

Supplemental Data

Clonal Origin and Evolution of a Transmissible Cancer

Claudio Murgia, Jonathan K. Pritchard, Su Yeon Kim, Ariberto Fassati, and Robin A. Weiss

Supplemental Experimental Procedures

Microsatellite Analysis

The markers adopted were previously used by Parker et al., (2004), representing 73 different sites in the canine genome. The primer sequences can be found at the Dog Genome Project website (http://research.nhgri.nih.gov/dog_genome/). To optimise the amplifications, a gradient PCR with different primer concentrations was performed. All PCR amplifications were carried out in 10 μ l containing 25ng of DNA. PCR products were run on an ABI 377 sequencer (PE Applied Biosystems). The alleles were analysed using GeneScan software (PE Applied Biosystems). To determine the specific tumor genotype, normal contaminating alleles were excluded.

Supplemental References

Bandelt, H. J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16, 37-48.

Choi, Y., Ishiguro, N., Shinagawa, M., Kim, C. J., Okamoto, Y., Minami, S., and Ogihara, K. (1999). Molecular structure of canine LINE-1 elements in canine transmissible venereal tumor. *Anim Genet* 30, 51-53.

Kennedy, L. J., Angles, J. M., Barnes, A., Carter, S. D., Francino, O., Gerlach, J. A., Happ, G. M., Ollier, W. E., Thomson, W., and Wagner, J. L. (2001). Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: Second report of the ISAG DLA Nomenclature Committee. *Tissue Antigens* 58, 55-70.

Kennedy, L. J., Barnes, A., Happ, G. M., Quinnell, R. J., Courtenay, O., Carter, S. D., Ollier, W. E., and Thomson, W. (2002). Evidence for extensive DLA polymorphism in different dog populations. *Tissue Antigens* 60, 43-52.

Ponchel, F., Toomes, C., Bransfield, K., Leong, F. T., Douglas, S. H., Field, S. L., Bell, S. M., Combaret, V., Puisieux, A., Mighell, A. J., et al. (2003). Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnol* 3, 18.

Posada, D., and Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817-818.

Venta, P. J., and Cao, Y. (1999). A PCR-RSP Csp6I site in the canine DLA-DQA1 gene. *Anim Genet* 30, 67.

Wagner, J. L., Creer, S. A., and Storb, R. (2000). Dog class I gene DLA-88 histocompatibility typing by PCR-SSCP and sequencing. *Tissue Antigens* 55, 564-567.

Table S1. Specimens

Dogs Providing Fresh Tumor and Blood Samples						
Sample	Date	Breed	Sex	Age (Yrs)	Tumor Location	Country
A	2001	mixed	male	2	penis	Italy (Catania)
B	2001	mixed	male	1	penis	Italy (Catania)
C	2001	mixed	male	10	penis, eye, skin, lymph nodes	Italy (Catania)
D	2001	mixed	male	2	penis	Italy (Catania)
E	2001	mixed	female	5	vagina	Italy (Catania)
F	2003	mixed	male	3	penis	India (Kolkata)
G	2003	mixed	male	2	penis	India (Kolkata)
H	2003	mixed	female	2	vagina	India (Kolkata)
I	2003	mixed	female	3	vagina	India (Kolkata)
L	2003	mixed	male	nd	penis	Kenya (Nairobi)
M	2003	mixed	male	nd	penis	Kenya (Nairobi)
N	2005	mixed	female	7	vagina	Italy (Messina)
P	2005	mixed	female	8	vulva	Italy (Messina)
Q	2005	mixed	female	2	vulva	Italy (Messina)
R	2005	mixed	female	3	vagina	Italy (Messina)
S	2005	mixed	female	2	vagina	Italy (Messina)
Paraffin-Embedded Samples Used for Microdissection of Tumors						
Sample	Date	Breed	Sex	Age (Yrs)	Tumor Location	Country
2	nd	nd	nd	nd	nd	Turkey
3	nd	nd	nd	nd	nd	Turkey
4	nd	nd	nd	nd	nd	Turkey
5	nd	nd	nd	nd	nd	Turkey
6	nd	nd	nd	nd	nd	Turkey
8	nd	nd	nd	nd	nd	Turkey
9	nd	nd	nd	nd	nd	Turkey
11	nd	nd	nd	nd	nd	Turkey
12	nd	nd	nd	nd	nd	Turkey
14	1985	mixed	female	3	vagina	Spain
17	1985	mixed	female	2	vagina	Spain
18	1986	mixed	female	2	vagina	Spain
19	1986	mixed	female	3	vagina	Spain
20	1999	mixed	male	1	penis	USA (Georgia)
21	1999	beagle	male	2	penis	USA (Georgia)
25	1999	husky	female	4	vulva	Brazil
27	1997	husky	male	1	penis	Brazil
29	1999	mixed	male	4	penis/skin	Brazil
30	1995	mixed	female	nd	cervix	Brazil
32	1995	maremmano	male	2	penis	Italy (Sardinia)
33	1995	mixed	male	nd	penis	Italy (Sardinia)
35	1992	mixed	female	nd	vagina	Italy (Sicily)
36	1983	mixed	female	nd	vagina	Italy (Sicily)
37	1976	mixed	female	nd	vagina	Italy (Sicily)

Tumor and normal (blood) tissues from Sicily were brought to the UK with permission of the Department of Environment, Foods and Rural Affairs and were tested for the absence of rabies by RT-PCR at the Veterinary Laboratory Agency (Weybridge, UK) prior to use. DNA was extracted from tumours and blood samples of Indian and Kenyan specimens on site. The Messina samples of matched tumor and normal tissues were analyzed for the LINE-1/c-myc insertion only. nd = not determined.

Table S2. PCR Primers and Conditions for DNA Amplification**A. PCR for LINE-1/c-myc 3' Insertion**

Primers	Forward	Reverse	Size (bp)
LINE-1A	GGTGGGGCAGGGAGACAACATTTTA	ATCCTAGAGAAGAACACAGGCAACAC	390
LINE-1B	GGTGAGGCTTCCCATCCTT	CTTCTTGCAAGATACATCCA	150

The PCR programmes were: 94°C for 5 min, 30 cycles of: 30 sec at 94, 30 sec at 60°C, for LINE-1A primers and 30 sec at 50°C for LINE-1B primers, 1 min at 72°C, followed of 5 min at 72°C. LINE-1A primers for fresh material were as previously reported (Choi et al., 1999); LINE-1B for paraffin-embedded tissue were newly designed.

B. Intron Primer Sequences Encompassing Exons of DLA Loci

Gene	Forward	Reverse	Size
DLA-88 exon 1-2-3	AGTCCAGCGGCGACGGCCAGTGT CCCCGGA	AGCCCTCCCTAGTGGAGGGCGAGATC GGGGA	1,100
DRB1 exon 2	CCGTCCCCACAGCACATTTT	TGTGTACACACCTCAGCACCA	350
DQB1	TCACTGGCCCCGGCTGTCTCC	GGTGCCTCACCTCGCCGCT	350
DQA1	CTCAGCTGACCATGTTGC	GGACAGATTCAAGTGAAGAGA	300
Sequencing			
DLA-88 exon 2	TTCACCCGTCGGCTCCGCAG	GATGGGGTCTGTGCCCTGGCC	350
DLA-88 exon 3	ATTGGCGGCTGTCCGG	AGGCGAGATCGGGGAGGC	350

Both strands of the class II genes were sequenced using the vector specific primers T7 and Sp6. Sequencing reactions were carried out using the CEQ 2000 Dye Terminator Cycle Sequencing Kit (BECKMAN COULTER) according to the manufacturer's instructions. Cycle sequencing program consisted of 94°C 3 min followed of 30 cycles of 20 sec at 96°C, 20 sec at 50°C, 4 min at 60°C. The sequences were analysed using Sequencer software and subsequently were compared using the Bioedit software. Derived from Kennedy et al., 2001; Kennedy et al., 2002; Wagner et al., 2000.

C. DLA Primers and Conditions for Tumor-Specific Alleles

DLA Locus	Forward	Reverse	Size
CTVT DLA-88 α	CCTTCAAGGAGACCGCACGAGGG	CTCCAGGTAGTTCCTTTTCGTGC	420
CTVT DLA-88 β	CCATCAAGGAGACCGCACAGGTG	CTCCAGGTAGTTCCTTTTCGTGC	420
CTVT DRB1	CGGTTTCTGGCGAGAAGCA	TCCACCGCGGCCCGCTCCTG	150
CTVT DQB1	GGCTTCTGGCGAGAGACATC	CGCCTCTGCTCCAAGAGCT	150
CTVT DQA1	GAATTTGATGGCGATGAGTT	TCAGGATGTTCAAGTTTTGTTTTAT	150

PCR Steps	DLA 88 33	DLA 88 35	DRB1	DQB1	DQA1
Denaturation	94 (1 min)	94 (1 min)	94 (45 sec)	94 (45 sec)	94 (45 sec)
Annealing	65 (2 min)	64 (1 min)	60 (45 sec)	58 (45 sec)	50 (45 sec)
Extension	72 (45 sec)	72 (45 sec)	72 (45 sec)	72 (45 sec)	72 (45 sec)
Final extension	72 (5 min)	72 (5 min)	72 (5 min)	72 (5 min)	72 (5 min)

D. qPCR for Gene Dosage

Gene	Forward	Reverse	Size
GAPDH	GGCGGGGCCAAGAGGGTCA	TCTTGAGGGAGTTGTCAT	120
B-Actin	CTCCATCATGAAGTGTGACGTTG	CGATGATCTTGATCTTCATTGTGC	150
DQA1	CTCAGCTGACCATGTTGC	CACAGGCAGCCGCCAGAC	150
DQA1†	TAAGGTTCTTTTCTCCCTCTGT	TGCTAGGGAGGAAGGGGAAAG	389
DQB1	TCACTGGCCCCGGCTGTCTCC	CTCCCCACGTCGCTGTC	150
DRB1	CCGTCCCCACAGCACATTTTC	TGTGTACACACCTCAGCACCA	350
DLA-88	TCTCACCCGTCGGCTCCGCAG	GATGGGGGTCGTGCCCTGGCC	350
DRA	CATCCAAACCCAGTGCTCC	ACCCCTGTGGAAGTGGGAGAG	212

In order to determine the gene dosage of the DLA genes, real-time PCR based on SYBR-Green I fluorescence (Ponchel et al., 2003) was used. Quantification was based on the kinetic method, which requires a standard curve. To generate the standard curve, target and reference genes were amplified and cloned into the pGEM T vector. A dilution series containing 10^2 - 10^9 copies was used to construct the standard curve. For each gene tumor and normal copy number was determined against a standard curve run in parallel with the samples, to obtain absolute quantification. Samples were run in triplicate. PCR was carried out in 50µl using 50ng and 100ng of genomic DNA for fresh tissue and for paraffin-embedded tumour tissue respectively, using the SYBR-Green PCR Core Reagents kit (Applied Biosystems) according to the manufacturer's instructions. Quantitative PCR was performed using the ABI Prism® 7000 Sequence Detection System (SDS). PCR was optimised to avoid the amplification of the unspecific product and reduce the primer dimers formation. Amplification of the correct sequence was confirmed by dissociation curve analysis. The gene dosage was determined using the following formula:

$$N = \frac{\text{Copy number of target gene}}{\text{Copy number of reference gene}}$$

The percentage of the tumour cell within the tumor tissue was determined by histological analysis and all tumour samples contained more than 70% tumor cells (Fig. 1d). To determine the cut-off point in the tumor samples, 4 unrelated normal samples were tested. Although all micro-dissected samples were informative for the LINE-1/*c-myc* and DLA analyses, samples 3, 17 and 30 were excluded from the DLA quantification analysis because their dissociation curve show non-specific amplification.

†To test the possibility that the apparent hemizygous status of DQA1 was due to a mutation in the intron primers, we used different intron primers (Venta and Cao, 1999) designed outside of the first set which confirmed the hemizygous result.

E. Primers and Conditions for RT-PCR of DLA Expression

Gene	Forward	Reverse	Size
C DLA-88 (29)	CGCCAAGGAGACCGCACAGGTGT	CCTCAGGTGCCCTGCATCACCT	420bp
C DLA-88 (42)	CCATCAAGGAGACCGCACAGAGG	CCTCAGGTGCCCTGCATCACCT	420
C DLA-DRB1	CGTTTCGTGGAAGATACA	CAATCACCCGTAGTTGTG	169

PCR Steps	C DLA 88 (29)	C DLA 88 (42)	C DLA DRB1
Initial Denaturation	95 (3min)	95 (3min)	95 (3min)
Denaturation	94 (1min)	94 (1min)	94 (1min)
Annealing	65 (2min)	64 (1min)	61 (1min)
Extension	72 (45sec)	72 (45sec)	72 (45sec)

F. Amplification of the mtDNA Control Region

Primer	Forward	Reverse	Size
H15422/L16106	CTCTTGCTCCACCATCAGC	AAACTATATGTCCTGAAACC	722*
H15422/H15710	CTCTTGCTCCACCATCAGC	GCATGGTGATTAAGCCCTTAT	290†

*Fresh tissue

†Paraffin-embedded tissue.

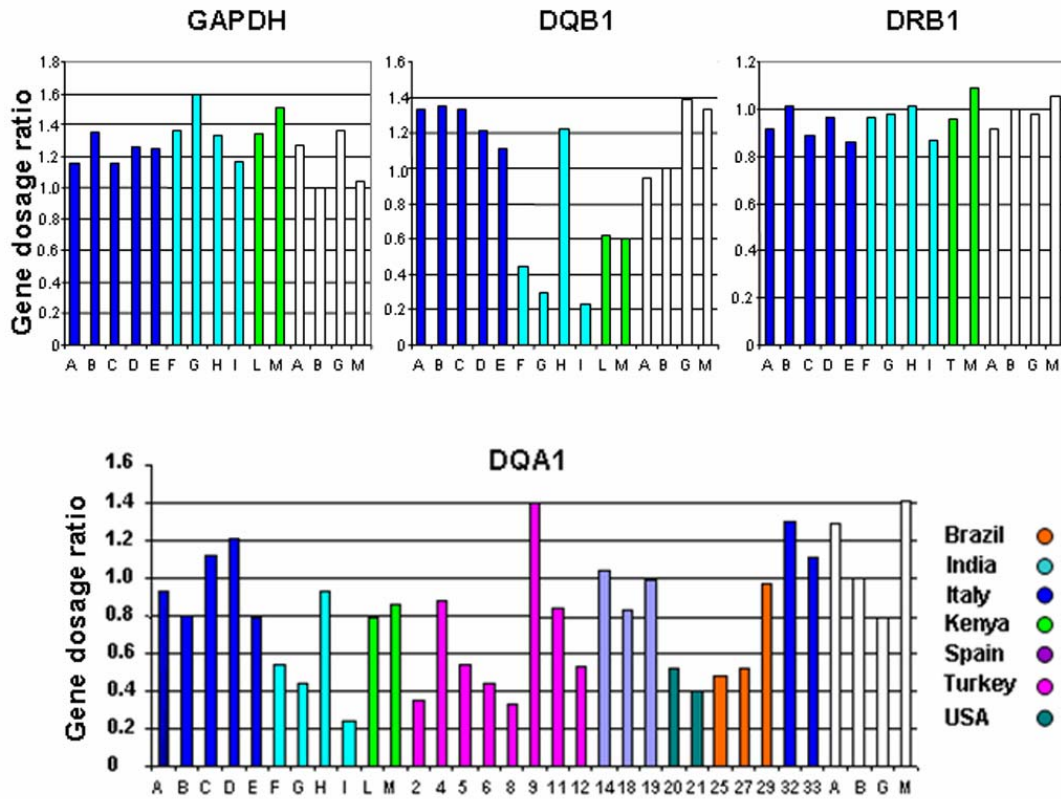


Figure S1. Gene Dosage of DLA Class II Alleles in Fresh and Microdissected Tumor Samples

The copy number of the target DLA genes was compared to the reference gene β -actin using qPCR with SYBR-Green and the kinetic method (standard curve) so that the diploid β -actin gene was calibrated as 1.0. β -actin was compared to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as a second standard for diploid gene dosage. Fresh tumors are denoted by letter and paraffin extracted tumors by number; normal tissues from matched animals A, B, G and M are shown as white bars. Ratios of 0.8 and above were scored as diploid and of 0.6 or below as haploid.

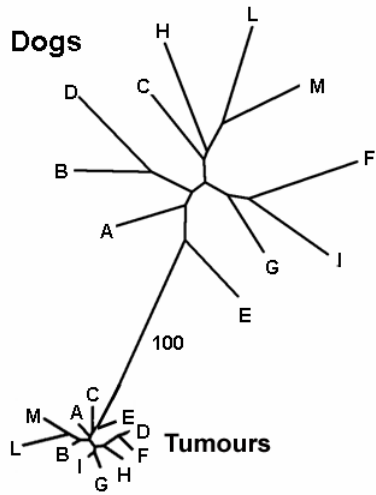


Figure S2. Comparison of 21 Microsatellites in Matched Tumor and Normal Tissues from 11 Dogs
Neighbor-joining tree based on proportional shared alleles.

intermediates in a minimum spanning network. The size of squares is proportional by length to the number of tumors sharing the same haplotype; yellow circles indicate the published haplotypes of clade A (Savolainen et al., 2002), called Clade 1 by Vila et al (1997). Some host tissue haplotypes from tumor-bearing dogs belong to clades B and C. For each tumor, 7-10 clones derived from PCR amplified mtDNA were sequenced and some variation in mtDNA sequence is seen within tumors.



Figure S4. Maximum-Likelihood Tree of mtDNA Sequences

The tree was constructed from 21 CTVT samples and 45 additional dogs, wolves and a coyote outgroup, without assuming that the tumor sequences are monophyletic. Given the short (257bp) amplified region from archival samples, there may be uncertainty about the true position on the tree of the two sequences that group quite separately from the main tumor clusters; moreover they may represent contaminating host mtDNA.

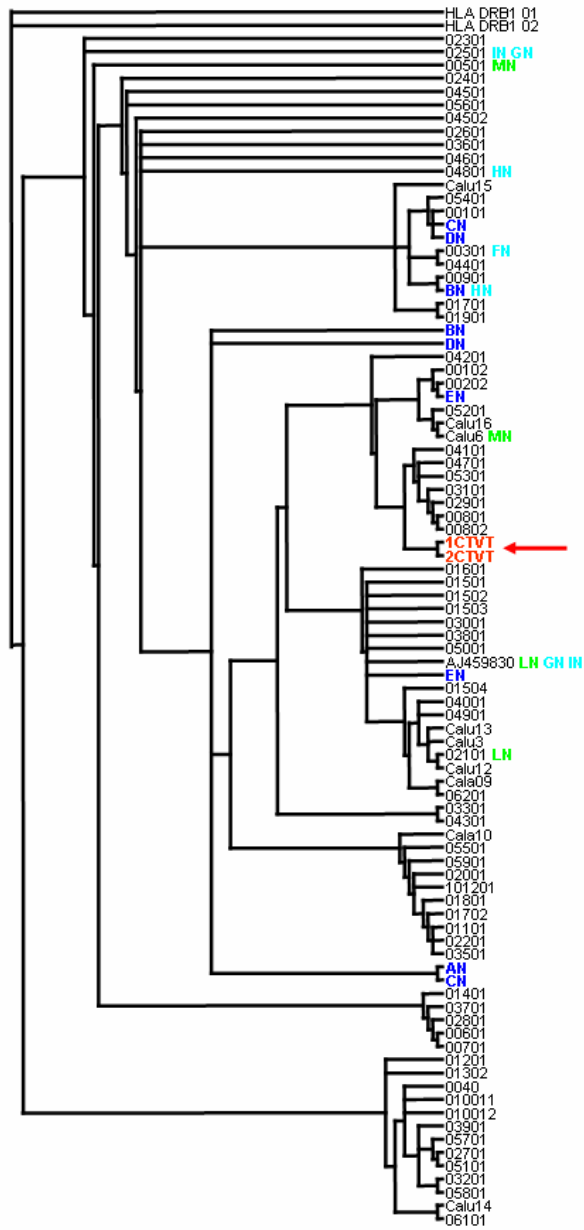


Figure S5. DLA Maximum-Likelihood Tree for DRB1

CTVT samples (red arrow) were included in a tree constructed from previously published data (Kennedy et al 2002b). DRB1 alleles for normal tissue of dogs bearing CTVT are shown as dark blue (Italy), pale blue (India), green (Kenya). Analysis of the other DLA genes also grouped CTVT with wolves and 'old' dog breeds (see text).

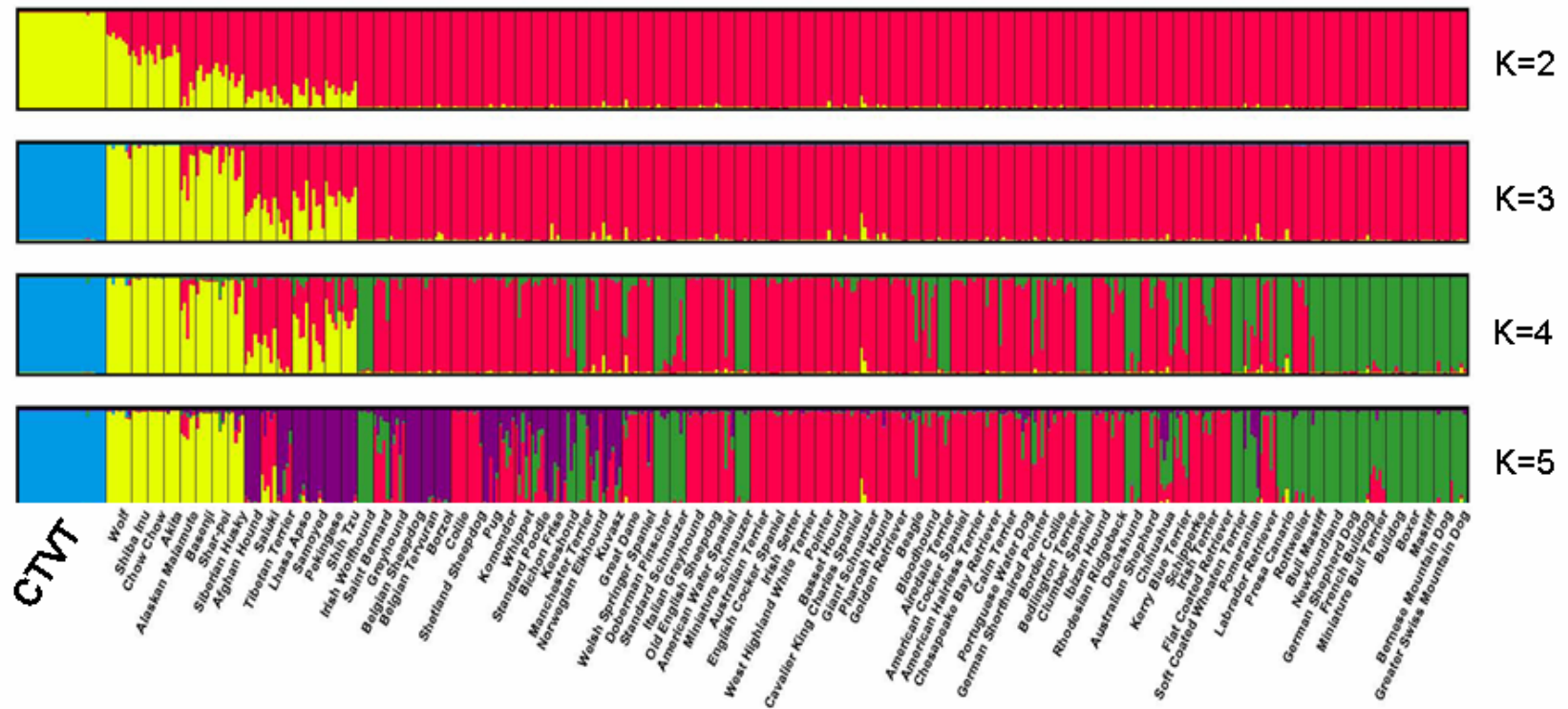


Figure S6. CTVT in Relation to Wolves and 85 Dog Breeds

Data are based on 73 microsatellite loci. Each panel shows the results from a model-based clustering algorithm, Structure, that assigns sampled individuals, based on their genotypes, to a prespecified number, K , of clusters. Each tumor sample, or individual dog, is represented by a vertical bar, with colored segments indicating the proportion of that individual's membership in each cluster. At $K=2$, the tumors cluster clearly with wolves and 'old' dog breeds; for larger, more stringent K values, the tumors form a distinct group, indicating a common origin.