infections, Robert Koch Institute, Wernigerode, Germany, **2** Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany, **3** Post Graduate Training in Applied Epidemiology, Robert Koch Institute, Berlin, Germany, **4** European Programme for Intervention Epidemiology Training (EPIET), European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden, **5** Max Planck Institute for Molecular Genetics, Berlin, Germany, **6** Berlin-Brandenburg Center for Regenerative Therapies, Charité University Medical Center, Berlin, Germany, **7** Institute of Medical Virology, University of Zürich, Zürich, Switzerland, **8** Department of Neonatology, Charité University Medical Center, Berlin, Germany, **9** Institute of Hygiene and Environmental Medicine, Charité University Medical Center, Berlin, Germany

Acknowledgements

BCM-HGSC

Diane I. Scaduto – UTMDA

Xiao-Mei Liu

Wade C. Haaland *

Donna Muzny

Eric Boerwinkle

Richard A. Gibbs

U of Texas

Jeremy M. Brown

Derrick J. Zwickl

David Hillis

U of Michigan

David P. Mindell

Roger G. Ptak