HIV Origins, Entry and Evolution

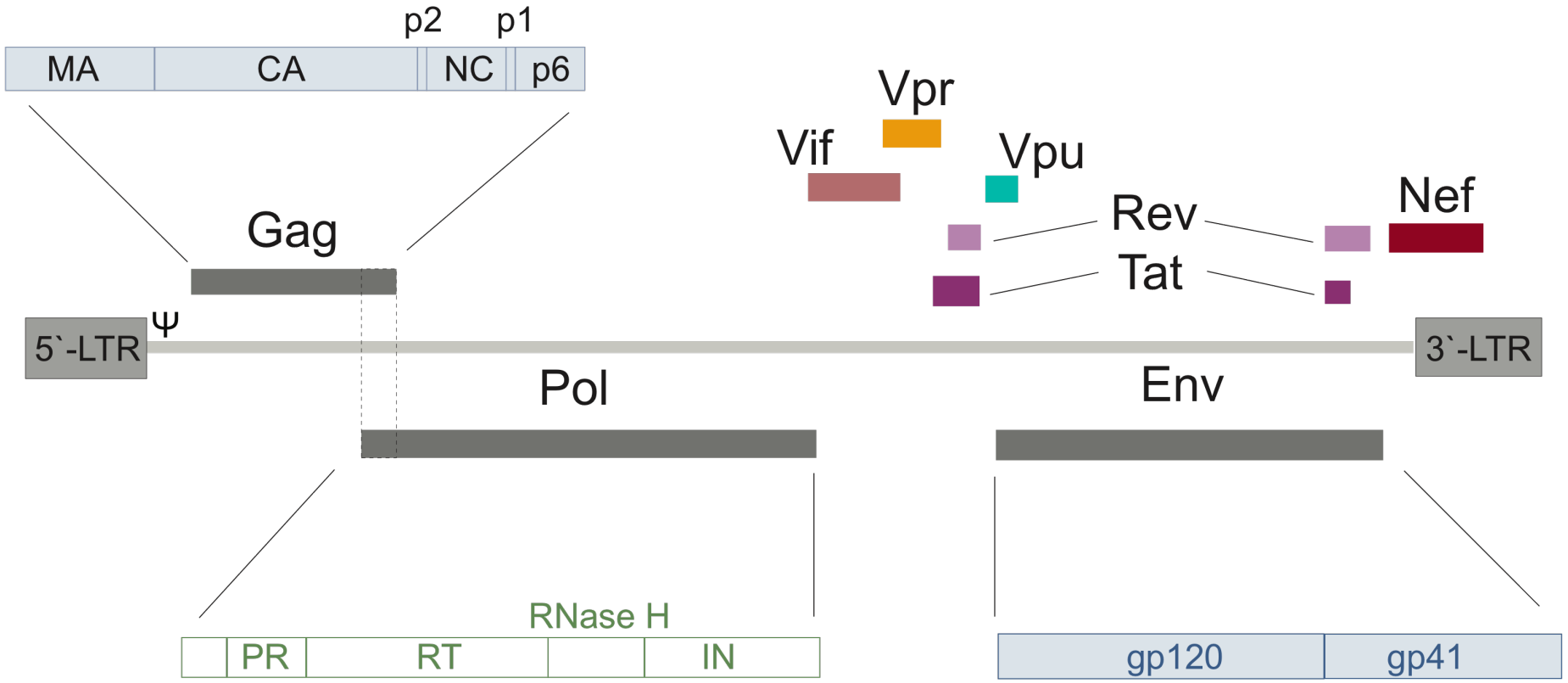
Conor Meehan
conor.meehan@dal.ca

What is HIV?

HIV Taxonomy

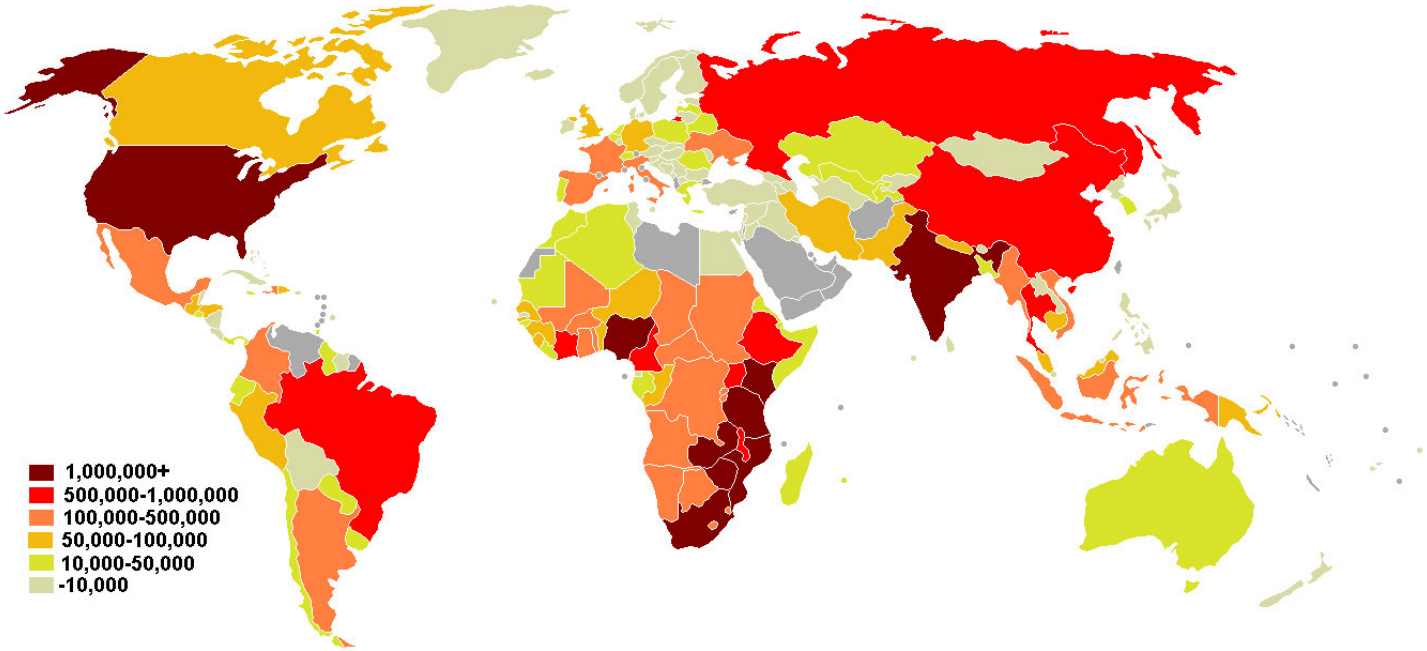
- Retro-transcribing viruses (Group VI)
  - Single stranded RNA with reverse transcriptase
- Family: Retroviridae
  - Uses reverse transcriptase to produce DNA from RNA
- Genus: Lentivirus
  - Long incubation time
- Subgenus: Primate lentivirus
  - Infect primates
- Species: Human Immunodeficiency Virus
  - Types: HIV-1 and HIV-2

HIV Genome


p2
p1
MA
CA
NC
p6
Vpr
Vif
Vpu
Rev
Nef
Tat
Gag
Ψ
5`-LTR
3`-LTR
Pol
Env
RNase H
PR
RT
IN
gp120
gp41

HIV Infection Statistics (UNAIDS 2011) [1]

- 34.2 million people infected worldwide
  - 23.25 million of these are in sub-Saharan Africa
  - 2,100 in Czech Republic
  - 2.7M were infected in 2010
- 1.8 million people died from AIDS
  - Half of these are in sub-Saharan Africa



Where did HIV come from?

Simian Immunodeficiency Virus

- Also called African Green Monkey Virus
- Primate lentivirus in monkeys and apes
- Estimated to be a primate virus for up to 32,000 years[2]
- Mainly non-lethal to host

SIV to Humans

- Most zoonotic transmission thought to have occurred in the 1930's[3]
  - From chimpanzees to humans (HIV-1) [4]
  - From sooty mangabeys to humans (HIV-2) [5]

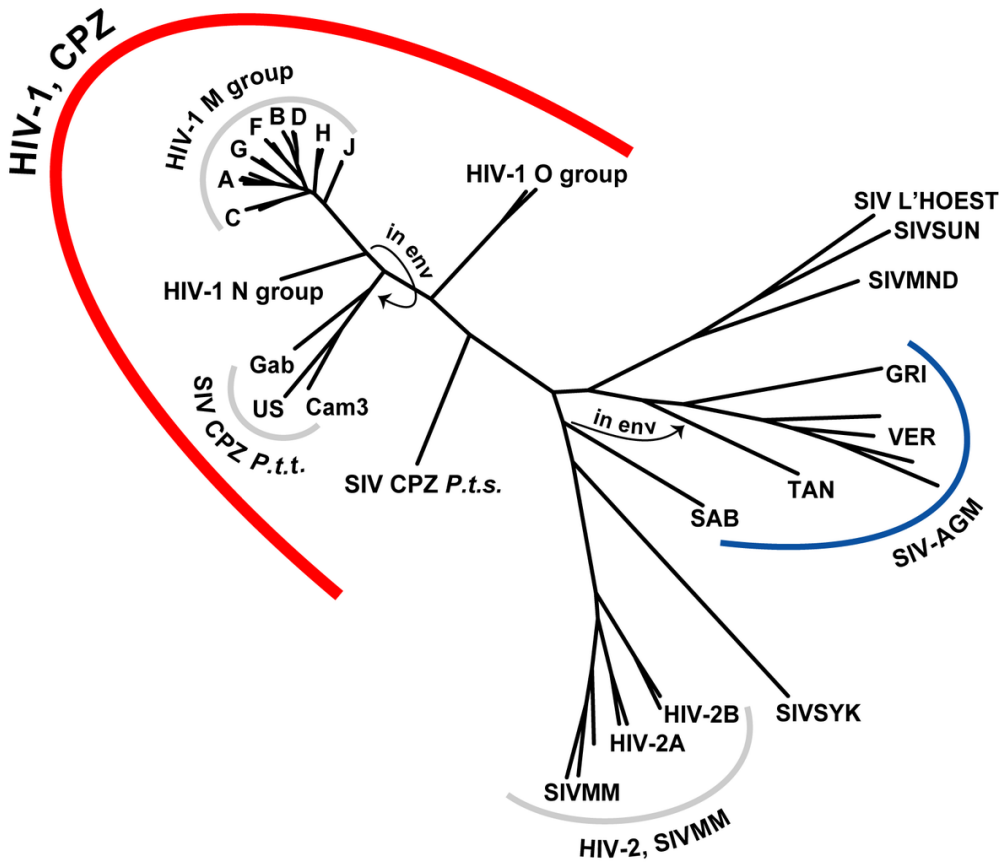


Image from ref 6

How did SIV get into Humans

- Transmission thought to be through cuts on hunters butchering bush meat[7]
- Likely multiple transmissions of SIV to human hosts
  - Not all successful
  - At least one transmission lead to each HIV-1 group and HIV-2 group
- Certain human factors may reduce susceptibility
  - Trim5α blocks retrovirus infection [8]
  - Certain SIV strains less resistant to human trim5α[8]
  - Likely these strains that were successful cross-species

How does HIV get from one host to another?

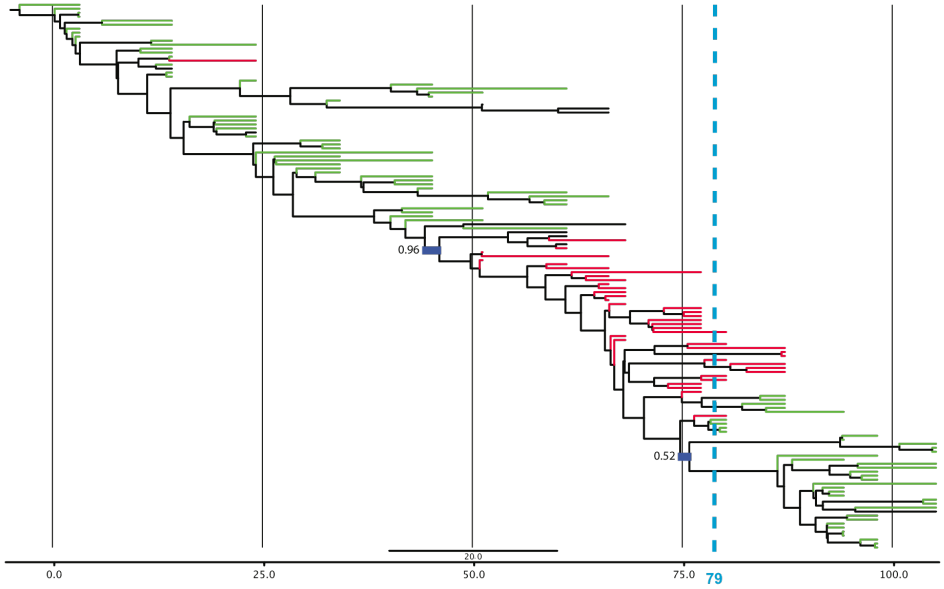
Quiz time!

Host Entry

- Although several virions (virus particles) may enter the body it is likely only 1 or a small few survive[9]
  - Genetic bottleneck
  - New diversity in each person
  - Not always the donors dominant strain that infects new host[10]
- Infects cells of the immune system
  - Macrophages
  - T-cells
- Binds first CD4 (primary receptor) [11] and then a chemokine receptor (coreceptor) [12,13]
  - CCR5 on macrophages
  - CXCR4 on T-cells

Coreceptor usage modulates within a host

- CCR5 usage almost always observed upon initial infection[14]
  - People who have a deletion in their CCR5 gene (termed CCR5Δ32) are resistant to infection
- CXCR4 usage evolves in ~50% of patient[15]
  - Separate evolution event in each host
  - Switch extremely rare in SIV
- Coreceptor switch is bi-directional[16]
  - CXCR4 usage may represent reduced fitness
  - Switch made to evade immune system



CCR5-using
CXCR4-using

How and why does HIV mutate?

Reverse Transcriptase

- DNA $\longrightarrow$ RNA $\longrightarrow$ Protein
- RNA $\longrightarrow$ DNA (via reverse transcriptase)

- Mainly found in reverse-transcribing viruses
  - Also in some eukaryotic retrotransposons and some bacterial cells

- Is a good drug target due to its absence from central human cell function

Errors in RT

- No proof-reading
  - DNA polymerase reverses direction by 1 base pair when error is detected
  - RT does not have this function
- This allows mutations to occur in the resulting DNA
  - More likely to occur in ‘hotspots’
  - Hotspots contain homopolymeric nucleotide runs
- ~3.4 x 10^{-5} number of mutations per site occurs within each replication cycle (~1/day) [17]
  - Human rate: ~2.5×10^{-8} per base per generation (~1/20 years) [18]
  - ~1.36 times higher in HIV with a generation time 7,300 times higher than humans

HIV Divergence

- HIV sequences can be very different from each other
- HIV nomenclature designed to represent genetic similarity and clustering patterns
- HIV-1 segregated into groups and subtypes[19]
  - 4 main groups (M, N, O, P)
    - separate SIV transmission events
  - Group M further split into subtypes: A, B, C, D, F, G, H, J, K
    - 1 SIV transmission, subsequent divergence in humans
  - A and F further split into sub-subtypes: A1, A2; F1, F2
  - Subtypes may combine and create recombinant forms
    - One host infected multiple times with different subtypes
    - Referred to as a super-infection

HIV Subtypes

- 3 gene trees generally agree on subtype relationships
- Segregation could be due to sampling bias

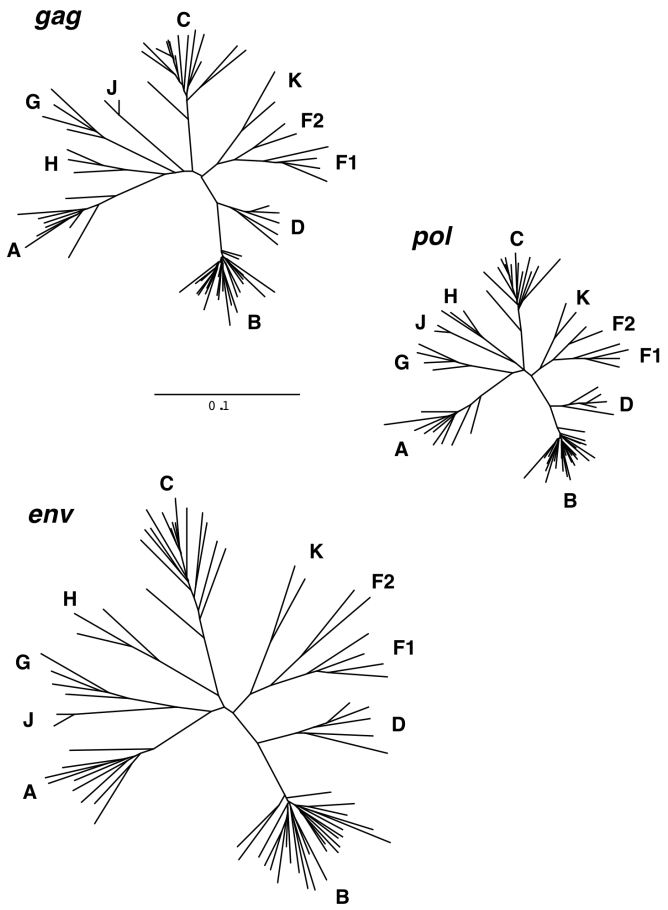
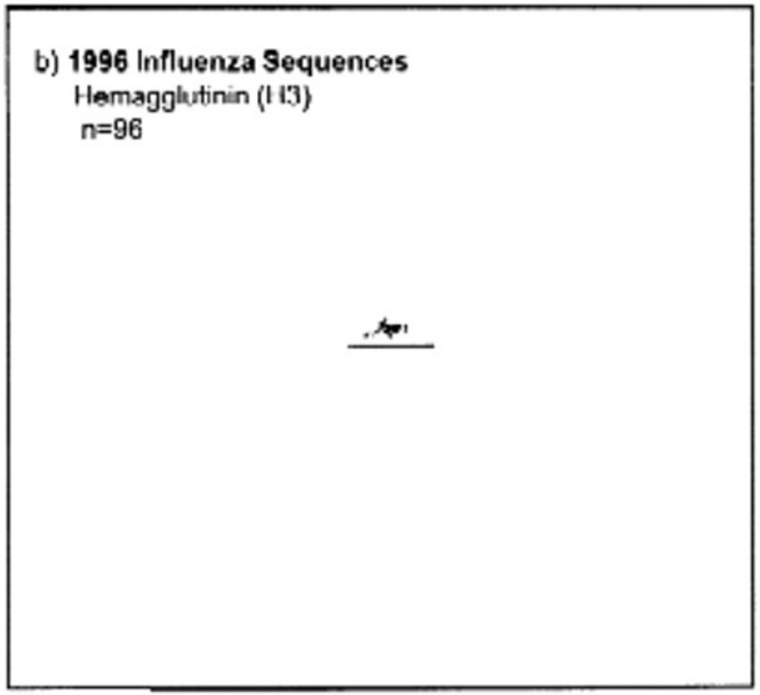
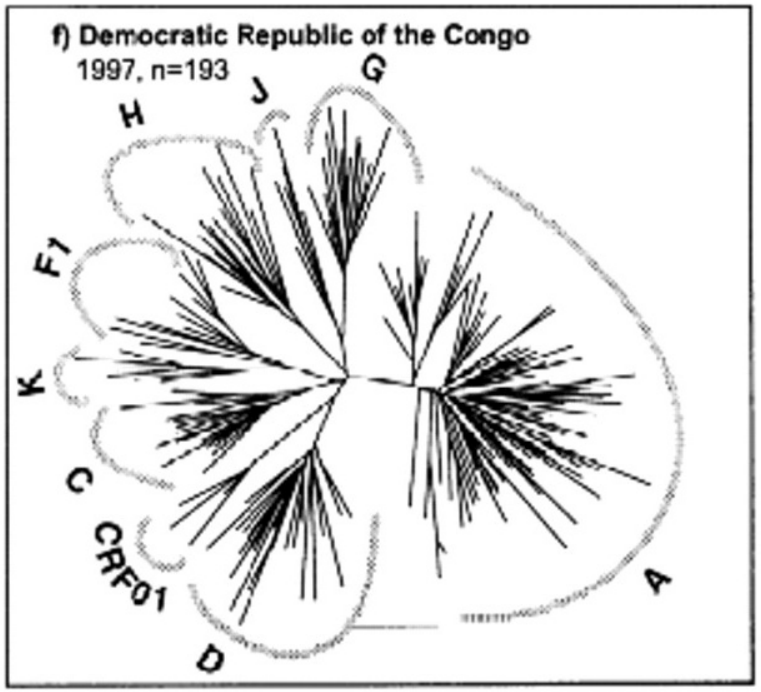


Image from ref 19

Difference in Influenza and HIV divergence

- Compare all Influenza sequences up until 1996 to all HIV sequences in one country in 1997
- Place trees on same scale[20]





What benefit does rapid mutation give?

- Changes in cell type infectivity
  - Switch coreceptor usage to infect new cells
  - Depletion of cells necessitate a change
- Host immune system evasion
  - Change epitopes
  - Glycan shield
- Drug resistance
  - Point mutations

How does HIV evade the immune system and drug treatments?

Epitopes

- An antigen recognised by the immune system
- Specific protein motifs
- Mutation to these motifs reduce recognition
- Modulate frequently during infection

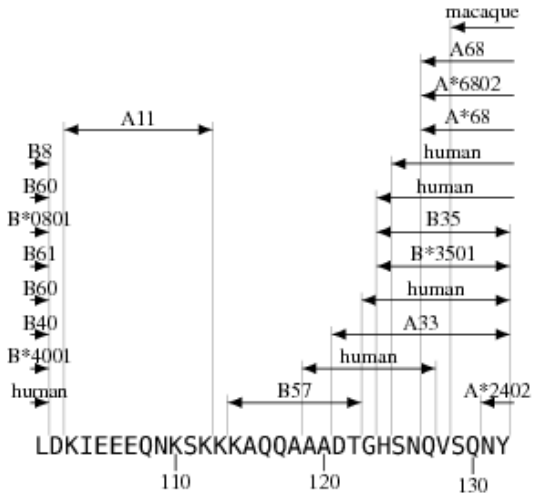


Image from ref 21

Glycan Shield

- Sugars from the host added to outside of HIV envelope[22]
- Inhibits recognition of virus as immune system sees virus as ‘self’
- Sugars also cover epitopes on envelope
- Some antibodies recognise specific glycosylation patterns
- Modulation of shield during infection to evade such antibodies

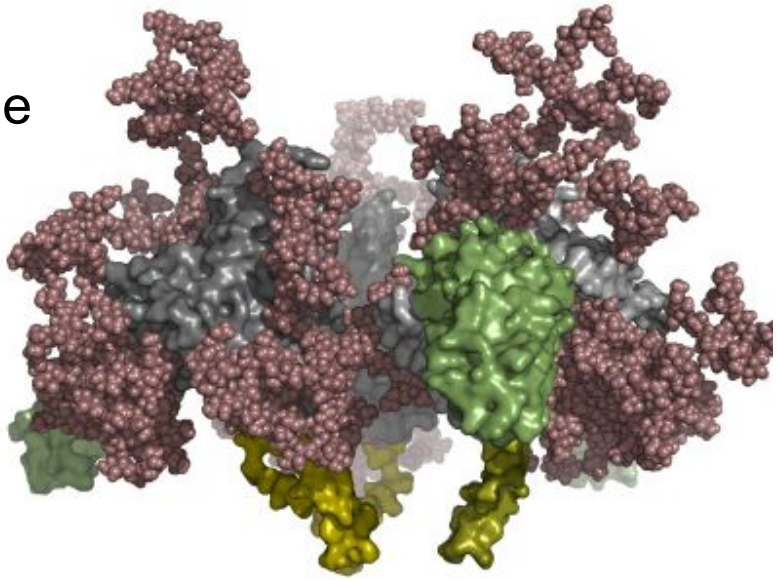


Image made with PyMOL

Drug Targets

- Most drugs target specific epitopes
- Specific point mutations reduce effectiveness of drugs[24]
- Point mutations can affect multiple drugs[24]
- Constant drug pressure selects for viruses with point mutations, despite likely reduced fitness
- Flexible proteins allow for such mutations to occur without affecting overall function

	NVP	DLV	EFV
Wild Type	0.7	0.6	0.6
103N	49	27	20
181C	123	33	1.2
103N,181C	400	225	29
100I,103N	69	190	400

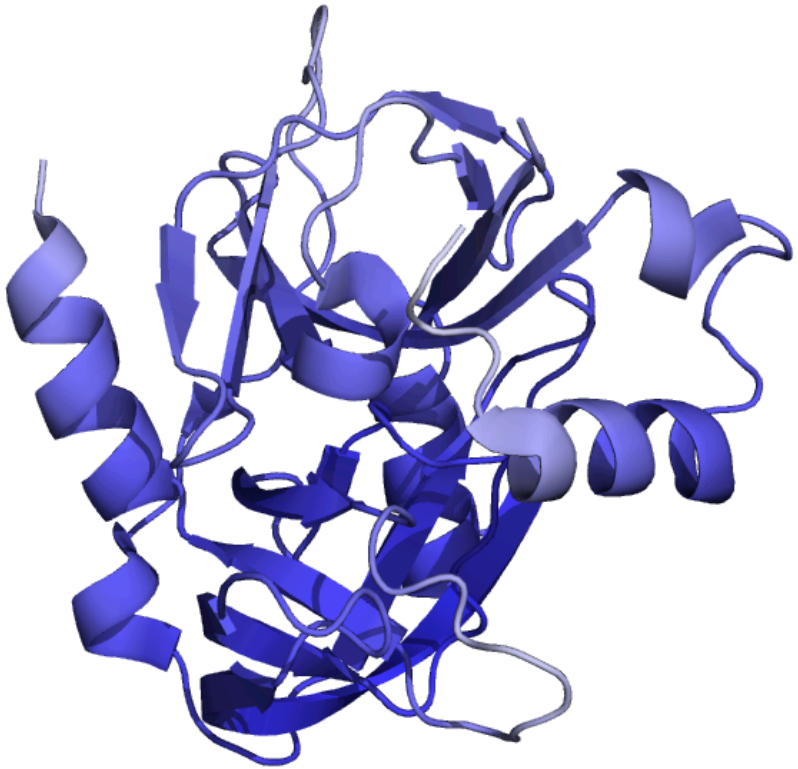
Data from ref 23

HIV protein structures

Viral Structural Flexibility[28]

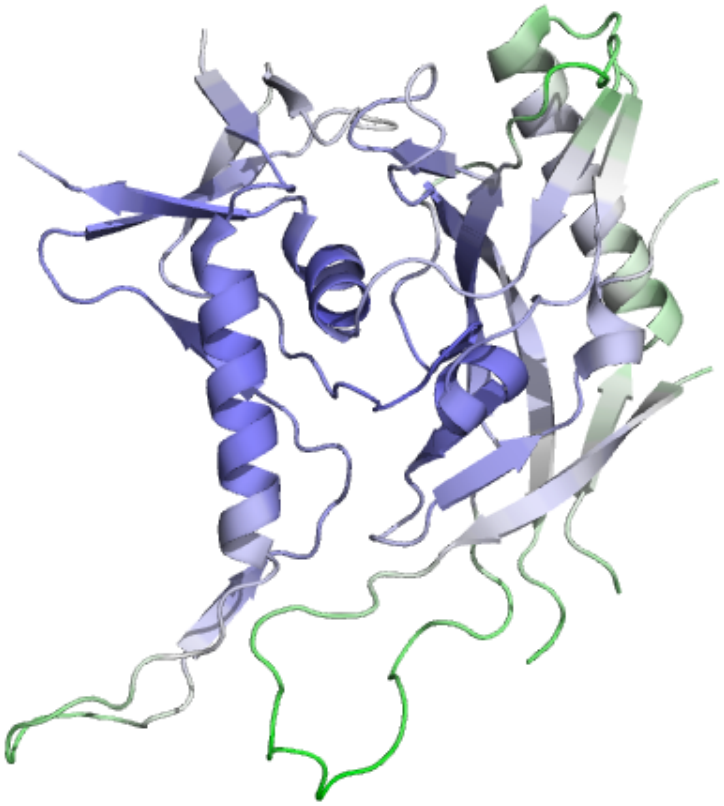
- Eukaryotic/Bacterial proteins usually densely packed
  - High number of interacting partners
  - Mostly contain secondary structures (alpha helices, beta sheets etc)
  - Evolved to allow for higher thermostability
    - Mutations buffered by overall stability
- Viral proteins are loosely packed
  - Less contacting partners
  - More disordered sections
  - Allows for high mutation rates
    - Do not destabilise protein as easily

GP120 Structural Flexibility



HtrA (Heat Shock Protein)
1L1J

B-factor range:
20-100



Gp120
3HI1

B-factor range:
51-237

B-factor colouring:
Blue: compact
White: intermediate
Green: flexible

HIV evolution

- Evolved from SIV
  - Gains entry to hosts through blood
- One of the most diverse organisms known
  - Mutation rate close to theoretical maximum
  - Overcome this by population level evolution
- Similar evolutionary paths taken within different hosts
- Utilises several strategies to evade immune system and drug treatments
  - Epitope modification
  - Glycan shield modularisation
- Highly flexible both at protein sequence and structure level
  - No error correction in RT
    - Large numbers of mutations
  - Structures loosely packed
    - Allow for large mutation rate without destabilising protein
- Overall virus has evolved to allow rapid evasion and adaptation
- All accomplished with a 9 gene genome

Convicting and Curing

The use of phylogenetics in the courts

Phylogenetics and criminal prosecution of HIV transmission

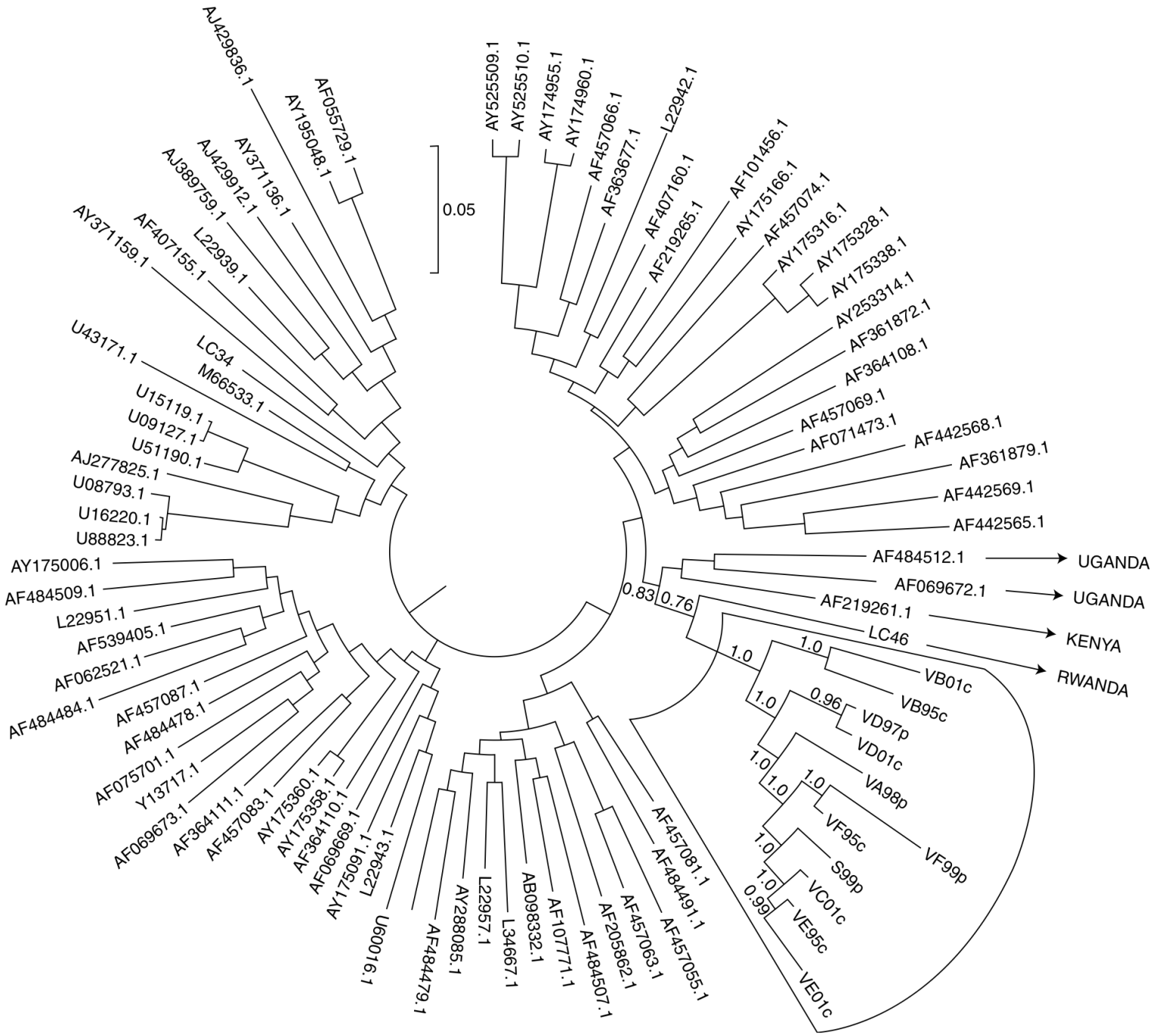
- Intentional or negligent transmission of HIV can result in charge of assault, manslaughter or murder in several countries
- Two things often must be proven for this:
  - The defendant was reckless
  - The defendant infected the complainant
- In the UK it was required that scientific evidence must be used to prove infection, even if a plea of ‘guilty’ was entered
  - Phylogenetics is often used in this step
- Phylogenetics is often required to prove recklessness too
  - Time of infection must be after the defendant became aware of their status and before the complainant became aware of the defendant’s status

Phylogenetics and criminal prosecution of HIV transmission

- First used in 1990 in a case of a dentist infecting several patients[30] though this case never went to court
- First used in a criminal case in Sweden in 1992, though directionality was not determined[31]
- In 2002 phylogenetic analysis was used to uphold a conviction during appeal by a gastroenterologist in the 2nd degree murder charge of his girlfriend after it had been found to meet standards of evidence admissibility[32]

An example

- Lemey *et al.* “Molecular testing of multiple HIV-1 transmissions in a criminal case”, AIDS 19(15), 2005[33]
- One suspect and six victims
- 2 samples from each person, anonymously labelled and sequenced for *pol* and *env* fragments
- 30 controls taken from local hospital fitting as closely to the age, risk and geographical parameters as the suspect/victims and from around the same time of alleged transmission as possible
- Phylogenetic trees built under ML using 3 methods and also using Bayesian inference
  - Sites known to infer drug resistance were excluded to prevent clustering based on drug regimes


AJ429836.1
AF055729.1
AY195048.1
AY371136.1
AJ429912.1
AJ389759.1
L22939.1
AF407155.1
AY371159.1
LC34
M66533.1
U43171.1
U15119.1
U09127.1
U51190.1
AJ277825.1
U08793.1
U16220.1
U88823.1
0.05
AY525509.1
AY525510.1
AY174955.1
AY174960.1
AF457066.1
AF363677.1
L22942.1
AF407160.1
AF219265.1
AF101456.1
AY175166.1
AF457074.1
AY175316.1
AY175328.1
AY175338.1
AY253314.1
AF361872.1
AF364108.1
AF457069.1
AF071473.1
AF442568.1
AF361879.1
AF442569.1
AF442565.1
AF484512.1
UGANDA
AF069672.1
UGANDA
AF219261.1
KENYA
LC46
RWANDA
0.83
0.76
1.0
1.0
VB01c
VB95c
1.0
0.96
VD97p
VD01c
1.0
1.0
VA98p
1.0
VF95c
VF99p
1.0
S99p
1.0
VC01c
0.99
VE95c
VE01c
AF457081.1
AF484491.1
AF457055.1
AF457063.1
AF205862.1
AF484507.1
AF107771.1
AB098332.1
L34667.1
L22957.1
AY288085.1
AF484479.1
U60019.1
L22943.1
AY175091.1
AF069669.1
AF364110.1
AY175358.1
AY175360.1
AF457083.1
AF364111.1
AF069673.1
Y13717.1
AF075701.1
AF484478.1
AF457087.1
AF484484.1
AF062521.1
AF539405.1
L22951.1
AF484509.1
AY175006.1

An example

- Lemey *et al* “Molecular testing of multiple HIV-1 transmissions in a criminal case”, AIDS 19(15), 2005[33]
- One suspect and six victims
- 2 samples from each person, anonymously labelled and sequenced for *pol* and *env* fragments
- 30 controls taken from local hospital fitting as closely to the age, risk and geographical parameters as the suspect/victims and from around the same time of alleged transmission as possible
- Phylogenetic trees built under ML using 3 methods and also using Bayesian inference
  - Sites known to infer drug resistance were excluded to prevent clustering based on drug regimes
- Demonstrated grouping of suspect and victim samples, monophyletic to the exclusion of controls
  - No inference was made about directionality (usually indicated by paraphyletic relationship of source sequences around recipient sequences)
  - Cannot rule out case of both suspect and victim infected by a 3rd person or suspect infecting a person who infected victims

Is there a cure?

The Berlin Patient

- A man, Timothy Brown from Seattle, Washington, USA was living in Berlin, Germany in the 1990's
- In 1995 he was diagnosed with HIV
- In 2006 he was diagnosed with myeloid leukaemia
- In 2007 he underwent a stem cell transplant to cure his leukaemia
- It was found post-transplant that he had no detectable HIV within his system

How was he cured?

- Mr. Browns oncologist Dr. Gero Hütter knew of the CCR5Δ32 resistance to HIV
- He found a donor who had this genetic deletion
- He tested the coreceptor usage of Mr. Browns HIV population and found it was a CCR5-using population
- Transplantation of the donor stem cells resulted in truncated CCR5 receptors in Mr. Brown
  - HIV could no longer complete cell entry
  - Viral strains died out and the patient was cured
- Reported in Hütter *et al*, N Engl J Med 2009[35], confirmed in Blood 2011 paper[36]

Can this approach be widely used?

- In short: No
- Each infected person would require a stem cell transplant
  - Requires irradiation of the body
  - Requires full replacement of immune system
  - Not recommended unless absolutely necessary
- The stem cell donor would have to have the CCR5Δ32 deletion
  - 5-15% prevalence in Caucasians[37]
  - <1% prevalence in Africans, Asians and South Americans[37]
- The HIV infected individual would have to have only CCR5-using strains
  - CXCR4-usage appears in at least 50% of cases
  - Detection is not perfect, minor population of CXCR4-using strains may be missed in screening

References

1 http://www.unaids.org/globalreport/global_report.htm
2 http://www.sciencemag.org/content/329/5998/1487.full
3 http://www.sciencemag.org/content/288/5472/1789.full
4 http://www.nature.com/nature/journal/v397/n6718/full/397436a0.html
5 http://www.nature.com/nature/journal/v339/n6223/abs/339389a0.html
6 http://en.wikipedia.org/wiki/HIV
7 http://www.nature.com/nrg/journal/v5/n1/full/nrg1246.html
8 http://www.sciencedirect.com/science/article/pii/S0966842X11000540
9 http://www.sciencemag.org/content/254/5034/963.abstract
10
11 http://www.nature.com/nature/journal/v312/n5996/abs/312763a0.html
12 http://www.nature.com/nature/journal/v381/n6584/abs/381667a0.html
13 http://www.sciencemag.org/content/272/5263/872.short
14 http://jvi.asm.org/content/67/6/3345.short
15 http://jvi.asm.org/content/73/12/10489.short
16 http://onlinelibrary.wiley.com/doi/10.1002/jmv.21922/full
17 http://jvi.asm.org/content/69/8/5087.abstract
18 http://www.genetics.org/content/156/1/297.full
19 http://www.hiv.lanl.gov/content/immunology/pdf/1999/4/nomenclature.pdf
20 http://bmb.oxfordjournals.org/content/58/1/19.full
21 http://www.hiv.lanl.gov/content/immunology/maps/ctl/p17.html
22 http://www.nature.com/nm/journal/v9/n4/full/nm0403-393.html
23 http://hivdb.stanford.edu/
24 http://nar.oxfordjournals.org/content/31/1/298.short
25 http://aac.asm.org/content/49/11/4721.short
26 http://jvi.asm.org/content/80/10/4909.short
27 http://jvi.asm.org/content/81/5/2359.short
28 http://www.sciencedirect.com/science/article/pii/S0968000408002533
29 http://onlinelibrary.wiley.com/doi/10.1111/j.1468-1293.2007.00486.x/full
30 http://www.cdc.gov/MMWR/preview/mmwrhtml/00001679.htm
31 http://jvi.asm.org/content/68/9/5918.long
32 http://www.pnas.org/content/99/22/14292.abstract
33 http://journals.lww.com/aidsonline/Fulltext/2005/10140/Molecular_testing_of_multiple_HIV_1_transmissions.12.aspx
34 http://www.jiasociety.org/index.php/jias/article/view/17407
35 http://www.nejm.org/doi/full/10.1056/NEJMoa0802905
36 http://bloodjournal.hematologylibrary.org/content/117/10/2791.short
37 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1377146/

Will there ever be a vaccine?

Challenges to vaccine production

- Reviewed in 34
- Most vaccines work by eliciting a small set of neutralising antibodies against a few viral surface proteins
- The high HIV mutation rate creates constantly changing epitopes
  - Drug targets are not present in every strain
- The likely candidate for antibody binding is gp120
  - Structural flexibility of gp120 makes conserved epitopes to bind difficult to find
  - Conformational shifts during host receptor binding hide any such conserved epitopes from host immune system
  - The glycan shield further block any epitopes from being bound
- Post transmission, spread through the body is fast
  - HIV virions may remain dormant in several cell types
  - Such dormant particles create reservoirs for resurgence of infection
  - HIV can persist in several compartments of the body
  - Targeting all compartments with a vaccine can be difficult

Advances in vaccine production

- Early trials using recombinant gp120 as a vaccine showed no marked improvements in resistance to infection
- Later trials used a Gag/Pol/Nef combination to test for T-cell effectiveness in infection control
  - Found to be ineffective with possible increased susceptibility in some groups
- The RV144 trial found that a live vector coupled with gp120 showed some protection (31%) though for a limited period of time
  - Was dependant on eliciting an IgG, not an IgA based response
- The SAV CT 01 trial (phase I) used a GM killed whole virus and found increased HIV antibody production. Made by Sumagen Canada.
- One broadly neutralising antibody VRC01 has been found to be effective against 90% of circulating strains
  - Binds the CD4 contact site of gp120
  - Only produced years post-HIV infection
  - A vaccine would need to somehow elicit the production of these antibodies pre-infection