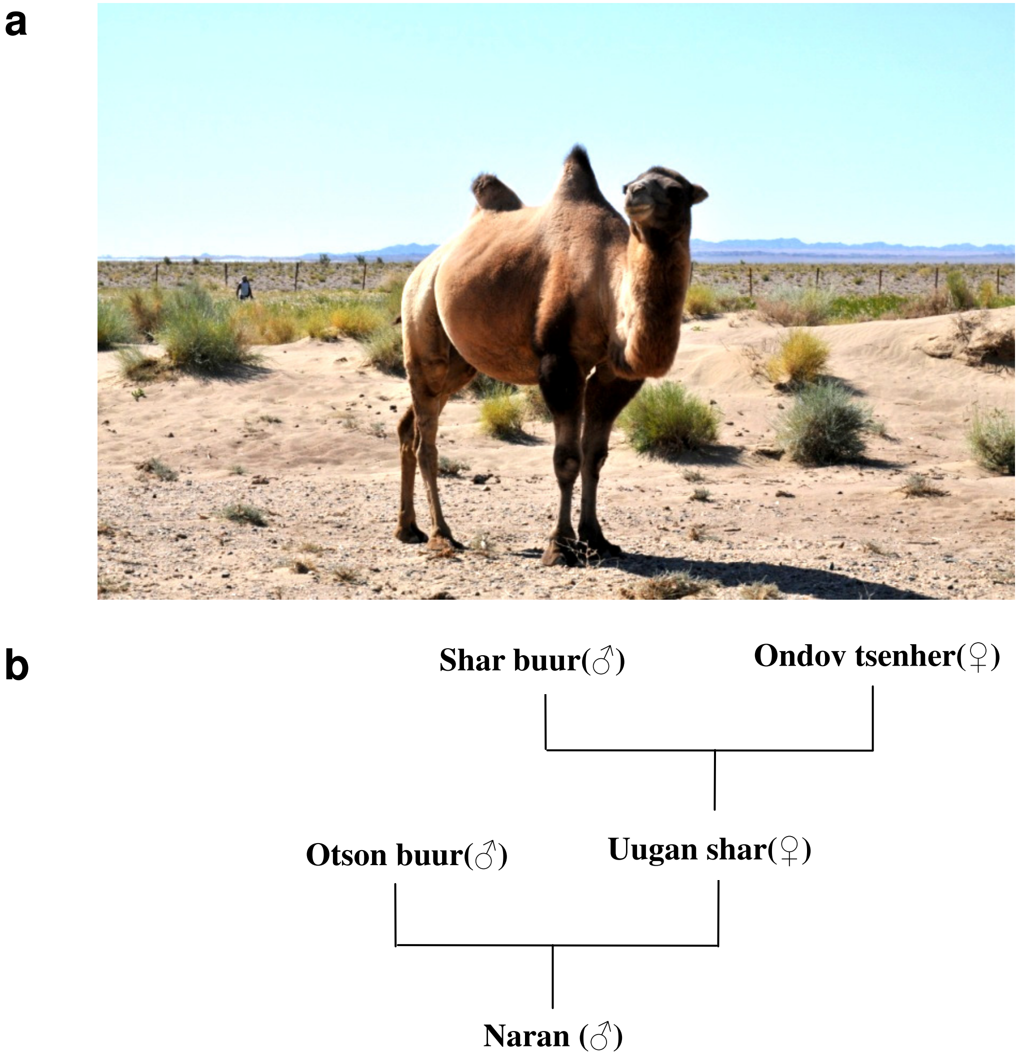
Supplementary Figures and Tables

for

Genome sequences of wild and domestic bactrian camels

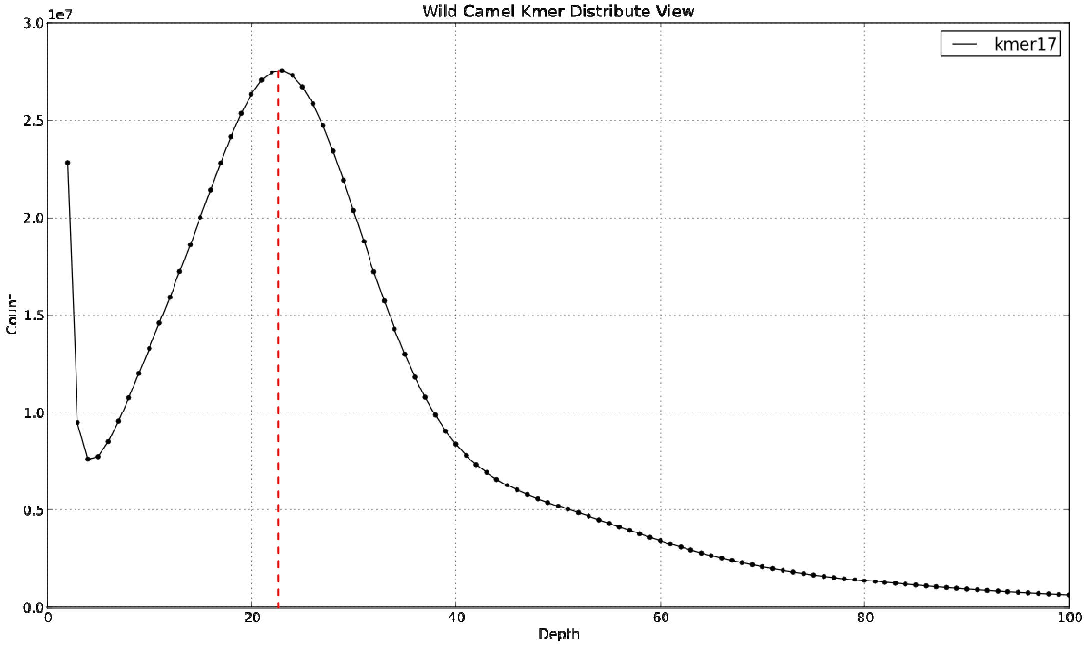
The Bactrian camels Genome Sequencing Consortium

Supplementary Figures

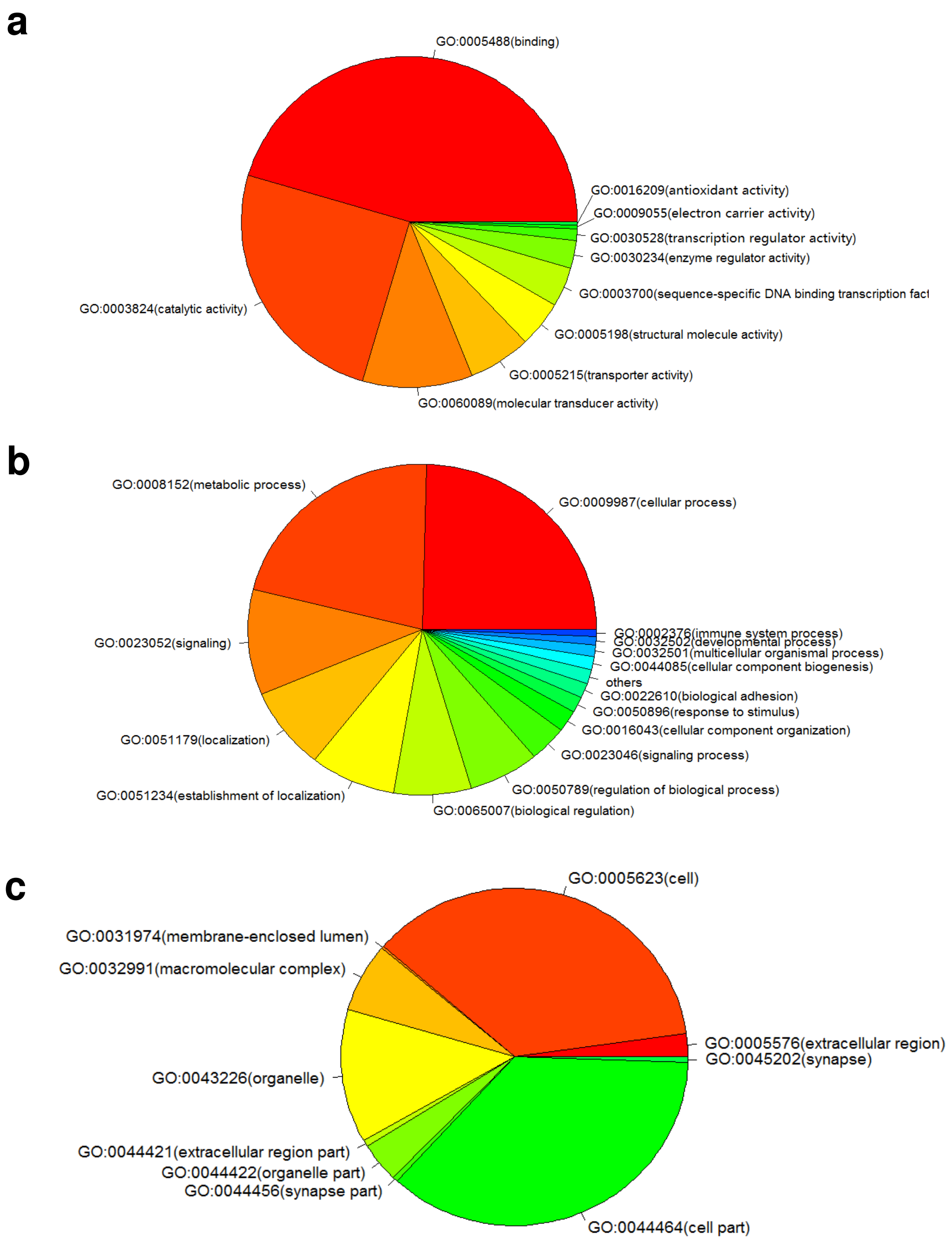


Supplementary Figure S1. The wild bactrian camel named "Naran" sequenced in this study. a) Picture of "Naran". b) Pedigree of "Naran". "Naran" was born in Mongolian Wild Camel Protection Area in 2002, then raised in pens. "Uugan shar" was born in Mongolian Wild Camel Protection Area in 1995, then raised in pens until died in 2004. "Otson buur" was captured from Great Gobi field in 1991, then raised in pens until back to field in 2001. "Ondov tsenher" was captured from Great Gobi field as a baby camel in 1990, and died in 2001. "Shar buur" was captured from Great Gobi field in 1990, then raised in pens until back to field in 1995.

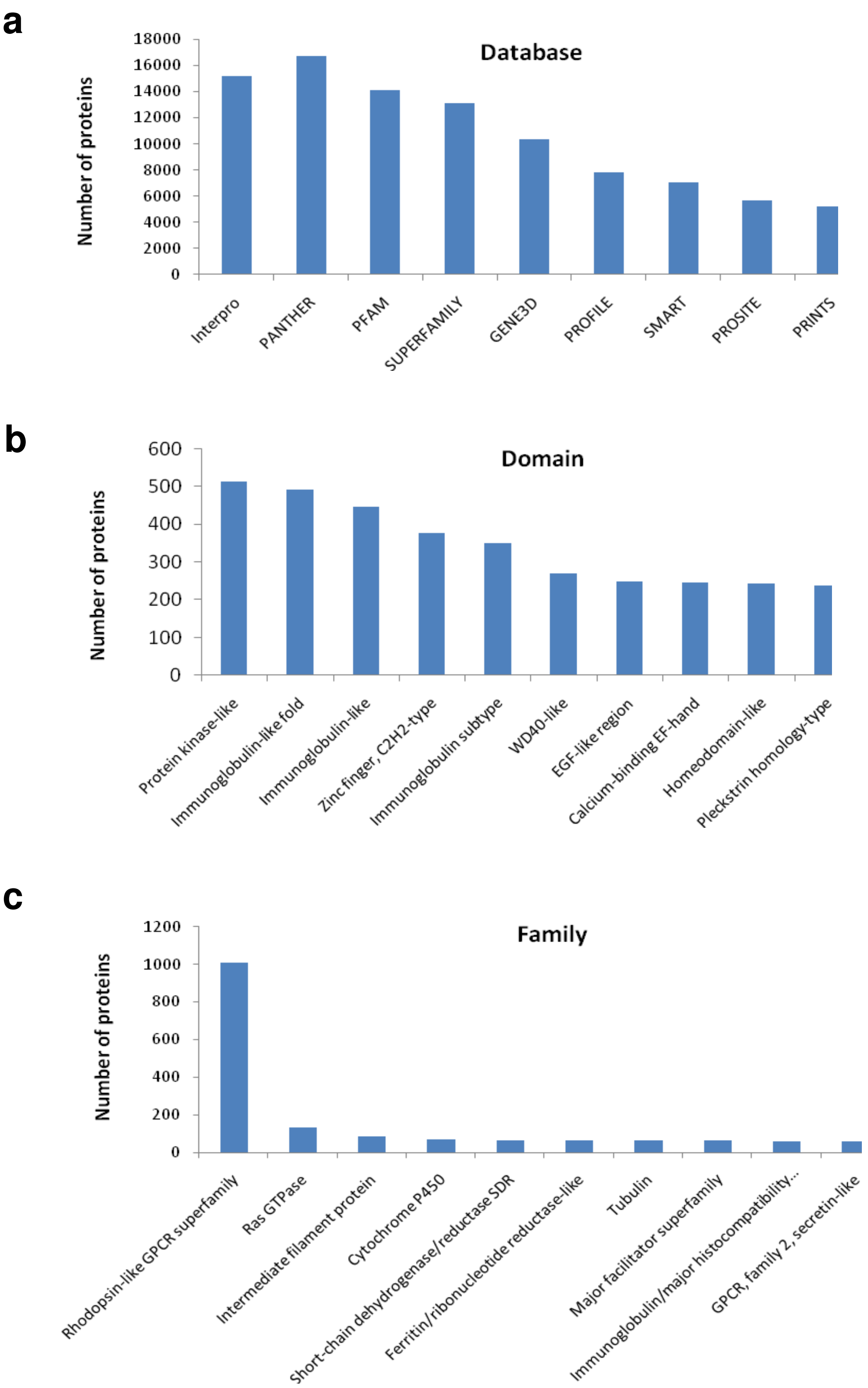
Supplementary Figure S2. Picture of the domestic bactrian camel sequenced in this study.



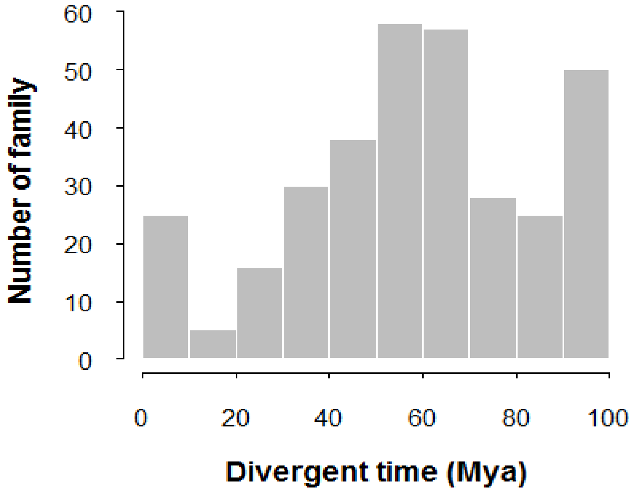
Supplementary Figure S3. Distribution of 17-mer frequency in the raw reads. The depth at the peak of the frequency distribution is 23×.



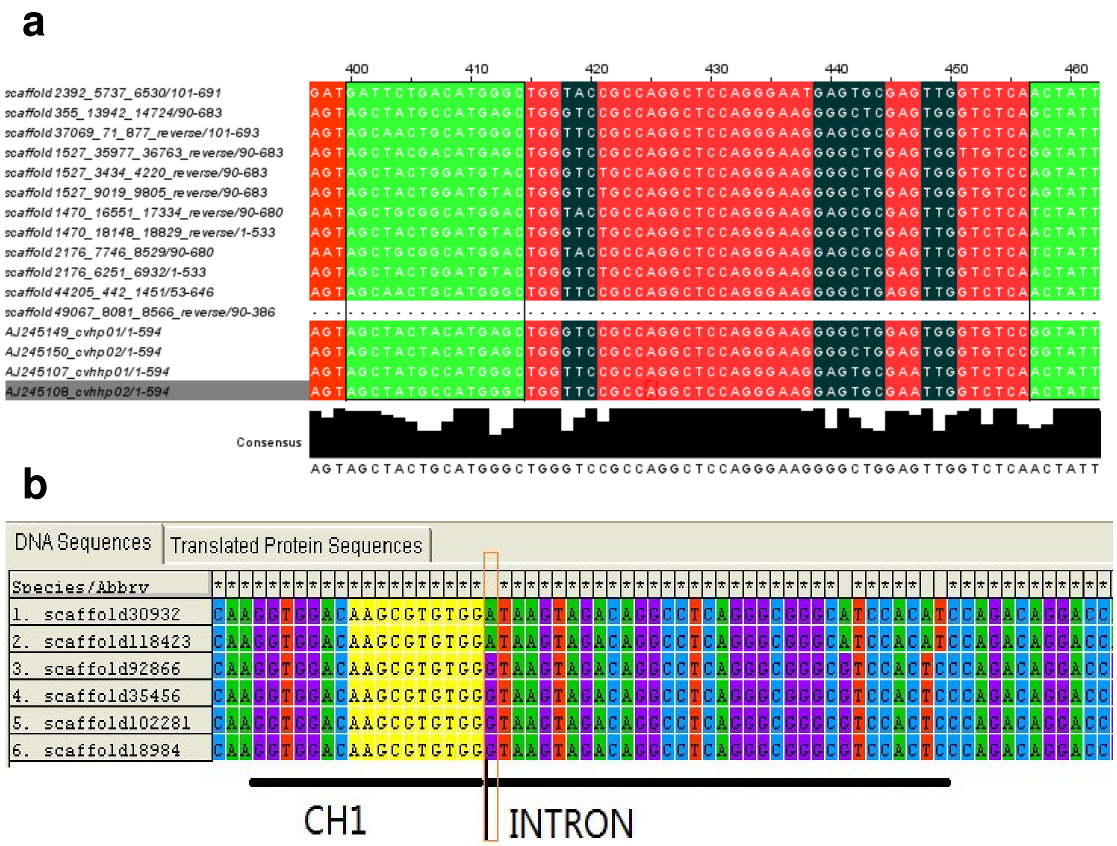
Supplementary Figure S4. Gene Ontology (GO) assignment for the camel genes. There are 12,050 genes corresponding to at least one GO term, which is represented by the second layer in the GO hierarchy. a) Molecular functions (MF). b) Biological processes (BP). c) Cellular components (CC).



Supplementary Figure S5. Number of camel proteins annotated by InterProScan. a) Databases searched by InterProScan. b) Top 10 protein domains found in the camel genome. c) Top 10 protein families found in the camel genome.



Supplementary Figure S6. Distribution of the divergence time between camel and cattle. The divergence time was estimated based on 332 constantly evolving ortholog families in mammals. The distribution indicates that cattle and camel lineages were separated about 57.5 ± 1.5 (mean ± standard error) million years ago (Mya).



Supplementary Figure S7. Gene segments of heavy-chain antibodies (HCAbs). a) Multiple alignment of FR2 regions in VH (heavy-chain variable region of conventional antibodies) and VHH (heavy-chain variable region of HCAbs) sequences. Codons in black denote the site mutations of V37F(Y), G44E, L45R and W47G, which are considered to be critical for the formation of HCAbs. Note that we defined the V regions in scaffold 44205 and scaffold 49067 to be VHH, though only two of the four mutation sites occur. Two VH (AJ245149, AJ245150) and two VHH (AJ245107, AJ245108) sequences from Genbank were used to locate the FR2 region and mutation sites. b) Multiple alignment of Cγ gene sequences. Nucleotide in red frame indicates a splice site mutation of "G to A" at the CH1/INTRON boundary, thus the CH1 region is completely removed in HCAbs.

Supplementary Tables

Supplementary Table S1. Differences between wild and domestic bactrian camels.

Body index/Figure	Domestic bactrian camel		Wild bactrian camel	
	Male(n=15)	Female(n=15)	Male(n=1)	Female(n=4)
Weight (kg)	166.0±0.30	164.0±0.3	189.5	180.3±5.9
Body length (cm)	148.3±.27	146.0±0.4	146.0	142.2±1.9
Chest width (cm)	46.5±0.05	41.4±0.4	46.0	43.8±0.8
Chest depth (cm)	82.8±0.21	78.8±0.3	85.0	78.6±1.9
Ankle perimeter (cm)	23.0±0.40	19.1±0.1	22.0	17.4±0.2
Leg length (cm)	89.8±0.10	87.2±0.1	104.5	101.7±0.2
Hump	hypertrophy and plump		underneath is round Pointed in top	
Figure	short and obesity		tall and big	
Color	claret, black, gray, pale yellow, white		gray	
Tufts	dense		very sparse	
Chaetae	fluffy and long		thick and short	
Limbs	stout		slender	
Sole	round		oval	
Toenail	bald		special pointed	
Running speed	slow		fast	

Supplementary Table S2. Summary of raw data.

Library Type	Platform	Animal	Insert size (bp)	Avg read length (bp)	Total bases (GB)
500bp paired-end	Illumina Genome Analyzer	domestic	430	94.5	49
500bp paired-end	Illumina Genome Analyzer	wild	430	95.5	145
3kb mate-pair	Illumina Genome Analyzer	wild	2800	93.4	26
10kb mate-pair	Applied Biosystems SOLiD	wild	8900	35.0	27
20kb mate-pair	Roche Genome Sequencer FLX System	wild	18800	177.7	0.6

Supplementary Table S3. Sequencing depth and coverage for the assembled scaffolds.

Library Type	Platform	Animal	Reserved base (GB)	Depth	Coverage (%)
500bp paired-end	Illumina Genome Analyzer	domestic	47	24	97.29
500bp paired-end	Illumina Genome Analyzer	wild	114	57	97.66
3kb mate-pair	Illumina Genome Analyzer	wild	16	9	83.71
10kb mate-pair	Applied Biosystems SOLiD	wild	10	8	62.0
20kb mate-pair	Roche Genome Sequencer FLX System	wild	0.3	2	7.8

Supplementary Table S4. Genome size estimation based on C values.

	Reported C value (pico)	Reported Genome size (Gb)	Gb/pico	Gb of 2.41 pico	Gb of 2.86 pico
Human	3.5	3.2	0.91	2.20	2.62
Rat	3.1	2.8	0.90	2.18	2.58
mouse	3.3	2.5	0.76	1.83	2.17
Cattle	3.7	2.87	0.78	1.87	2.22
Pig	3.2	2.72	0.85	2.05	2.43
Average value of estimated genome size for camel (Gb)				2.02	2.40

Supplementary Table S5. Summary of repeat sequences.

Elements	Number of elements	Length occupied (bp)	Percentage of sequence
SINEs:	560273	79896785	3.92%
Alu/B1	0	0	0.00%
MIRs	460330	68890993	3.38%
LINEs:	883074	394112453	19.32%
LINE1	552674	302328505	14.82%
LINE2	280625	79958902	3.92%
L3/CR1	38504	8972347	0.44%
RTE	10913	2766794	0.14%
LTR elements:	321595	118418136	5.81%
ERVL	76625	33162713	1.63%
ERVL-MaLRs	133362	47148327	2.31%
ERV_class I	77025	33253620	1.63%
ERV_class II	23095	1938661	0.10%
DNA elements:	282696	61204314	3.00%
hAT-Charlie	175700	34335816	1.68%
TcMar-Tigger	44496	13820703	0.68%
Unclassified:	6050	1318304	0.06%
Small RNA:	101947	11120002	0.55%
Satellites:	3319	379039	0.02%
Simple repeats:	244141	11307451	0.55%
Low complexity:	225901	9079366	0.45%

Supplementary Table S6. Syntenic blocks between camel and other mammals.

Species	Number of syntenic blocks	Coverage in the camel genome (Mb)	Number of covered orthologs
Human	1,044	1,058	12,009
Mouse	1,070	1,025	11,653
Cattle	1,121	1,008	11,197
Dog	1,005	1,005	10,929

Supplementary Table S7. KEGG pathway enrichment for rapidly evolving genes in the camel genome.

Pathway		Total genes	Rapidly evolving genes	Odds ratio	P-value	FDR
ko00908	Zeatin biosynthesis	1	1	7.63	0.00E+00	0.00E+00
ko03015	mRNA surveillance pathway	63	25	3.05	2.66E-08	3.83E-06
ko05200	Pathways in cancer	242	61	1.94	1.09E-07	6.28E-06
ko04120	Ubiquitin mediated proteolysis	92	31	2.59	9.08E-08	6.54E-06
ko04141	Protein processing in endoplasmic reticulum	124	38	2.36	8.65E-08	8.30E-06
ko03013	RNA transport	113	35	2.38	1.87E-07	8.97E-06
ko04310	Wnt signaling pathway	110	34	2.37	2.85E-07	1.17E-05
ko04721	Synaptic vesicle cycle	49	19	2.97	1.31E-06	4.72E-05
ko05166	HTLV-I infection	189	48	1.95	1.56E-06	4.99E-05
ko05220	Chronic myeloid leukemia	53	19	2.75	5.56E-06	1.60E-04
ko03040	Spliceosome	102	29	2.18	1.16E-05	2.79E-04
ko05215	Prostate cancer	73	23	2.41	1.09E-05	2.85E-04
ko04961	Endocrine and other factor-regulated calcium reabsorption	36	14	2.98	2.06E-05	4.23E-04
ko04666	Fc gamma R-mediated phagocytosis	57	19	2.55	1.97E-05	4.36E-04
ko05211	Renal cell carcinoma	45	16	2.72	2.73E-05	5.24E-04
ko04722	Neurotrophin signaling pathway	83	24	2.22	3.90E-05	7.01E-04
ko05210	Colorectal cancer	42	15	2.73	4.15E-05	7.03E-04
ko04144	Endocytosis	151	37	1.88	4.51E-05	7.21E-04
ko03022	Basal transcription factors	27	11	3.12	6.47E-05	8.88E-04
ko04910	Insulin signaling pathway	100	27	2.07	5.97E-05	9.05E-04
ko05223	Non-small cell lung cancer	39	14	2.75	6.31E-05	9.09E-04
ko05214	Glioma	48	16	2.55	7.12E-05	9.32E-04
ko00020	Citrate cycle (TCA cycle)	28	11	3.00	9.98E-05	1.20E-03
ko04664	Fc epsilon RI signaling pathway	49	16	2.50	9.58E-05	1.20E-03

ko04012	ErbB signaling pathway	60	18	2.30	1.54E-04	1.78E-03
ko05221	Acute myeloid leukemia	42	14	2.55	1.69E-04	1.87E-03
ko00051	Fructose and mannose metabolism	30	11	2.80	2.20E-04	2.35E-03
ko04520	Adherens junction	57	17	2.28	2.37E-04	2.44E-03
ko03050	Proteasome	39	13	2.55	2.60E-04	2.58E-03
ko04150	mTOR signaling pathway	36	12	2.55	4.01E-04	3.85E-03
ko05162	Measles	80	21	2.01	4.67E-04	4.34E-03
ko05212	Pancreatic cancer	46	14	2.33	5.24E-04	4.72E-03
ko03010	Ribosome	97	24	1.89	5.81E-04	5.07E-03
ko04725	Cholinergic synapse	88	22	1.91	7.67E-04	6.31E-03
ko04530	Tight junction	104	25	1.84	7.51E-04	6.36E-03
ko04930	Type II diabetes mellitus	34	11	2.47	8.44E-04	6.48E-03
ko04962	Vasopressin-regulated water reabsorption	34	11	2.47	8.44E-04	6.48E-03
ko05217	Basal cell carcinoma	43	13	2.31	8.11E-04	6.49E-03
ko04662	B cell receptor signaling pathway	49	14	2.19	1.10E-03	8.12E-03
ko05213	Endometrial cancer	40	12	2.29	1.26E-03	9.06E-03
ko05100	Bacterial invasion of epithelial cells	50	14	2.14	1.38E-03	9.27E-03
ko00290	Valine, leucine and isoleucine biosynthesis	5	3	4.58	1.32E-03	9.28E-03
ko04070	Phosphatidylinositol signaling system	65	17	2.00	1.36E-03	9.30E-03
ko04010	MAPK signaling pathway	190	39	1.57	1.51E-03	9.88E-03
ko00310	Lysine degradation	41	12	2.24	1.62E-03	1.01E-02
ko04340	Hedgehog signaling pathway	41	12	2.24	1.62E-03	1.01E-02
ko00720	Carbon fixation pathways in prokaryotes	15	6	3.06	1.62E-03	1.04E-02
ko05164	Influenza A	106	24	1.73	2.26E-03	1.31E-02
ko05168	Herpes simplex infection	106	24	1.73	2.26E-03	1.31E-02
ko00230	Purine metabolism	135	29	1.65	2.37E-03	1.34E-02
ko04712	Circadian rhythm - plant	3	2	5.09	2.25E-03	1.35E-02
ko04210	Apoptosis	58	15	1.98	2.56E-03	1.40E-02

ko04620	Toll-like receptor signaling pathway	58	15	1.98	2.56E-03	1.40E-02
ko04916	Melanogenesis	75	18	1.84	3.12E-03	1.63E-02
ko05120	Epithelial cell signaling in Helicobacter pylori infection	54	14	1.98	3.20E-03	1.65E-02
ko04350	TGF-beta signaling pathway	59	15	1.94	3.09E-03	1.65E-02
ko04111	Cell cycle - yeast	49	13	2.03	3.27E-03	1.65E-02
ko04110	Cell cycle	92	21	1.75	3.35E-03	1.66E-02
ko03018	RNA degradation	60	15	1.91	3.71E-03	1.80E-02
ko04914	Progesterone-mediated oocyte maturation	60	15	1.91	3.71E-03	1.80E-02
ko05130	Pathogenic Escherichia coli infection	36	10	2.12	5.00E-03	2.36E-02
ko04660	T cell receptor signaling pathway	73	17	1.78	5.37E-03	2.49E-02
ko04960	Aldosterone-regulated sodium reabsorption	32	9	2.15	6.14E-03	2.81E-02
ko04810	Regulation of actin cytoskeleton	156	31	1.52	6.32E-03	2.85E-02
ko04920	Adipocytokine signaling pathway	53	13	1.87	7.03E-03	3.12E-02
ko04510	Focal adhesion	158	31	1.50	7.67E-03	3.30E-02
ko05110	Vibrio cholerae infection	43	11	1.95	7.66E-03	3.34E-02
ko04370	VEGF signaling pathway	54	13	1.84	8.37E-03	3.55E-02
ko05222	Small cell lung cancer	66	15	1.74	9.92E-03	4.14E-02
ko00561	Glycerolipid metabolism	35	9	1.96	1.21E-02	4.98E-02

Supplementary Table S8. Number of SNPs in the whole genome.

Genotype	Heterozygous SNP wild camel	Heterozygous SNP domestic camel	Homozygous SNP
A/G	689,145	731,046	599,458
T/C	688,645	729,530	600,690
A/T	148,035	167,097	125,867
A/C	161,266	176,459	136,587
T/G	162,547	177,265	136,858
C/G	136,782	148,045	111,572
Transition (ts)	1,377,790	1,460,576	1,200,148
Transversion (tv)	608,630	668,866	510,884
Total	1,986,420	2,129,442	1,711,032
Ts / tv ratio	2.26	2.18	2.35

Supplementary Table S9. Number of small indels in the whole genome.

Indel length	Heterozygous indel wild camel	Heterozygous indel domestic camel	Homozygous indel
1bp	93,883	99,411	123,023
2bp	63,950	62,961	42,703
3bp	22,107	24,797	19,718
4bp	23,526	22,937	20,056
5-10bp	23,265	21,316	20,525
>10bp	12,374	9,856	10,476
Total	239,105	241,278	236,501

Supplementary Table S10. Distribution of SNPs in different genomic regions.

Region	Total length	Heterozygous SNP wild camel	Heterozygous SNP domestic camel
Genome	2,010,007,732	1,986,420	2,129,442
Gene interval	1,539,141,966	1,570,921	1,705,319
Exon	27,521,659	19,350	16,586
Intron	472,747,974	396,149	407,537
Synonymous	9,173,886	8,958	7,773
Nonsynonymous	18,347,773	10,126	8,718

Supplementary Table S11. Genomic blocks where the heterozygosity of the domestic camel is significantly lower than the wild camel.

Scaffold[a]	Start	End	Wild.SNP[b]	Dome.SNP[c]	Gene[d]	Gene annotation
scaffold1022	621001	730001	514	2	evm.model.scaffold1022.8_1	
scaffold1045	384001	406001	91	2	evm.model.scaffold1045.2_1	ribosomal protein L3
scaffold1095	5458001	5471001	48	1	evm.model.scaffold1095.67_1	chondroitin sulfate N-acetylgalactosaminyltransferase 1-like
scaffold1108	631001	649001	62	1	evm.model.scaffold1108.5_1	mKIAA1821 protein
scaffold1108	3982001	4154001	777	14	evm.model.scaffold1108.73_1	protein phosphatase 1 regulatory inhibitor subunit 16B-like
scaffold1108	3982001	4154001	777	14	evm.model.scaffold1108.74_1	Protein FAM83D
scaffold1108	3982001	4154001	777	14	evm.model.scaffold1108.75_1	probable ATP-dependent RNA helicase DHX35 isoform 2
scaffold1113	97001	135001	143	2	evm.model.scaffold1113.3_1	
scaffold114	392001	431001	175	2	evm.model.scaffold114.15_1	MAD1-like 1 protein
scaffold1141	1316001	1340001	80	0	evm.model.scaffold1141.14_1	60S ribosomal protein L9
scaffold122	171001	191001	71	0	evm.model.scaffold122.10_1	olfactory receptor 51I2-like
scaffold122	171001	191001	71	0	evm.model.scaffold122.11_1	olfactory receptor 51I2-like
scaffold122	225001	249001	91	2	evm.model.scaffold122.15_1	olfactory receptor, family 52, subfamily I, member 2
scaffold122	225001	249001	91	2	evm.model.scaffold122.16_1	Ssu72 RNA polymerase II CTD phosphatase homolog
scaffold1221	105001	119001	66	14	evm.model.scaffold1221.7_1	olfactory receptor, family 2, subfamily G, member 6-like
scaffold123	732001	756001	88	4	evm.model.scaffold123.11_1	
scaffold125	622001	636001	46	1	evm.model.scaffold125.19_1	vasoactive intestinal peptide receptor 2
scaffold125	163001	217001	246	2	evm.model.scaffold125.8_1	
scaffold1250	356001	373001	53	1	evm.model.scaffold1250.9_1	CTSB protein
scaffold1260	9001	23001	53	8	evm.model.scaffold1260.1_1	cOR8V11 olfactory receptor family 8 subfamily V-like
scaffold1260	9001	23001	53	8	evm.model.scaffold1260.2_1	olfactory receptor 1087-like

scaffold1309	475001	532001	216	5	evm.model.scaffold1309.7_1	ribosomal protein L32-like
scaffold1322	490001	549001	218	5	evm.model.scaffold1322.7_1	
scaffold137	160001	189001	139	18	evm.model.scaffold137.1_1	Rho GDP dissociation inhibitor (GDI) alpha
scaffold137	212001	237001	92	4	evm.model.scaffold137.2_1	intersectin 1
scaffold138	1905001	1920001	50	1	evm.model.scaffold138.53_1	
scaffold1408	1426001	1445001	67	0	evm.model.scaffold1408.25_1	hypothetical protein
scaffold1413	2030001	2148001	509	5	evm.model.scaffold1413.26_1	secreted frizzled-related protein 2-like
scaffold1413	2030001	2148001	509	5	evm.model.scaffold1413.27_1	ring finger protein 175
scaffold1413	2030001	2148001	509	5	evm.model.scaffold1413.28_1	Ring finger protein 175 isoform 5
scaffold1413	2030001	2148001	509	5	evm.model.scaffold1413.29_1	toll-like receptor 2 precursor
scaffold1428	4455001	4466001	49	5	evm.model.scaffold1428.65_1	hypothetical protein
scaffold1428	4833001	4848001	45	0	evm.model.scaffold1428.74_1	eukaryotic translation initiation factor 3 subunit L-like isoform 2
scaffold1430	363001	400001	195	20	evm.model.scaffold1430.13_1	hypothetical protein
scaffold1431	577001	591001	48	1	evm.model.scaffold1431.9_1	lipopolysaccharide-induced TNF factor
scaffold1437	627001	645001	75	5	evm.model.scaffold1437.12_1	
scaffold1437	747001	794001	189	8	evm.model.scaffold1437.17_1	
scaffold146	1130001	1143001	44	1	evm.model.scaffold146.19_1	coiled-coil domain-containing protein 3-like
scaffold1481	267001	292001	94	5	evm.model.scaffold1481.23_1	olfactory receptor, family 11, subfamily H, member 1 isoform 2
scaffold1481	267001	292001	94	5	evm.model.scaffold1481.24_1	olfactory receptor 11H6-like
scaffold1506	987001	1008001	81	2	evm.model.scaffold1506.12_1	emerin-like
scaffold1533	2300001	2323001	102	5	evm.model.scaffold1533.25_1	TRIM5/CypA fusion protein
scaffold1533	1009001	1022001	41	1	evm.model.scaffold1533.8_1	5-hydroxytryptamine receptor 5B-like
scaffold1533	1009001	1022001	41	1	evm.model.scaffold1533.9_1	5-hydroxytryptamine receptor 5B-like
scaffold1561	386001	422001	127	7	evm.model.scaffold1561.9_1	Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI) (Endozepine) (EP)
scaffold1577	359001	376001	55	4	evm.model.scaffold1577.5_1	C-type lectin domain family 4 member K (Langerin)

						(CD207 antigen)
scaffold1578	62001	89001	100	1	evm.model.scaffold1578.3_1	arylsulfatase E
scaffold161	289001	364001	291	7	evm.model.scaffold161.5_1	platelet endothelial aggregation receptor 1-like
scaffold161	289001	364001	291	7	evm.model.scaffold161.6_1	platelet endothelial aggregation receptor 1
scaffold1679	316001	344001	95	3	evm.model.scaffold1679.5_1	
scaffold1683	4510001	4524001	62	7	evm.model.scaffold1683.42_1	Epithelial membrane protein-1 (EMP-1) (Tumor-associated membrane protein)
scaffold1683	999001	1023001	91	1	evm.model.scaffold1683.6_1	capping protein alpha 3
scaffold1683	6077001	6103001	113	0	evm.model.scaffold1683.71_1	natural killer cell receptor Ly49P2-like
scaffold1692	62001	86001	104	1	evm.model.scaffold1692.1_1	hypothetical protein
scaffold1749	1109001	1123001	42	0	evm.model.scaffold1749.11_1	hypothetical protein
scaffold175	1732001	1746001	35	0	evm.model.scaffold175.37_1	somatostatin receptor type 3
scaffold1753	1	21001	104	0	evm.model.scaffold1753.1_1	syntrophin 5
scaffold176	79001	91001	37	0	evm.model.scaffold176.4_1	Alpha-sarcoglycan precursor (Alpha-SG) (Adhalin) (50 kDa dystrophin-associated glycoprotein) (50DAG) (Dystroglycan-2)
scaffold176	79001	91001	37	0	evm.model.scaffold176.5_1	spermatid-specific linker histone H1-like protein
scaffold1891	608001	631001	79	0	evm.model.scaffold1891.3_1	
scaffold200	327001	341001	64	8	evm.model.scaffold200.5_1	tripartite motif-containing protein 29-like
scaffold216	1336001	1358001	84	5	evm.model.scaffold216.40_1	
scaffold216	1555001	1582001	98	1	evm.model.scaffold216.50_1	fibroblast growth factor
scaffold218	266001	376001	637	23	evm.model.scaffold218.11_1	GDP-mannose 4,6 dehydratase-like
scaffold218	584001	620001	148	10	evm.model.scaffold218.18_1	zonadhesin-like
scaffold218	8001	41001	189	18	evm.model.scaffold218.2_1	
scaffold218	854001	888001	165	11	evm.model.scaffold218.28_1	Serpin B9 (Cytoplasmic antiproteinase 3) (CAP-3) (CAP3) (Proteinase inhibitor 9)
scaffold218	8001	41001	189	18	evm.model.scaffold218.3_1	
scaffold218	162001	258001	496	31	evm.model.scaffold218.6_1	

scaffold218	162001	258001	496	31	evm.model.scaffold218.7_1	GDP-mannose 4,6 dehydratase
scaffold218	266001	376001	637	23	evm.model.scaffold218.9_1	
scaffold2195	95001	115001	78	1	evm.model.scaffold2195.3_1	potassium inwardly-rectifying channel J15 isoform 3
scaffold233	808001	843001	165	5	evm.model.scaffold233.30_1	
scaffold233	808001	843001	165	5	evm.model.scaffold233.31_1	rearranged T-cell receptor delta-chain/ Vdelta2-Ddeltas-Jdelta1 - pig (fragment) dbj|BAA08523.1| rearranged T-cell receptor delta-chain/ Vdelta2-Ddeltas-Jdelta1
scaffold234	721001	751001	148	2	evm.model.scaffold234.21_1	mas-related G-protein coupled receptor MRG-like
scaffold2519	5585001	5618001	168	7	evm.model.scaffold2519.35_1	hypothetical protein
scaffold253	528001	546001	63	5	evm.model.scaffold253.12_1	D-amino-acid oxidase
scaffold25522	645001	662001	69	8	evm.model.scaffold25522.13_1	hypothetical protein
scaffold265	550001	575001	83	0	evm.model.scaffold265.12_1	
scaffold2781	11001	31001	72	1	evm.model.scaffold2781.2_1	vomeronasal 1 receptor bosTauV1R403
scaffold293	2196001	2219001	95	3	evm.model.scaffold293.28_1	
scaffold303	1157001	1213001	241	9	evm.model.scaffold303.7_1	transmembrane protease serine 5-like
scaffold306	1951001	1967001	51	0	evm.model.scaffold306.36_1	small proline-rich protein 1B
scaffold309	756001	782001	98	0	evm.model.scaffold309.21_1	transcription elongation regulator 1-like protein-like
scaffold312	485001	535001	240	15	evm.model.scaffold312.5_1	prominin-1-like
scaffold322	518001	533001	43	0	evm.model.scaffold322.15_1	olfactory receptor 281-like
scaffold331	1379001	1407001	120	3	evm.model.scaffold331.27_1	
scaffold335	940001	952001	39	1	evm.model.scaffold335.13_1	
scaffold339	2427001	2438001	48	2	evm.model.scaffold339.29_1	olfactory receptor MOR276-1
scaffold339	2918001	2936001	61	1	evm.model.scaffold339.50_1	olfactory receptor Olr12
scaffold339	2918001	2936001	61	1	evm.model.scaffold339.51_1	olfactory receptor, family 6, subfamily F, member 1-like
scaffold345	1058001	1098001	140	1	evm.model.scaffold345.35_1	
scaffold347	448001	464001	73	9	evm.model.scaffold347.11_1	fibromodulin precursor

scaffold350	357001	393001	168	12	evm.model.scaffold350.12_1	
scaffold350	54001	109001	193	0	evm.model.scaffold350.3_1	
scaffold350	3310001	3375001	283	3	evm.model.scaffold350.69_1	ubiquitin specific peptidase 17
scaffold350	3310001	3375001	283	3	evm.model.scaffold350.70_1	Protein FAM90A1
scaffold350	3409001	3504001	504	9	evm.model.scaffold350.71_1	neuronal membrane glycoprotein M6-a-like
scaffold351	514001	563001	254	22	evm.model.scaffold351.12_1	olfactory receptor 10J1-like
scaffold351	514001	563001	254	22	evm.model.scaffold351.13_1	olfactory receptor, family 10, subfamily J, member 1, partial
scaffold351	514001	563001	254	22	evm.model.scaffold351.14_1	olfactory receptor 10J4
scaffold351	566001	586001	74	3	evm.model.scaffold351.15_1	olfactory receptor 10J1-like
scaffold363	590001	608001	60	1	evm.model.scaffold363.12_1	Calsyntenin-1 precursor (Alcadein-alpha) (Alc-alpha) (Alzheimer-related cadherin-like protein) (Non-classical cadherin XB31alpha)
scaffold365	4490001	4533001	179	0	evm.model.scaffold365.92_1	olfactory receptor 859 (predicted)
scaffold365	4490001	4533001	179	0	evm.model.scaffold365.93_1	olfactory receptor MOR258-6-like
scaffold365	4490001	4533001	179	0	evm.model.scaffold365.94_1	olfactory receptor 2A12-like
scaffold372	814001	843001	112	2	evm.model.scaffold372.18_1	
scaffold372	814001	843001	112	2	evm.model.scaffold372.19_1	
scaffold38411	309001	350001	123	1	evm.model.scaffold38411.5_1	protein phosphatase 1L-like
scaffold38411	656001	667001	36	0	evm.model.scaffold38411.7_1	otolin-1-like
scaffold385	418001	437001	59	0	evm.model.scaffold385.11_1	
scaffold385	745001	772001	100	1	evm.model.scaffold385.27_1	olfactory receptor 523
scaffold388	408001	450001	181	1	evm.model.scaffold388.2_1	protein arginine methyltransferase 8
scaffold390	1191001	1208001	61	3	evm.model.scaffold390.39_1	
scaffold400	1242001	1265001	83	3	evm.model.scaffold400.22_1	olfactory receptor 10A5-like
scaffold400	1242001	1265001	83	3	evm.model.scaffold400.23_1	olfactory receptor 10A5 (HP3) (Olfactory receptor-like protein JCG6)
scaffold41	541001	572001	110	0	evm.model.scaffold41.11_1	p2X purinoceptor 1-like

scaffold41513	48001	64001	60	2	evm.model.scaffold41513.1_1	Hypothetical BAF-like protein C20orf179
scaffold418	590001	610001	75	1	evm.model.scaffold418.27_1	olfactory receptor 7E24
scaffold418	653001	690001	131	2	evm.model.scaffold418.30_1	olfactory receptor 7A10-like
scaffold444	1043001	1059001	58	1	evm.model.scaffold444.7_1	60S ribosomal protein L13
scaffold449	97001	114001	53	0	evm.model.scaffold449.10_1	matrix metalloproteinase-19-like
scaffold449	97001	114001	53	0	evm.model.scaffold449.9_1	transmembrane protein 198-like
scaffold465	85001	96001	36	0	evm.model.scaffold465.1_1	transcription cofactor vestigial-like protein 4
scaffold465	1421001	1440001	69	0	evm.model.scaffold465.24_1	prothyroliberin-like
scaffold466	577001	607001	105	0	evm.model.scaffold466.13_1	small subunit ribosomal protein S2e
scaffold475	557001	600001	155	1	evm.model.scaffold475.3_1	laforin isoform a
scaffold486	451001	462001	36	0	evm.model.scaffold486.13_1	
scaffold493	1267001	1279001	37	0	evm.model.scaffold493.4_1	40S ribosomal protein SA-like
scaffold497	315001	362001	164	1	evm.model.scaffold497.4_1	Sodium channel protein type 7 subunit alpha (Sodium channel protein type VII subunit alpha) (Putative voltage-gated sodium channel subunit alpha Nax) (Sodium channel protein cardiac and skeletal muscle subunit alpha)
scaffold524	143001	164001	71	0	evm.model.scaffold524.4_1	coproporphyrinogen-III oxidase-like
scaffold539	390001	428001	142	3	evm.model.scaffold539.1_1	RING finger protein 26-like
scaffold539	390001	428001	142	3	evm.model.scaffold539.2_1	RING finger protein 26-like
scaffold539	390001	428001	142	3	evm.model.scaffold539.3_1	Rnf26 protein
scaffold547	248001	283001	136	10	evm.model.scaffold547.7_1	hypothetical protein
scaffold548	759001	801001	148	2	evm.model.scaffold548.14_1	
scaffold55	5001	24001	113	18	evm.model.scaffold55.1_1	Olfactory receptor 1F12 (Hs6M1-35P)
scaffold565	1598001	1611001	41	0	evm.model.scaffold565.3_1	hypothetical protein
scaffold573	17001	50001	117	8	evm.model.scaffold573.2_1	cystatin-C-like
scaffold60782	7229001	7250001	73	0	evm.model.scaffold60782.88_1	G-protein coupled receptor 183-like
scaffold637	449001	467001	63	1	evm.model.scaffold637.3_1	

scaffold639	179001	195001	60	1	evm.model.scaffold639.5_1	gastricsin-like
scaffold650	21001	51001	118	4	evm.model.scaffold650.3_1	hypothetical protein
scaffold650	306001	325001	59	0	evm.model.scaffold650.30_1	olfactory receptor Olr78
scaffold650	306001	325001	59	0	evm.model.scaffold650.31_1	Matrix metalloproteinase-26 precursor (MMP-26) (Matrilysin-2) (Endometase)
scaffold650	395001	414001	68	2	evm.model.scaffold650.37_1	olfactory receptor, family 51, subfamily T, member 1
scaffold650	395001	414001	68	2	evm.model.scaffold650.38_1	putative olfactory receptor 51H1-like
scaffold650	395001	414001	68	2	evm.model.scaffold650.39_1	olfactory receptor 555
scaffold650	21001	51001	118	4	evm.model.scaffold650.4_1	olfactory receptor 614
scaffold650	21001	51001	118	4	evm.model.scaffold650.5_1	olfactory receptor 611
scaffold672	473001	502001	115	0	evm.model.scaffold672.8_1	DCN1-like protein 3
scaffold681	2188001	2201001	37	0	evm.model.scaffold681.19_1	sorting nexin-20
scaffold69	342001	365001	73	0	evm.model.scaffold69.15_1	
scaffold692	1506001	1541001	143	8	evm.model.scaffold692.21_1	Uncharacterized protein C1orf131
scaffold698	546001	561001	63	4	evm.model.scaffold698.6_1	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3
scaffold698	546001	561001	63	4	evm.model.scaffold698.7_1	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 precursor isoform 1
scaffold698	546001	561001	63	4	evm.model.scaffold698.8_1	procollagen-lysine,2-oxoglutarate 5-dioxygenase 3-like
scaffold699	2220001	2352001	518	2	evm.model.scaffold699.23_1	haloacid dehalogenase-like hydrolase domain-containing protein 2
scaffold70	5001	16001	69	16	evm.model.scaffold70.1_1	Leukocyte immunoglobulin-like receptor subfamily B member 3 precursor (Leukocyte immunoglobulin-like receptor 3) (LIR-3) (Immunoglobulin-like transcript 5) (ILT-5) (Monocyte inhibitory receptor HL9) (CD85a antigen)
scaffold70	150001	162001	45	4	evm.model.scaffold70.11_1	osteoclast-associated receptor-like
scaffold705	125001	139001	44	0	evm.model.scaffold705.3_1	high mobility group protein B4-like

scaffold719	976001	987001	47	4	evm.model.scaffold719.10_1	claudin-20-like
scaffold726	4093001	4121001	111	2	evm.model.scaffold726.62_1	bradykinin B1 receptor
scaffold726	4093001	4121001	111	2	evm.model.scaffold726.63_1	bradykinin receptor B1
scaffold728	106001	152001	176	1	evm.model.scaffold728.2_1	calponin-2
scaffold743	781001	806001	97	0	evm.model.scaffold743.33_1	free fatty acid receptor 2-like
scaffold761	692001	714001	83	1	evm.model.scaffold761.14_1	60S ribosomal protein L6-like
scaffold782	39001	57001	56	0	evm.model.scaffold782.2_1	hypothetical protein
scaffold803	3878001	3902001	83	2	evm.model.scaffold803.38_1	T-cell receptor gamma chain
scaffold803	3905001	3920001	55	0	evm.model.scaffold803.40_1	TCR gamma V9-1-like
scaffold804	2888001	2908001	81	2	evm.model.scaffold804.40_1	T cell receptor beta
scaffold804	2888001	2908001	81	2	evm.model.scaffold804.41_1	hypothetical protein
scaffold804	2888001	2908001	81	2	evm.model.scaffold804.42_1	V_segment translation product
scaffold804	2888001	2908001	81	2	evm.model.scaffold804.43_1	putative T cell receptor beta chain variable region
scaffold823	231001	263001	110	4	evm.model.scaffold823.8_1	probable palmitoyltransferase ZDHHC19-like
scaffold823	231001	263001	110	4	evm.model.scaffold823.9_1	zinc finger, DHHC-type containing 19
scaffold824	194001	212001	81	6	evm.model.scaffold824.1_1	
scaffold829	102001	113001	35	0	evm.model.scaffold829.12_1	olfactory receptor 652
scaffold841	1343001	1355001	36	0	evm.model.scaffold841.18_1	Seizure related 6 homolog (mouse)-like
scaffold849	2645001	2671001	98	2	evm.model.scaffold849.68_1	
scaffold851	41001	92001	182	1	evm.model.scaffold851.2_1	
scaffold857	32001	44001	82	20	evm.model.scaffold857.2_1	olfactory receptor Olr1014
scaffold86	155001	171001	60	5	evm.model.scaffold86.5_1	histone H2A type 3
scaffold86314	149001	192001	210	11	evm.model.scaffold86314.5_1	Fc receptor-like protein 1-like
scaffold875	3272001	3310001	115	0	evm.model.scaffold875.44_1	cytochrome P450, family 24, subfamily A, polypeptide 1-like isoform 2
scaffold880	2620001	2662001	215	19	evm.model.scaffold880.66_1	
scaffold88230	160001	184001	99	4	evm.model.scaffold88230.13_1	olfactory receptor, family 51, subfamily V, member 1 (predicted)

scaffold88230	190001	206001	62	4	evm.model.scaffold88230.15_1	olfactory receptor, family 52, subfamily A, member 1 (predicted)
scaffold88230	190001	206001	62	4	evm.model.scaffold88230.16_1	olfactory receptor, family 52, subfamily A, member 5
scaffold896	88001	127001	148	0	evm.model.scaffold896.4_1	
scaffold921	534001	552001	60	0	evm.model.scaffold921.9_1	glyceraldehyde-3-phosphate dehydrogenase
scaffold931	7845001	7863001	61	2	evm.model.scaffold931.92_1	Adaptor-related protein complex AP-1, sigma 3
scaffold96841	24001	39001	64	3	evm.model.scaffold96841.3_1	olfactory receptor 10A7-like
scaffold977	1087001	1131001	191	6	evm.model.scaffold977.11_1	tetratricopeptide repeat domain 15
scaffold98824	480001	499001	83	5	evm.model.scaffold98824.11_1	60S ribosomal protein L28-like
scaffold98824	480001	499001	83	5	evm.model.scaffold98824.12_1	Serine protease HTRA1 precursor (L56) (Serine protease 11)

[a] Serial number of scaffolds.

[b] Number of SNPs in the wild camel.

[c] Number of SNPs in the domestic camel.

[d] Genes contained in the block, represented by serial number.

Supplementary Table S12. GO enrichment for the genes in Supplementary Table S11.

GO ID	Ontology	Term	Selected genes	Total genes	Log odds ratio	FDR
GO:0009987	biological_process	cellular process	73/180	5025/19629	0.66	8.66E-04
GO:0007154	biological_process	cell communication	49/180	1787/19629	1.58	3.15E-10
GO:0007165	biological_process	signal transduction	48/180	1704/19629	1.62	2.64E-10
GO:0007166	biological_process	cell surface receptor signaling pathway	47/180	1235/19629	2.05	1.86E-14
GO:0007186	biological_process	G-protein coupled receptor signaling pathway	47/180	1162/19629	2.14	3.82E-15
GO:0023052	biological_process	signaling	48/180	1721/19629	1.60	3.15E-10
GO:0050789	biological_process	regulation of biological process	49/180	2057/19629	1.38	3.14E-08
GO:0050794	biological_process	regulation of cellular process	48/180	2028/19629	1.37	6.01E-08
GO:0050896	biological_process	response to stimulus	48/180	1963/19629	1.41	2.23E-08
GO:0051716	biological_process	cellular response to stimulus	48/180	1773/19629	1.56	8.32E-10
GO:0065007	biological_process	biological regulation	49/180	2153/19629	1.31	1.34E-07
GO:0016020	cellular_component	membrane	56/180	3213/19629	0.93	7.90E-05
GO:0016021	cellular_component	integral to membrane	48/180	1852/19629	1.50	3.49E-09
GO:0031224	cellular_component	intrinsic to membrane	48/180	1854/19629	1.50	3.49E-09
GO:0044425	cellular_component	membrane part	50/180	2013/19629	1.44	4.88E-09
GO:0004871	molecular_function	signal transducer activity	44/180	1337/19629	1.84	1.69E-11
GO:0004872	molecular_function	receptor activity	45/180	1314/19629	1.90	2.57E-12
GO:0004888	molecular_function	transmembrane signaling receptor activity	44/180	1109/19629	2.11	4.03E-14
GO:0004984	molecular_function	olfactory receptor activity	38/180	742/19629	2.48	3.82E-15
GO:0038023	molecular_function	signaling receptor activity	44/180	1193/19629	2.01	4.66E-13
GO:0060089	molecular_function	molecular transducer activity	44/180	1337/19629	1.84	1.69E-11

Supplementary Table S13. Copy number of P450 genes in camel and other mammals.

Family	Subfamily	Cattle	Horse	Human	Camel
CYP1		4	3	3	3
	CYP1A	3	2	2	2
	CYP1B	1	1	1	1
	CYP1C	0	0	0	0
CYP2		23	31	20	27
	CYP2A	1	3	4	1
	CYP2B	1	3	2	1
	CYP2C	8	9	4	4
	CYP2D	2	7	2	2
	CYP2E	1	1	1	2
	CYP2F	1	1	1	2
	CYP2G	1	1	1	0
	CYP2J	4	1	1	11
	CYP2K	0	0	0	0
	CYP2R	1	1	1	1
	CYP2S	1	1	1	1
	CYP2T	0	1	0	0
	CYP2U	1	1	1	1
	CYP2W	1	1	1	1
CYP3		3	7	5	6
	CYP3A	3	7	5	6
CYP4		13	15	12	7
	CYP4A	3	3	2	1
	CYP4B	1	3	1	2
	CYP4F	7	7	6	2
	CYP4V	1	1	1	1
	CYP4X	1	1	1	1
	CYP4Z	0	0	1	0
CYP5		1	1	1	0
	CYP5A	1	1	0	0
CYP7		1	2	1	2
	CYP7A	0	1	0	1
	CYP7B	1	1	1	1
CYP8		2	2	2	2
	CYP8A	0	1	0	0
	CYP8B	0	1	0	2
CYP11		3	3	3	2
	CYP11A	1	1	1	1
	CYP11B	2	2	2	1
CYP17		1	1	1	1

	CYP17A	1	1	1	1
CYP19		1	1	1	1
	CYP19A	1	1	1	1
CYP20		0	1	0	1
	CYP20A	0	1	0	1
CYP21		2	1	6	1
	CYP21A	2	1	6	1
CYP24		0	1	1	1
	CYP24A	0	1	1	1
CYP26		3	3	3	3
	CYP26A	1	1	1	1
	CYP26B	1	1	1	1
	CYP26C	1	1	1	1
CYP27		3	3	3	3
	CYP27A	1	1	1	1
	CYP27B	1	1	1	1
	CYP27C	1	1	1	1
CYP39		0	1	0	1
	CYP39A	0	1	0	1
CYP46		1	1	1	1
	CYP46A	1	1	1	1
CYP51		1	1	1	1
	CYP51A	1	1	1	1

Supplementary Table S14. Scaffolds involved in generation of antibodies.

Serial number of scaffold	Type and number of antibody gene
102281	C(1)
118423	C(1)
189423	C(1)
30932	C(1)
35456	C(1)
92866	C(1)
13147	V(1)
1470	V(3)
15220	V(4)
1527	V(3)
2392	V(1)
37069	V(1)
44205	V(1)
49067	V(1)
355.1	V(2), D(7), J(6), C(2)
355.2	C(2)

Supplementary Table S15. VH/CH gene annotation in detail.

Serial number of scaffold	Start	End	Gene type
2392	6430	5837	VHH3
355.1	17759	18350	VH3
355.1	13993	14617	VH3
37069	171	777	VHH3
1527	36077	36663	VH3
1527	3534	4120	VH3
1527	9119	9705	VH3
1470	16651	17234	VHH3
1470	18248	18729	VH3
1470	7777	7944	VH4
13147	1061	705	VH4
15220	2111	2404	VH4
15220	7693	7955	VH4
15220	17036	17291	VH4
15220	18404	18692	VH4
44205	1351	542	VHH
49067	8181	8466	VHH
102281	1978	292	Cγ
118423	496	68	Cγ
18984	465	1	Cγ
30932	597	4	Cγ
35456	2430	1625	Cγ
92866	24	523	Cγ
355.1	70147	75177	Cμ
355.1	79388	82732	Cδ
355.2	138826	142993	Cε
355.2	149874	153819	Cα