

OMTM, Volume 4

Supplemental Information

Mechanism of Deletion Removing All Dystrophin

Exons in a Canine Model for DMD Implicates

Concerted Evolution of X Chromosome Pseudogenes

D. Jake VanBelzen, Alock S. Malik, Paula S. Henthorn, Joe N. Kornegay, and Hansell H. Stedman

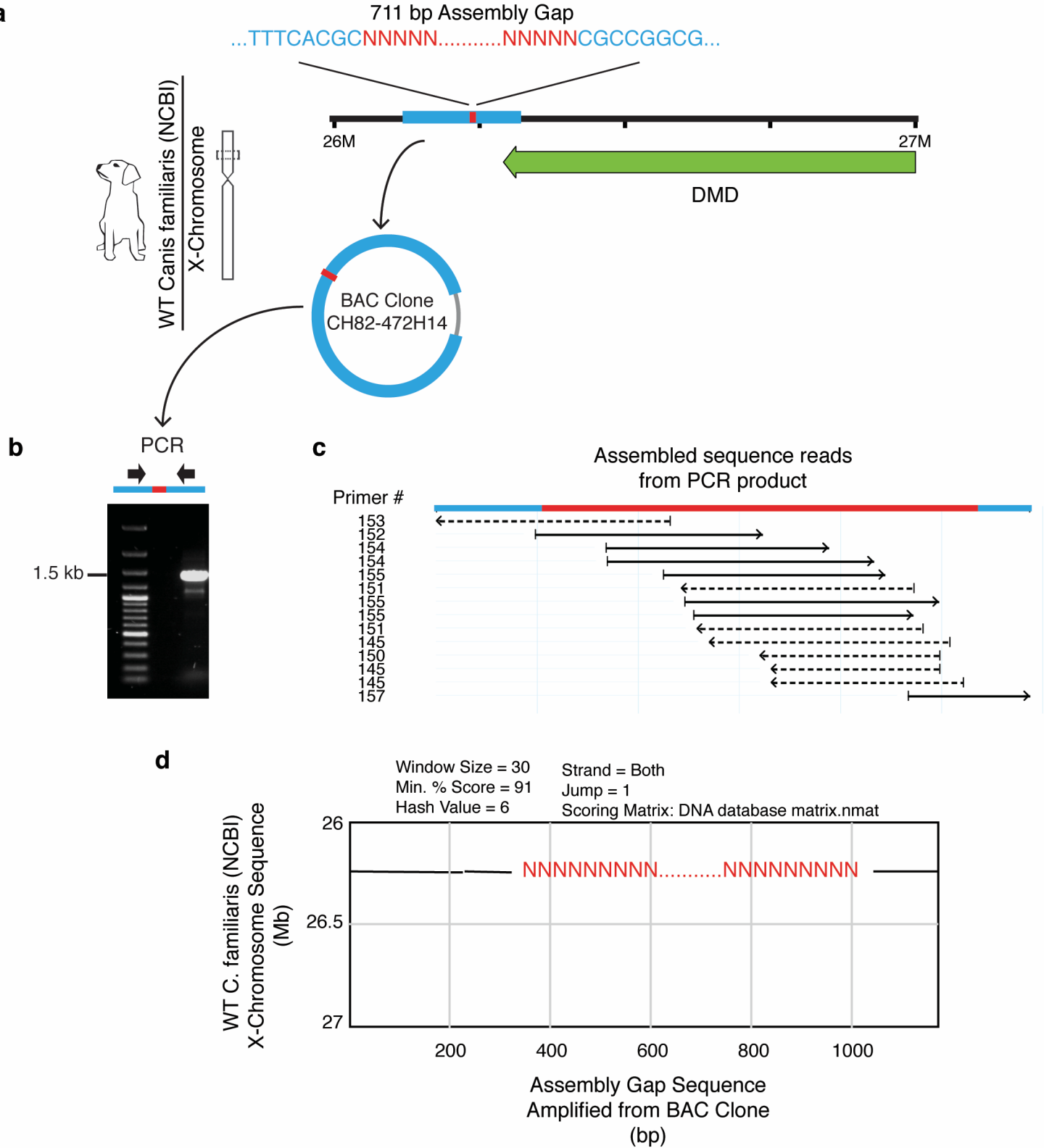


Figure S1. Sequencing of assembly gap in dog reference genome near telomeric breakpoint of GSHPMD deletion. **(a)** A region of the reference dog X-chromosome is shown. The region of the X-chromosome contained within the library BAC clone is highlighted in blue. The assembly gap is shown in red. BAC, bacterial artificial chromosome. N, any nucleotide. **(b)** Primers, depicted by arrows, that anneal outside opposite ends of the assembly gap were used to PCR amplify across the assembly gap using the BAC clone as template. PCR products are displayed on an agarose gel following electrophoresis. **(c)** The gel-purified PCR product was Sanger sequenced using the indicated primers, and overlapping reads were assembled into a single contig. **(d)** Pustell matrix comparing the sequenced PCR product to the indicated region of the dog X-chromosome reference sequence. Internal break in homology is expected and represents the assembly gap in the dog reference sequence, which is depicted by a stretch of red N's.

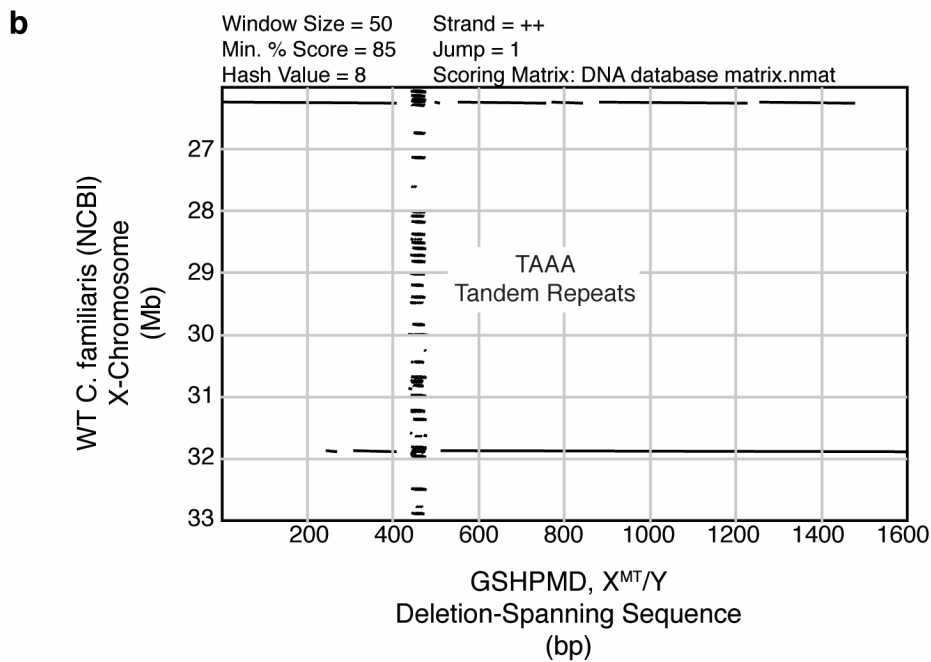
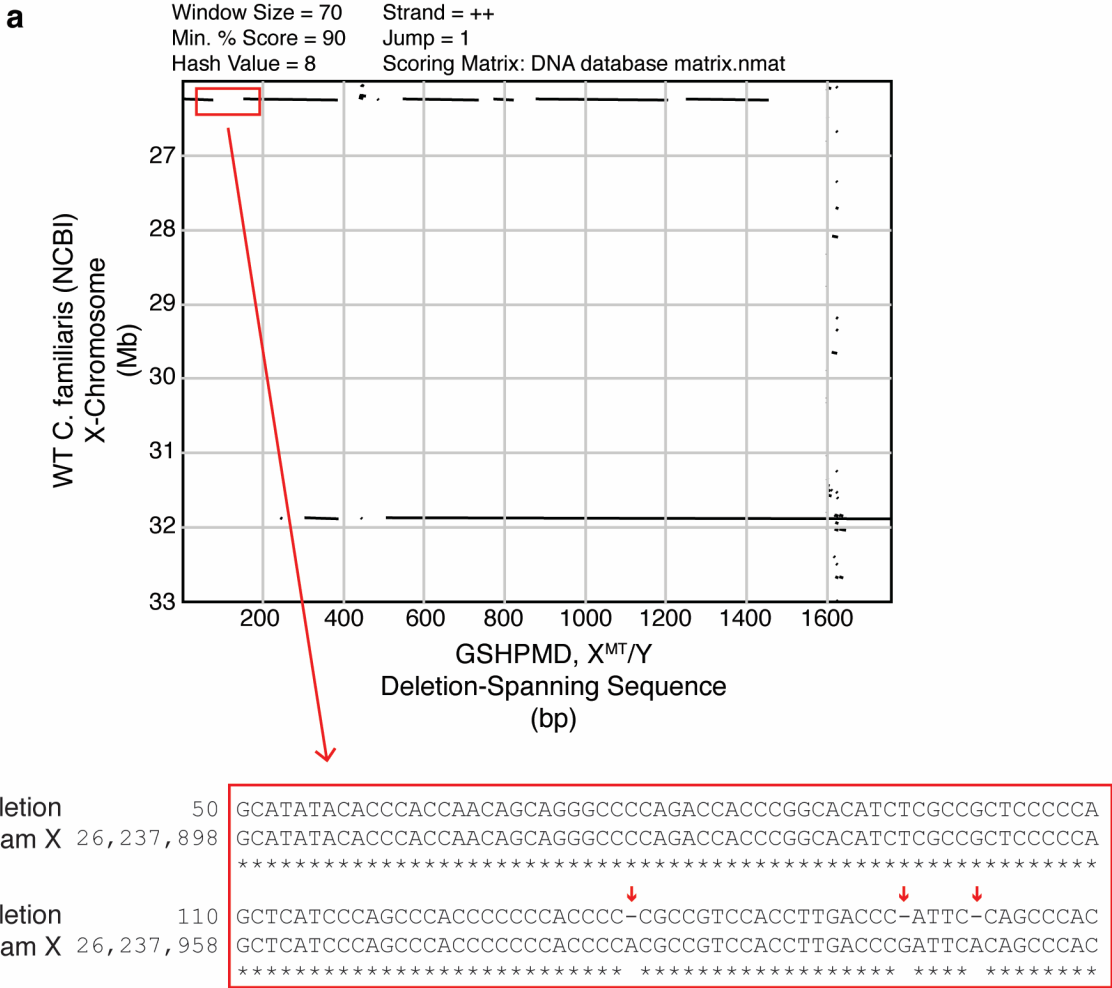


Figure S2. Investigation of breaks in homology between GSHPM deletion-spanning sequence and dog genome reference sequence. **(a)** Pustell matrix comparing deletion-spanning sequence from GSHPM to a region of the dog X-chromosome reference sequence. Red box highlights a break in sequence homology near the telomeric breakpoint, and a ClustalW alignment of the corresponding sequences is shown below. Red arrows indicate three indels responsible for the observed lapse in homology. **(b)** Pustell matrix, with reduced specificity parameters, depicting the presence of a TAAA tandem repeat present in this region of the dog X-chromosome.

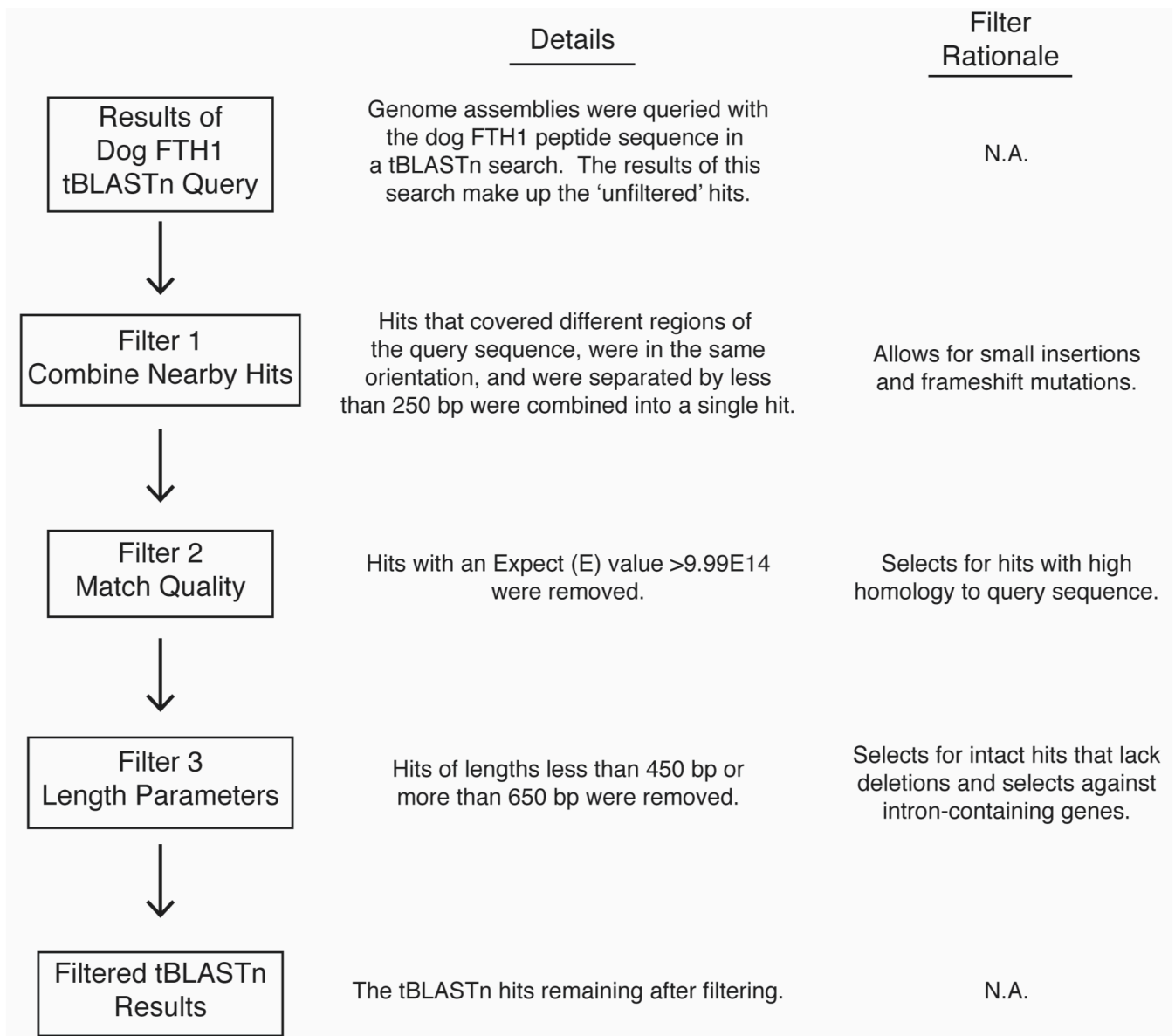


Figure S4. Schematic of filtering process of tBLASTn hits. The reference genomes of several species were queried with the dog FTH1 peptide sequence using tBLASTn. The returned hits from this search were filtered as indicated, with the goal of selecting for FTHL loci.

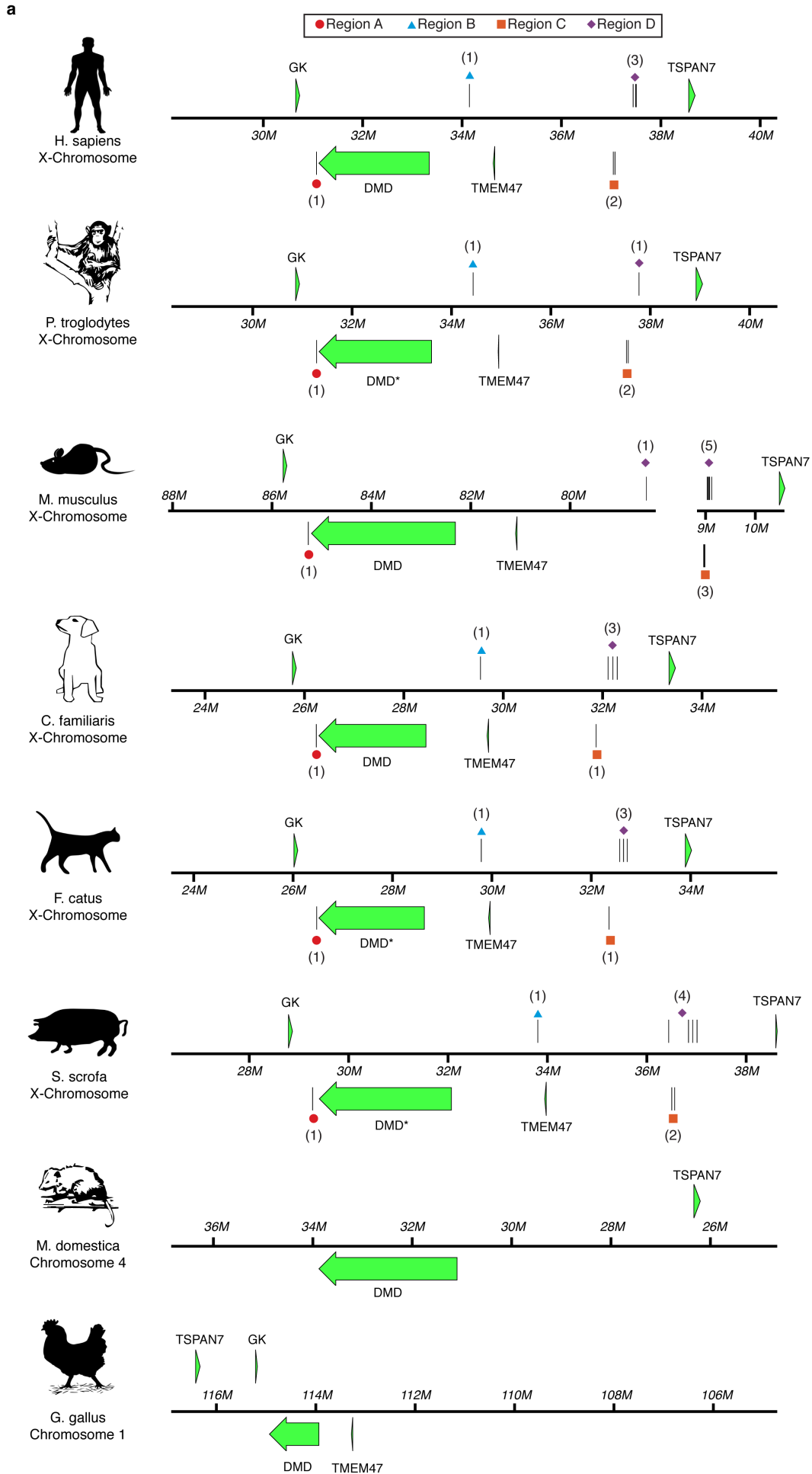


Figure S5. Phylogenetic analysis of identified FTHL loci from tBLASTn search (expanded to additional species). **(a)** Comparison of syntenic portion of dystrophin-containing chromosome from each species showing the grouping of FTHL loci to four regions, labeled A, B, C, and D. The opossum and chicken lack FTHL loci in these regions. Number of pseudogenes present in reach region is indicated in parenthesis.

b

Method: Neighbor Joining; Best Tree; tie breaking = Systematic
 Distance: Jukes-Cantor; Gamma correction = Off
 Gaps distributed proportionally

● Region A ▲ Region B ■ Region C ◆ Region D

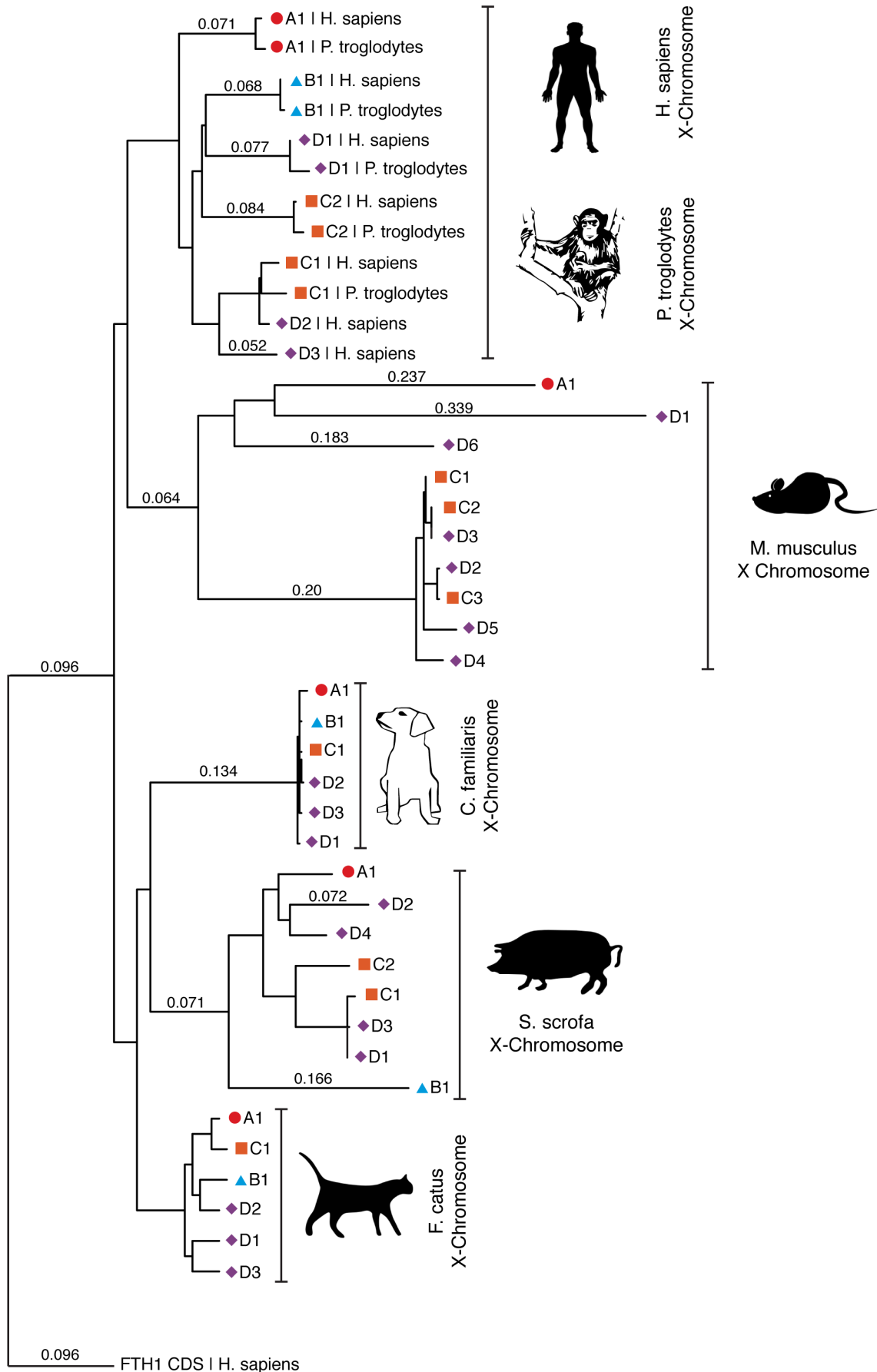


Figure S5. Phylogenetic analysis of identified FTHL loci from tBLASTn search (expanded to additional species). (b) Phylogeny of identified FTHL loci. The human FTH1 CDS is used as an outgroup. Branches with values less than 0.05 are not displayed. Tree-building parameters are provided as text in figure. CDS, coding DNA sequence.

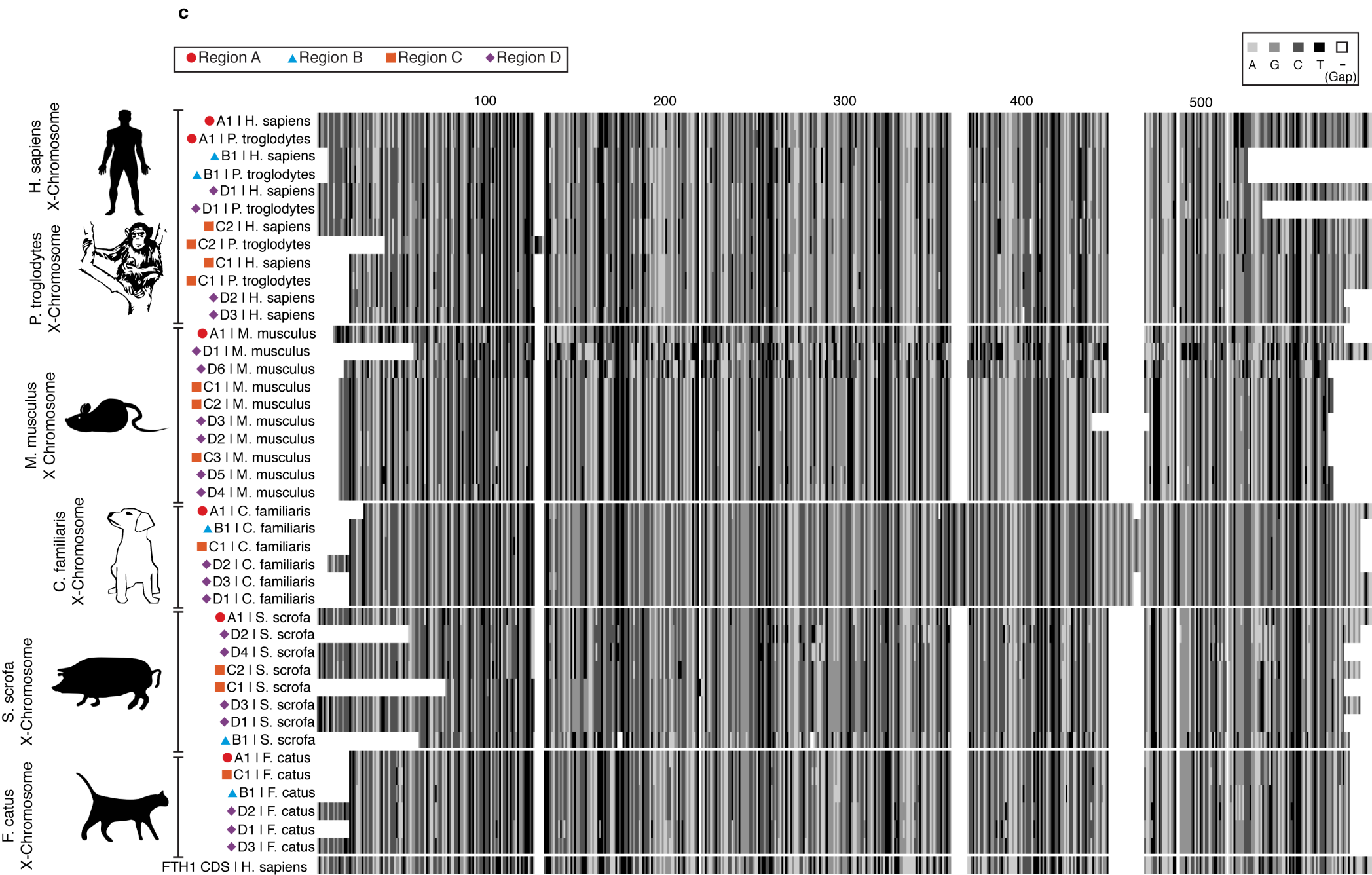


Figure S5. Phylogenetic analysis of identified FTHL loci from tBLASTn search (expanded to additional species).
(c) ClustalW multiple DNA sequence alignment of identified FTHL loci. Note sequence homology among FTHL loci from all regions is unique to the dog.

Sequence: AGap Sequence Range: 1 to 1173

```

AGap Sequence      10      20      30      40      50      60      70      80      90     100
TCTTGAAGGTTTCATGCAAGGCAATACTGCACATGCACGCTCGGGGGGGCCCAAGGAAAGTGCCTCGGGCTCCCGGGGAGGTGTCCGTCC

CfaX +
26238379 26238389 26238399 26238409 26238419 26238429 26238439 26238449 26238459 26238469
A.....>
AGap Sequence      110     120     130     140     150     160     170     180     190     200
AGGTCAAGTTCCTTTGTGCTGTGCGAGGGTGAAGCTGTGTAAGGGTACTTGGCAGGCCGCTTCGGGGGCCCCACAGCTGCAGAGCTGTGTCGG

CfaX +
26238479 26238489 26238499 26238509 26238519 26238529 26238539 26238549 26238559 26238569
.....A..G.....>
AGap Sequence      210     220     230     240     250     260     270     280     290     300
TAGCTCCACAGAGCTCTTGATGGCTTGCCTCTGCTGCTGCTGCTCACGGAGGTAGGGGCTCCAGGAGTGCAGAGCTGGAGCTGCTTCTGCTC

CfaX +
26238579 26238589 26238599 26238609 26238619 26238629
.....G..T>
AGap Sequence      310     320     330     340     350     360     370     380     390     400
GGTGGACAGCCGCTGAGGTAGGTGGAGGACAGAGCTGTGGTTCAGAGCTTCTTCAGGTGCAGAGGAGGCTGCTGTGGCTCGGGGCTCTCCAG

CfaX +
26238689 26238699 26238709 26238719 26238729
.....T...T...>
AGap Sequence      410     420     430     440     450     460     470     480     490     500
GGCTGGGGTGGGCTCTTAGAGCTGGGACAGAGGATGGGGCCCCGGGGCTGTGTGAGCTCACAGAGCATCTGGGCTGCTGCTCTCTCCGGG

AGap Sequence      510     520     530     540     550     560     570     580     590     600
CCTGGGCTGGAGAAAGGGGCAAGTTCCTCAGGGGCCCTGCTGGGCTCAAGAGGAAAGGCTGAGCTGTAGAGCTAGAGAGGAGCAGAGCTCCAG

AGap Sequence      610     620     630     640     650     660     670     680     690     700
GCTGATCGGCTTGTGAGGGGCTCTGCAGTCTGGGCTAGTCTGGGAACTGGGAAATGGGCGGCGGCAAGGCTGGGGCCGGGGGGGGGGGG

AGap Sequence      710     720     730     740     750     760     770     780     790     800
GGGGGGGAGGGGGGGGGGGGGGAGGGGAGGGGGCCACGGCGGAGAGAGGCTCGGGGCGCARTGGCGGTGGGGGAGTCTCGGTTGTTGT

AGap Sequence      810     820     830     840     850     860     870     880     890     900
CCAGGCTGGAGGCCAGAGAGGCTCTGCGGCTCGCTGGGCTGTGATGCGGAGAGGAGATCCCTTAAAGTCCGTTGTTGAGGTGGAGGTGA

AGap Sequence      910     920     930     940     950     960     970     980     990     1000
ATCCCTTAGCGGGGGGGGGGGTCAAGTCAAGCAGAGGGGGGGGAGGTTAGGGGGGGGGGCTGCTGCGGGGAGCGGGGAGGGGGCTGGCAGGGA

>=510p
AGap Sequence      1810    1820    1830    1840    1850    1860    1870    1880    1890    1180
GTGGCCCGGGGGGGGGGGCAGAGGTGTGGAGGGCTGGGAGGGCTGGCCGTTGGCGGGGCGAGGGCTGGCCGGCGGGGGCTGGGG

CfaX +
26239489 26239499 26239509 26239519 26239529 26239539
|.....>
AGap Sequence      1110    1120    1130    1140    1150    1160    1170
CGGGGGGGGACCGCTTCAAGGTTCCATCTGAGTCTGGGCGGGGAGAGGCAAGGGAGTCTGATCTTC

CfaX +
26239549 26239559 26239569 26239579 26239589 26239599 26239609
|.....>
AGap Sequence      1110    1120    1130    1140    1150    1160    1170
CGGGGGGGGACCGCTTCAAGGTTCCATCTGAGTCTGGGCGGGGAGAGGCAAGGGAGTCTGATCTTC

```

Figure S8. Sequence alignment of BAC-derived PCR product and assembly gap region of the dog X-chromosome reference sequence. Alignment corresponds to output from Pustell matrix in Figure S10.

Primer Pair	Primer Number	Sequence (5'-3')	X-Chromosome Annealing Region		Annealing Temp. (°C)	Amplicon Size	PCR Results (Yes/No)	
			Start	Stop			WT-Hsa	GNPM-Hsa
1	1	5'GTGGTCTGATGCTGCTTCTTC	23152290	23152313	61	231	Y	Y
	2	3'GCCAGGAAATTTCCAGGAAG	23152497	23152520				
	3	5'TGGTGGTCTGCAACACAC	23960538	23960560				
	4	4'TGGCCAGGACAGTGGGGAG	23960689	23960710				
	5	5'AAGGCGATGCTGCTAGGGTGTG	24750036	24750058				
	6	4'TAGAGCTGCTGCTGGCTC	24750057	24750076				
	7	5'TGGTGAAGCCACACTAGC	25330610	25330630				
	8	3'CACTCCCACTGCTTGAAGCC	25330825	25330847				
	9	5'CTCTCTGTGCTGGAG	26346126	26346147				
	10	10'CTCATGGGTGCTCCAGTT	26345431	26345470				
	11	11'CTCATGTAAGAGGGGG	26176440	26176501				
	12	12'GTAGAGTCTGTGCTGTGG	26174609	26174638				
	13	13'CTCAAGAGCACTGCCA	26384245	26384264				
	14	14'GTGGAGGAGATGCTGAGC	26384380	26384399				
	15	15'CACTTCCCTTGGCTCTC	26194195	26194214				
	16	16'GAGAAACCGGACTTGTCTCA	26349922	26349942				
	17	17'CTCTTGGCTATTGTTCTCA	26204548	26204587				
	18	18'ACAGATCCCAAGACCTCT	26204974	26204993				
	19	19'GGAGACGTGGAGTGTGACT	26214222	26214241				
	20	20'AGGATGCTGCTGACATGTA	26214897	26214917				
	21	21'CACTCCACACAAAGCCGA	26224113	26224132				
	22	22'ACTGTAGTCCCAAGACAA	26224713	26224732				
	23	23'CTCATCTAAGAGCCAGAT	26232427	26232446				
	24	24'CACTTTTGTGCTGTCAAA	26233390	26233409				
	25	25'AACTCTGCTAGGAGACA	26233393	26233412				
	26	26'CTCTTTGTGAGTGTGGATGA	26234469	26234490				
	27	27'AGCAATGATGGCACTGT	26235187	26235206				
	28	28'TCCCTGTGAGACGCTGTG	26235531	26235550				
	29	29'TGAGACCCCGTACCATAG	26236154	26236173				
	30	30'ATGATAGTTCCGCTGGTC	26236525	26236544				
	31	31'CTGCTGGTGGAGTGTGAT	26237046	26237065				
	32	32'CTGCTGTGTGGGTGTAT	26237902	26237921				
	33	33'AGCTTCAGATGCTCACT	26239551	26239570				
	34	34'AAAGAGCTCAAGAGGACCT	26239826	26239845				
	35	35'GTCTCATATGATGGTGTACT	26241080	26241103				
	36	36'CACTCATTTTCTGACTCTG	26241553	26241576				
	37	37'GAGAGAGGACCAAGCAATA	26242082	26242103				
	38	38'TTGATGATAGGGAGAGAG	26242661	26242682				
	39	39'CACTTCAGTCCCTCTC	26244699	26244718				
	40	40'TCTGTGTGCTGTGTATC	26245535	26245554				
	41	41'CAACCACTTTTACTTCTCTC	26246775	26246797				
	42	42'CTCAATATAGAGCAATATC	26247117	26247140				
	43	43'TAGATAGCTTCCAGCAAGG	26250951	26250974				
	44	44'GGGAGCACTATGATCAAGAG	26251835	26251858				
	45	45'CATAGAGTCCAGCATAAATG	26252880	26252913				
	46	46'CACTGTACATAGGAGCAATCTC	26259338	26259361				
	47	47'GACTGTAGTGGAGTATATAG	26267941	26267964				
	48	48'CTATCTGAGGAGAAATCTAC	26267431	26267454				
	49	49'AAGGCACTGTGAGTGTG	26243207	26243226				
	50	50'CTCATATGCTGCTCCCA	26263840	26263863				
	51	51'TGGCTTAGAAGTGGACAA	26444213	26444232				
	52	52'AAGGCGAAGTGTGTGAG	26444967	26444986				
	53	53'CCGCAAGAGAGTGTG	26444127	26444146				
	54	54'TCCAGAGAGCTCTCA	26444928	26444947				
	55	55'CAATGTGTAACTTGTTCAC	26444271	26444295				
	56	56'TGGAGAGACTTCCAGG	26444336	26444355				
	57	57'ATCTGCAACAGCAAGCCCT	26244140	26244159				
	58	58'CACTGTTATGCTCACTCTC	26244438	26244457				
	59	59'CAATAGCATATCACTCTC	27046481	27046507				
	60	60'CTCTGTGCTGTGATATC	27047343	27047362				
	61	61'CCCAAGCACTGACTCTCT	27350531	27350554				
	62	62'CATGCACTGTGTAGGAGAG	27351345	27351364				
	63	63'CACTTCACTCTCTCACTG	27360252	27360273				
	64	64'GGGATTAATCTGTGTATG	27361138	27361157				
	65	65'GTGAAATCTGCTTTAAGGG	27367034	27367055				
	66	66'GAGTGAATCTCTGAGTGTG	27367494	27367513				
	67	67'CAATGACTTATGCTCAATG	28244026	28244045				
	68	68'CCCTAATGAAAGCAAGCTAG	28245728	28245747				
	69	69'CCGCAATGAGCAAGCCCA	28207181	28207200				
	70	70'ACAGAGGCTGATAGTCAAA	28307728	28307747				
	71	71'TGAAATGCAACTACTGGGA	28408368	28408387				
	72	72'TGTAGCTGCTGATGGCA	28408748	28408767				
	73	73'ACAGAGAGGAGATCACTT	28508480	28508509				
	74	74'CACTGCAATGTGTGTGAA	28509045	28509064				
	75	75'CAAGCAAGTTTGAAGCAAT	28612113	28612132				
	76	76'TGCTCCCTCTGCTGTGA	28613007	28613026				
	77	77'GGTCTCAACTGCTGCA	28812182	28812201				
	78	78'CACTGCACTAGAGGCTTG	28813040	28813059				
	79	79'GGGATGATGCTATGAGGGGA	28914218	28914239				
	80	80'GATCTGAGAGCACTTACG	28914710	28914730				
	81	81'ACTGCTGAGCAATGAGAT	29046624	29046645				
	82	82'CAAGCAAGGATTTGAGAAGCA	29016999	29017023				
	83	83'ATTTTCCAGAGCACTGTGGA	29244273	29244294				
	84	84'CTGCTGCTCTTGAAGC	29245428	29245448				
	85	85'ATGCTGCTCTCTAGAGAG	29242426	29242436				
	86	86'CTCTGCTCTCTAGAGCT	29225530	29225550				
	87	87'CTGCACTTGTGAGTGTG	29226541	29226560				
	88	88'AGAGAGGATGATCACTCTCTG	29237146	29237169				
	89	89'CTCAAGAGCACTCTTGAGC	29823215	29823233				
	90	90'CACTGCTCTCTCTATGATA	29613046	29613025				
	91	91'CTCTTGGTGAATGTTCTC	29626056	29626027				
	92	92'CTCTGGAATGATAGAG	29621274	29621255				
	93	93'AGTAAAGCTCTCACTCTG	30227121	30227142				
	94	94'CACTTACTGCTGATAGCTG	30227444	30227423				
	95	95'TTCTCACTCTACTCA	30530417	30530438				
	96	96'CCCTGATTAAGACAGAGAG	30531022	30531001				
	97	97'CTGAAATCTGATGACTCT	30849888	30849869				
	98	98'CTGAAATGAGTCAAGATC	30827170	30827149				
	99	99'ACTTCTGAGATCACTG	31134433	31134454				
	100	100'GAACTGCAATGTTGATGAT	31135070	31135049				
	101	101'GCTGCAAGTTGGGGAGC	31450263	31450282				
	102	102'GGTGGAGCACTGATCTCA	31440945	31440964				
	103	103'GAGAGGAGCACTCTCTC	31553528	31553548				
	104	104'CCCAAGCACTTGAAGGGA	31553865	31553885				
	105	105'CACTCAAGCTTTTACTC	31653058	31653077				
	106	106'CACTCACTGATGAGAGG	31653881	31653900				
	107	107'TGCTAGCTGTGATGCTCT	31734113	31734132				
	108	108'TTCTGCTGATGAGGGCT	31733773	31733792				
	109	109'TCCCACTGTAAAGTGAAT	31814038	31814059				
	110	110'CCCTTATAGAGCAAGAGC	31814405	31814424				
	111	111'GCTCTTCTGAGAGGCTT	31835312	31835331				
	112	112'AGGGATAGAGAAATGCTC	31835750	31835769				
	113	113'ACTGGATTCAGGGGGGA	31835378	31835397				
	114	114'GGGATTTCTGAGTGGC	31835841	31835860				
	115	115'CAACCAATGAGCTGGGAG	31843357	31843376				
	116	116'AGGCAATCACTGATGCCA	31845032	31845051				
	117	117'CTTTAGGGAGAGGCTT	31858072	31858091				
	118	118'AGGASCTTCAAGATAGCA	31858641	31858660				
	119	119'CAATTTGACCAACTGCTCT	31859222	31859244				
	120	120'AAAGCAAGGATGCTG	31859790	31859779				
	121	121'GGTGAAGCTGTAGTCTCT	31860531	31860550				
	122	122'GAGAGAGATGAATGGGTT	31861070	31861092				
	123	123'TCTTAAACAAAGCAAGTGTG	31864729	31864748				
	124	124'TCAAGAGGCTTCTCTGTG	31865152	31865171				
	125	125'TCCCAAGAGTGGAGCC	31867380	31867359				
	126	126'CCCTGAATGTAAGCAGTGA	31867961	31867980				
	127	127'CTCTCTCTCTGAGTGA	31871007	31871026				
	128	128'AGGATGCTCAAGAGGCA	31871204	31871224				
	129	129'CTGGGAGATGCTATGAGCT	31872250	31872273				
	130	130'TCTTGGAGCACTTACACAA	31872733	31872755				
	131	131'ACTTTTCTGATCATCTTT	31876063	31876076				
	132	132'ATCAAGCACTTTTGAAGAG	31876779	31876801				
	133	133'GGGAGAGCACTACTGG	31880639	31880637				
	134	134'GCTATTCTGAACATGAGCA	31880668	31880688				
	135	135'TAGACTCAGAGATGGGCA	31901007	31901027				
	136	136'CACTGTCTCTCAAGCT	31901033	31901047				
	137	137'GGAGAGCTCTCTCAAA	31913379	31913398				
	138	138'CTCCACAGCACTTACAG	31913914	31913933				
	139	139'CCCTTCACTCTCTCTC	32013183	32013180				
	140	140'AGGCTCTTGAAGAGCTG	32013719	32013738				
	141	141'GCACTTTCTCATGATTC	32124473	32124494				
	142	142'CACTCACTGATTTGATCT	32124487	32124497				
	143	143'CCCTTAGAAGCACTCGGT	32124343	32124362				
	144	144'GTCCCTCTACAGCAATGT	32124733	32124752				
	145	145'CAACACTTCTATGATCTTAGG	26237849	26237872				
	146	146'CTCGAATCTTCAACGATCTCAC	31870309	31870332				
	147	147'CTGAGGATTCAGTGGAGCAAT	26238779	26238803				
	148	148'AAAGAGCTCAAGGACTGTGAGC	26239821	26239845				
	149	149'CTAGCTTACAGTGTAGGCTTC						
	150	150'CACTGATGAGTGGATGAGG						
	151	151'CGGCTAAGGATCCACTCTC						
	152	152'GACTTTAGGGCACTCTCTG						
	153	153'GTGAGCTCTCTGAGCACTGAG						
	154	154'GGTCTCTGCGAGGCTCTCT						
155	155'TGGAGAGGCGCCAGGCGAGG							
156	156'CTGAACTCTCAAGCACTCAC							
157	157'GAGATGGCCAGGCTGGGCTGGA							
158	158'CAACACTATGATGATCTTAGG							
159	159'TTCTGGCAACTGGAGATGG							
160	160'GAGGCACTGACTGTGAGCACTGAG							

Figure S9. Detailed primer information. Primer numbers, sequences, annealing locations, annealing temperatures, expected amplicon size, and amplification results are provided.

Data S1. Results from HMMER search of reference proteomes using exon 2 peptide sequence of dog LOC612257 as query. Available as separate .txt file.

Data S2. tBLASTn output for all queried species. Each species is provided in a separate sheet. Detailed information for each hit is provided, and hits remaining after each filtering criteria was applied are indicated in separate columns. Available as separate .xls file.

Region	Figure Designation	Species	tBLASTn Hit Location on X-Chromosome		Orientation	NCBI Gene Symbol	NCBI Gene Type	NCBI Gene Description
A	A1	H. sapiens	31071953	31071402	-	FTHL17	protein coding	ferritin, heavy polypeptide-like 17
	A1	P. troglodytes	31285760	31285209	-	FTHL17	protein coding	ferritin, heavy polypeptide-like 17
	A1	M. musculus	85249677	85270291	+	Fthl17a	protein coding	ferritin, heavy polypeptide-like 17, member A
	A1	C. familiaris	26238702	26238481	-	N.A.: region is within a genome assembly gap		
	A1	F. catus	26476187	26475666	-	FTHL17	protein coding	ferritin, heavy polypeptide-like 17
	A1	S. scrofa	29274687	29275602	-	FTHL17	protein coding	ferritin, heavy polypeptide-like 17
B	B1	H. sapiens	34147040	34147516	+	FTHP14	pseudo	ferritin, heavy polypeptide 1 pseudogene 14
	B1	P. troglodytes	34435114	34435590	+	LOC473555	pseudo	ferritin heavy polypeptide-like 17
	B1	C. familiaris	29541841	29542399	+	LOC102153989	pseudo	ferritin, heavy polypeptide 1 pseudogene
	B1	F. catus	29786714	29787235	+	LOC101099617	pseudo	ferritin heavy chain pseudogene
	B1	S. scrofa	33806868	33807345	+	LOC100156789	pseudo	uncharacterized LOC100156789
C	C1	H. sapiens	37043556	37043023	-	FTHP18	protein coding	ferritin, heavy polypeptide 1 pseudogene 18
	C2		37078401	37077851	-	LOC442445	pseudo	ferritin, heavy polypeptide-like 17 pseudogene
	C1		37523369	37522836	-	FTHP18	protein coding	ferritin, heavy polypeptide 1 pseudogene 18
	C2	P. troglodytes	37558036	37557519	-	LOC473865	pseudo	ferritin heavy polypeptide-like 17
	C1	M. musculus	8962820	8962302	-	Gm5634	protein coding	predicted gene 5634; also known as Fthl17L1
	C2		8976404	8975886	-	Gm14511	protein coding	predicted gene 14511; also known as Fthl17L2
	C3		8986586	8986071	-	Gm14458	protein coding	predicted gene 14458; also known as Fthl17L3
	C1	C. familiaris	31869618	31869060	-	LOC612257	protein coding	ferritin heavy chain-like
	C1	F. catus	32359520	32358999	-	LOC101085694	pseudo	ferritin heavy chain pseudogene
	C1	S. scrofa	36502806	36502341	-	LOC100624935	pseudo	ferritin heavy chain-like
	C2		36562089	36561544	-	LOC100624737	protein coding	ferritin heavy chain-like
D1	37441523		37442074	+	LOC100420326	pseudo	ferritin, heavy polypeptide 1 pseudogene	
D	D2	H. sapiens	37492021	37492539	+	FTHP19	pseudo	ferritin, heavy polypeptide 1 pseudogene 19
	D3		37505334	37505854	+	FTHP27	pseudo	ferritin, heavy polypeptide 1 pseudogene 27
	D1		P. troglodytes	37772599	37773090	+	LOC737664	protein coding
	D1	M. musculus	78470555	78470058	-	Prrg1	protein coding	proline rich Gla (G-carboxyglutamic acid) 1
	D2		9033647	9034162	+	Fthl17	protein coding	ferritin, heavy polypeptide-like 17
	D3		9043736	9044239	+	Gm14499	protein coding	predicted gene 14499; also known as Fthl17L4
	D4		9063176	9063694	+	Gm5635	protein coding	predicted gene 5635; also known as Fthl17L5
	D5		9080053	9080571	+	Gm6826	pseudo	predicted gene 6826; also known as Fthl17L6
	D6		9123467	9123979	+	Gm5753	pseudo	ferritin heavy chain 1 pseudogene
	D1	C. familiaris	32108753	32109313	+	LOC100687930	pseudo	ferritin, heavy polypeptide 1 pseudogene
	D2		32205689	32206259	+	FTHP18	protein coding	ferritin heavy chain-like
	D3		32293541	32294101	+	LOC612281	pseudo	ferritin, heavy polypeptide 1 pseudogene
	D1	F. catus	32571487	32572008	+	LOC102901113	protein coding	ferritin heavy chain-like
	D2		32650295	32650834	+	LOC101082216	pseudo	ferritin heavy chain pseudogene
	D3		32724236	32724775	+	LOC101082471	pseudo	ferritin heavy chain pseudogene
	D1	S. scrofa	36440389	36440925	+	LOC102167173	protein coding	ferritin heavy chain-like
	D2		36839450	36839944	+	LOC100623926	protein coding	transmembrane gamma-carboxyglutamic acid protein 1-like
	D3		36922815	36923360	+	LOC106506938	protein coding	leucine-rich repeat extensin-like protein 5
	D4		37010756	37011292	+	LOC100625618	protein coding	ferritin heavy chain-like

Table S1. Summary of returned tBLASTn hits from multiple species.