eLIFE
elifesciences.org

SHORT REPORT

Figures and figure supplements

Mitochondrial genetic diversity, selection and recombination in a canine transmissible cancer

Andrea Strakova *et al*

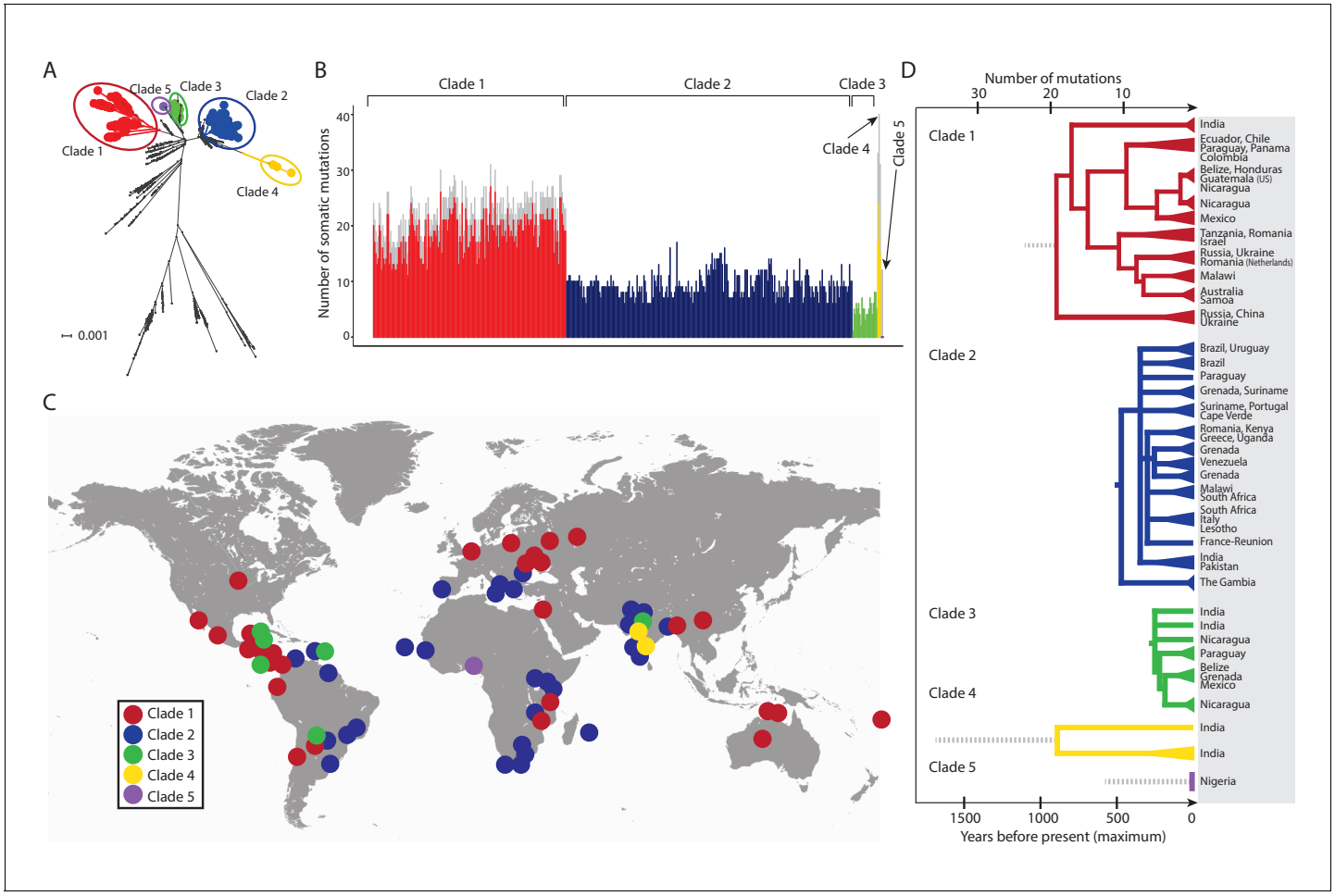

Figure 1. CTVT has acquired mtDNA by horizontal transfer at least five times. (**A**) Maximum likelihood phylogenetic tree constructed with complete mtDNA sequences from 449 CTVT tumours and 590 dogs. Coloured and black dots represent CTVT and dog mtDNA respectively. Scale bar indicates base substitutions per site. (**B**) Number of somatic substitution mutations per CTVT tumour. Coloured bars indicate somatic mutations acquired by each tumour since mtDNA capture. Grey bars indicate substitutions absent from normal dog mtDNA haplotypes but common to all tumours within a clade; thus the early somatic or rare germline status of these variants is unknown. (**C**) Geographical distribution of clades. Coloured dots represent locations from which one or more CTVT tumours were collected. (**D**) Simplified representation of maximum likelihood phylogenetic trees for each clade. Trees illustrate nodes with bootstrap support >60, and shaded triangles represent coalescence of individual branches within each group. Two tumours were collected in the United States and the Netherlands respectively from dogs imported from Guatemala and Romania. Discontinuous grey lines represent contributions of substitutions absent from normal dog mtDNA haplotypes but common to all tumours within a clade. Assuming a constant accumulation of mutations within and between clades, approximate number of somatic mutations and estimated timing is shown. Maximum likelihood trees upon which these representations are based are found in ***Figure 1—source data 2***.

DOI: 10.7554/eLife.14552.003

The following source data is available for figure 1:

Source data 1. Maximum likelihood phylogenetic tree of CTVT mtDNA.

DOI: 10.7554/eLife.14552.004

Source data 2. Maximum likelihood phylogenetic trees for CTVT clades 1 to 5.

DOI: 10.7554/eLife.14552.005

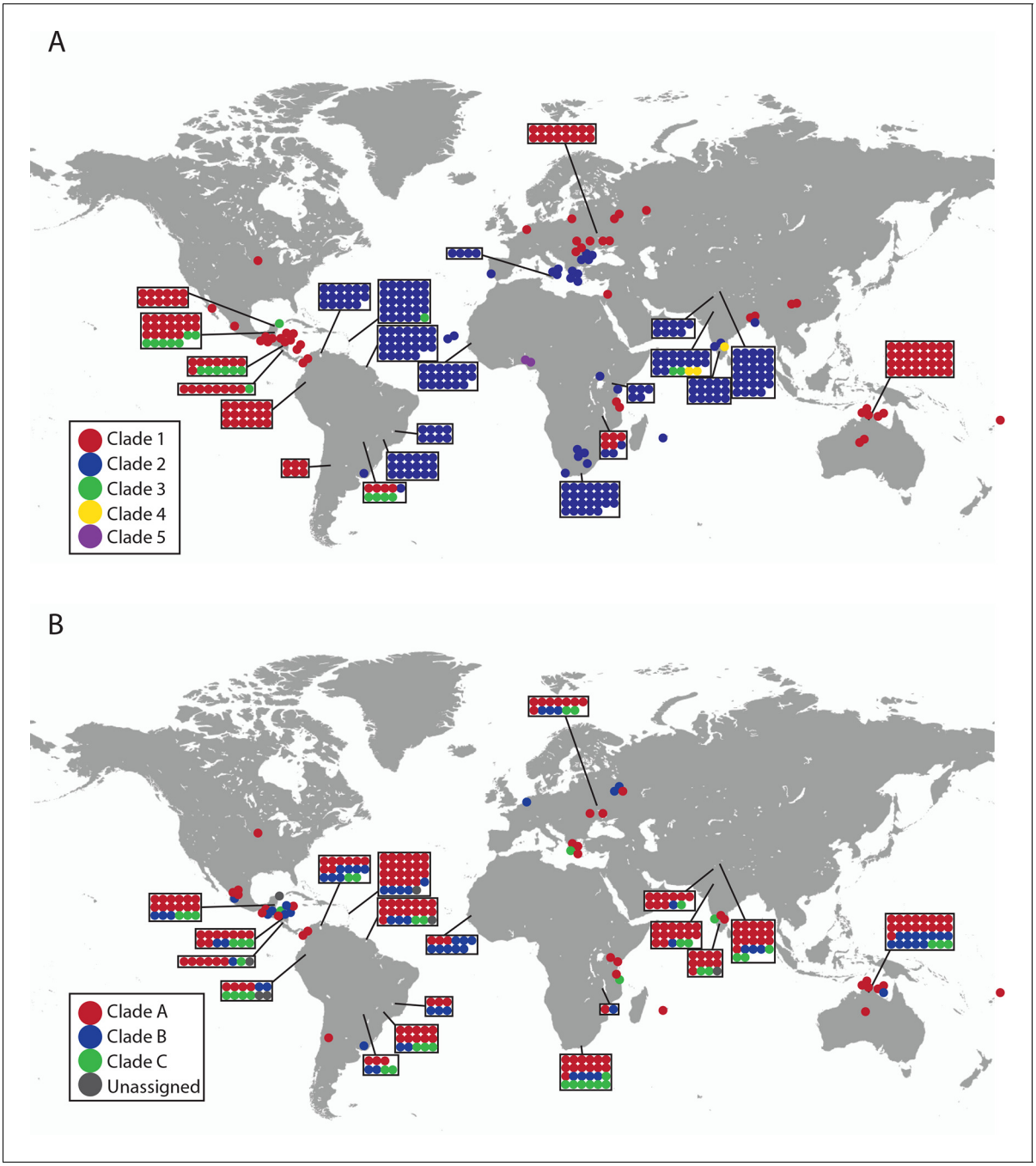

Figure 1—figure supplement 1. Geographical locations and mtDNA clades for CTVT tumours and hosts. Each dot represents the location of (**A**) CTVT tumours, coloured by CTVT mtDNA clade; or (**B**) CTVT hosts, coloured by dog mtDNA clade.
DOI: 10.7554/eLife.14552.006

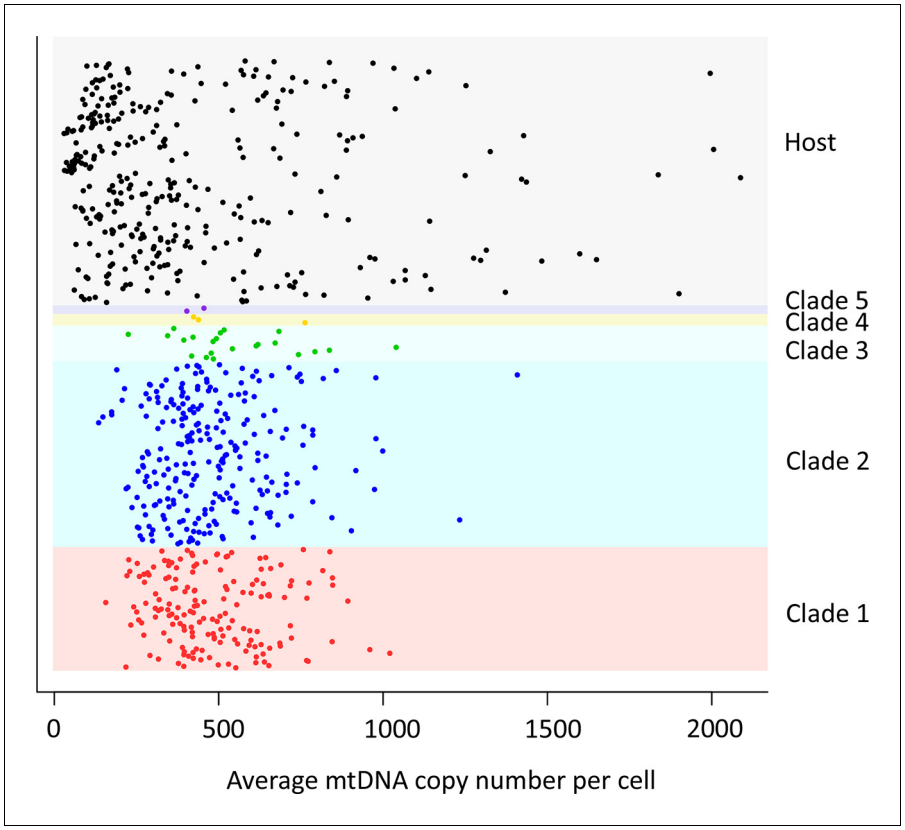

Figure 1—figure supplement 2. mtDNA copy number in CTVT. MtDNA copy number was estimated by normalising mtDNA sequence coverage to whole genome sequence coverage (***Supplementary file 2A***). Each point represents an individual tumour (labelled by clade) or host. MtDNA copy number in tumours was not normalised for host contamination. Host and tumour samples with average MT coverage >300X (see ***Supplementary file 2A***) were excluded from the analysis and from calculation of average number of mtDNA copies per cell.

DOI: 10.7554/eLife.14552.007

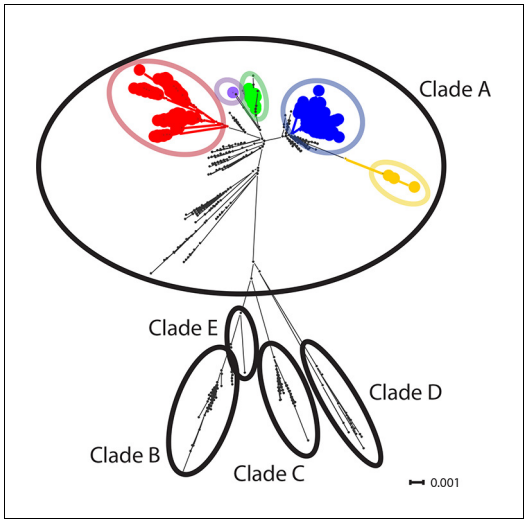

Figure 1—figure supplement 3. CTVT mtDNA clades 1 to 5 all arose from dog mtDNA clade A. Maximum likelihood phylogenetic tree constructed with complete mtDNA sequences from 449 CTVT tumours and 590 dogs. Coloured and black dots represent CTVT and dog mtDNA respectively (CTVT mtDNA clade colours are represented as in ***Figure 1A***). Dog mtDNA clades A to E are labelled (***Savolainen et al., 2002***; ***Vila et al., 1997***). Scale bar indicates base substitutions per site.

DOI: 10.7554/eLife.14552.008

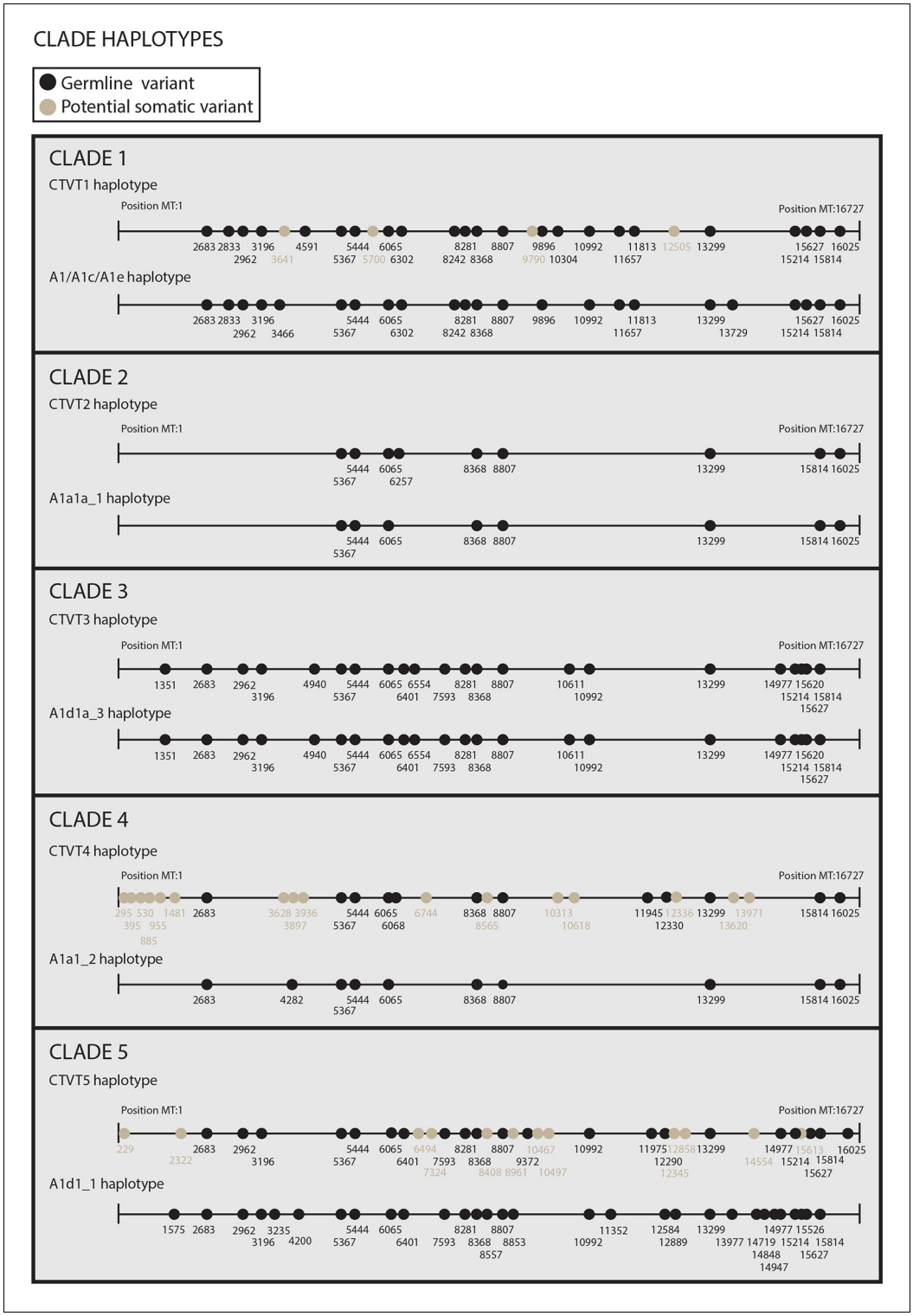

Figure 1—figure supplement 4. Reconstructed donor haplotypes for CTVT mtDNA clades 1 to 5. Diagrams representing the likely donor haplotype for each of the CTVT mtDNA clades 1 to 5. The coordinates for each substitution variant position are shown, and substitutions are colour-coded either as 'germline' (i.e. they are present in all tumours within a clade and are found in the most closely related dog mtDNA haplotype, which is represented below each of the clade diagrams or they are found in the most closely related dog mtDNA haplotype only); or 'potential somatic' (i.e. they are present in all tumours within a clade but are not found in the most closely related dog mtDNA haplotype).

Figure 1—figure supplement 4 continued on next page

Figure 1—figure supplement 4 continued

DOI: 10.7554/eLife.14552.009

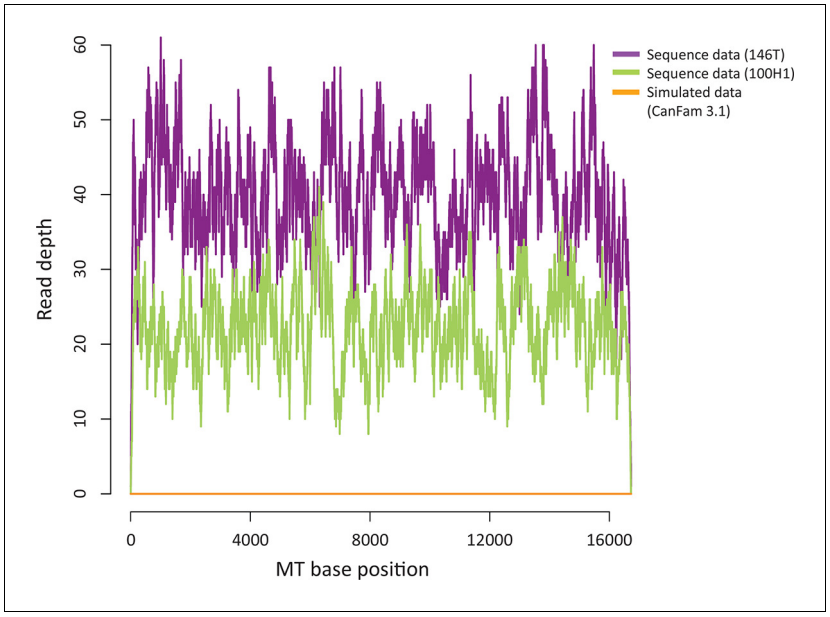

Figure 1—figure supplement 5. Sequence contribution of nuclear-encoded mtDNA (NuMTs). Sequence read depth across the MT genome for a representative CTVT tumour (146T) and host (100H1) sequenced in this study to ~0.3X whole genome average coverage. This is compared with sequence read depth for simulated reads from CanFam3.1 (excluding the MT chromosome); reads were simulated to ~0.3X whole genome average coverage.
DOI: 10.7554/eLife.14552.010

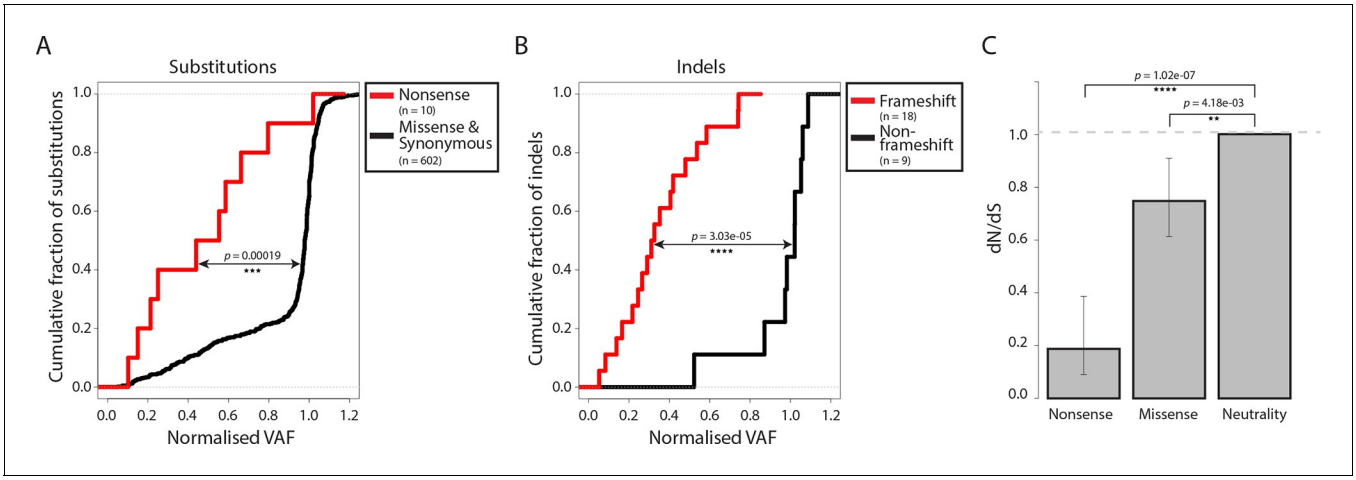

Figure 2. Negative selection operates to prevent the accumulation of gene-disrupting mutations in CTVT. Cumulative distribution functions for variant allele fraction (VAF) for gene-disrupting (**A**) substitutions and (**B**) indels. *P*-values were calculated using two-sample Kolmogorov-Smirnov tests. (**C**) dN/dS for somatic nonsense and missense substitutions. *P*-values were calculated using a likelihood ratio test with parameters estimated using a Poisson model. Error bars indicate 95 percent confidence intervals.

DOI: 10.7554/eLife.14552.011

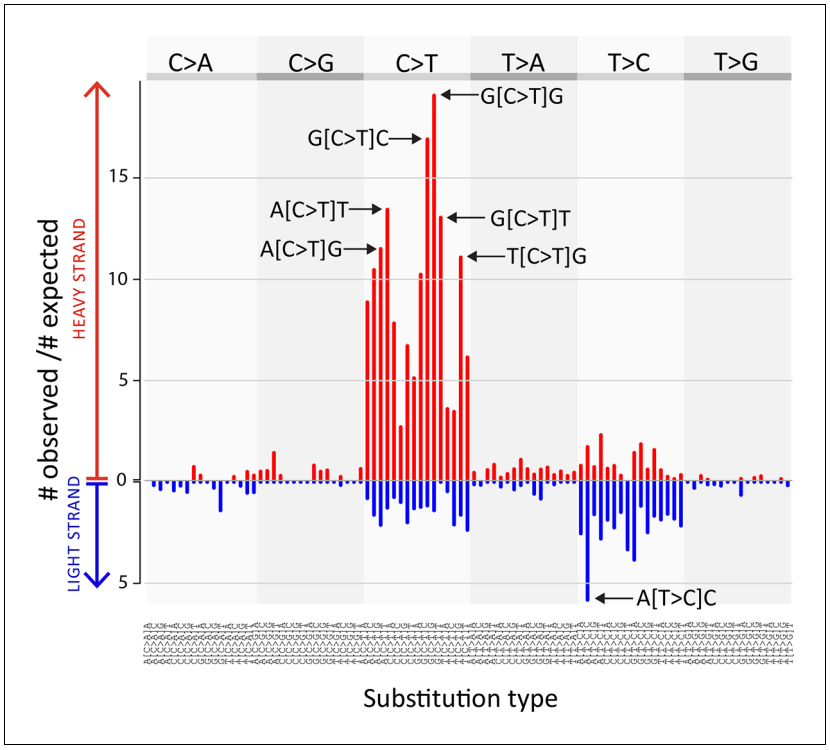

Figure 2—figure supplement 1. CTVT mtDNA somatic mutation spectrum. CTVT somatic mutations displayed by mutation type (in pyrimidine context) with 5′ and 3′ context and strand. Each of 96 mutation classes is displayed on the horizontal axis, with mutations occurring on the heavy strand displayed in red on the positive axis, and light strand mutations displayed in blue on the negative axis. The normalised substitution rate represents the (number of observed)/(number of expected) mutations, given mtDNA genome triplet content. Distinctive peaks are individually labelled. Only mutations on the 'conservative somatic list' were used (see Materials and methods and ***Supplementary file 4C***).

DOI: 10.7554/eLife.14552.012

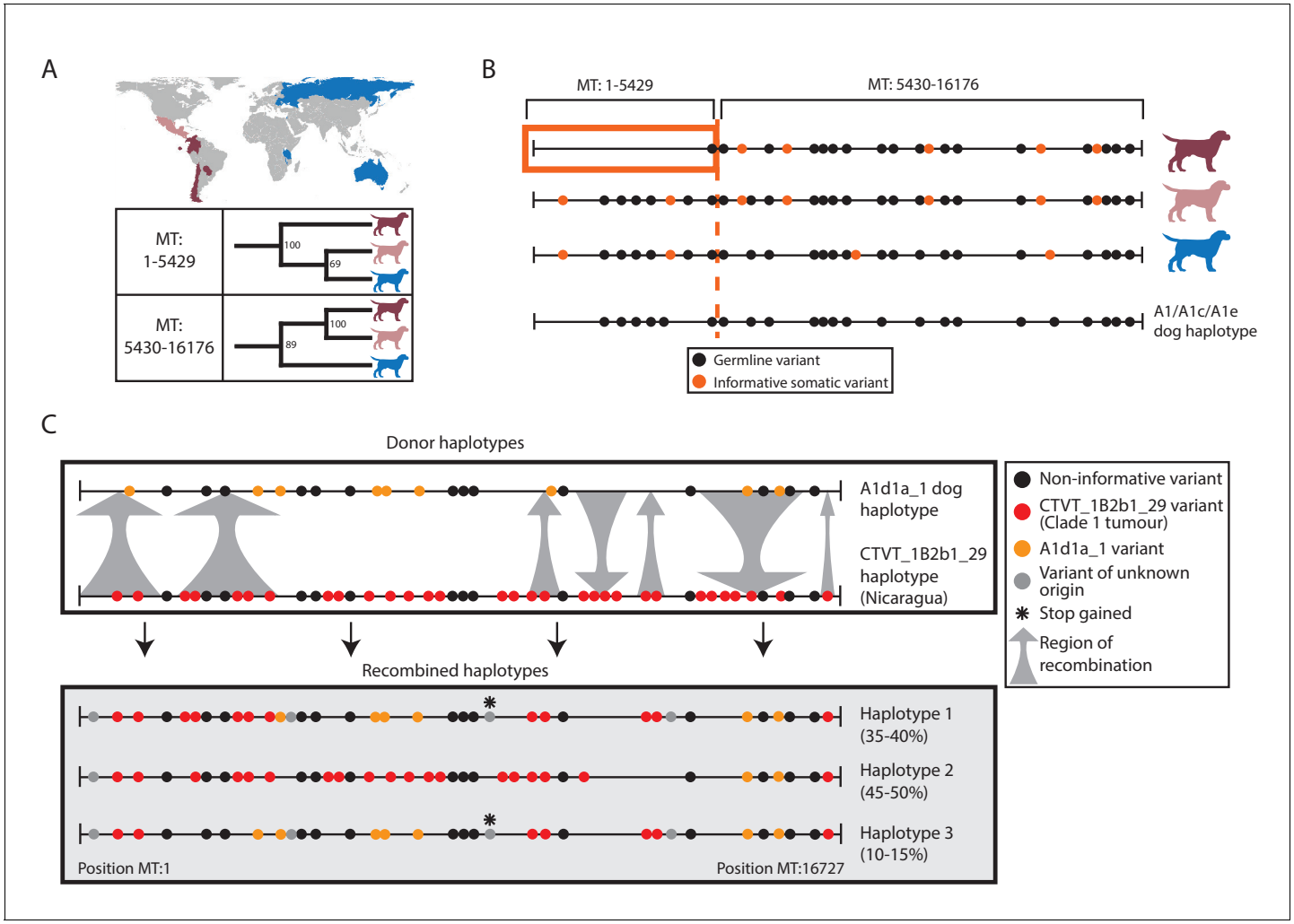

Figure 3. Ancient and modern mtDNA recombination in CTVT. (**A**) Maximum likelihood phylogenetic trees constructed using segments MT:1–5429 and MT:5430–16176 from clade 1 CTVT mtDNAs. Three clade 1 mtDNA haplotype groups are represented by coloured dog silhouettes, and their geographical distributions are colour-coded on the map. Bootstrap values were calculated from 100 iterations. Maximum likelihood trees upon which these representations are based are found in ***Figure 3—source data 1***. (**B**) Simplified haplotype diagrams for clade 1 CTVT mtDNAs derived from groups shown in (**A**). Germline variants were present in the donor mtDNA that founded clade 1, represented by the A1/A1c/A1e dog haplotype (see ***Figure 1—figure supplement 4***). Region putatively replaced by recombination is outlined with orange box. (**C**) Recombination detected in tumour 559T (Nicaragua). The estimated per cent contribution of each recombined haplotype to the mtDNA population within 559T CTVT cells is shown, and grey arrows indicate likely sites of recombination.

DOI: 10.7554/eLife.14552.013

The following source data is available for figure 3:

Source data 1. Ancient mtDNA recombination in CTVT clade 1.

DOI: 10.7554/eLife.14552.014