Microsoft Research

Faculty Summit

2014 15TH ANNUAL

Microsoft Research
Faculty
Summit
2014 15TH ANNUAL

Selection at the HIV Transmission Bottleneck

Jonathan Carlson, Microsoft Research

Microsoft

HIV research at MSR

People

David Heckerman

Nebojsa Jojic

Andrew Philips

Neil Daschau

Carl Kadie

Jennifer Listgarten

Nicolo Fusi

Research

Applying statistics and computation to better understand disease

10 yrs, >75 papers published in the HIV field

RESEARCH ARTICLE

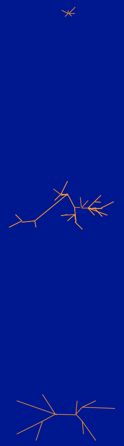
HIV TRANSMISSION

Selection bias at the heterosexual HIV-1 transmission bottleneck

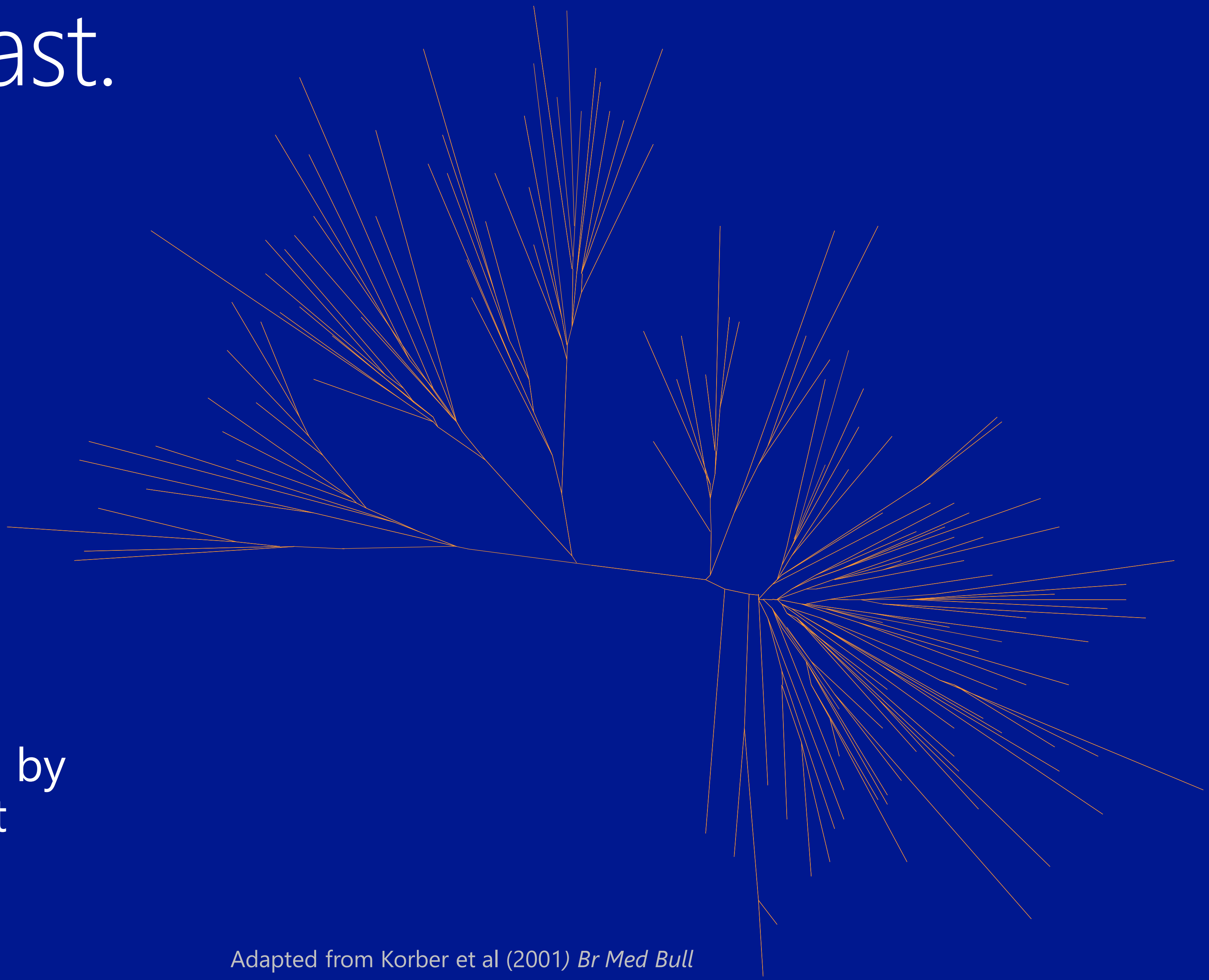
Jonathan M. Carlson,[1*†] Malinda Schaefer,[2†] Daniela C. Monaco,[2] Rebecca Batorsky,[3] Daniel T. Claiborne,[2] Jessica Prince,[2] Martin J. Deymier,[2] Zachary S. Ende,[2] Nichole R. Klatt,[2‡] Charles E. DeZiel,[1] Tien-Ho Lin,[1§] Jian Peng,[1¶] Aaron M. Seese,[3] Roger Shapiro,[4] John Frater,[5,6,7] Thumbi Ndung'u,[3,8,9,10] Jianming Tang,[11] Paul Goepfert,[11] Jill Gilmour,[12,13] Matt A. Price,[14,15] William Kilembe,[16] David Heckerman,[17] Philip J. R. Goulder,[8,18] Todd M. Allen,[3] Susan Allen,[16,19,20] Eric Hunter[2,16,19*]

July 11, 2014

HIV Mutates. Fast.

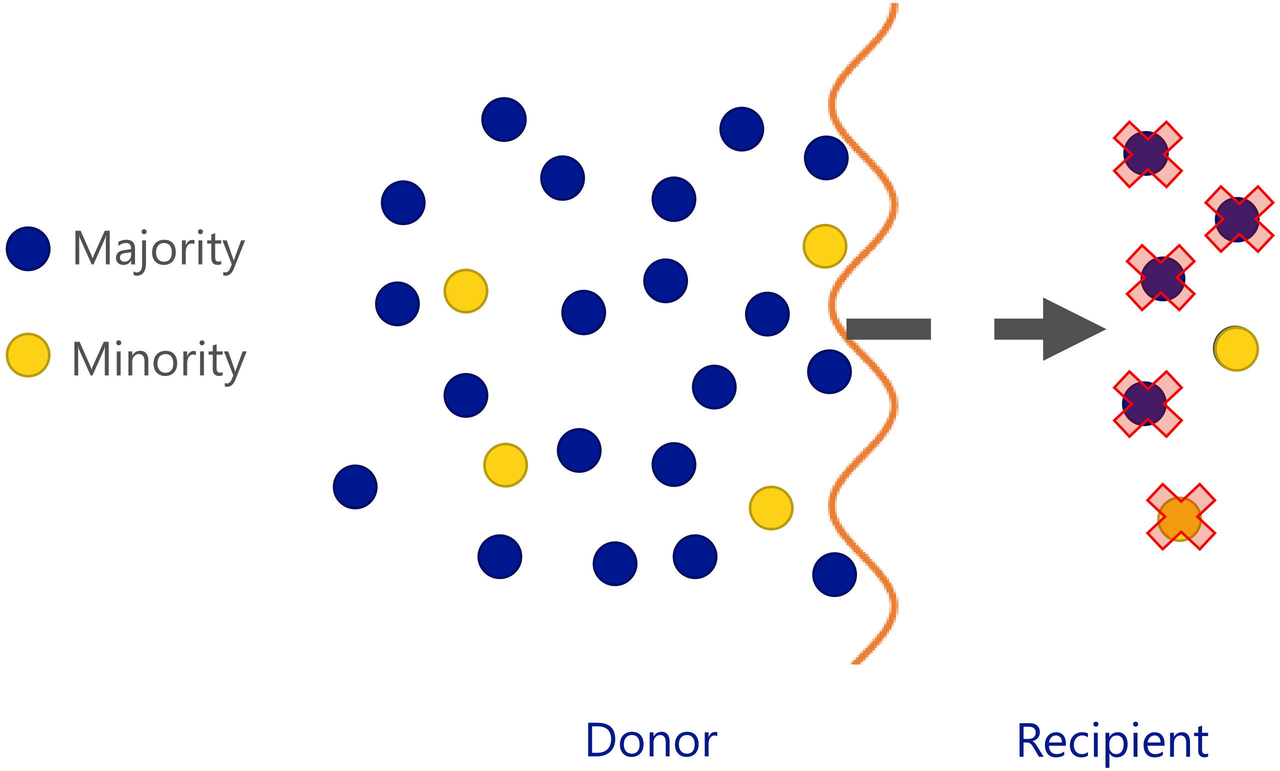


Infection is established by a single genetic variant



Adapted from Korber et al (2001) *Br Med Bull*

The transmission bottleneck



Transmission is inefficient

1/300 exposures result in infection

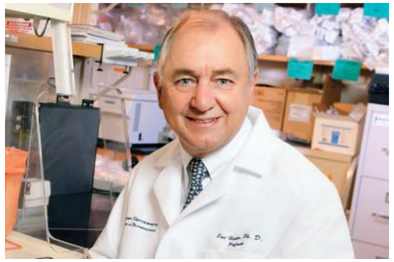
Women are at 2x risk relative to men

Men with STI's are at 2x risk

Hypothesis:

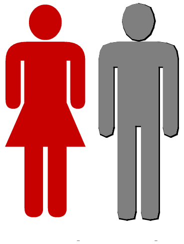
Could it be that many viruses make it across, but the replication rate is terminally low?

Susan Allen



Eric Hunter

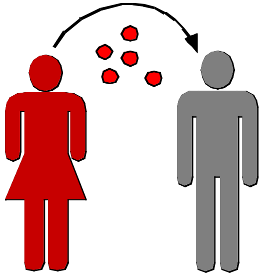
N>1,000



Couple is identified as HIV serodiscordant

Couples counseling and condoms reduce transmissions by 2/3

HIV negative partner is tested once per month



N=137

HIV negative partner seroconverts

Plasma collected from Donor and Recipient, median 45 days post estimated infection

HIV sequencing of both partners

Estimating fitness with statistical power

99.8% identity

	Donor	Recipient
ID1	...EPRGSDIAATTSNLQEQIGWMTSNPPIPV...	...EPRGSDIAGTTSNLQEQIGWMTSNPPIPV...
ID2	...DPRGSDIAGTTSNLQEQIGWMTNNPPIPV...	...DPRGSDIAGTTSTLQEQIGWMTNNPPIPV...
ID3	...EPRLSDIAGTTSNLQEQIGWMTNNPPIPV...	...EPRGSDIAGTTSNLQEQIGWMTNNPPIPV...
ID4	...EPRGSEIAGTTSTLQEQIAWMTNNPPIPV...	...EPRGSEIAGTTSTLQEQIGWMTNNPPIPV...
ID5	...EPRLSDIAGTTSNLQEQIGWMTNNPPIPV...	...EPRLSDIAGTTSNLQEQIGWMTNNPPIPV...
ID6	...EPRGSDIAGTTSTLQDQIGWMTNNPPIPV...	...EPRGSDIAGTTSTLQEQIGWMTNNPPIPV...
ID7	...DPRGSDIAGTTSNLQEQIAWMTHNPPVPV...	...DPRGSDIAGTTSNLQEQIAWMTNNPPVPV...
ID8	...EPRLSEIAGTTSTLQEQITWMTNNPPIPV...	...EPRLSEIAGTTSTLQEQITWMTNNPPIPV...

137

1712

Estimating fitness with statistical power

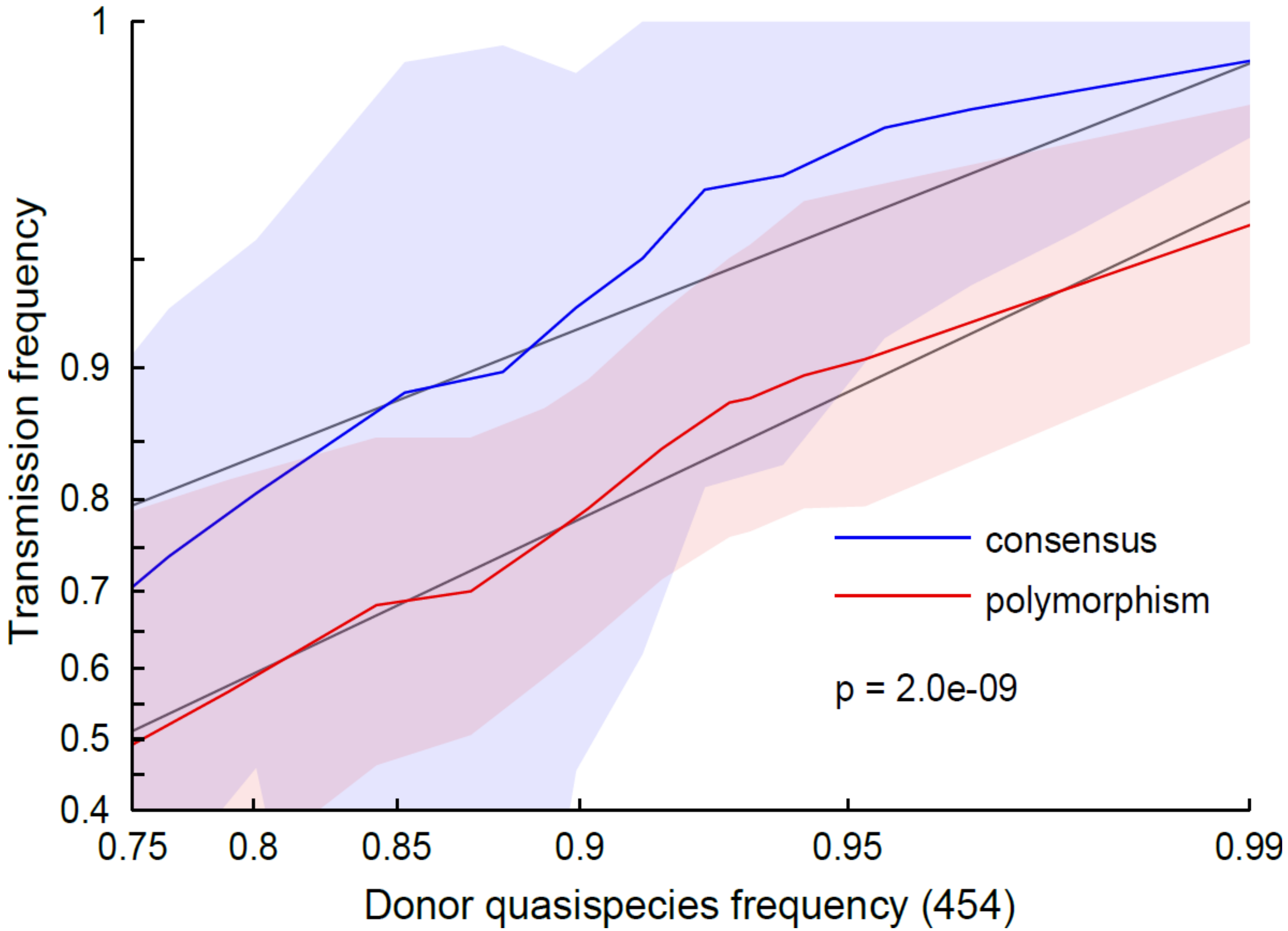
	D	R	T
ID1, pos1	E	E	1
ID1, pos2	P	P	1
…	…	…	…
ID2, pos11	N	T	0
ID2, pos12	S	S	1
ID2, pos13	T	T	1
…	…	…	…
ID8, pos1711	P	P	1
ID8, pos1712	V	V	1

N=228,362

Approach

$$\frac{\Pr(T=1)}{\Pr(T=0)} \approx \frac{f_{\mathrm{AA}}}{1-f_{\mathrm{AA}}} \times \frac{p_{\mathrm{AA}}}{p_{\overline{\mathrm{AA}}}}$$

$$\mathrm{logodds}(T=1) \approx \mathrm{logodds}(f_{\mathrm{AA}}) + \mathrm{bias}_{\mathrm{AA}}$$



Consensus $\coloneqq$ observed in $\geq 50\%$ of Zambian HIV+ individuals

Estimating fitness with statistical power

	D	R	T
ID1, pos1	E	E	1
ID1, pos2	P	P	1
…	…	…	…
ID2, pos11	N	T	0
ID2, pos12	S	S	1
ID2, pos13	T	T	1
…	…	…	…
ID8, pos1711	P	P	1
ID8, pos1712	V	V	1

N=228,362

Approach

$$\frac{\Pr(T=1)}{\Pr(T=0)} \approx \frac{f_{\mathrm{AA}}}{1-f_{\mathrm{AA}}} \times \frac{p_{\mathrm{AA}}}{p_{\overline{\mathrm{AA}}}}$$

$$\mathrm{logodds}(T=1) \approx \mathrm{logodds}(f_{\mathrm{AA}}) + \mathrm{bias}_{\mathrm{AA}}$$

Estimate probability of transmission
using a generalized linear mixed model

A fitness bias

AA features related to fitness

Conservation
Predicted impact on structure
Interaction network properties
Relationship to immune escape

A fitness bias

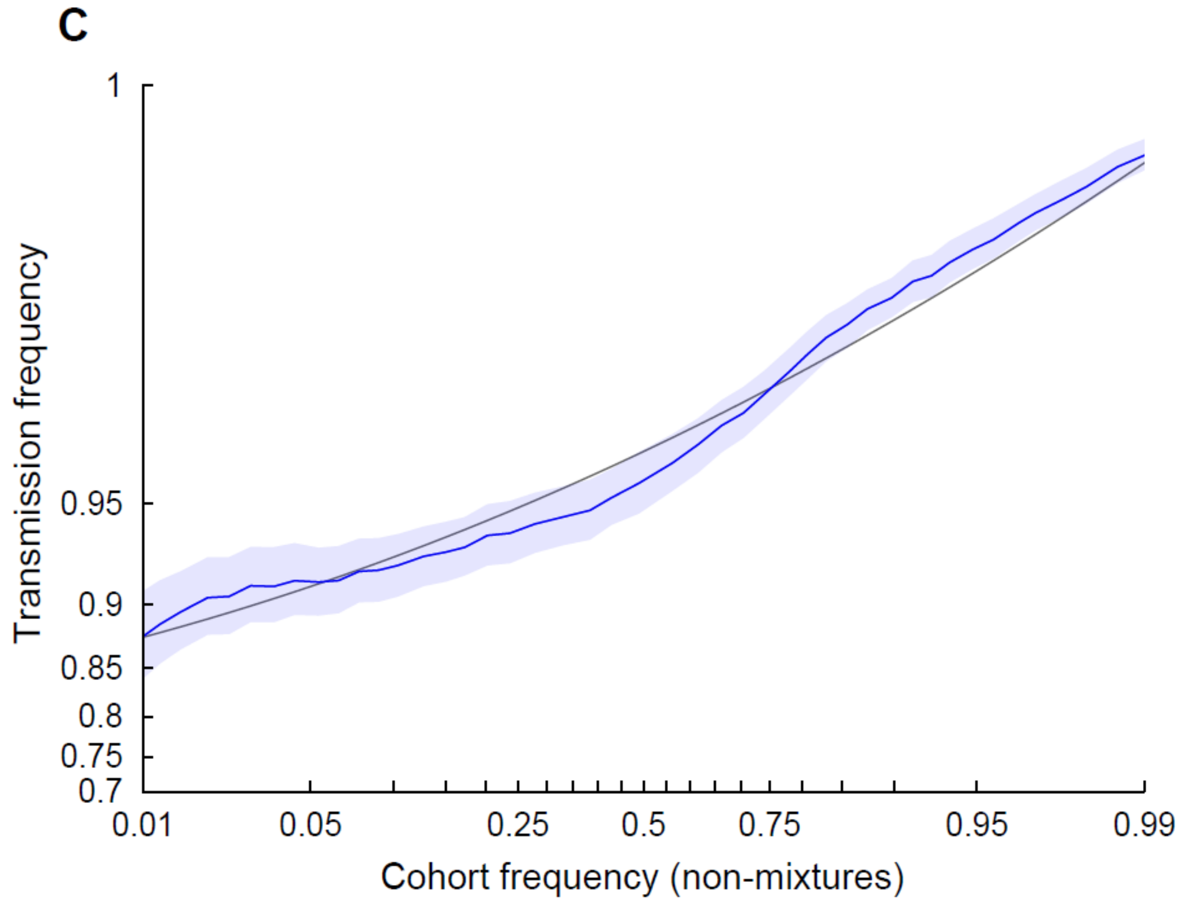
AA features related to fitness

Conservation

Predicted impact on structure

Interaction network properties

Relationship to immune escape



A fitness bias

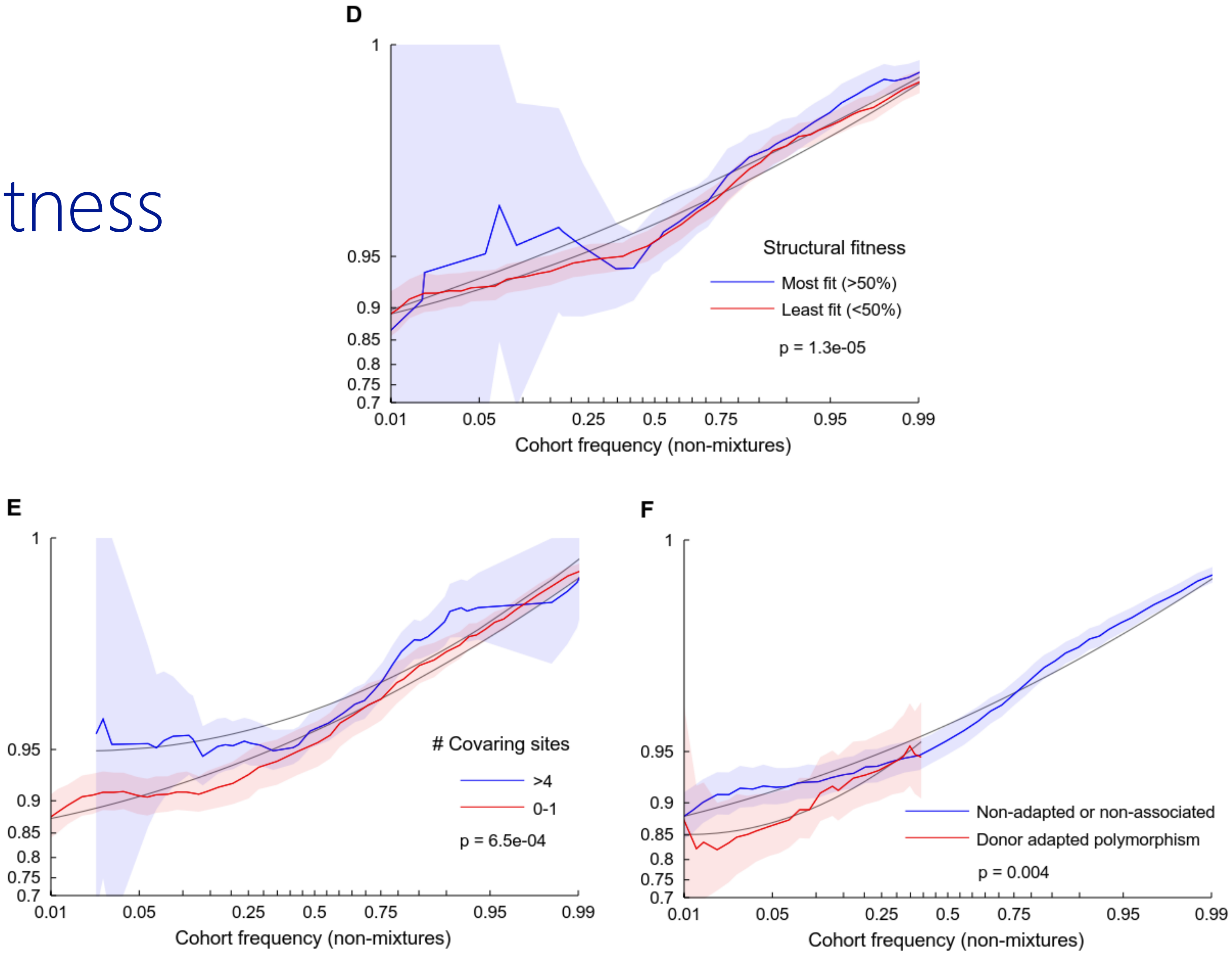
AA features related to fitness

Conservation

Predicted impact on structure

Interaction network properties

Relationship to immune escape

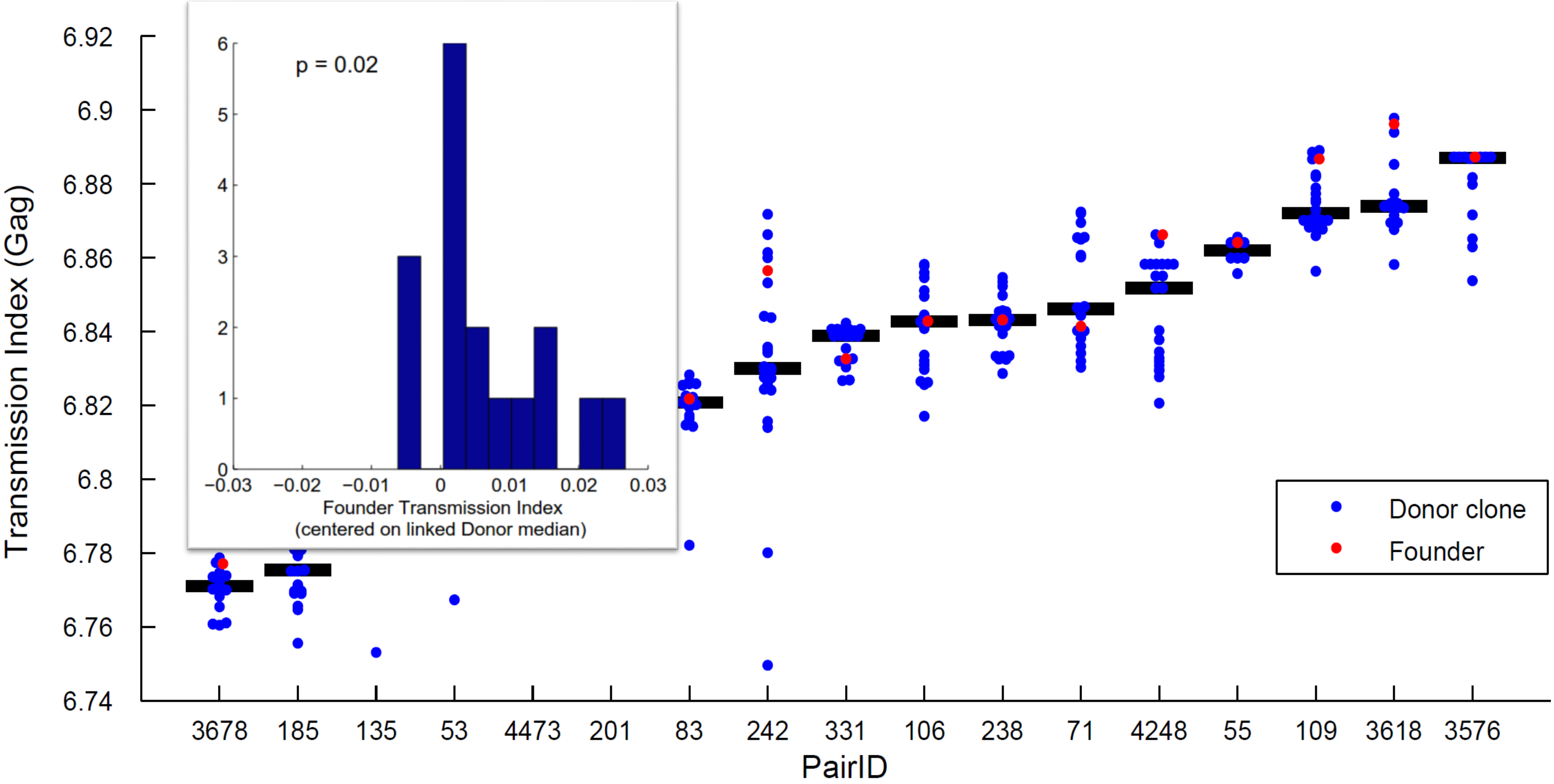


High-fitness amino acids are more likely to be transmitted

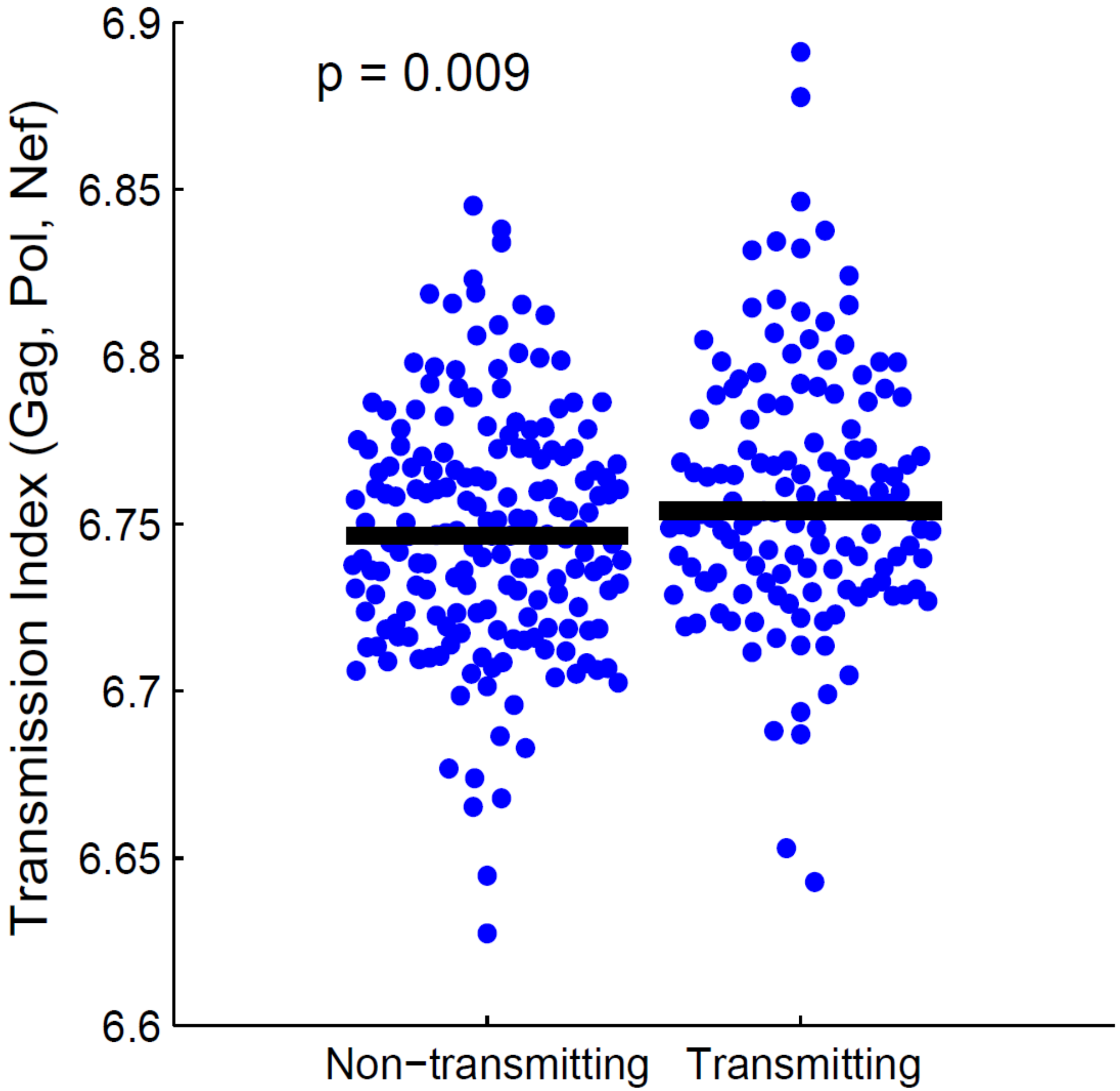
What about whole virus sequences?

Transmission Index

	D	R	T	Ln-odds (T)	Transmission Index
ID1, pos1	E	E	1	6.9	7.1
ID1, pos2	P	P	1	7.2	
…	…	…	…	…	…
ID2, pos11	N	T	0	4.6	6.3
ID2, pos12	S	S	1	5.3	
ID2, pos13	T	T	1	6.8	
…	…	…	…	…	…
ID8, pos1711	P	P	1	8.3	6.8
ID8, pos1712	V	V	1	6.3	


Transmission Index (Gag)
6.92
6.9
6.88
6.86
6.84
6.82
6.8
6.78
6.76
6.74
PairID
3678
185
135
53
4473
201
83
242
331
106
238
71
4248
55
109
3618
3576
Donor clone
Founder
p = 0.02
6
5
4
3
2
1
0
−0.03
−0.02
−0.01
0
0.01
0.02
0.03
Founder Transmission Index
(centered on linked Donor median)

Individuals with weak viral populations are less likely to transmit to their partners



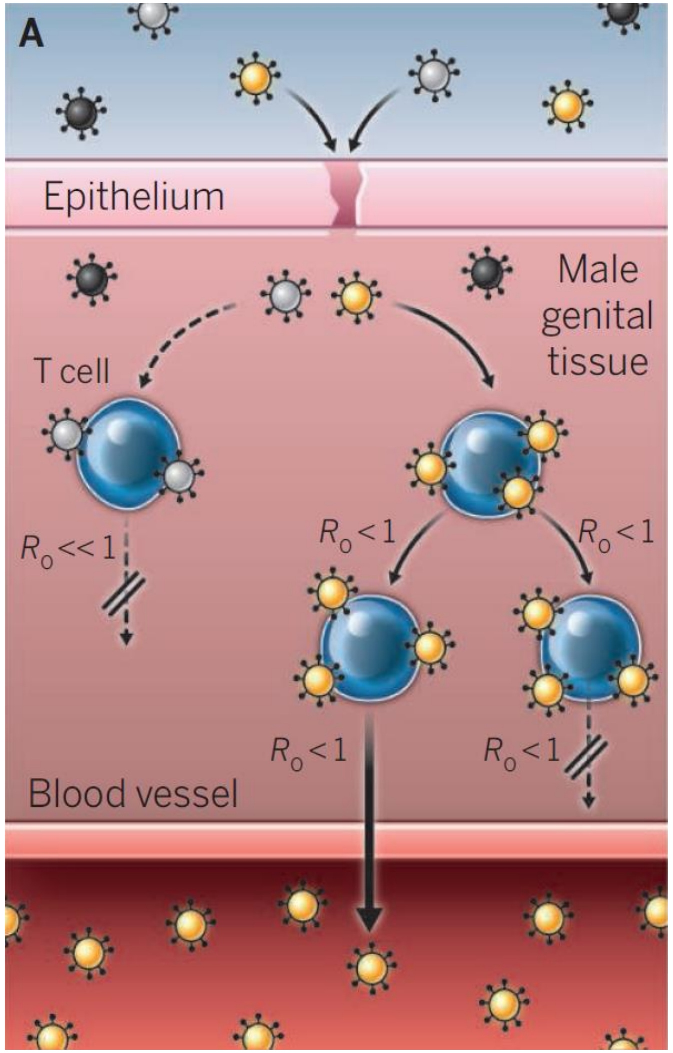
A new model of transmission

Frequent non-productive infection

Provides a window of opportunity for drugs and vaccines

What about risk factors?

$$\frac{\Pr(T=1)}{\Pr(T=0)} \approx \frac{f_{\mathrm{AA}}}{1-f_{\mathrm{AA}}} \times \frac{p_{\mathrm{AA}}+c}{p_{\overline{\mathrm{AA}}}+c} \xrightarrow[c \to \infty]{} \frac{\Pr(T=1)}{\Pr(T=0)} \approx \frac{f_{\mathrm{AA}}}{1-f_{\mathrm{AA}}}$$



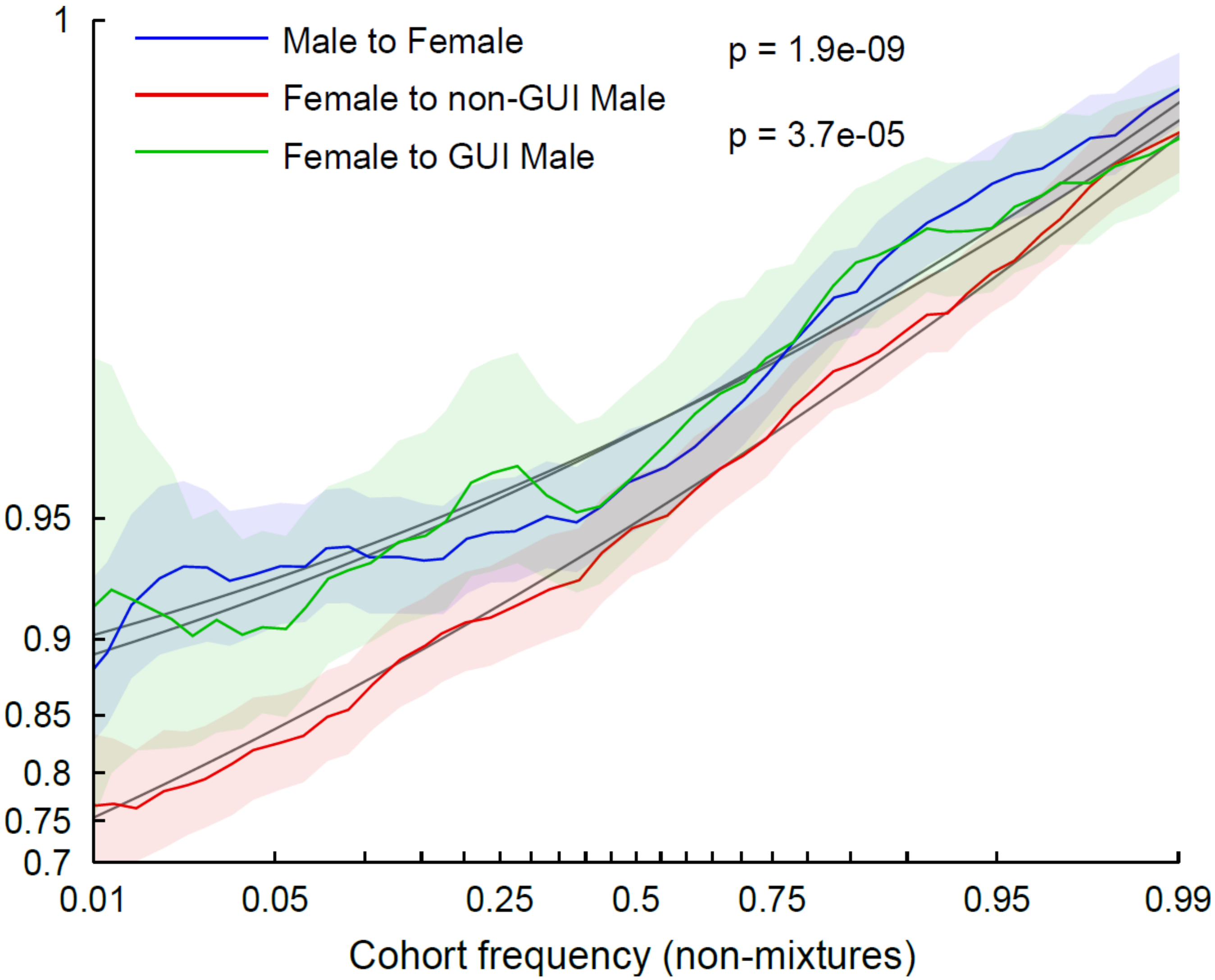
Joseph & Swanstrom, *Science* 2014

Men vs Women

Stronger bias in men

Less likely to be infected

More likely to be infected with a strong virus



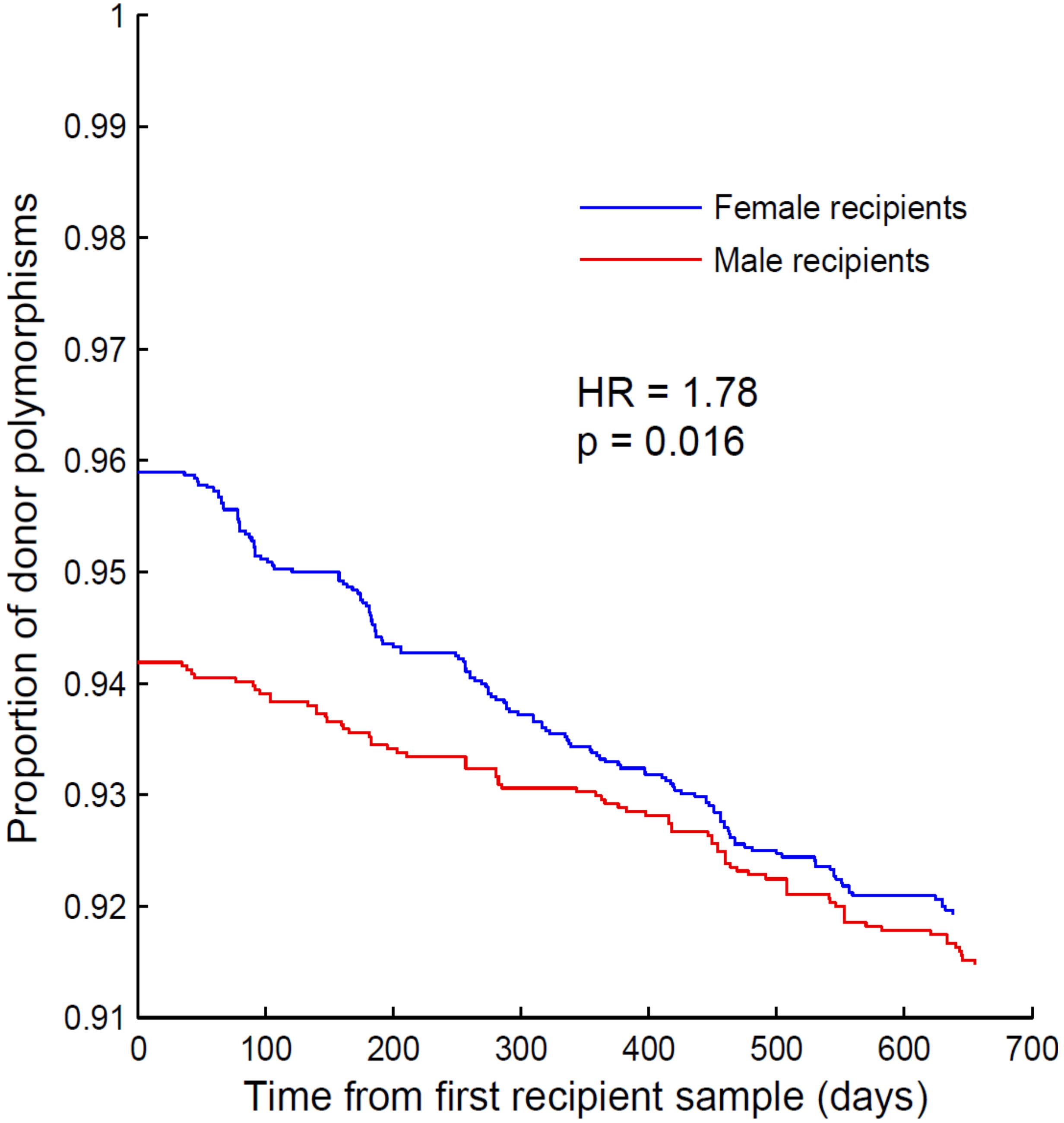
Men vs Women

Stronger bias in men

Less likely to be infected

More likely to be infected with a strong virus

More evolutionary pressure on women’s virus



Men vs Women

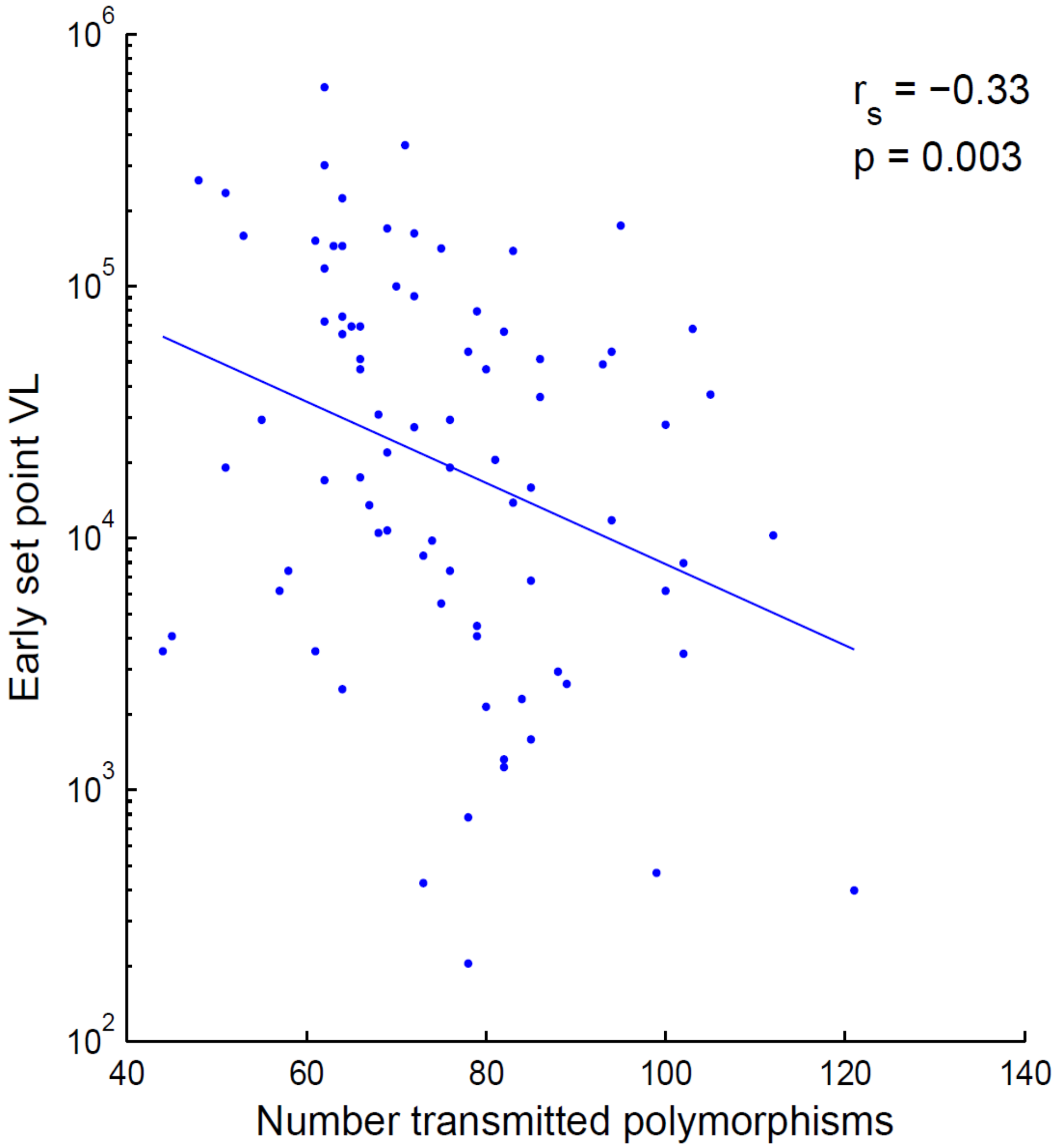
Stronger bias in men

Less likely to be infected

More likely to be infected with a strong virus

More evolutionary pressure on women's virus

Stronger transmitted viruses lead to more severe disease



A paradigm:

Low biological risk leads to lower odds of infection,

but more severe disease when infection happens

Summary

Fitness bottleneck at transmission

Features consistent with viral fitness predict transmission

Viral sequences and populations with high predicted fitness are more likely to establish infection

Risk factors mitigate the fitness bottleneck

A mechanism for increased risk is a reduction in the fitness bottleneck

Possibly due to more target cells, or higher activation state of target cells

Clinical consequences

Anything that weakens the virus will reduce transmission rates

Vaccines and drugs that protect individuals from transmission may lead to more severe disease when breakthrough infection occurs

Acknowledgments

Eric Hunter Lab, Emory

Malinda Schaefer, Daniela Mónaco

Dan Claiborne, Jessica Prince

Martin Deymier, Zachary Ende, Nichole Klatt

ZEHRP

Susan Allen

William Kilembe

Mubiana Inambao

Staff & study participants

Microsoft Research

David Heckerman

Henry Lin

Jian Peng

Charlie DeZiel

Others

Todd Allen, Rebecca Batorsky & Aaron Seese, Ragon

Jill Gilmour & Matt Price, IAVI

Philip Goulder & John Frater, Oxford

Thumbi Ndungu, U KwaZulu Natal

Roger Shapiro, Beth Israel Deaconess MC

James Tang & Paul Goepfert, UAB

Save the planet and return
your name badge before you
leave (on Tuesday)

Microsoft