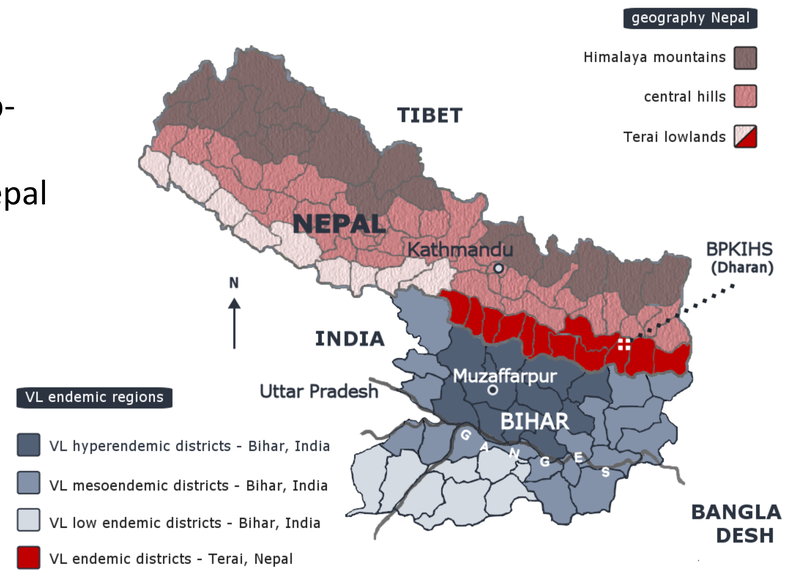
Genomic ancestry blocks decipher population gene flow and admixture in monomorphic *Leishmania donovani*

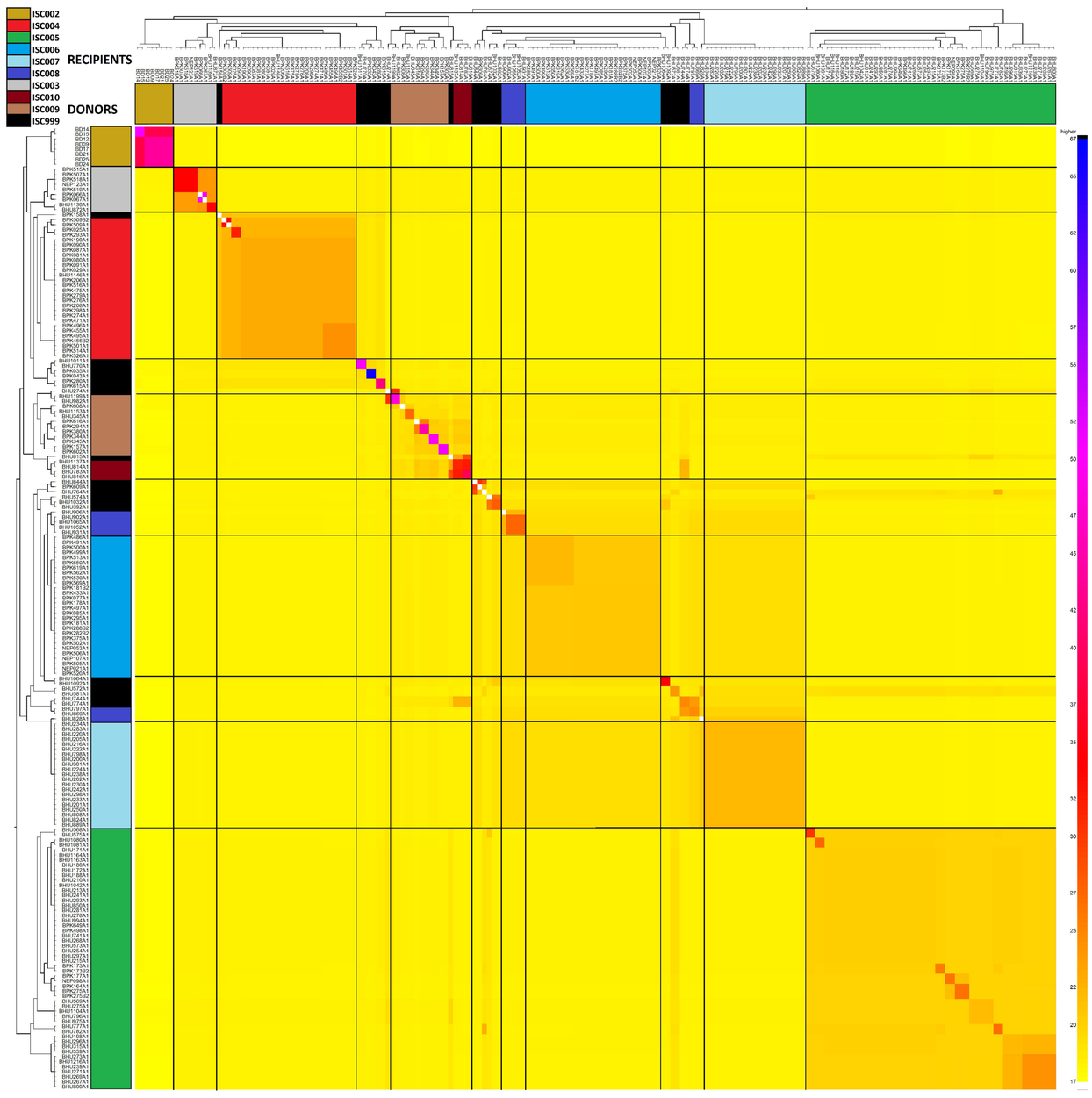
Downing T[1], Imamura H[2,3], Vanaerschot M[2], Mannaert A[2], Schonian G[4], Dujardin JC[2], Berriman M[3], Cotton JA[3].

[1] College of Science, National University of Ireland Galway. [2] Institute of Tropical Medicine, Antwerp, Belgium. [3] Wellcome Trust Sanger Institute, Cambridge, UK. [4] Charitè University, Berlin, Germany.

VL (visceral leishmaniasis) is endemic in the ISC (Indian sub-continent), especially in Bihar (India) and the lowlands of Nepal (see map right).

We genome-sequenced 191 samples belonging to a single genetically homogeneous outbreak of *Leishmania donovani* isolated during 2002-11.





GENOME DIVERSITY: Just 2,418 SNPs were variable within this group across 26 Mb mappable sites in a 32 Mb genome, equivalent to 3.4 SNPs/Mb on average between pairs. Although 46.5% of SNPs were in genes, 62.6% of all SNPs were unique to single samples, and the ratio of nonsynonymous (N) to synonymous (S) SNPs (N/S=749/376=1.99) was far higher than the ancestral rate using *L. infantum* (1.06), consistent with a recent expansion of population size or range.

Allele frequency correlations with *Structure* identified 6 populations (ISC002-7) and 4 other genetically heterogenous groups (ISC008-999; see Table right). The number of samples taken (N) did not reflect directly the number of SNPs or haplotypes. This is shown for K=4, K=7 and K=9 populations below.

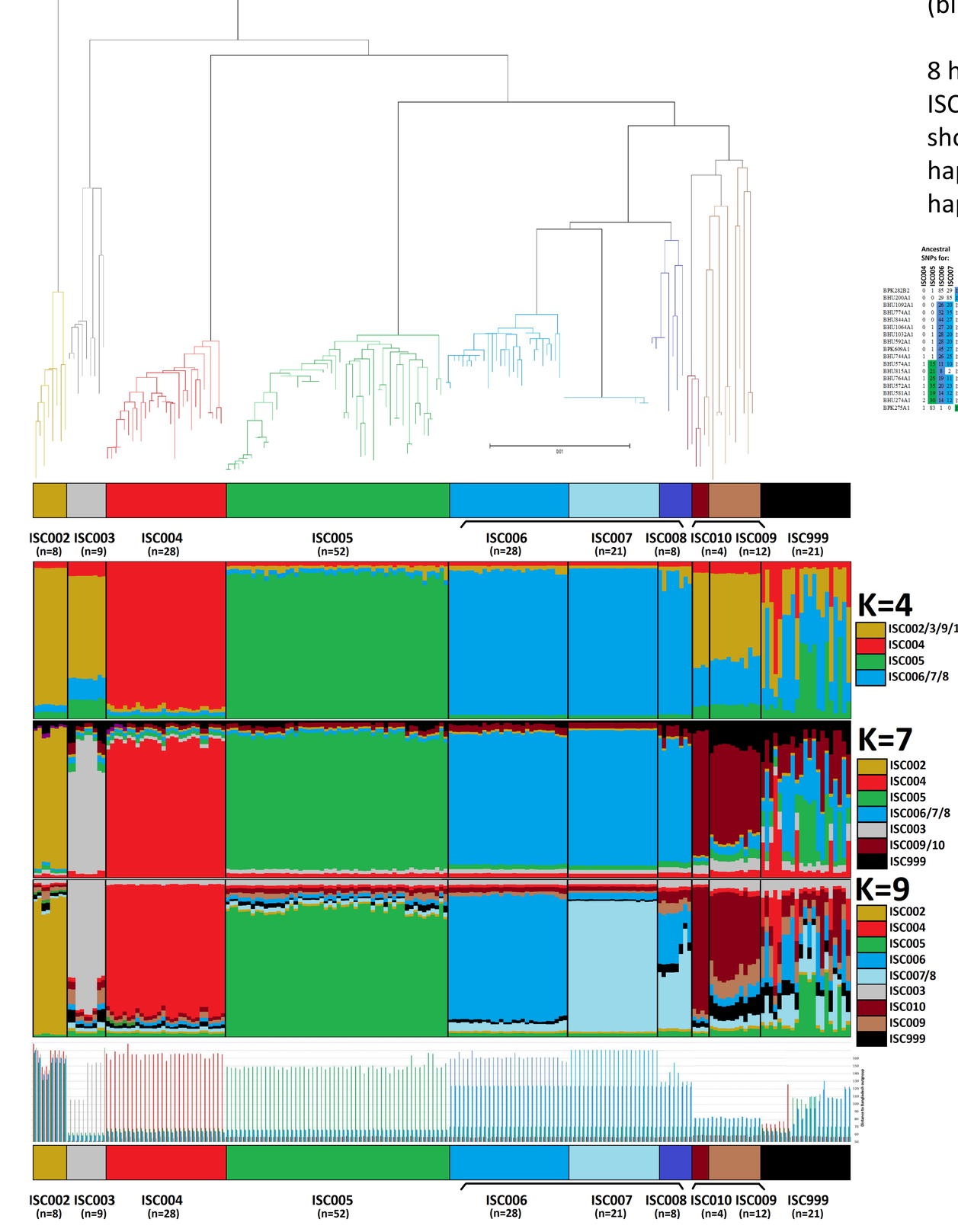
A *MEGA* maximum composite likelihood tree (below) was constructed (excluding ISC999).

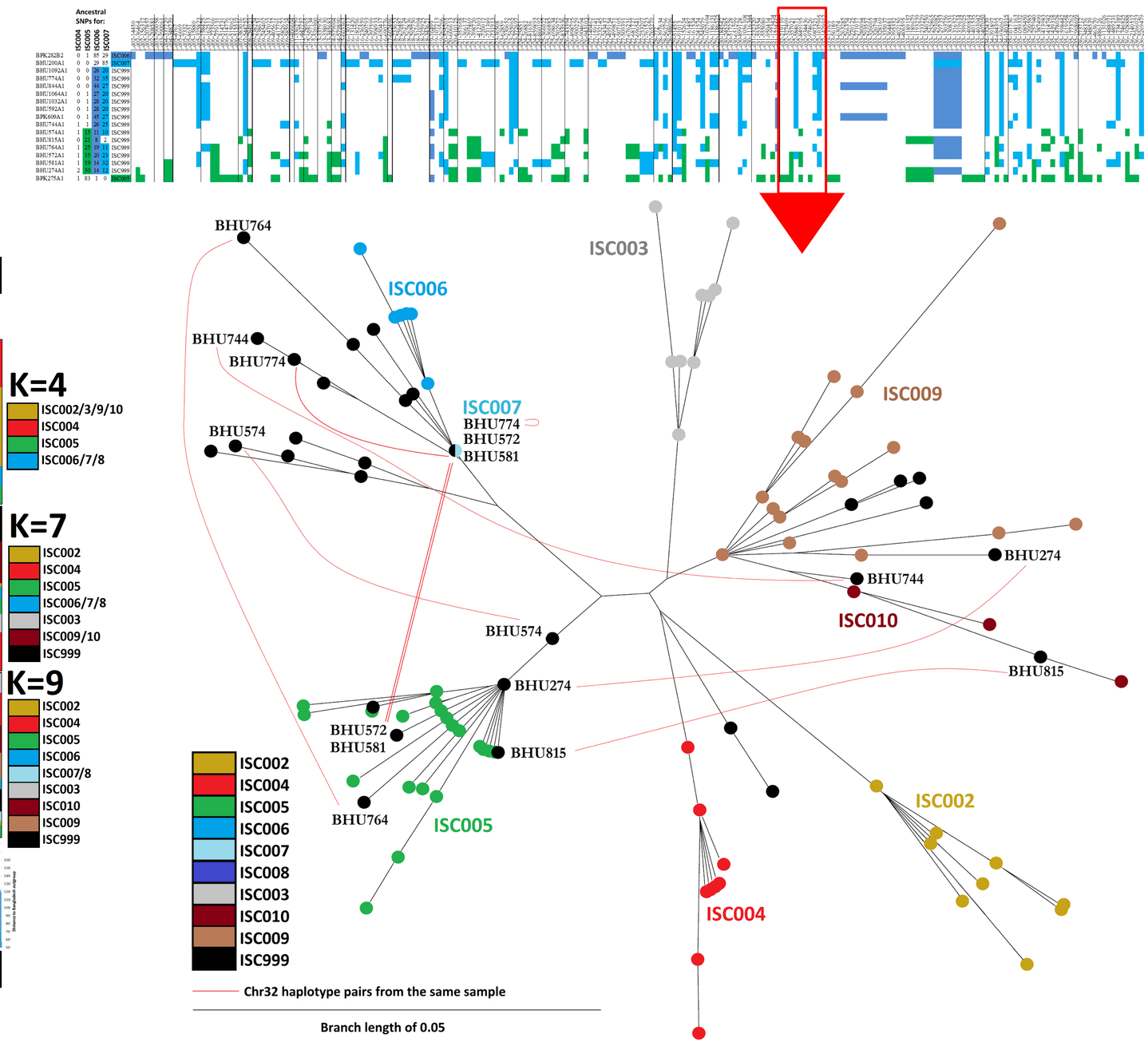
	N	SNPs	Haplotypes
ISC002	8	112	8
ISC004	9	227	9
ISC005	28	256	28
ISC006	52	284	41
ISC007	28	109	28
ISC008	21	2	3
ISC003	8	98	7
ISC009	4	49	4
ISC010	12	350	12
ISC999	21	502	21

HAPLOTYPE SHARING: The haplotypes shared by the 8 ISC999 hybrids and the main populations were discovered in a comparison of all samples with *FineStucture*. Numbers of haplotypes shared between pairs scales from low (yellow) to intermediate (red) to high (blue) shown for haplotypes donated (upper triangle) and received (lower triangle).

KEY FINDING: Our approach distinguished 8 samples in ISC999 that were hybrids of other populations from the 13 remaining, which were rare types.

8 hybrids were mixes of ISC005 (green) and either ISC006/7 (light & dark blue) or ISC009/10 (light & dark brown) using ancestry-informative SNPs from *ChromoPainter* shown in the table below for representative samples. BioNJ trees of the phased haplotypes for each chromosome (eg chr32 below) illustrated that the 8 hybrids had haplotypes from known populations and were not unique.





ANCESTRY: The proportions of ancestry were apparent from the genetic distances of 191 strains compared to 5 representative strains from the distinct populations (above): BPK067/0cl2 (ISC003, grey), BPK087/0cl11 (ISC004, red), BPK275/0cl18 (ISC005, green), BPK282/0cl4 (ISC006, dark blue), BHU200/0 (ISC006, light blue). These were scaled by BD09 from ISC002 (y-axis).

CONCLUSION: The genetic origins of new outbreaks of VL can be resolved: we discovered hybrids from as few as 60 genome-wide SNPs and tracked the spread of drug-resistant ISC005 genotypes in the ISC since 2002, and linked this to aneuploidy and structural variation.

See also P40 "Resistome of *Leishmania donovani*: multi-factorial genomic origin of clinical antimony resistance".
This work was funded by the Wellcome Trust Sanger Institute and FP7 Kaladrug-R project (grant 222895).

Email: Tim.Downing@nuigalway.ie

NUI Galway
OÉ Gaillimh