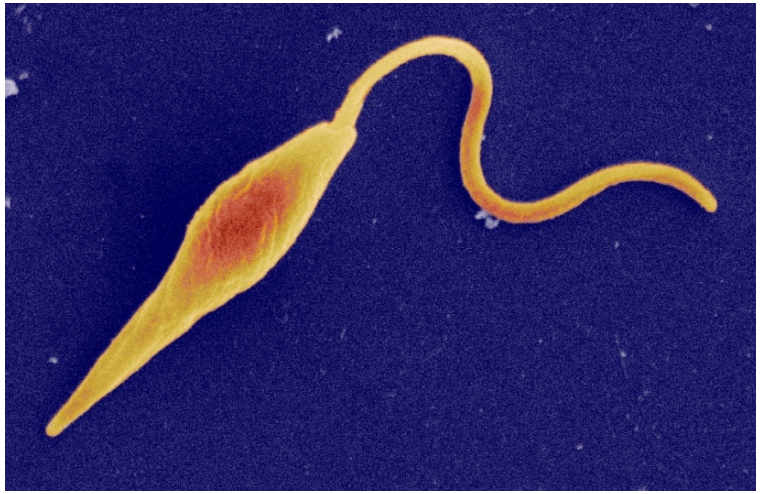
Extensive and adaptive genome structure variation drives the evolution and epidemiology of visceral leishmaniasis in the Indian subcontinent

2nd September 2014

Tim Downing



Main acknowledgements

Wellcome Trust Sanger Institute
Matthew Berriman
James Cotton
Mandy Sanders
Matthew Rogers

ITM Antwerp
Jean-Claude Dujardin
Hideo Imamura
An Mannaert
Saskia Decuypere
Manu Vanaeroschot
Simonne de Doncker

Strathclyde University
Katherine Carter
Julien Longchamp
Craig Shaw

Charité University Berlin
Olivia Stark
Gabriele Schönian

Syamal Roy
Suman Rijal
Poonam Salotra
Shyam Sundar
Debbie Smith
Petr Volf

www.leishrisk.net/kaladrug
EC-FP7-222895

WT 085775/Z/08/Z
and 076355

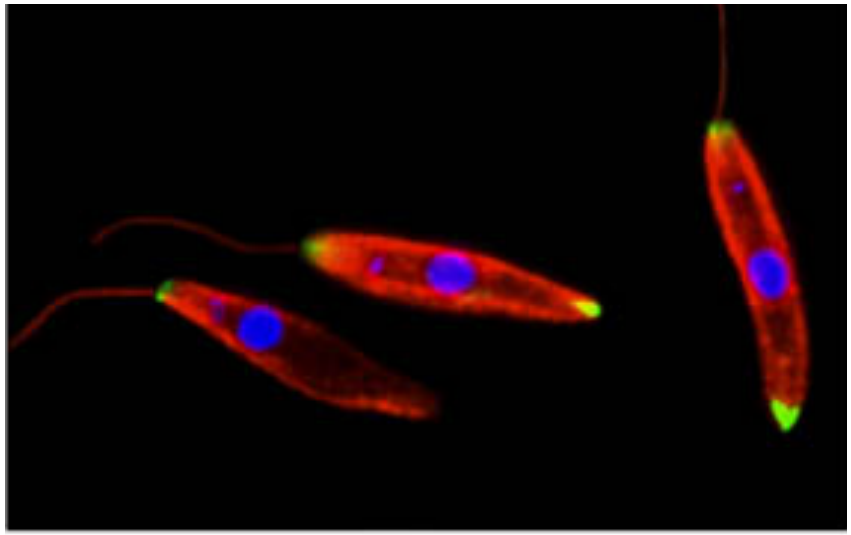
Introduction

1. The problem: Epidemiology of leishmaniasis
2. The context: *Leishmania* reproduction and biology
3. Genetics of *Leishmania donovani*
4. Population structure of leishmaniasis in India/Nepal
5. Genome plasticity and drug tolerance
6. Experimental evolution of resistance & mechanisms

Leishmania

Single celled

Promastigote in vector
Amastigote in host



Cells of the *Leishmania* parasite (*Leishmania tropica*). DNA is stained blue and two structural proteins, tubulin and vimentin, are stained red and green, respectively.

Vector = *Phlebotominus argentipes* (sandfly)

Leishmania donovani = Anthroponotic

Spleen aspirate microscopy

1. *Leishmania* epidemiology

350 million people at risk
12 million infections (Alvar et al 2012 Plos One)

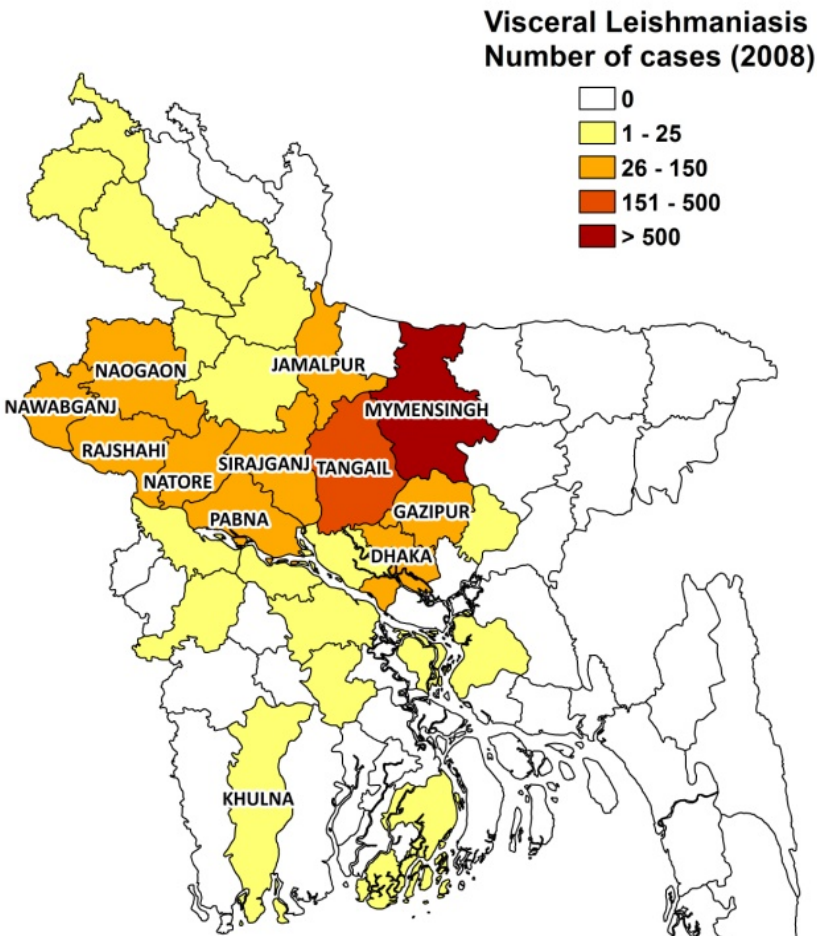
Cutaneous (CL) and visceral (VL) leishmaniasis

Each year: 200-400k new VL, 700-1,200k new CL

98 countries
90% VL: India/Bangladesh; Sudan/Ethiopia; Brazil
90% = asymptomatic (India/Nepal, Ostyn et al 2011)

Leishmania in Bangladesh

VL in 1824 in Jessore (Bern and Chowdhury 2006)
75k dead in 1824-7



1 million cases reported in Bengal in 1931-43

1968-80: 59 cases (Rahman & Islam 1983)
VL gone in 1970s?

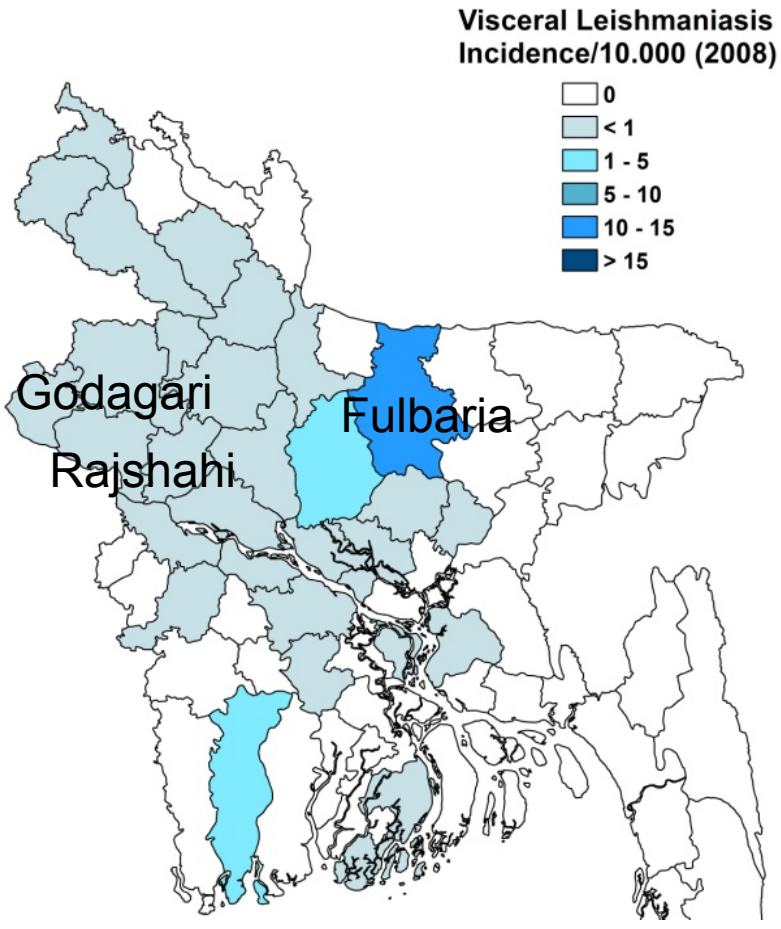
1994-2004: 74k cases
50% of cases in Mymensingh

Data from Alvar et al 2012 Plos One

Leishmania in Bangladesh

50% of cases in Mymensingh

2007: 137k active cases
but 5k actually reported



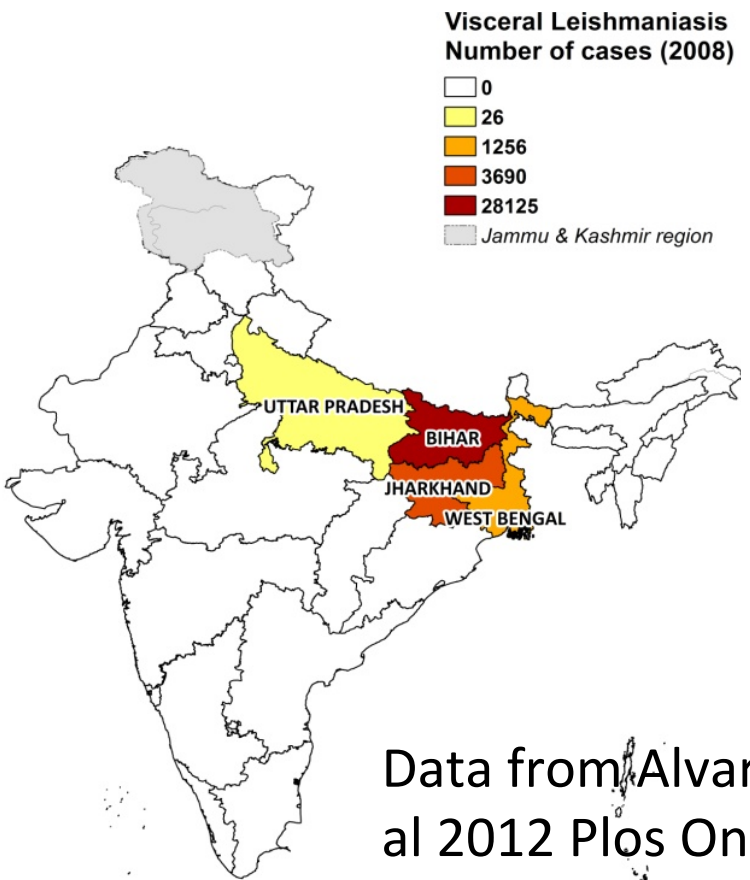
Fulbaria: 1.56% person years (Bern et al 2007)
Godagari/Rajshahi: 2.7% (Mondal et al 2009)

Miltefosine since 2008
Sodium stibogluconate (SSG)
2nd line: Amphotericin B

Data from Alvar et al 2012 Plos One

Leishmania in India

VL common until 1947 until 1960s Malaria Eradication Program & DDT spraying



Data from Alvar et al 2012 Plos One

1869 Garo Hills outbreak
1899 (Purnea in Bihar)

No reports in 1962/3

100k cases in 1977
40k in 1978

4.2-fold underestimation (Singh et al 2006)

Leishmania in India

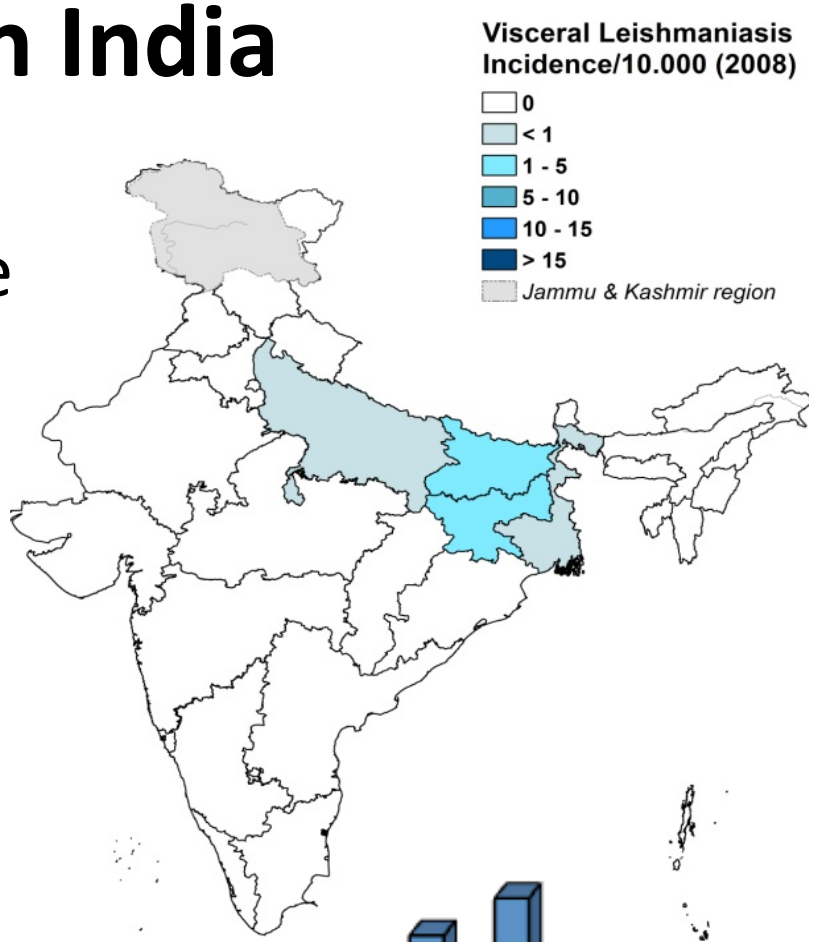
~271k in 2007 (45k cases)

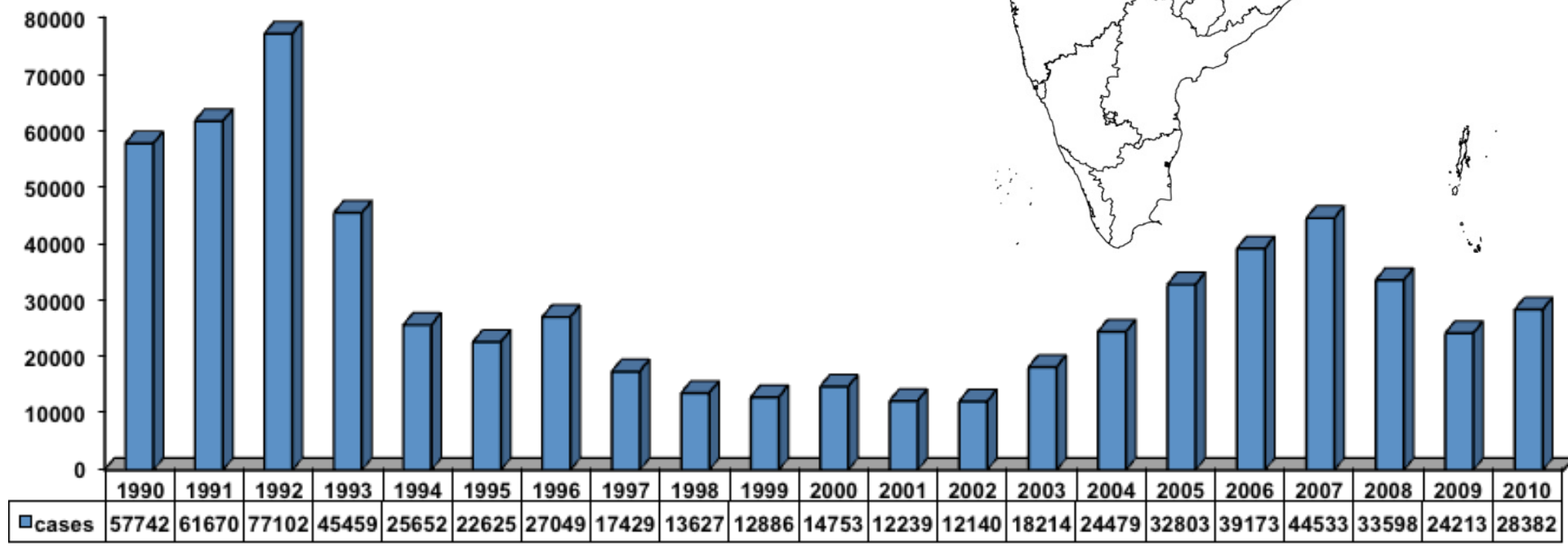
Miltefosine in 2012: 6% non-cure

2nd line: Amphotericin B: 2%

Antimonial-R up to 60% in Bihar

SSG used 50% => 40% 2nd course





	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
cases	57742	61670	77102	45459	25652	22625	27049	17429	13627	12886	14753	12239	12140	18214	24479	32803	39173	44533	33598	24213	28382

Data from Alvar et al 2012 Plos One

Leishmania in Nepal

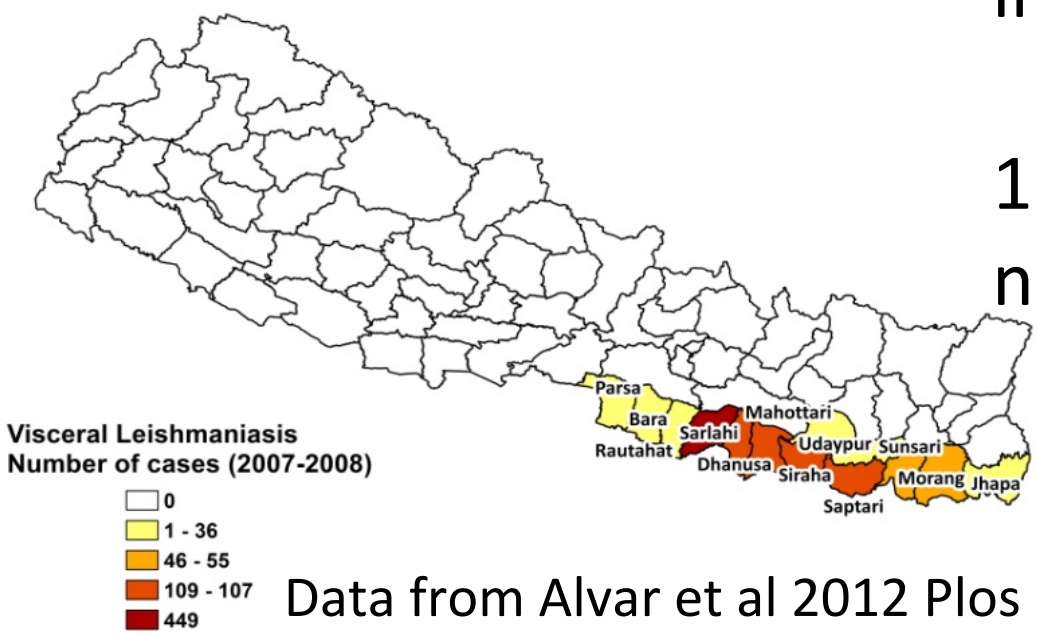
Common before 1958 until 1960s Malaria Eradication Program & DDT spraying

VL (CL is rare)

Found again 1980

Spread via 1977 outbreak in Bihar (Bista 1998)

1.5 higher than reported numbers (BPKHIS 2010)



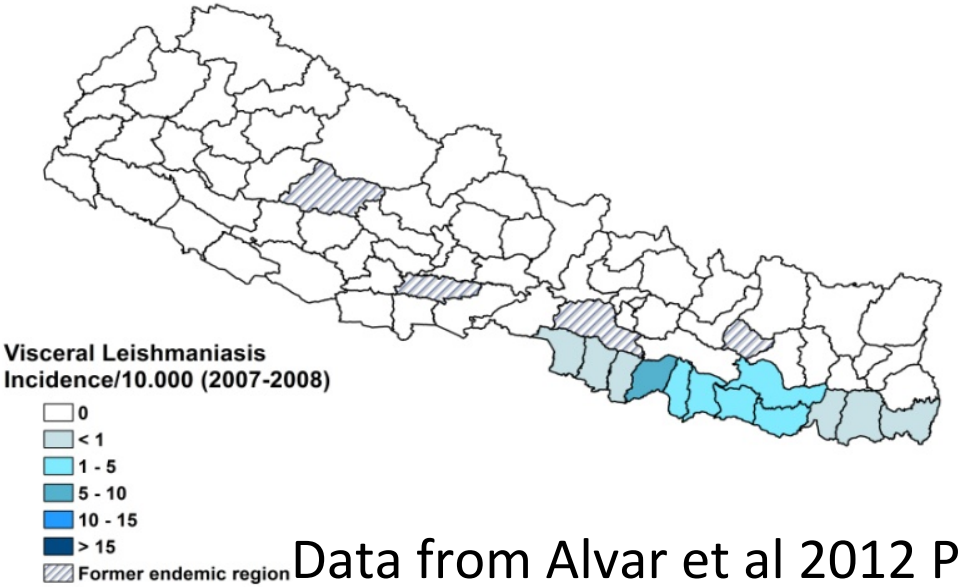
Data from Alvar et al 2012 Plos One

Leishmania in Nepal

1980-9 => 0.01% (years)
Now 0.043-55% (Schenkel et al 2006)

1980-2006: 26k cases in 16 districts

2007-10 VL in 26 districts



Data from Alvar et al 2012 Plos One

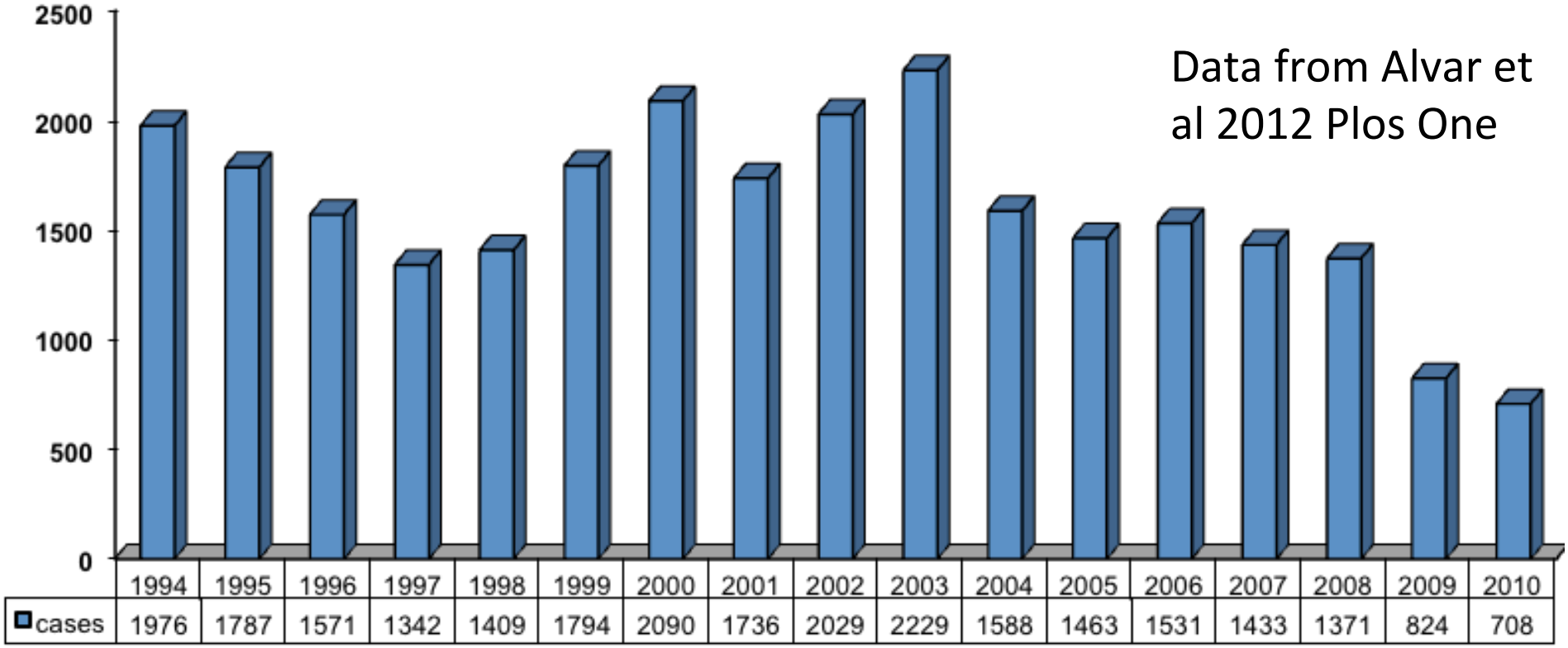
Leishmania in Nepal

Miltefosine in 2012: 11% relapse, 1.5% death

SSG still used in 4 districts

2nd line: Amphotericin B

23% of people use indigenous healers! (Mondal et al '09)



	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
cases	1976	1787	1571	1342	1409	1794	2090	1736	2029	2229	1588	1463	1531	1433	1371	824	708

2. *Leishmania* reproduction & biology

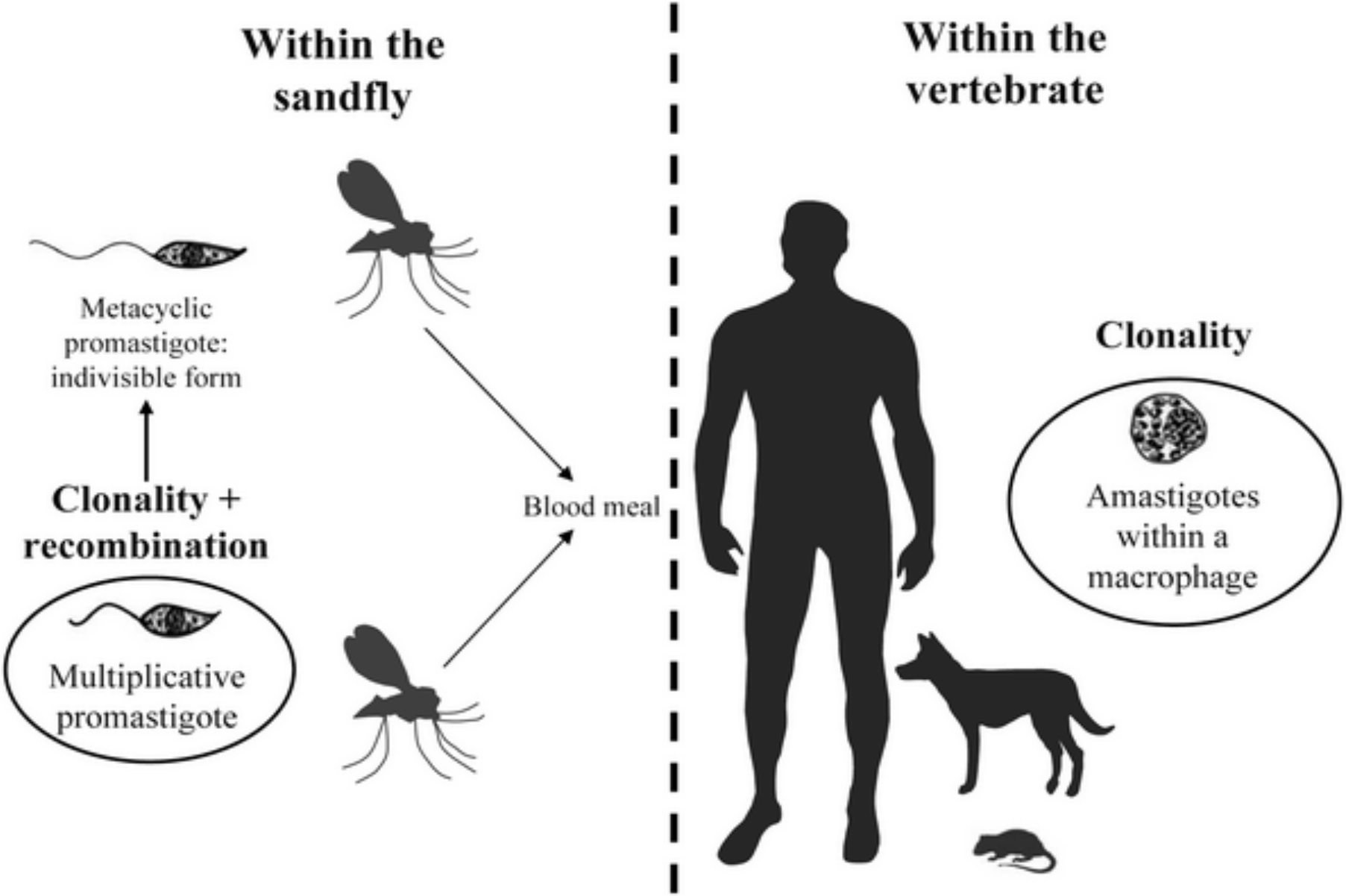
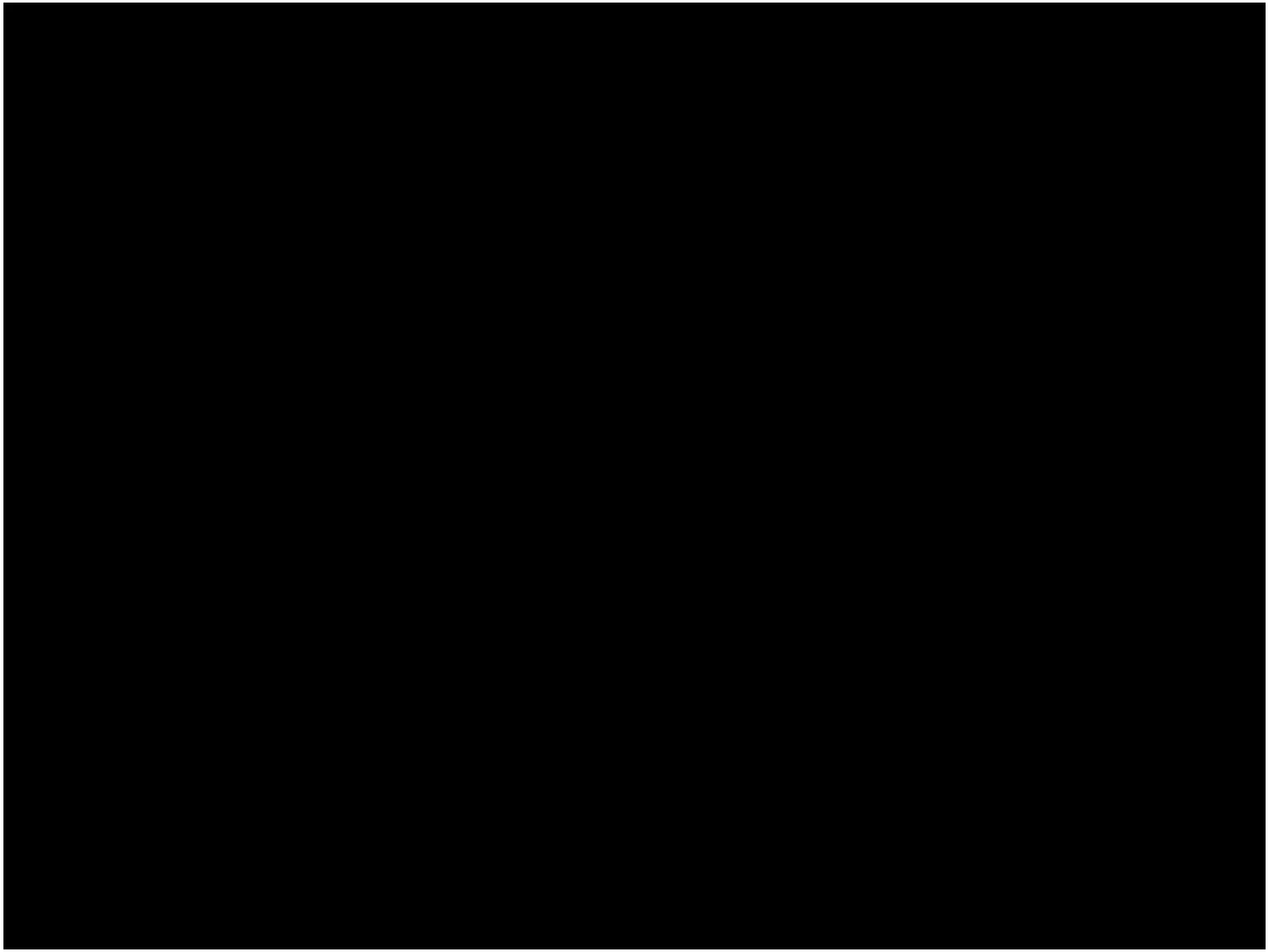


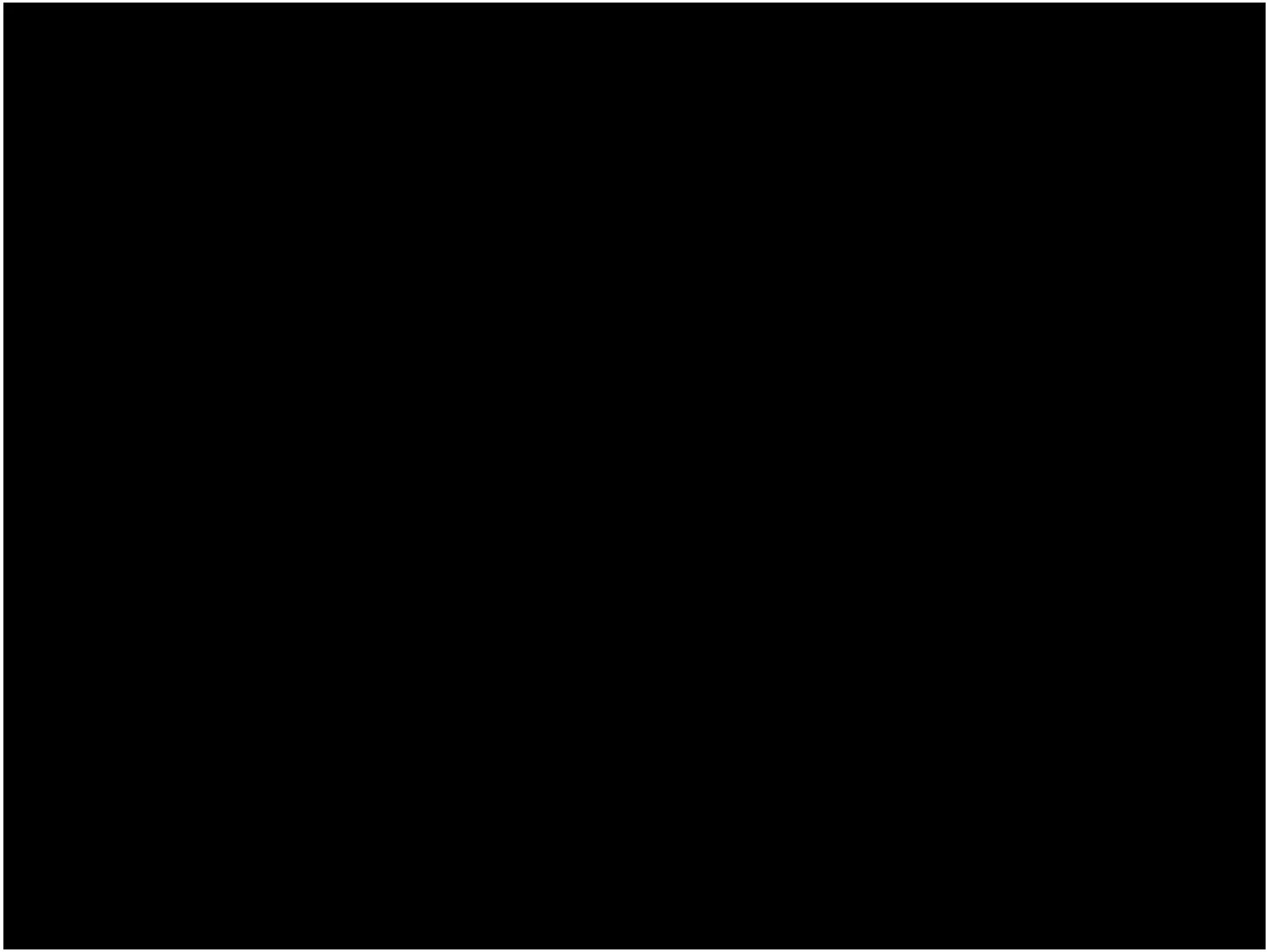
Figure 1 in Rougeron V, De Meeûs T, Kako Ouraga S, Hide M, et al. (2010) "Everything You Always Wanted to Know about Sex (but Were Afraid to Ask)" in Leishmania after Two Decades of Laboratory and Field Analyses. PLoS Pathog 6(8): e1001004.

2. *Leishmania* in the vector



Teixeira DE, Benchimol M, Rodrigues JCF, Crepaldi PH, et al. (2013) The Cell Biology of *Leishmania*: How to Teach Using Animations. PLoS Pathog 9(10): e1003594.

2. *Leishmania* in the host



Teixeira DE, Benchimol M, Rodrigues JCF, Crepaldi PH, et al. (2013) The Cell Biology of *Leishmania*: How to Teach Using Animations. PLoS Pathog 9(10): e1003594.

2. *Leishmania* reproduction & biology

Clonality = binary fission => heterozygosity fixed **WRONG** (Meselson effect)

Autogamous = selfing => mixed haplotypes **RIGHT**

Leishmania are sexual but vector co-infection is rare

Apparent "clonal" pattern

Parasexuality with aneuploidy, peculiar combinations

Akopyants et al 2009 Science

3. Genetics of *Leishmania donovani*

Sauro vs Old World vs Viannia species

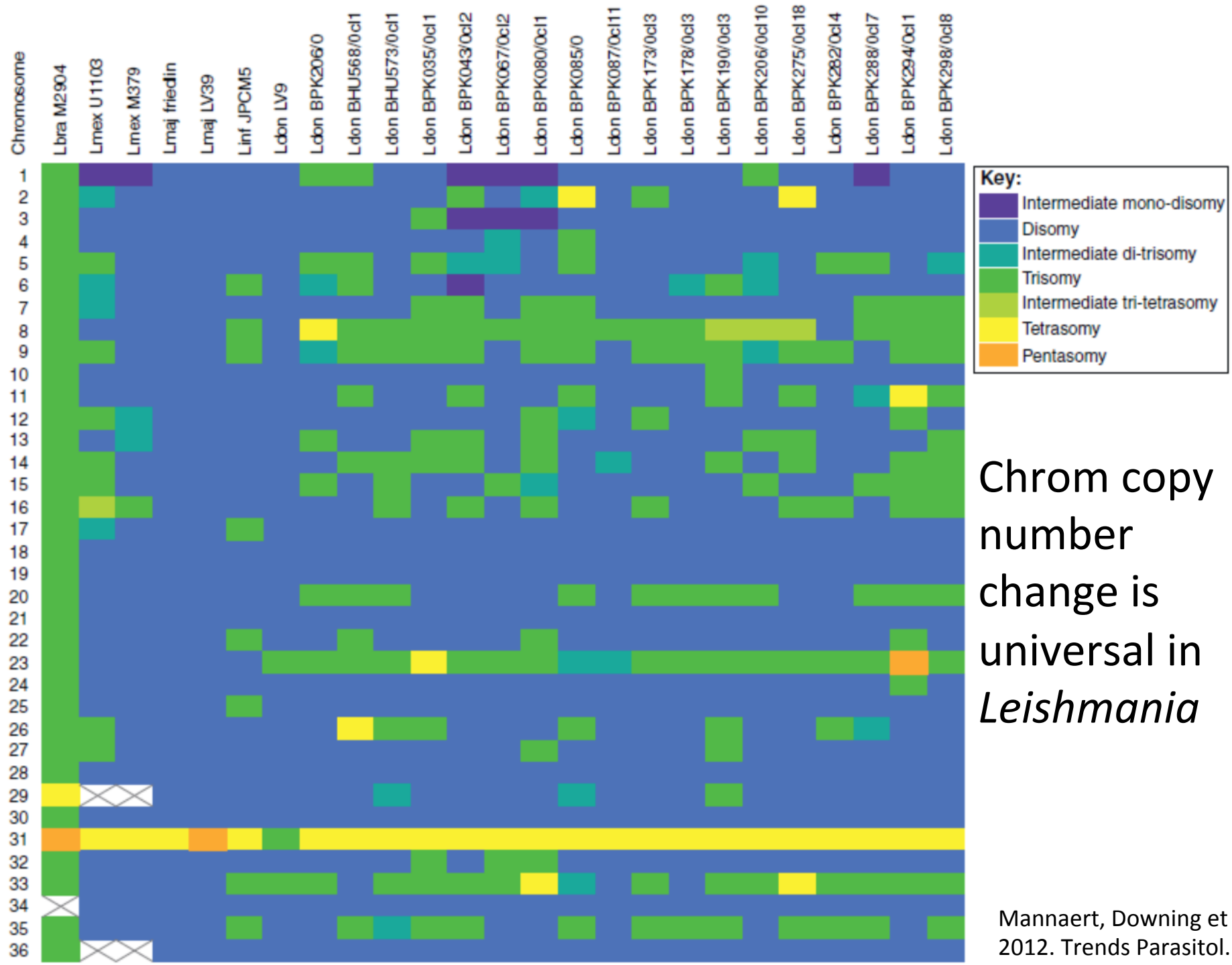
L. donovani:

36 chromosomes

8,252 chromosomal genes

Aneuploidy

Downing, Imamura et al 2011. Genome Research.



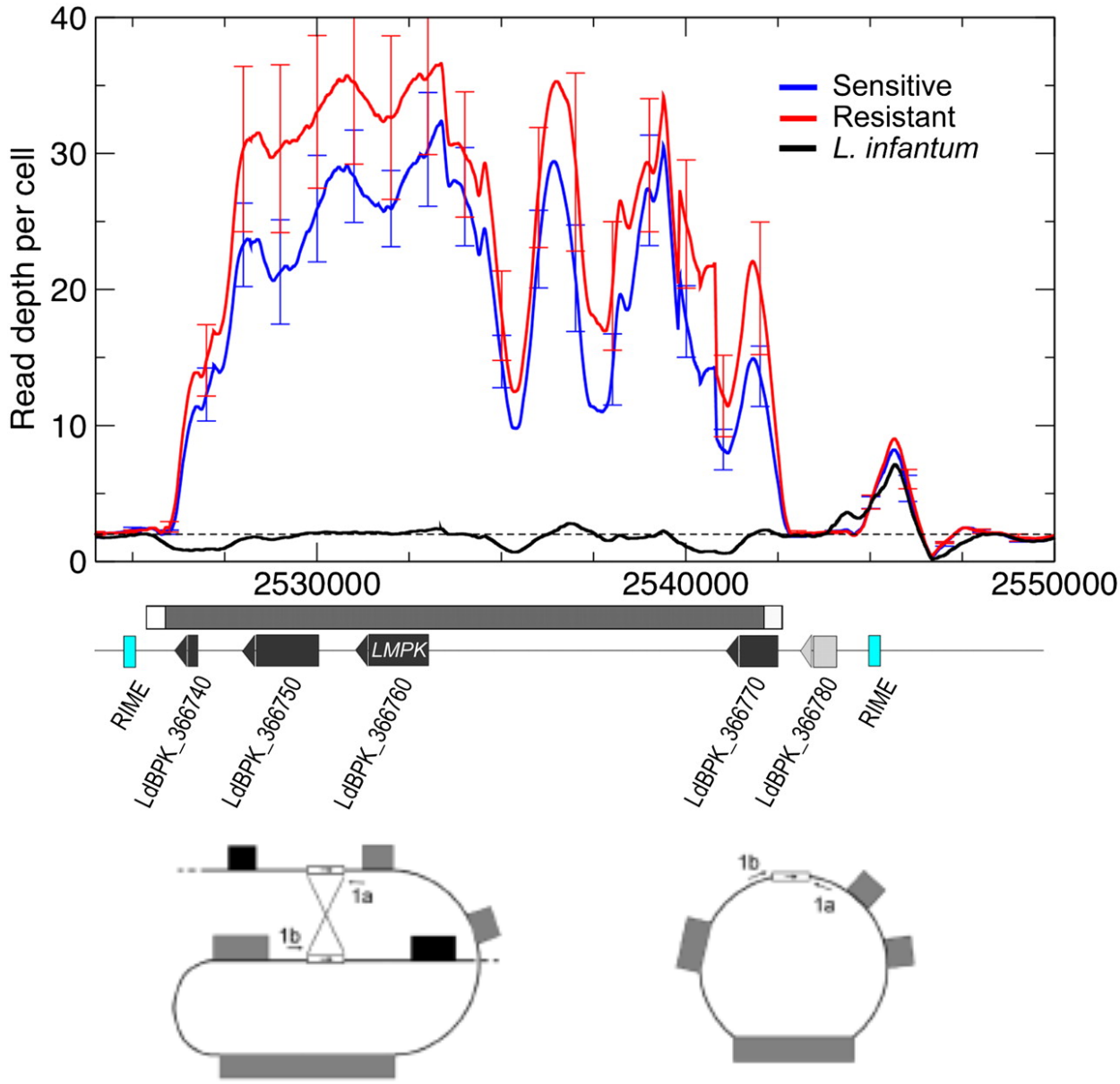
Chrom copy number change is universal in *Leishmania*

Mannaert, Downing et al. 2012. Trends Parasitol.

3. Genetics of *Leishmania donovani*

CNVs, episomes flanked by short direct repeats or ribosomal mobile elements

SSG: MAPK-locus (chr36)
822-bp - acid phosphatase genes



Downing, Imamura et al 2011. Genome Research.

3. Genetics of *Leishmania donovani*

3 SNPs to detect SSG treatment failure patients

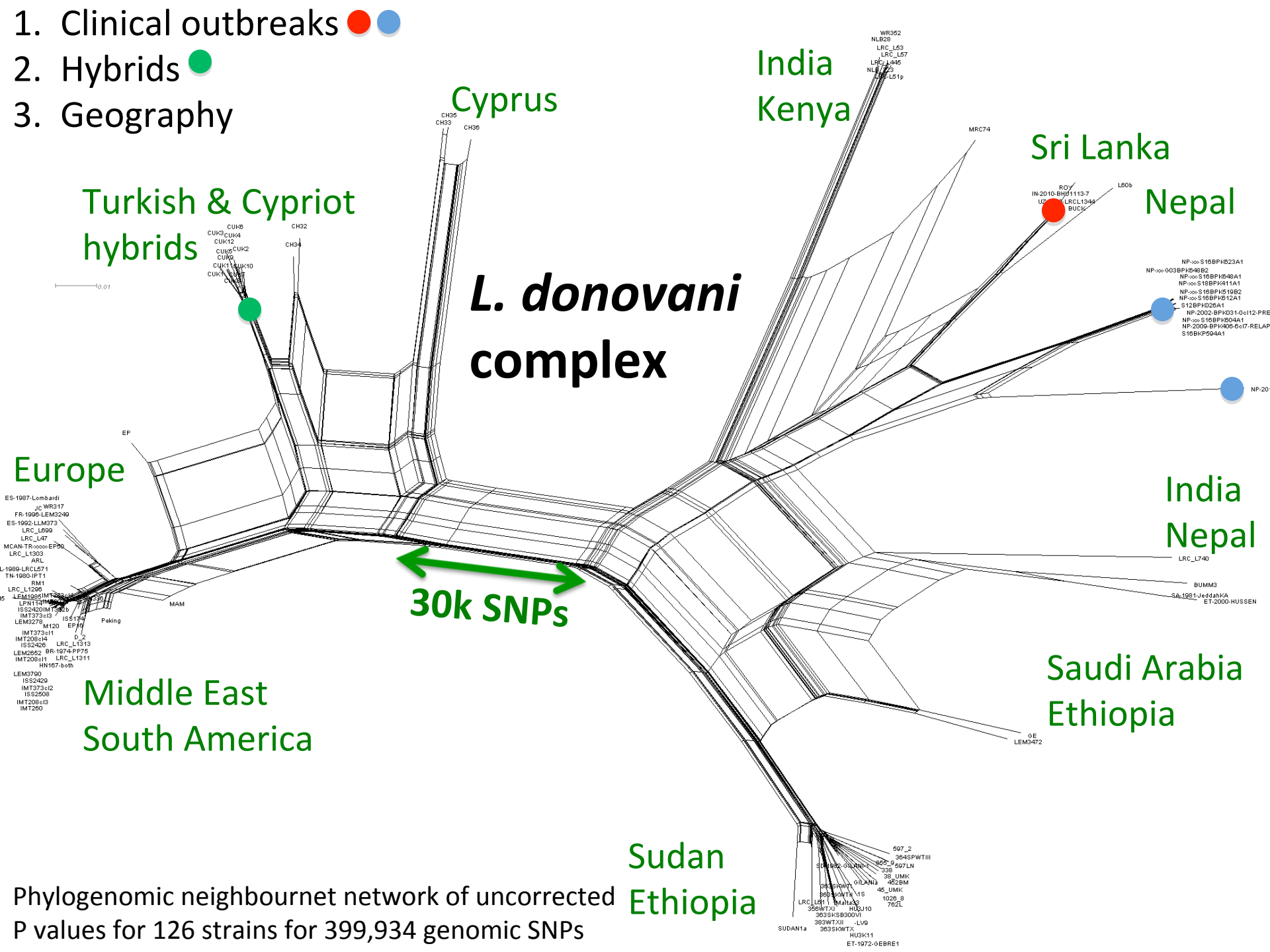
SNPs at chr13:442924, chr24:26882, chr35:1619958

Sensitivity of 77.8% (40.2%–96.1%)
Specificity of 100.0% (82.2%–100.0%)

Not SSG-R vs SSG-S: infrapopulation heterogeneity?

Vanaerschot, Decuypere, Downing et al. 2012. J Infect Dis.

1. Clinical outbreaks
2. Hybrids
3. Geography



Phylogenomic neighbournet network of uncorrected P values for 126 strains for 399,934 genomic SNPs

3. Genetics of *Leishmania donovani*

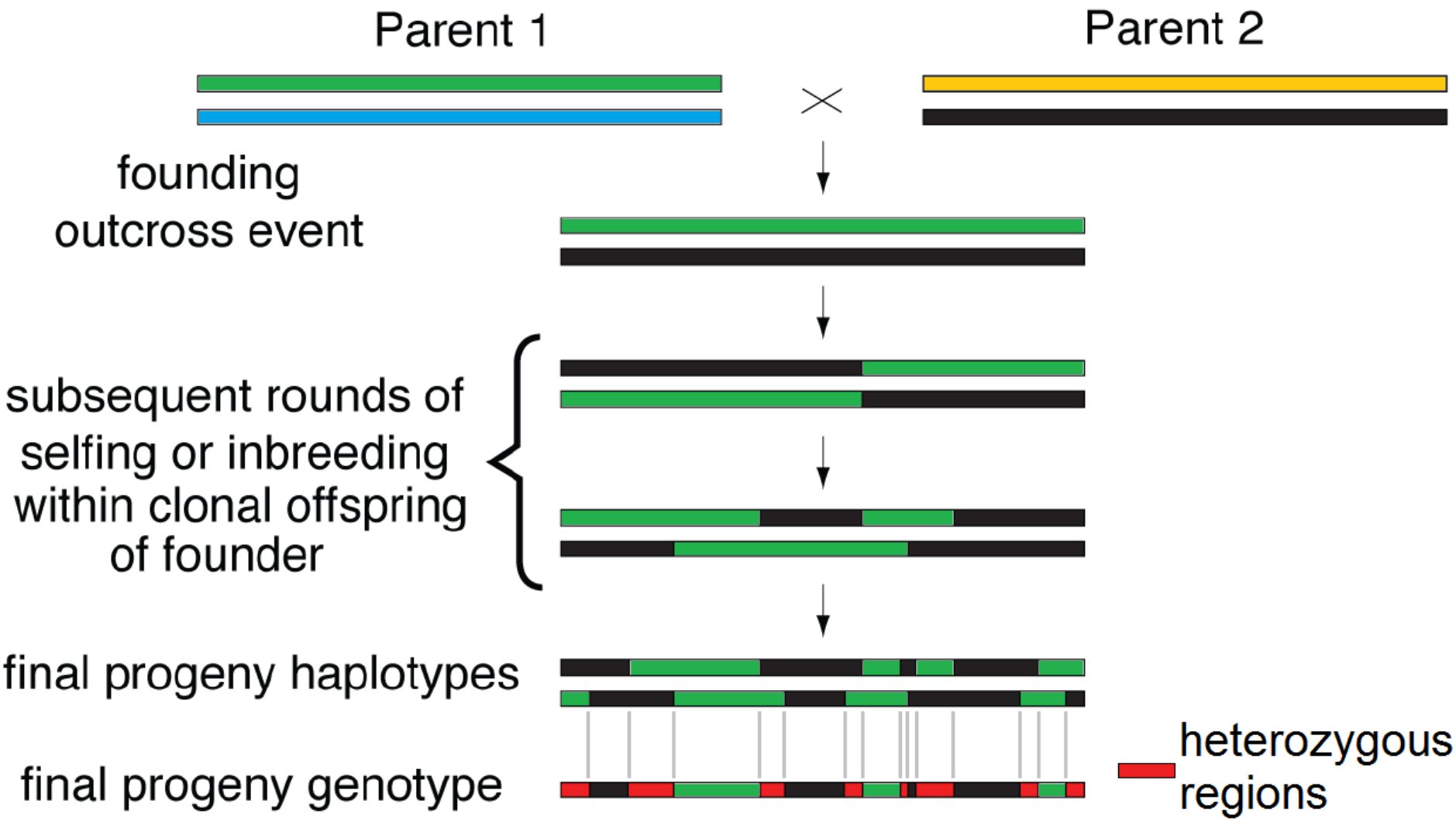
Genomes of 12 *Leishmania*:
11 from sand flies (*Phlebotomus tobbi*) and 1 from a patient (all isolates, cell population)

Single focus of CL in Çukurova (southeast Turkey)

Cross of two diverse strains + F1 selfing
Bimodal divergence compared to reference (related)
Phylogenetic incompatibility
Sexual and asexual reproduction within population

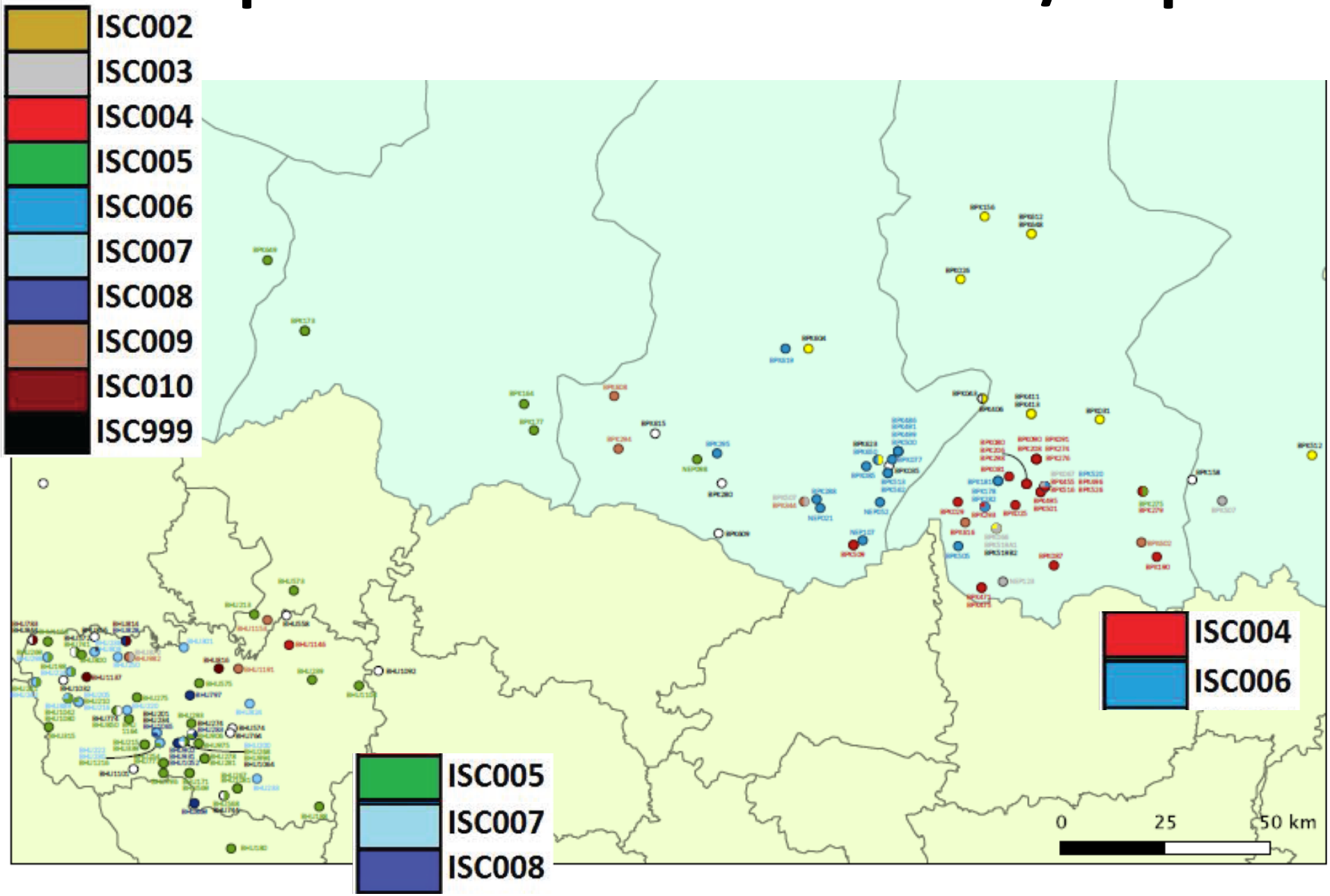
Adapted from Downing, Rogers et al 2014. Plos Genetics.

3. Genetics of *Leishmania donovani*



Adapted from Downing, Rogers et al 2014. Plos Genetics.

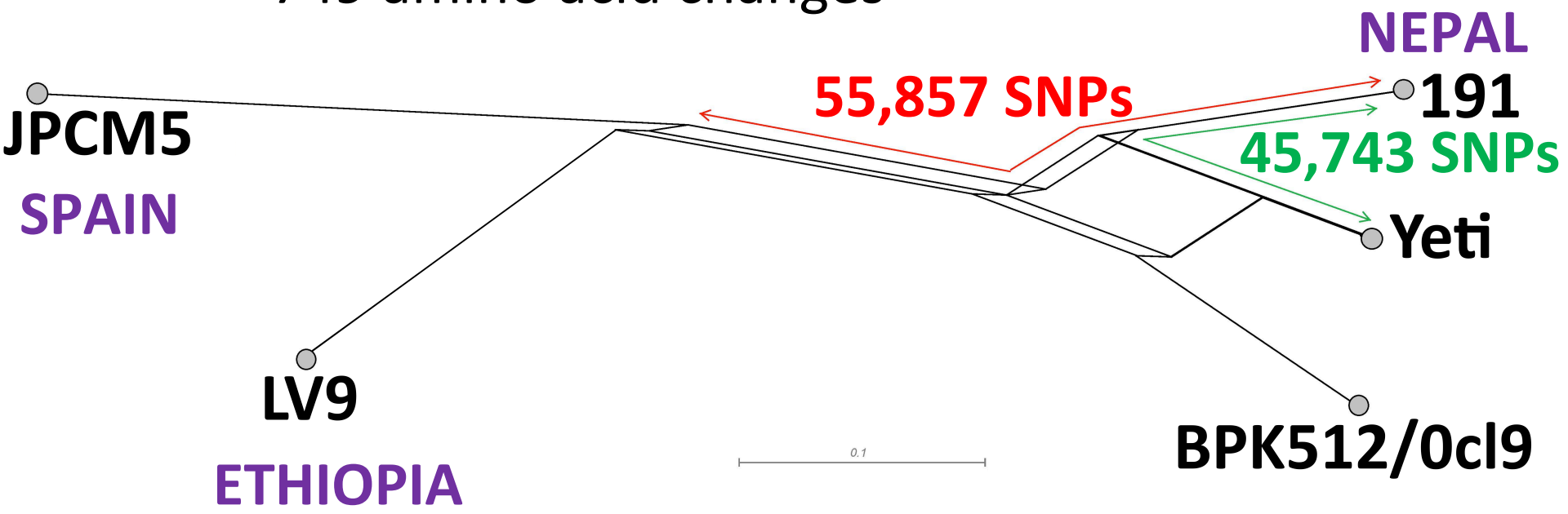
4. Population structure in India/Nepal



4. Population structure in India/Nepal

98 Bihar, 98 Terai, 8 Bangladesh

In 191: 2,418 SNPs (62% are in one sample only)
749 amino acid changes



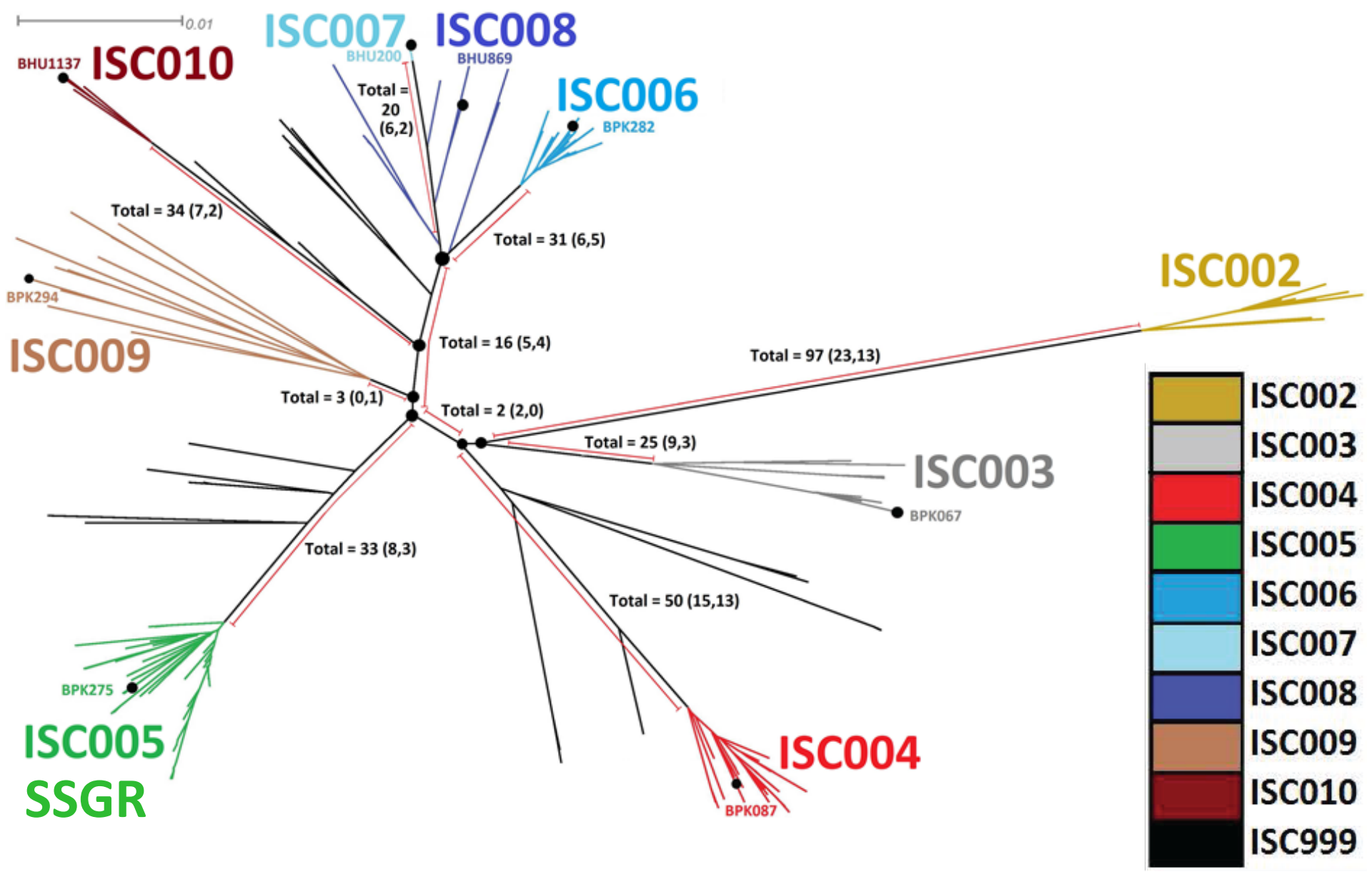
4. Population structure in India/Nepal

1 SNP per 100 kb
Mean 88 SNPs b/w pairs

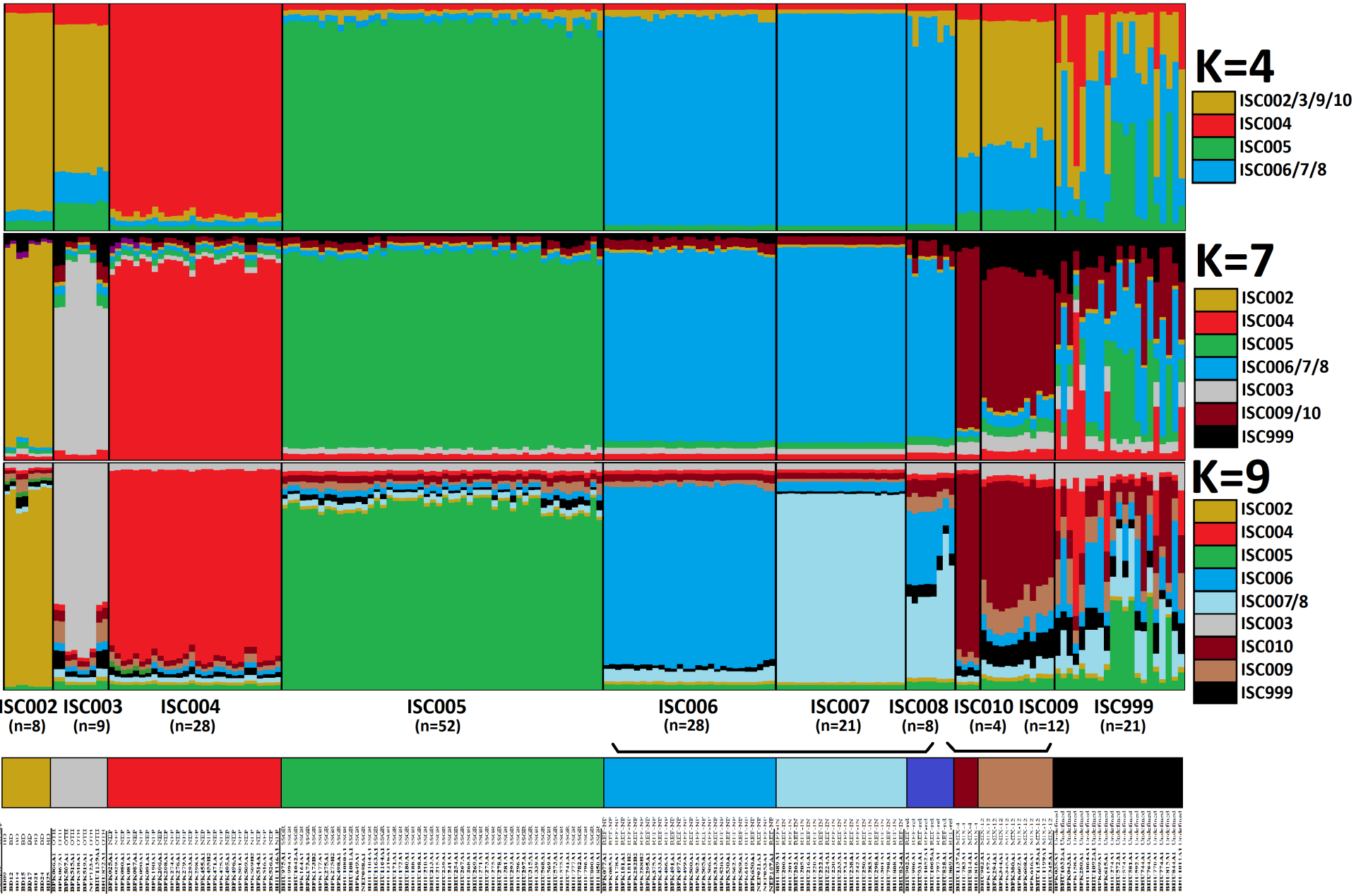
Only 566 genes have 1+ non-singleton SNPs
Only 63 genes have 2+ non-singleton SNPs
Only 12 genes have 3+ non-singleton SNPs

46.5% SNPs in CDS
Nonsyn/syn ~ 749/364 => recent population expansion

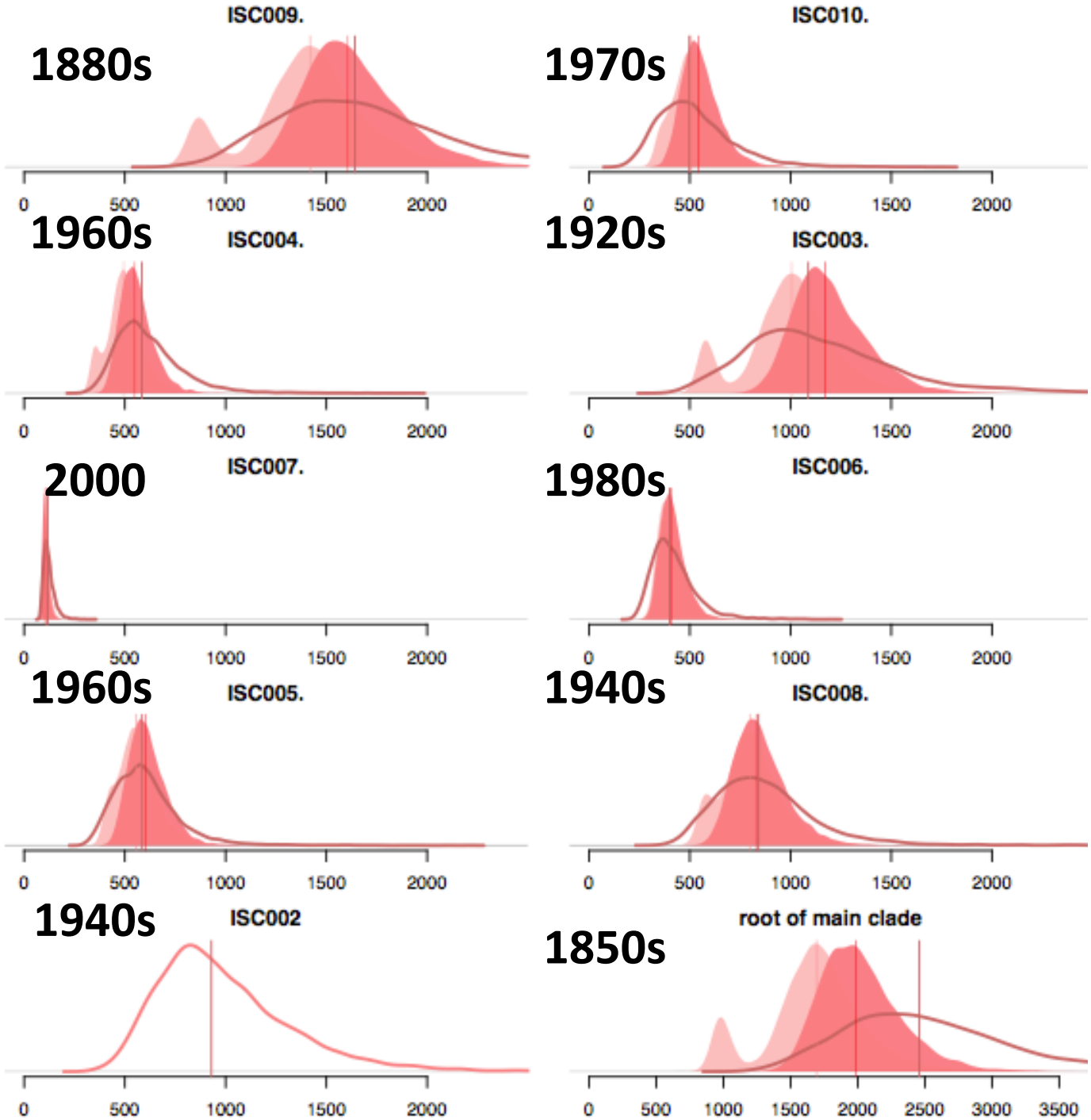
4. Population structure in India/Nepal



4. Population structure in India/Nepal



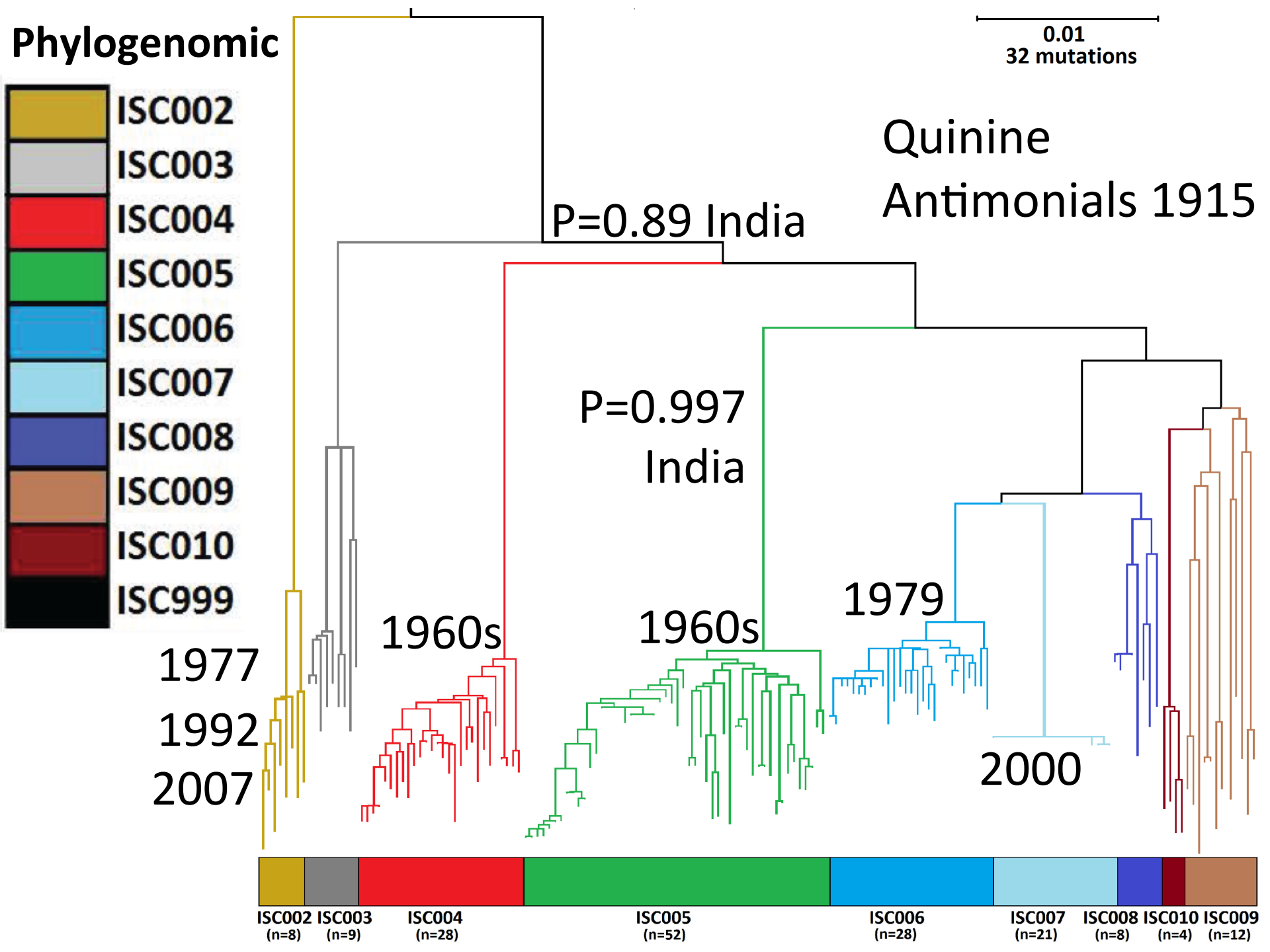
How old?



Full marginal posterior densities of ages of the oldest split in each group (months) 183 non-hybrid samples

red (strict) and pink (relaxed) lines = medians

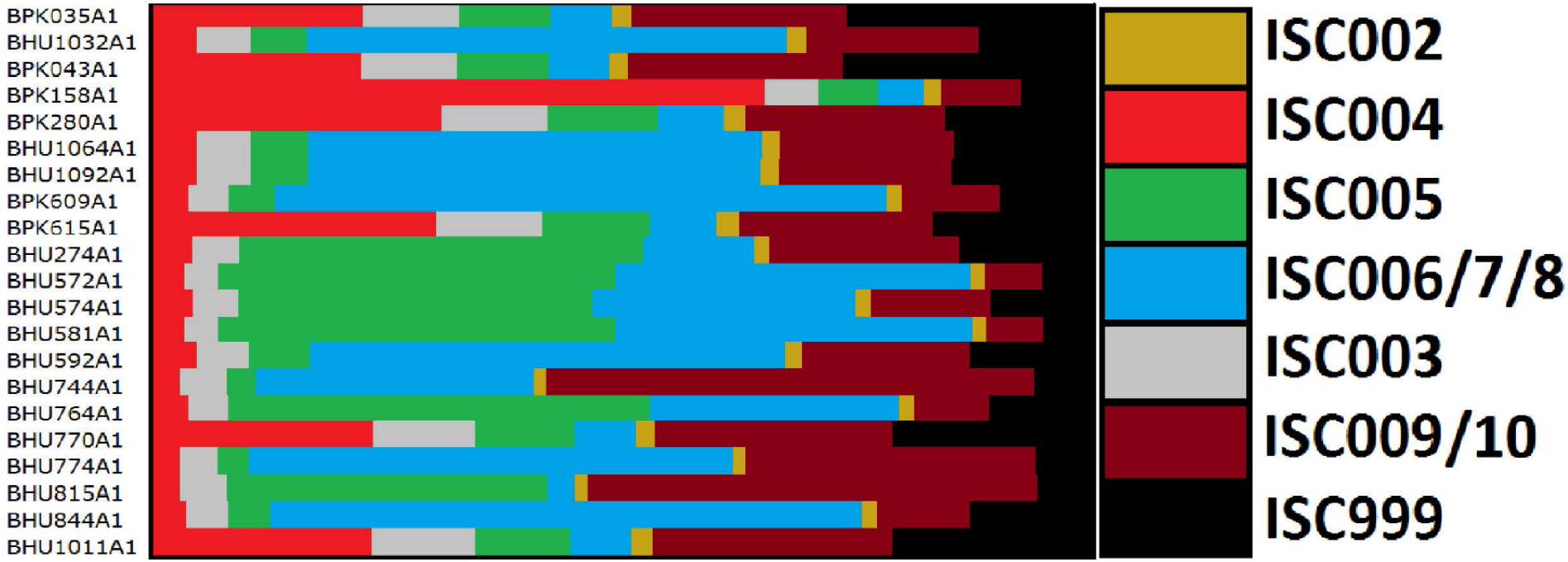
James Cotton (Sanger UK)


Phylogenomic
ISC002
ISC003
ISC004
ISC005
ISC006
ISC007
ISC008
ISC009
ISC010
ISC999
0.01
32 mutations
Quinine
Antimonials 1915
P=0.89 India
P=0.997
India
1977
1992
2007
1960s
1960s
1979
2000
ISC002 (n=8)
ISC003 (n=9)
ISC004 (n=28)
ISC005 (n=52)
ISC006 (n=28)
ISC007 (n=21)
ISC008 (n=8)
ISC010 (n=4)
ISC009 (n=12)

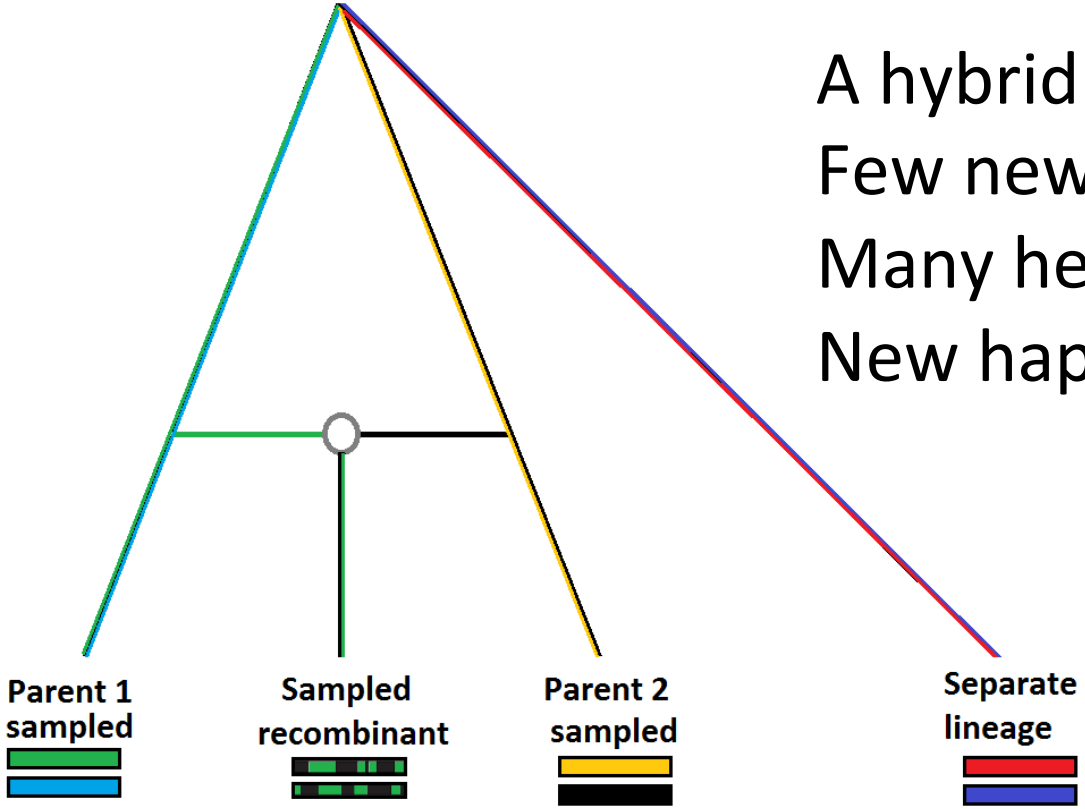
	N	SNPs	SSG-S	SSG-R
ISC002	8	112	-	-
ISC003	9	227	1	1
ISC004	28	256	11	4
ISC005	52	284	2	10
ISC006	28	109	4	3
ISC007	21	2	-	-
ISC008	8	98	-	-
ISC009	4	49	0	1
ISC010	12	350	1	1
ISC999	21	502	3	6

4. Population structure in India/Nepal

Unassigned samples: hybrids or rare?

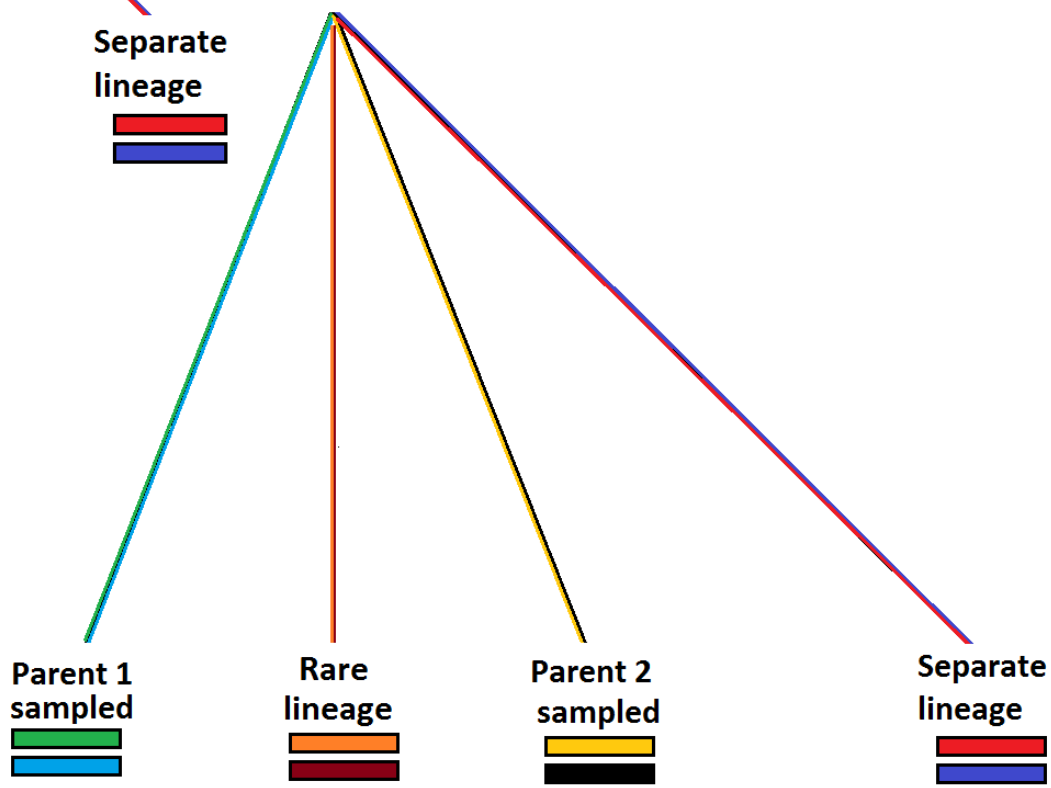


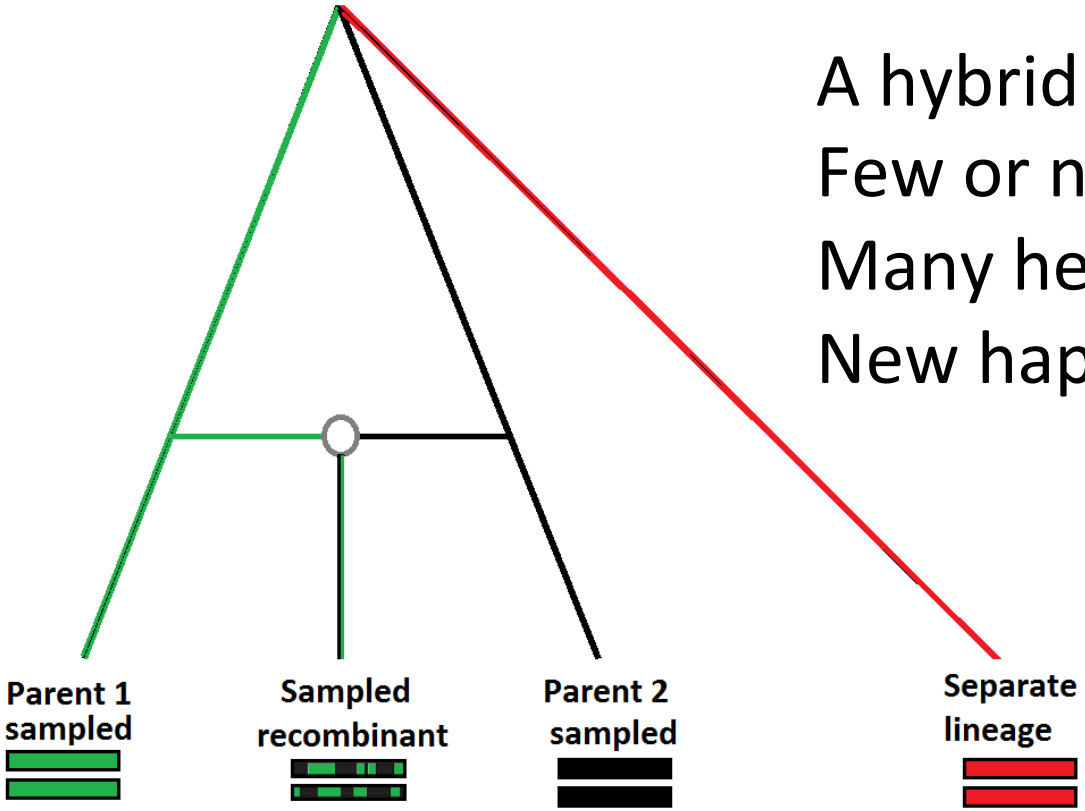
21 strains with no clear population membership shown with ancestry proportions for K=7 populations



A hybrid will have:
Few new SNPs
Many heterozygous SNPs
New haplotypes from known ones

A rare lineage will have:
Many new SNPs
Few heterozygous SNPs
New haplotypes

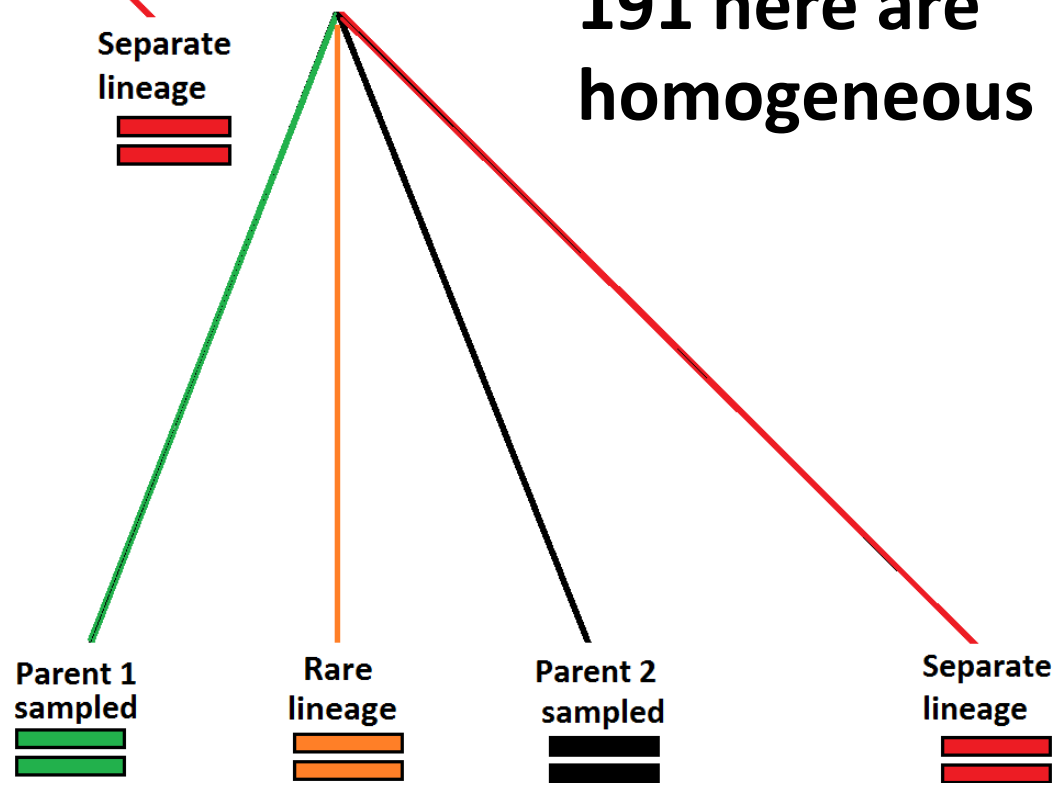




A hybrid will have:
Few or no new SNPs
Many heterozygous SNPs
New haplotypes from known ones

191 here are homogeneous

A rare lineage will have:
A few new SNPs
Few heterozygous SNPs
New haplotypes



Recombination

Rare

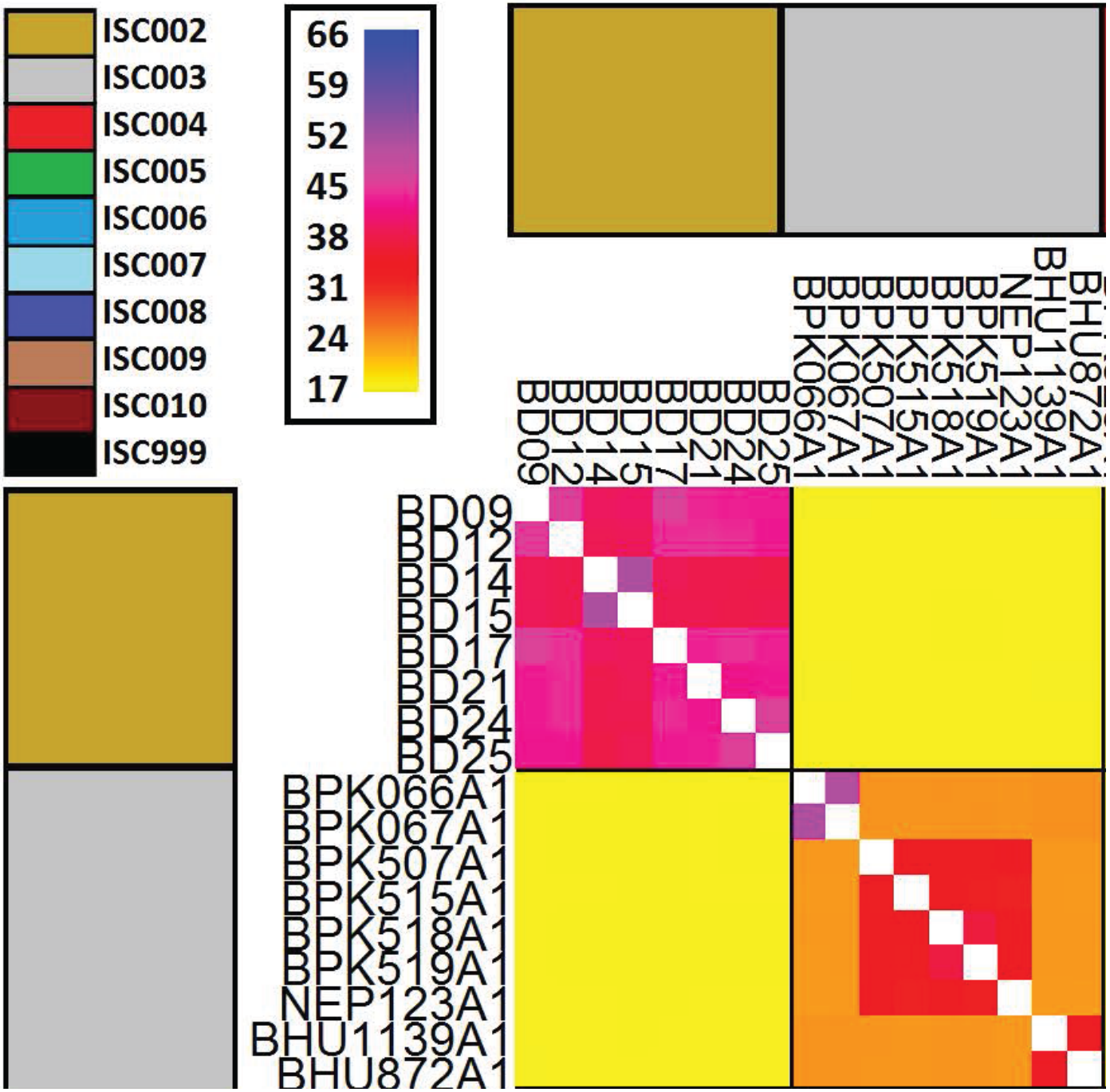
62+ events

four gamete test: 9,876 from 2,029,105 SNP pairs (0.5%)

Pervasive LD

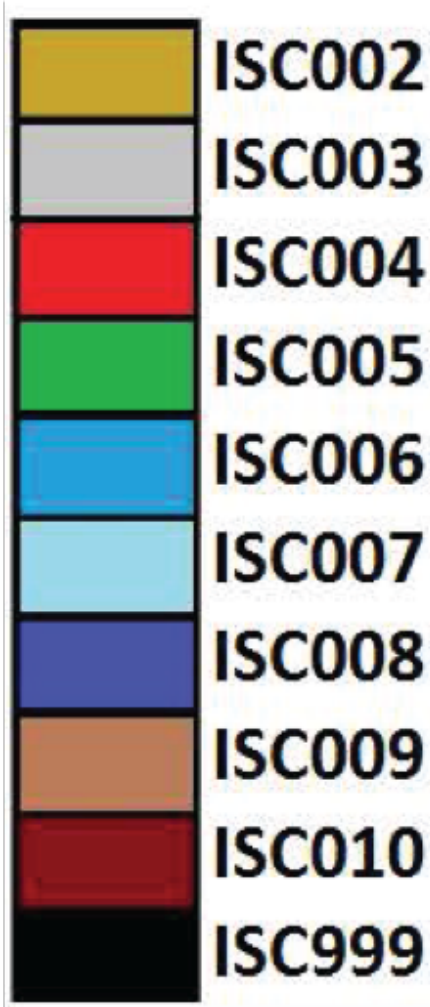
Pairwise haplotype sharing in 191 samples => right

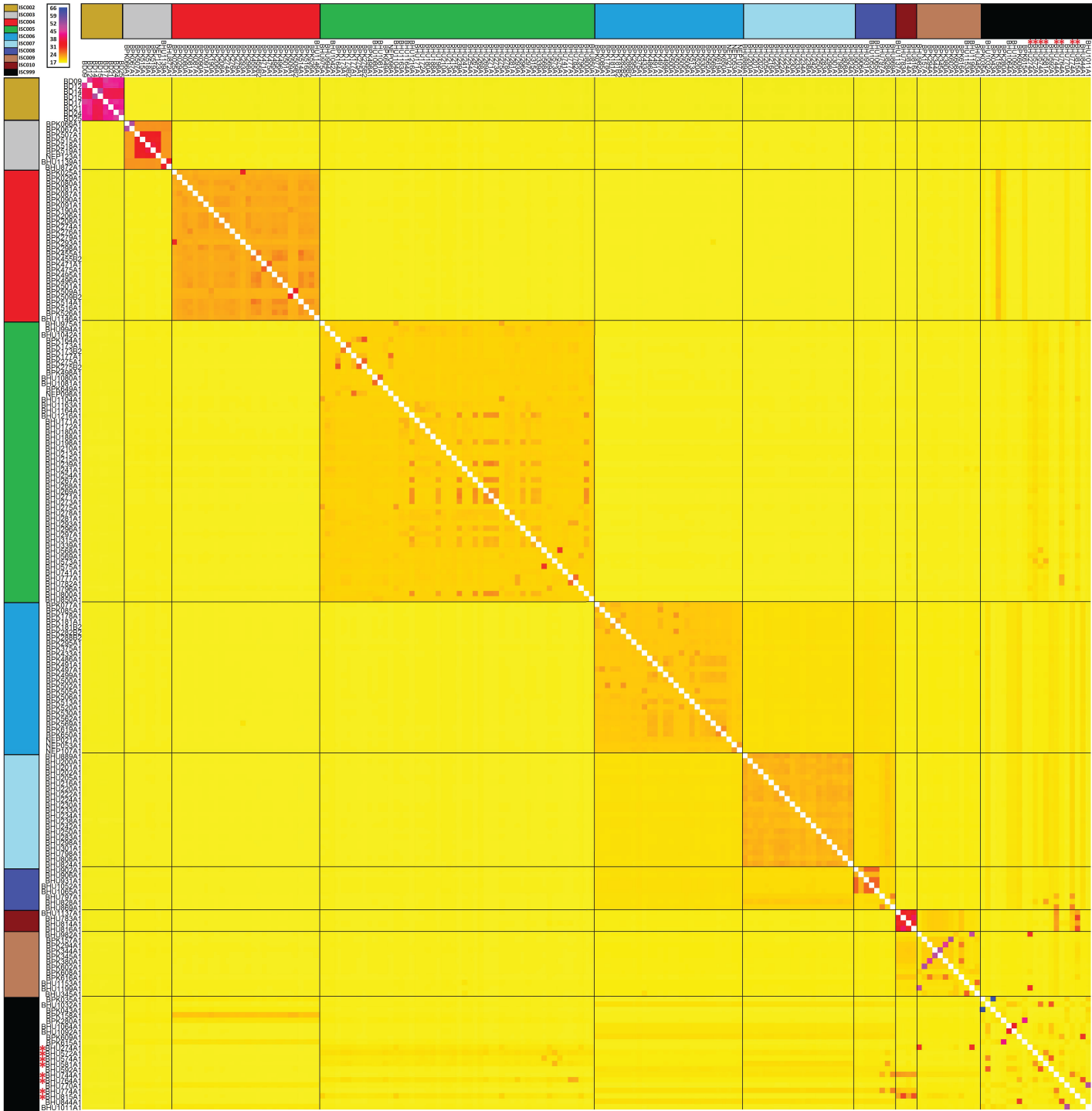
Lawson et al 2012



Pairwise haplotype sharing in 191 samples

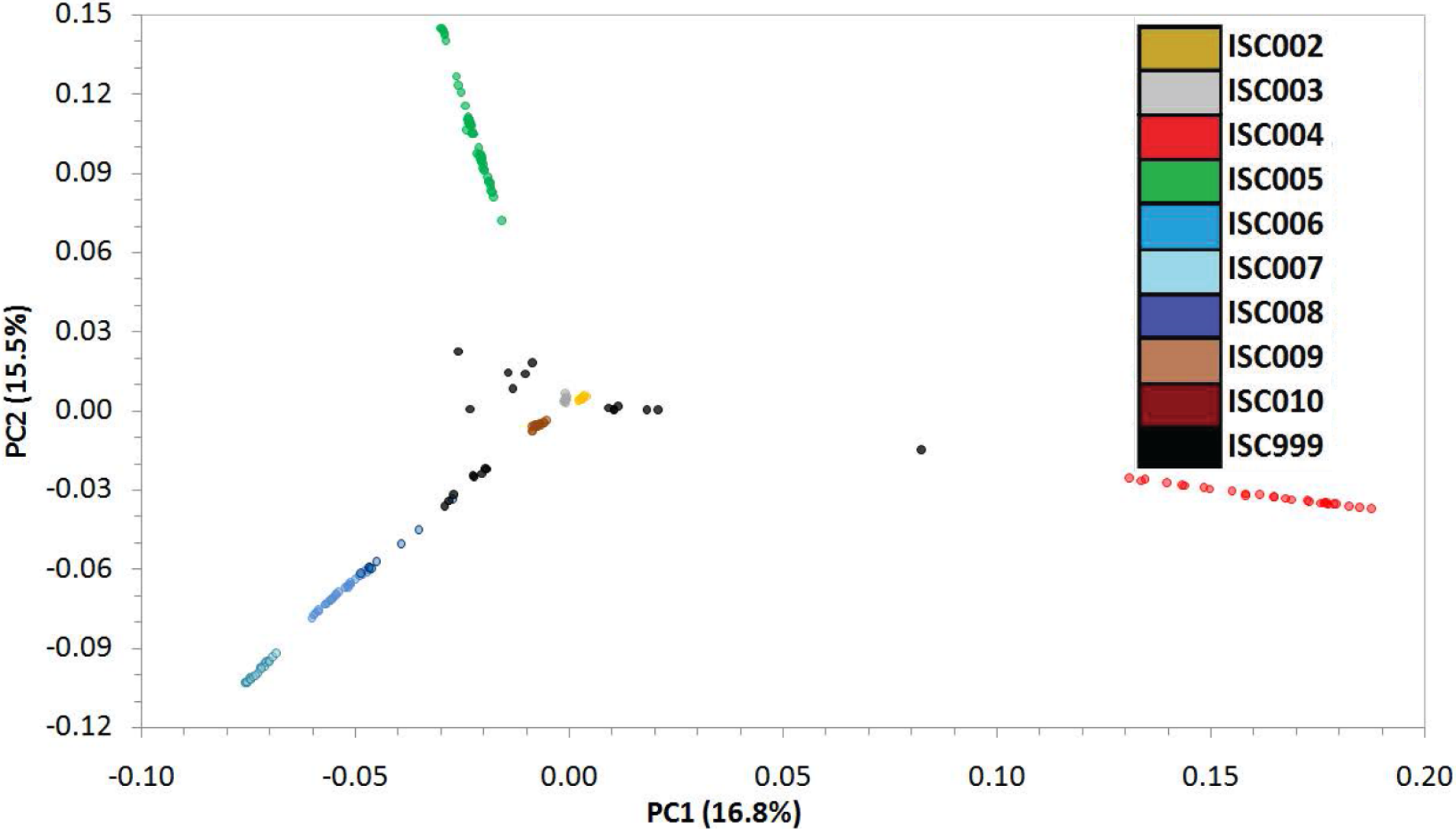
610 SNPs





4. Population structure in India/Nepal

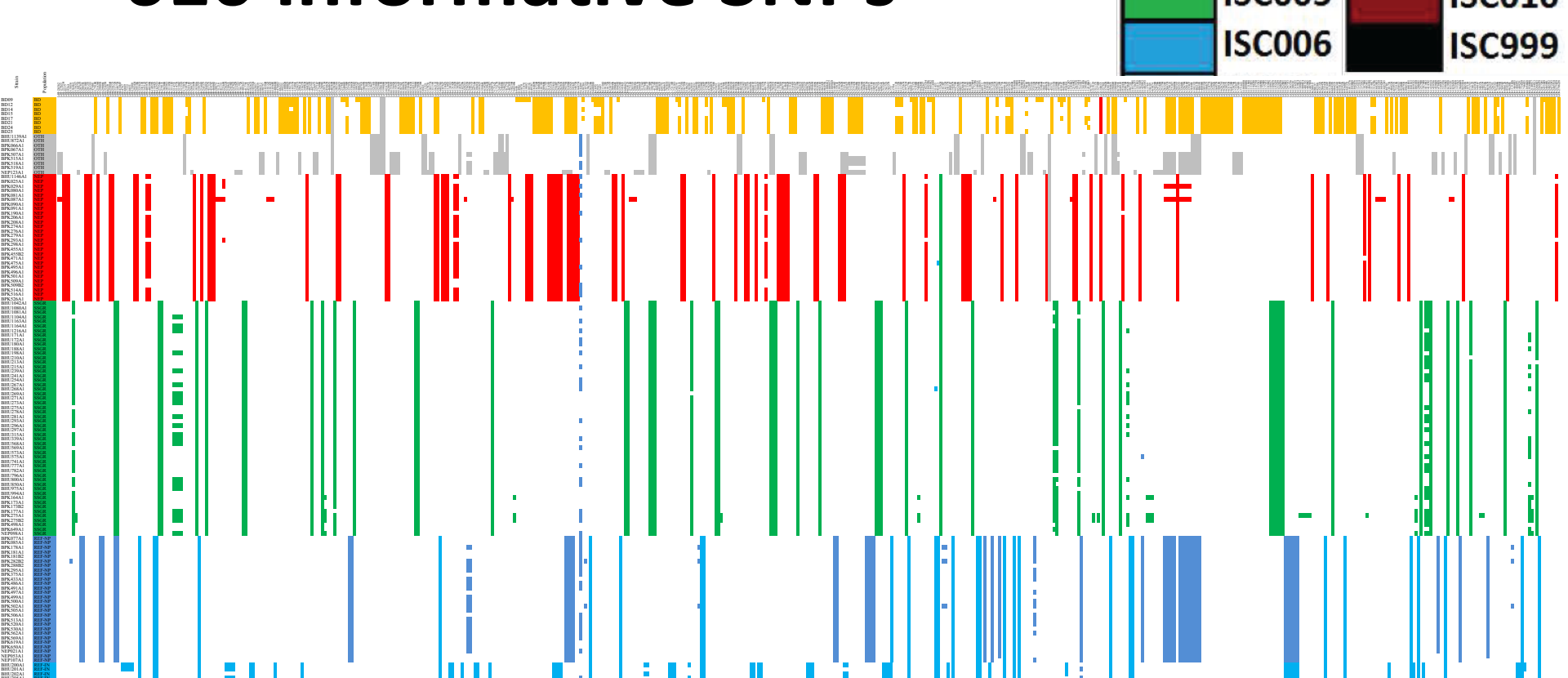
PCA of haplotype patterns scales across ISC004/5/6/7



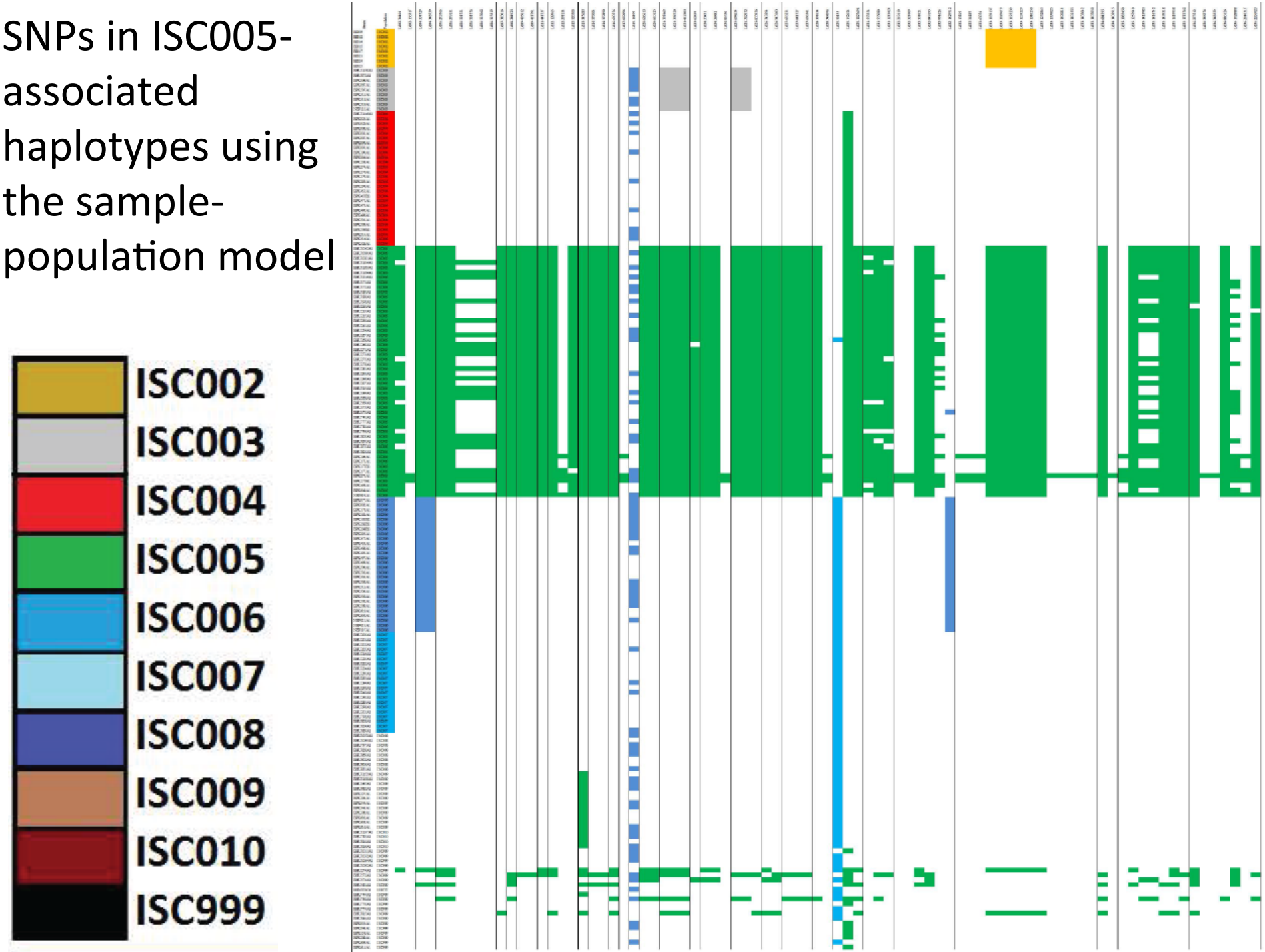
4. Population structure in India/Nepal

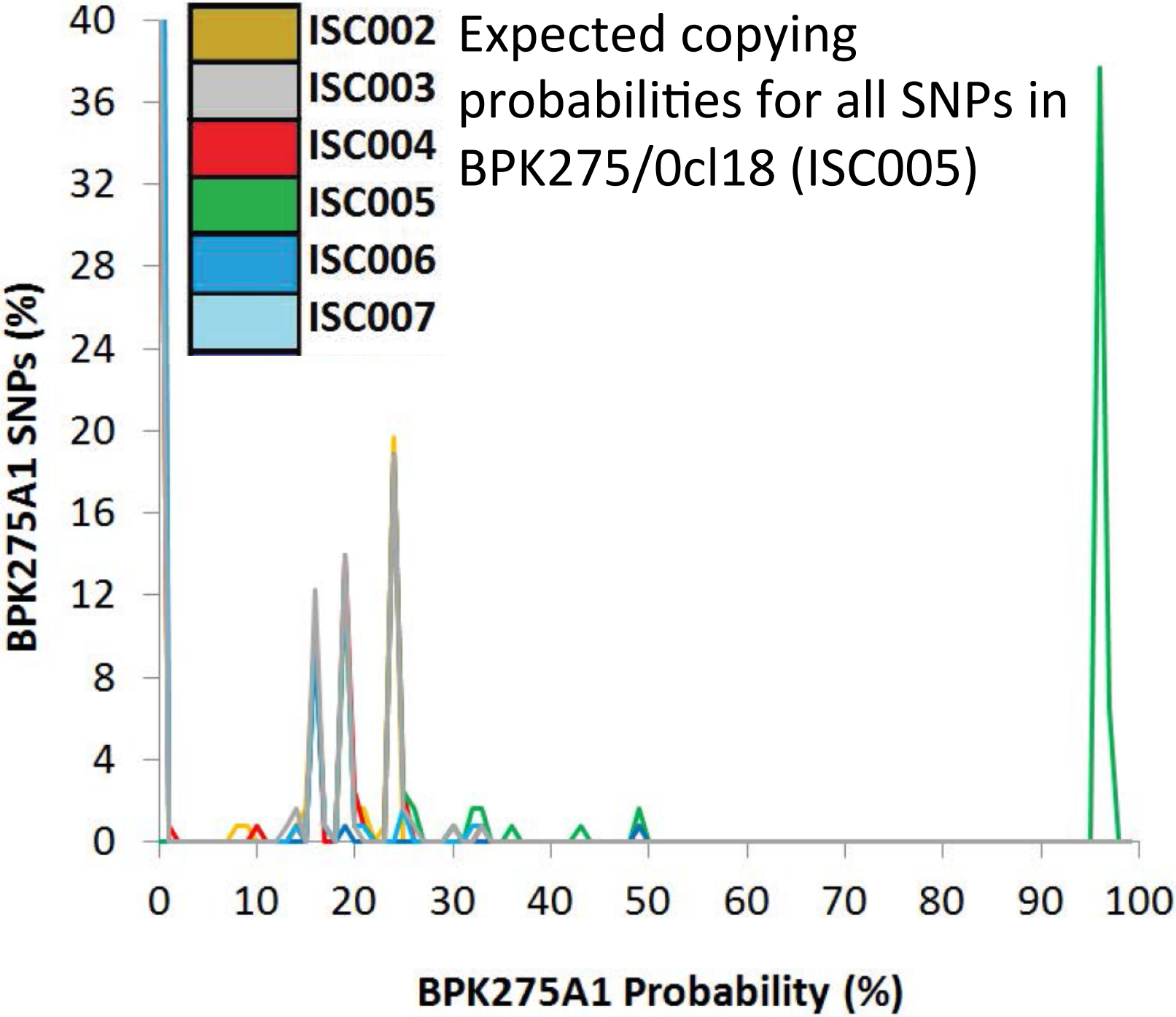
ISC group	Origin	002	High	003	High	004	High	005	High	006	High	007	High
BHU770A1	ISC004	1	0	2	1	13	8	1	0	0	0	0	0
BHU1011A1	ISC004	1	0	2	1	13	8	2	0	1	0	0	0
BPK035A1	ISC004	1	0	2	1	13	8	2	0	1	0	0	0
BPK043A1	ISC004	1	0	2	1	13	8	2	0	1	0	0	0
BPK280A1	ISC004	1	0	1	0	21	11	1	0	0	0	0	0
BPK615A1	ISC004	1	0	1	0	21	11	1	0	0	0	0	0
BPK158A1	ISC004	1	0	1	0	61	35	1	0	0	0	0	0
BHU815A1	ISC005/6	0	0	1	1	0	0	21	14	8	0	2	0
BHU574A1	ISC005/6	0	0	0	0	1	0	15	10	11	0	10	0
BHU581A1	ISC005/6	0	0	0	0	1	0	19	15	14	0	32	10
BHU764A1	ISC005/6	0	0	0	0	1	0	25	15	19	1	11	0
BHU572A1	ISC005/6	0	0	0	0	1	0	35	24	20	0	23	6
BHU274A1	ISC005/6	0	0	2	1	2	0	30	18	14	0	12	2
BHU1092A1	ISC006/7	0	0	0	0	0	0	0	0	26	0	20	0
BHU774A1	ISC006/7	0	0	0	0	0	0	0	0	32	0	35	6
BHU844A1	ISC006/7	0	0	0	0	0	0	0	0	44	2	27	0
BHU1064A1	ISC006/7	0	0	0	0	0	0	1	0	27	0	20	0
BHU1032A1	ISC006/7	0	0	0	0	0	0	1	0	28	1	20	0
BHU592A1	ISC006/7	0	0	0	0	0	0	1	0	28	1	20	0
BPK609A1	ISC006/7	0	0	0	0	0	0	1	0	45	2	27	0
BHU744A1	ISC006/7	1	0	1	1	1	0	1	1	26	1	25	4

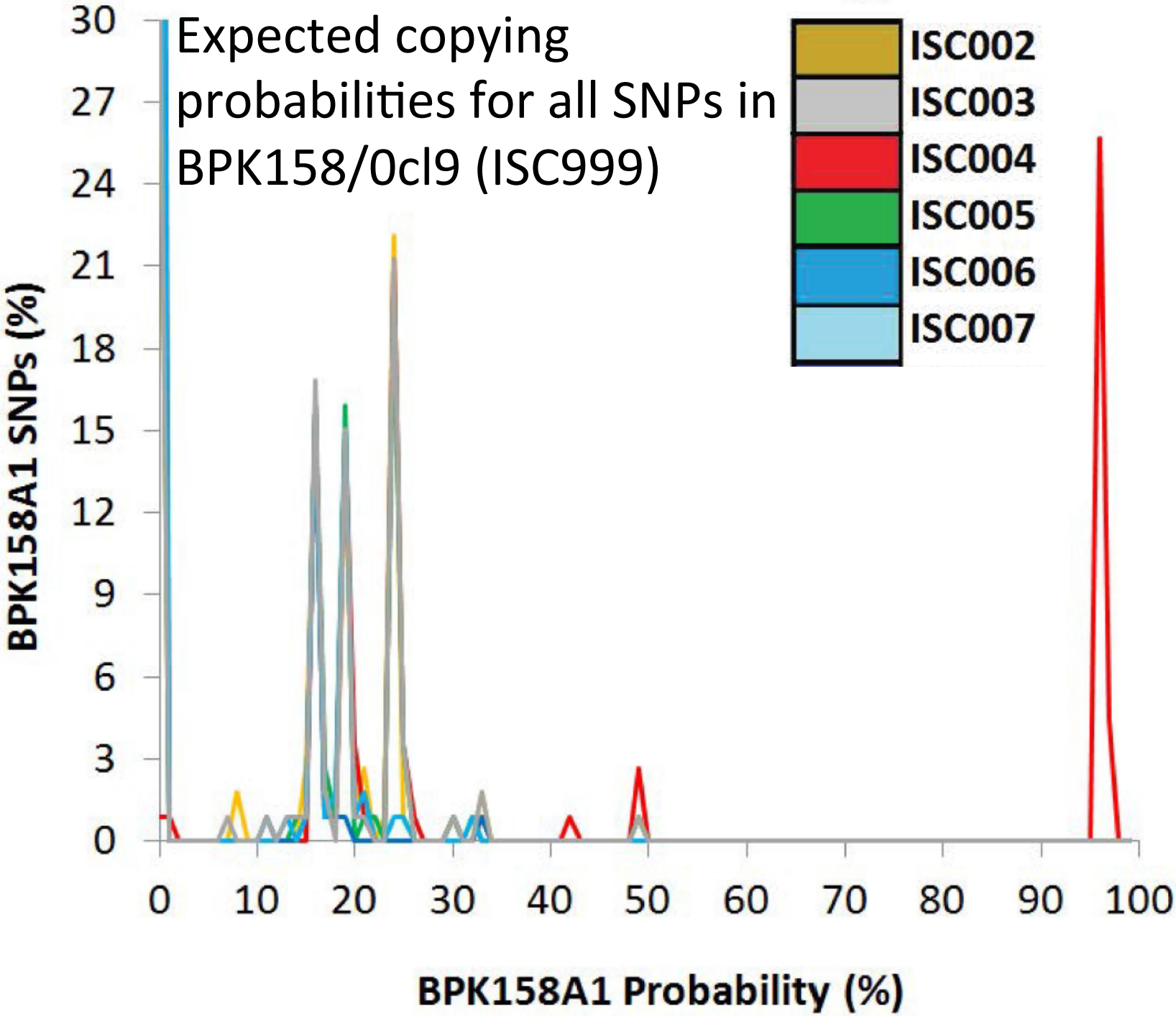
Painted genomes
610 informative SNPs

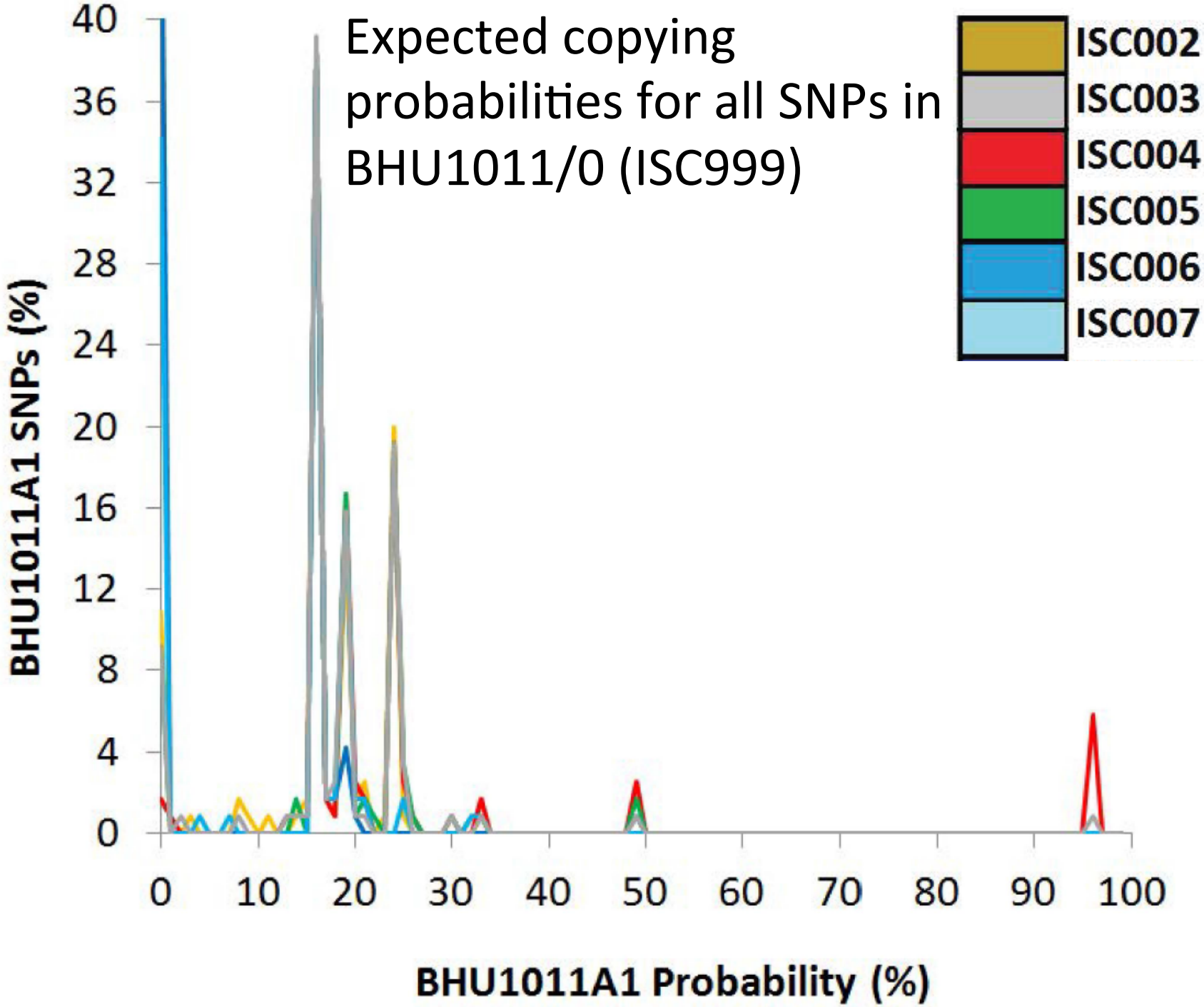


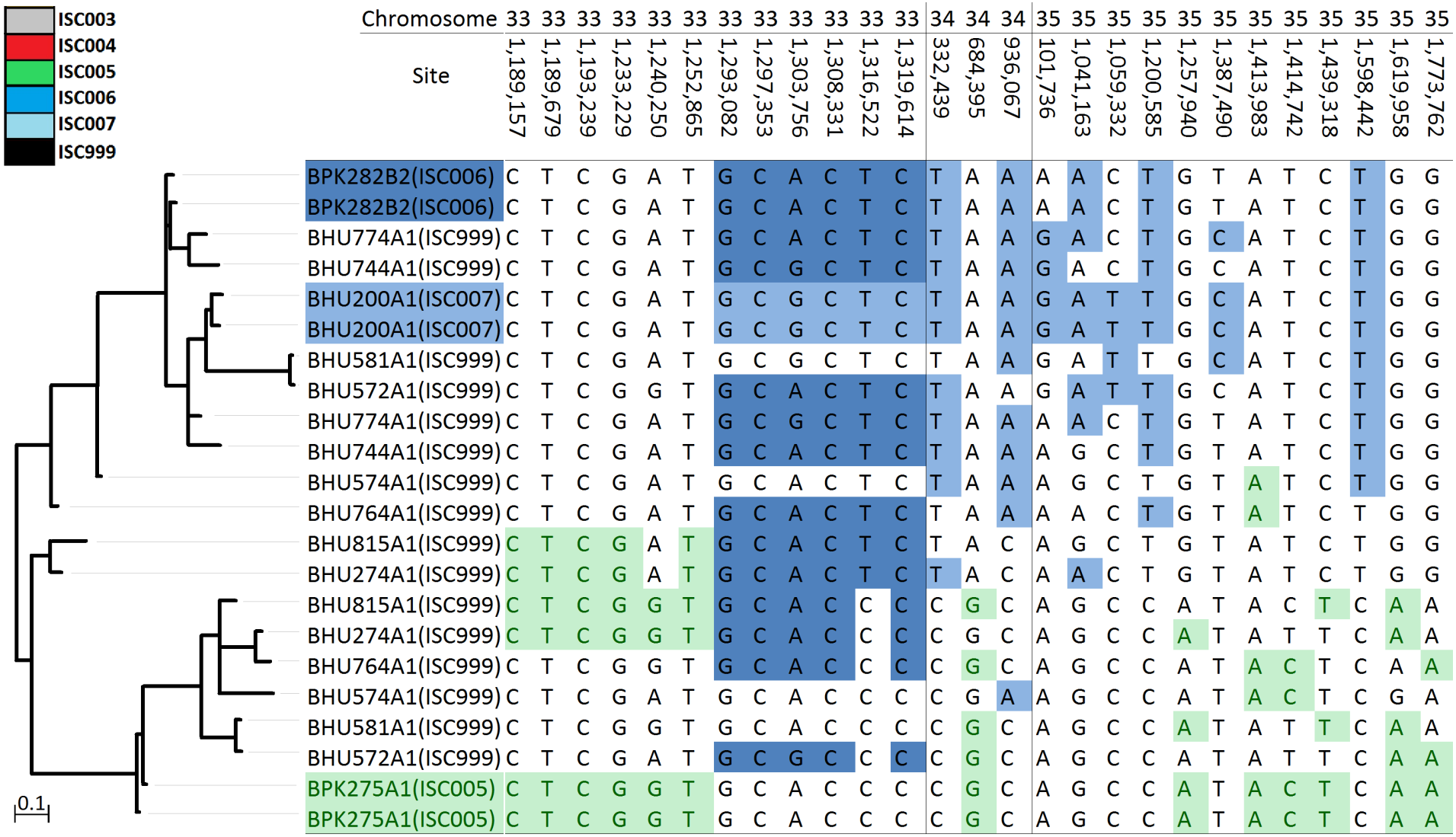
SNPs in ISC005-associated haplotypes using the sample-population model

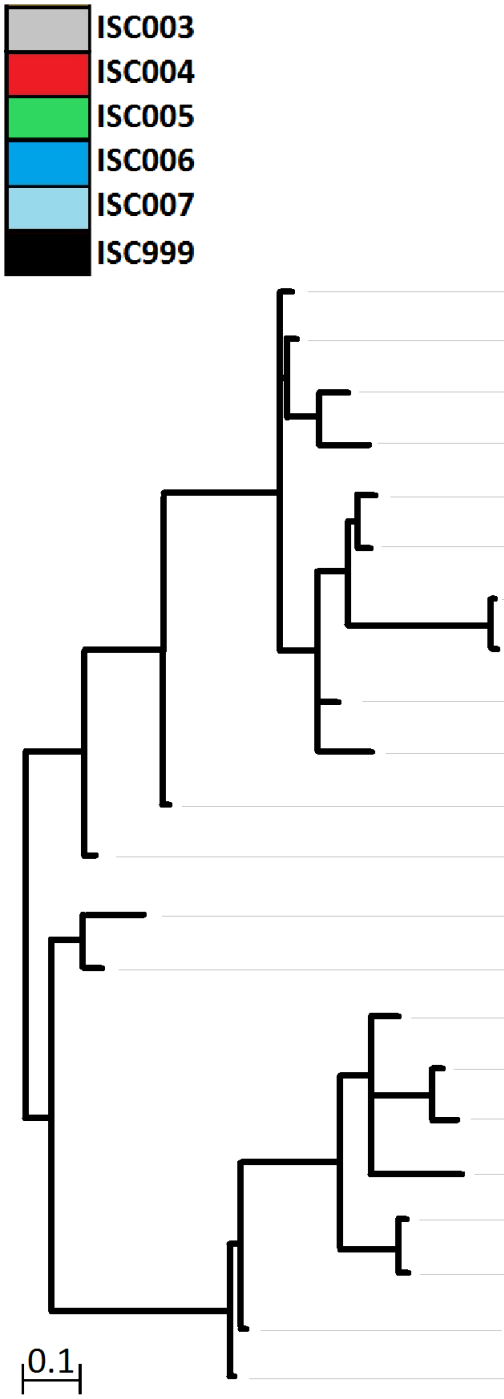



Expected copying probabilities for all SNPs in BPK275/0cl18 (ISC005)
ISC002
ISC003
ISC004
ISC005
ISC006
ISC007
BPK275A1 SNPs (%)
40
36
32
28
24
20
16
12
8
4
0
0
10
20
30
40
50
60
70
80
90
100
BPK275A1 Probability (%)


Expected copying probabilities for all SNPs in BPK158/0cl9 (ISC999)
ISC002
ISC003
ISC004
ISC005
ISC006
ISC007
BPK158A1 SNPs (%)
30
27
24
21
18
15
12
9
6
3
0
0
10
20
30
40
50
60
70
80
90
100
BPK158A1 Probability (%)


Expected copying probabilities for all SNPs in BHU1011/0 (ISC999)
ISC002
ISC003
ISC004
ISC005
ISC006
ISC007
BHU1011A1 SNPs (%)
40
36
32
28
24
20
16
12
8
4
0
0
10
20
30
40
50
60
70
80
90
100
BHU1011A1 Probability (%)


ISC003
ISC004
ISC005
ISC006
ISC007
ISC999
Chromosome 33 33 33 33 33 33 33 33 33 33 33 33 34 34 34 35 35 35 35 35 35 35 35 35 35 35 35 35
Site 1,189,157 1,189,679 1,193,239 1,233,229 1,240,250 1,252,865 1,293,082 1,297,353 1,303,756 1,308,331 1,316,522 1,319,614 332,439 684,395 936,067 101,736 1,041,163 1,059,332 1,200,585 1,257,940 1,387,490 1,413,983 1,414,742 1,439,318 1,598,442 1,619,958 1,773,762
BPK282B2(ISC006) C T C G A T G C A C T C T A A A A C T G T A T C T G G A T G G
BPK282B2(ISC006) C T C G A T G C A C T C T A A A A C T G T A T C T G G A T G G
BHU774A1(ISC999) C T C G A T G C A C T C T A A G A C T G C A T C T G G A T G G
BHU744A1(ISC999) C T C G A T G C G C T C T A A G A C T G C A T C T G G A T G G
BHU200A1(ISC007) C T C G A T G C G C T C T A A G A T T G C A T C T G G A T G G
BHU200A1(ISC007) C T C G A T G C G C T C T A A G A T T G C A T C T G G A T G G
BHU581A1(ISC999) C T C G A T G C G C T C T A A G A T T G C A T C T G G A T G G
BHU572A1(ISC999) C T C G G T G C A C T C T A A G A T T G C A T C T G G A T G G
BHU774A1(ISC999) C T C G A T G C G C T C T A A A A C T G T A T C T G G A T G G
BHU744A1(ISC999) C T C G A T G C A C T C T A A A G C T G T A T C T G G A T G G
BHU574A1(ISC999) C T C G A T G C A C T C T A A A G C T G T A A T C T G G A T G G
BHU764A1(ISC999) C T C G A T G C A C T C T A A A A C T G T A T C T G G A T G G
BHU815A1(ISC999) C T C G A T G C A C T C T A C A G C T G T A T C T G G A T G G
BHU274A1(ISC999) C T C G A T G C A C T C T A C A A A C T G T A T C T G G A T G G
BHU815A1(ISC999) C T C G G T G C A C C C C G C A G C C A A T C T G G A A A A
BHU274A1(ISC999) C T C G G T G C A C C C C G C A G C C A A T C T G G A A A A
BHU764A1(ISC999) C T C G G T G C A C C C C G C A G C C A A T C T G G A A A A
BHU574A1(ISC999) C T C G A T G C A C C C C G C A A A C A T C T G G A A A A
BHU581A1(ISC999) C T C G G T G C A C C C C G C G C A G C C A T A T T C A A A
BHU572A1(ISC999) C T C G A T G C G C C C C G C G C A G C C A T A T T C A A A
BPK275A1(ISC005) C T C G G T G C A C C C C G C G C A G C C A T A C T C A A A
BPK275A1(ISC005) C T C G G T G C A C C C C G C G C A G C C A T A C T C A A A
0.1



Chromosome	35	35	35	35	35	35	35	35	35	35	35	35
Site	101,736	1,041,163	1,059,332	1,200,585	1,257,940	1,387,490	1,413,983	1,414,742	1,439,318	1,598,442	1,619,958	1,773,762
BPK282B2(ISC006)	A	A	C	T	G	T	A	T	C	T	G	G
BPK282B2(ISC006)	A	A	C	T	G	T	A	T	C	T	G	G
BHU774A1(ISC999)	G	A	C	T	G	C	A	T	C	T	G	G
BHU744A1(ISC999)	G	A	C	T	G	C	A	T	C	T	G	G
BHU200A1(ISC007)	G	A	T	T	G	C	A	T	C	T	G	G
BHU200A1(ISC007)	G	A	T	T	G	C	A	T	C	T	G	G
BHU581A1(ISC999)	G	A	T	T	G	C	A	T	C	T	G	G
BHU572A1(ISC999)	G	A	T	T	G	C	A	T	C	T	G	G
BHU774A1(ISC999)	A	A	C	T	G	T	A	T	C	T	G	G
BHU744A1(ISC999)	A	G	C	T	G	T	A	T	C	T	G	G
BHU574A1(ISC999)	A	G	C	T	G	T	A	T	C	T	G	G
BHU764A1(ISC999)	A	A	C	T	G	T	A	T	C	T	G	G
BHU815A1(ISC999)	A	G	C	T	G	T	A	T	C	T	G	G
BHU274A1(ISC999)	A	A	C	T	G	T	A	T	C	T	G	G
BHU815A1(ISC999)	A	G	C	C	A	T	A	C	T	C	A	A
BHU274A1(ISC999)	A	G	C	C	A	T	A	T	T	C	A	A
BHU764A1(ISC999)	A	G	C	C	A	T	A	C	T	C	A	A
BHU574A1(ISC999)	A	G	C	C	A	T	A	C	T	C	G	A
BHU581A1(ISC999)	A	G	C	C	A	T	A	T	T	C	A	A
BHU572A1(ISC999)	A	G	C	C	A	T	A	T	T	C	A	A
BPK275A1(ISC005)	A	G	C	C	A	T	A	C	T	C	A	A
BPK275A1(ISC005)	A	G	C	C	A	T	A	C	T	C	A	A

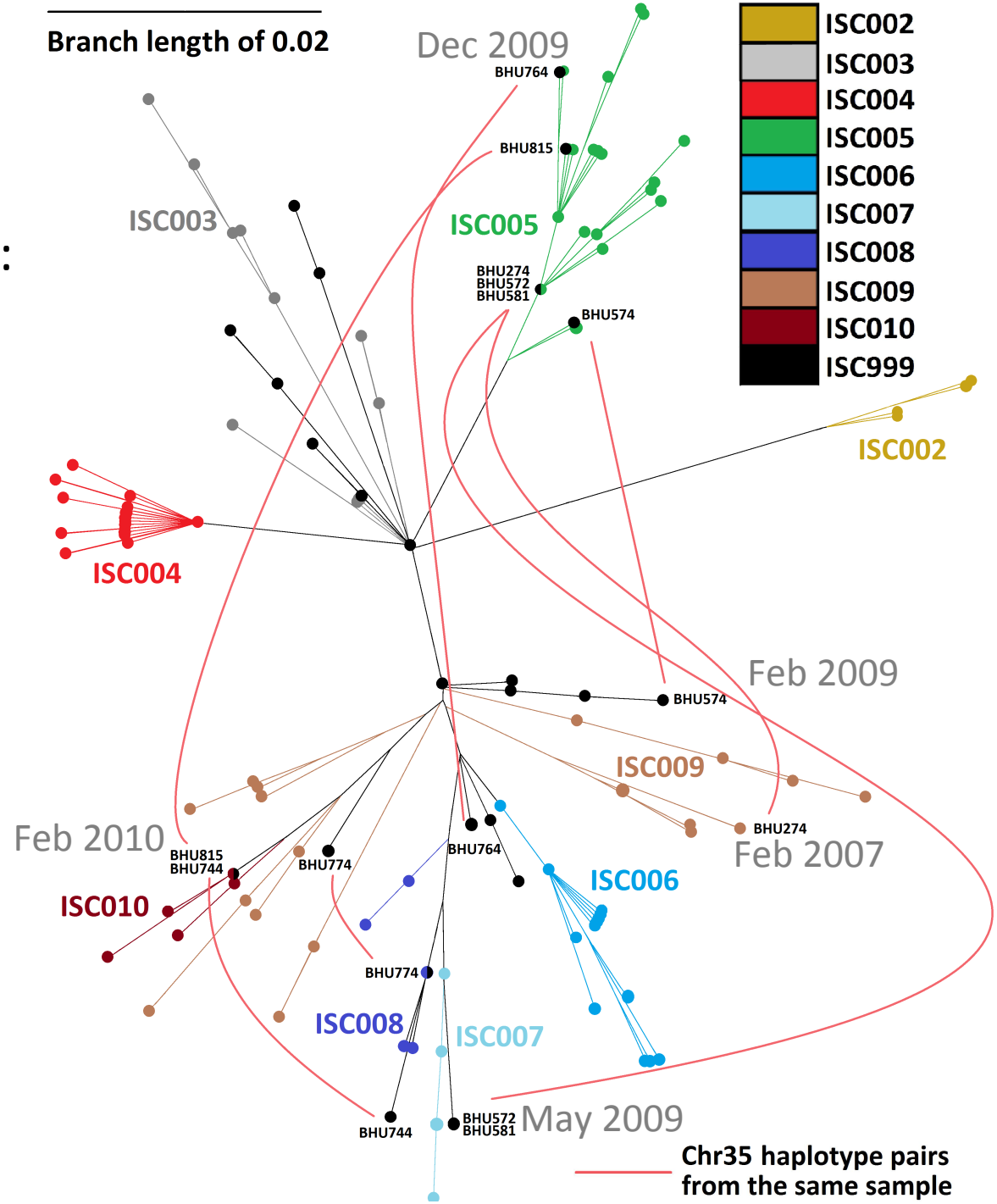
Phylogeny of phased chr35 haplotypes of 191 samples

We found hybrids between:

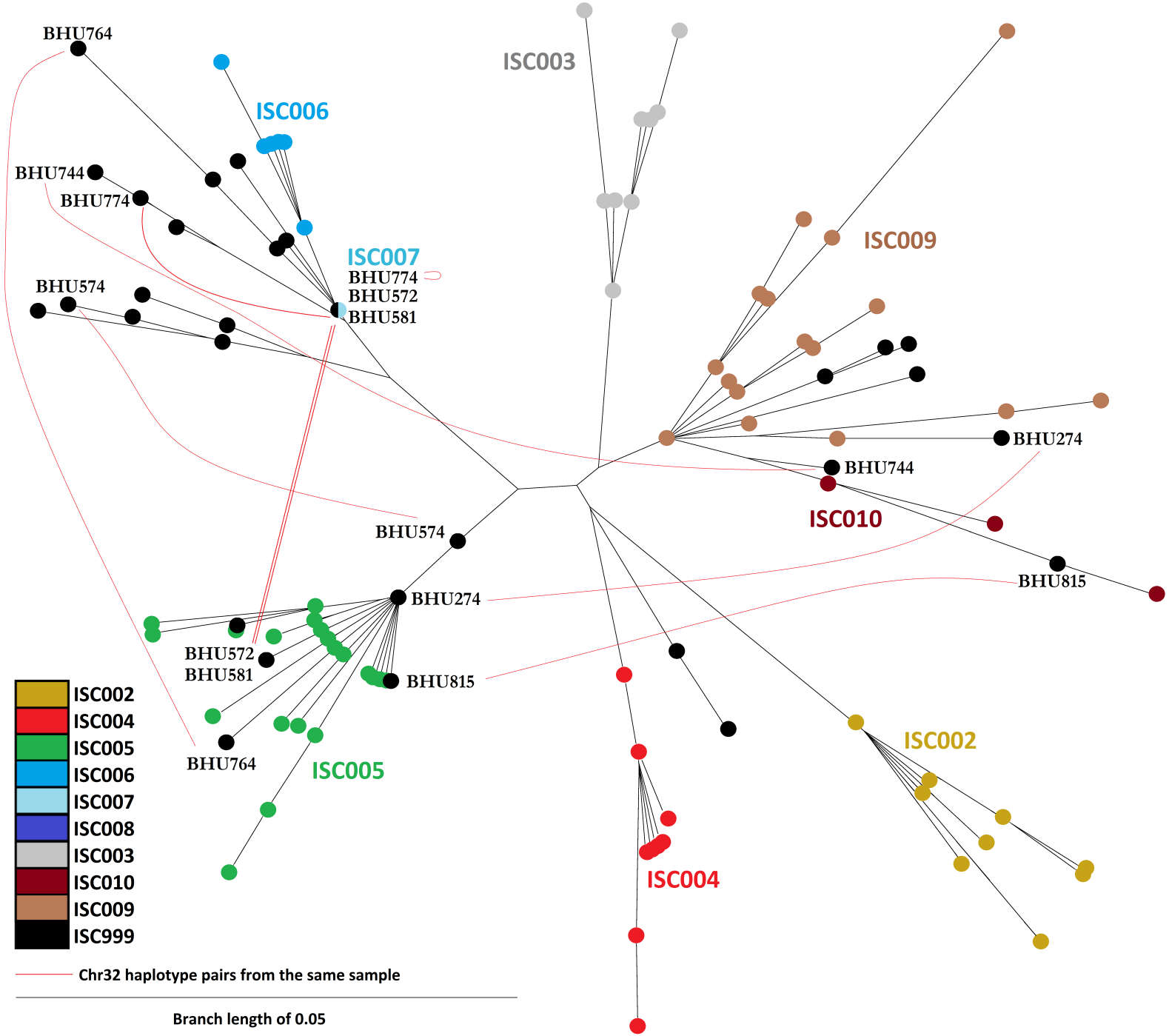
ISC005-ISC006
ISC005-ISC007
ISC005-ISC009
ISC005-ISC010
ISC008-ISC009
ISC008-ISC010

Many independent events
Many half-siblings

Recombination within groups



Phylogeny of phased chr32 haplotypes of 191 samples



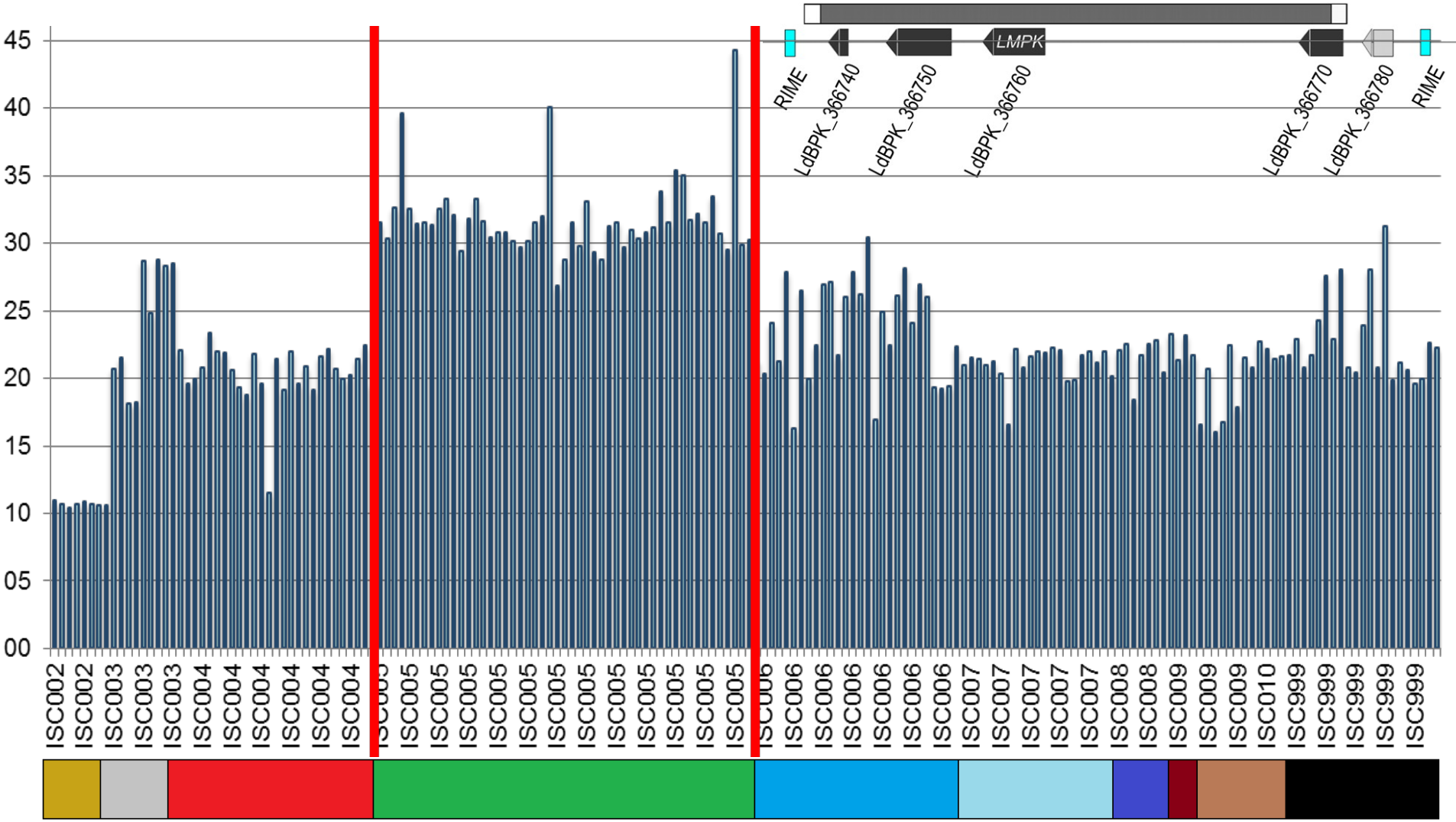
5. Genome plasticity and drug tolerance

MAPK (3) & H-locus (2) episomes have het SNPs

17 out of 167 small indels in CDS

CNVs duplications, tandem arrays, sub-telomeric
3.6 Mb (11% of genome)

5. Genome plasticity and drug tolerance MAPK-locus (chr36) and SSG



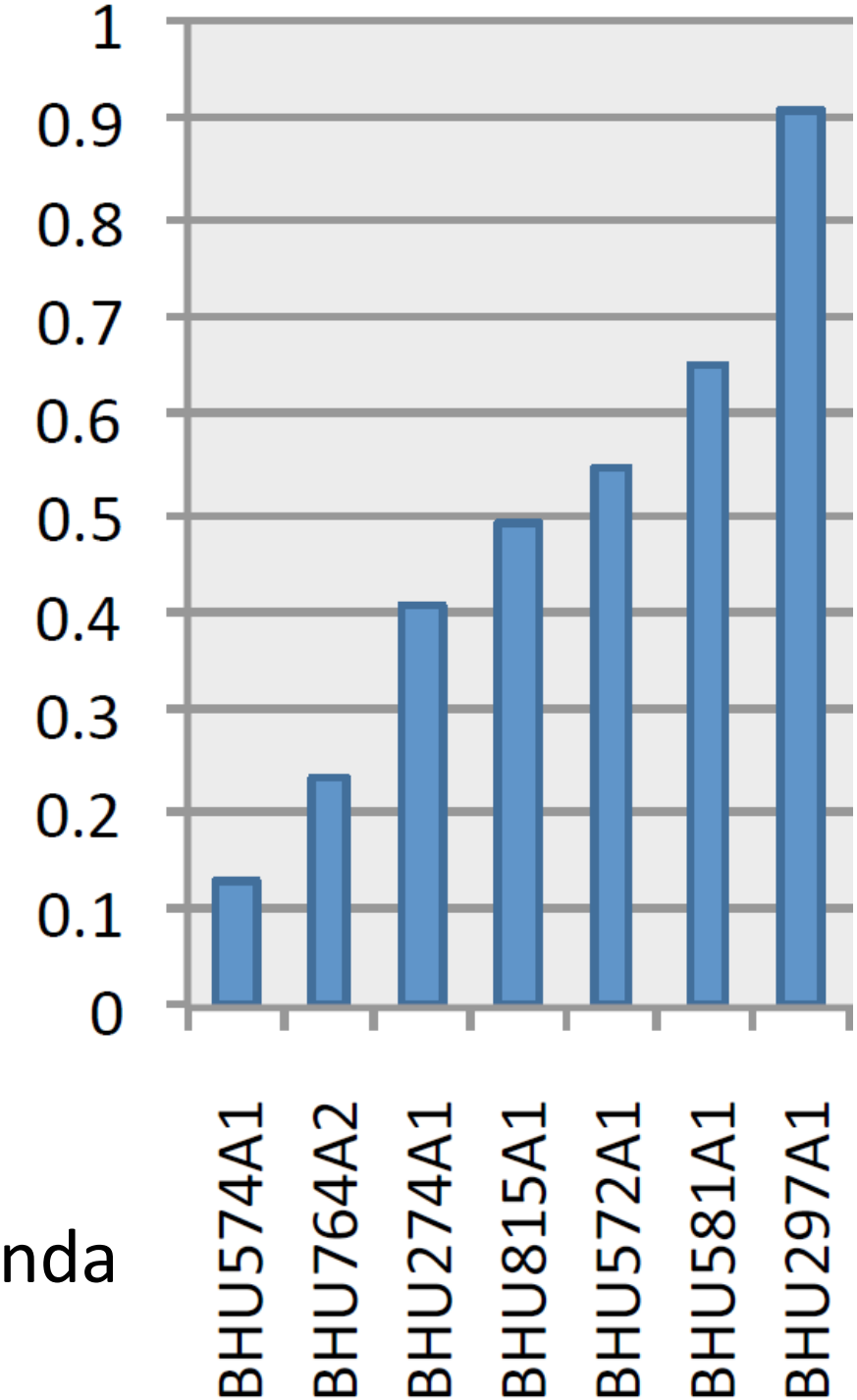
5. Genome plasticity and drug tolerance

2 bp insertion
=> premature stop
Aquaglyceroporin 1 (AQP1)

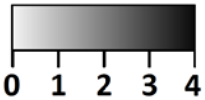
Antimonial transporter
Gourbal et al 2004 J Biol Chem

All ISC005 strains = near-fixed

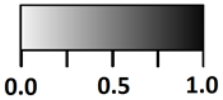
6 hybrids = kinda



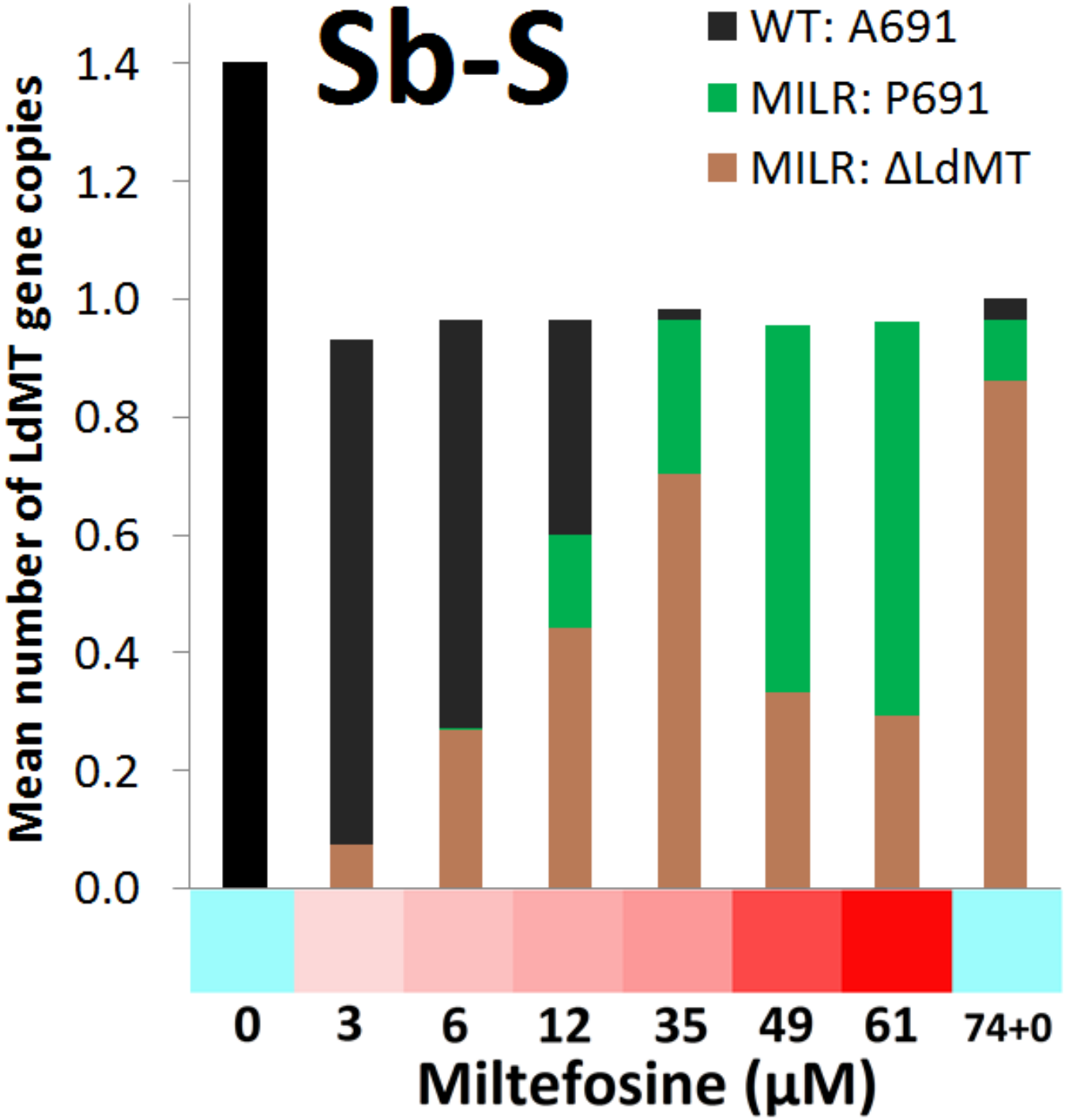
Region	Change		Strain
LdBPK_02	down	SOMY	Sb-R
LdBPK_08	down	SOMY	Sb-R
LdBPK_09	up	SOMY	Sb-R
LdBPK_13	down	SOMY	Sb-S
LdBPK_13	down	SOMY	Sb-R
LdBPK_131590	Δ	SV	Sb-S
LdBPK_131590	A691P	SNP	Sb-S
LdBPK_131590	E197D	SNP	Sb-R
LdBPK_131610	A1153	SNP	Sb-R
LdBPK_23	up	SOMY	Sb-S
LdBPK_272290-2440	up	SV	Sb-R
LdBPK_31	down	SOMY	Sb-R
LdBPK_310600-0870	down	SV	Sb-R
LdBPK_33	down	SOMY	Sb-S
LdBPK_33	down	SOMY	Sb-R
LdBPK_35	up	SOMY	Sb-R
LdBPK_350030-0700	up	SV	Sb-R
LdBPK_353570-5320	down	SV	Sb-R



Copy number of chromosome/variant



SNP allele frequency


Sb-S
WT: A691
MILR: P691
MILR: ΔLdMT
Mean number of LdMT gene copies
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
3
6
12
35
49
61
74+0
Miltefosine (µM)

Summary

① Leishmaniasis is an understudied problem

② *Leishmania* produce hybrids at every scale

③ Many mechanisms of drug tolerance and origins

④ Drug-R: Gene? Trunk? Genealogies? L-statistics?