Table S5. Assessing phylogenetic relationships among caballines and non-caballine equids with Kishino-Hasegawa and Shimodeira-Hasegawa tests

	Topology 1		Topology 2	L_1-L_2	KH p-value	SH p-value
Within non-Caballine horses: Topologies B,C,D and F, showing Sussemiones, Grevyi's zebras, mountains zebras or donkeys as independent monophyletic groups, are significantly more likely than topology A						
A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	B	(Out2,((STL4,HIP10,ONO5,CAB16),(Others37,SUS3)))	-34,3401	0,005*	0,005*
A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	C	(Out2,((STL4,HIP10,ONO5,CAB16),(Others36,GRE4)))	-82,8397	<0,0001*	<0,0001*
A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	D	(Out2,((STL4,HIP10,ONO5,CAB16),(Others36,ZEB4)))	-123,1452	<0,0001*	<0,0001*
A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	F	(Out2,((STL4,HIP10,ONO5,CAB16),(Others37,ASI3)))	-118,1673	0,02*	0,02*
Between hemiones and *E. hydruntinus*: Topology I is equivalent to topology H, except that samples TZ9 and CH561 (showing missing data) were removed from the (*E. hemionus*, *E. hydruntinus* and *E. kiang*) clade. Topologies A and I are significantly less likely than topology H, in agreement with taxonomic assignment of samples TZ9 and CH561 in *E. hydruntinus*						
H	(Out2,((STL4,HIP10,ONO5,CAB16),(Others30,(HYD_HEM_KIA10))))	A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	-94,1898	0,01*	0,009*
H	(Out2,((STL4,HIP10,ONO5,CAB16),(Others30,(HYD_HEM_KIA10))))	I	(Out2,((STL4,HIP10,ONO5,CAB16),(Others_32,(HYD_HEM_KIA8))))	-25,7459	0,009*	0,16
Among zebras: Topology E7 is significantly more likely than other topologies (except topology E6 according to the SH- test), suggesting that donkeys, mountain zebras, Grevyi's zebras and Plains zebras should be considered as independent monophyletic groups						
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40)))	-387,7301	0,001*	0,001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E	(Out2,((STL4,HIP10,ONO5,CAB16),(Others16,ZEB_GRE_QUA_CAP24)))	-269,5628	<0,0001*	<0,0001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E1	(Out2,((STL4,HIP10,ONO5,CAB16),(Others20,GRE_QUA_CAP20)))	-242,7296	<0,0001*	<0,0001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E2	(Out2,((STL4,HIP10,ONO5,CAB16),(Others20,ZEB_QUA_CAP20)))	-215,0070	<0,0001*	<0,0001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E3	(Out2,((STL4,HIP10,ONO5,CAB16),(Others24,QUA_CAP16)))	-188,4350	<0,0001*	<0,0001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E4	(Out2,((STL4,HIP10,ONO5,CAB16),(Others16,GRE4,ZEB_QUA_CAP20)))	-161,7811	<0,0001*	<0,0001*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E5	(Out2,((STL4,HIP10,ONO5,CAB16),(Others17,ASI_ZEB7,QUA_CAP16)))	-163,4433	<0,0001*	0,003*
E7	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,ZEB4,GRE4,ASI3,QUA_CAP16)))	E6	(Out2,((STL4,HIP10,ONO5,CAB16),(Others13,GRE4,ASI_ZEB7,QUA_CAP16)))	-75,9653	0,008*	0,207
Within non-Caballine horses: Except topology A (ie. null hypothesis showing no topological information within non-caballine Old World equids), all topologies receive rather similar support, suggesting that the exact branching order within non-caballine Old World equids cannot be deciphered thanks to the current HVR-1 dataset						
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	A	(Out2,((STL4,HIP10,ONO5,CAB16),Others40))	-421,3521	0,001*	0,001*
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	J1	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3,ZEB4),(GRE4,QUA_CAP16))))	0,0000	0,486	0,928
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	J2	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3,GRE4),(ZEB4,QUA_CAP16))))	0,0000	0,481	0,936
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	J3	(Out2,((STL4,HIP10,ONO5,CAB16),(HYD_HEM_KIA10,ASI3,SUS3,ZEB4,GRE4,QUA_CAP16)))	0,0000	0,431	0,968
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	J4	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3),SUS3,ZEB4,(GRE4,QUA_CAP16))))	0,0000	0,480	0,902
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	K	(Out2,((STL4,HIP10,ONO5,CAB16),(SUS3,(HYD_HEM_KIA10,ASI3,(ZEB4,GRE4,QUA_CAP16)))))	-0,8367	0,213	0,637
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	K1	(Out2,((STL4,HIP10,ONO5,CAB16),(SUS3,(HYD_HEM_KIA10,ASI3,ZEB4,(GRE4,QUA_CAP16)))))	-0,8367	0,213	0,637
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	K2	(Out2,((STL4,HIP10,ONO5,CAB16),(SUS3,(HYD_HEM_KIA10,ASI3,(ZEB4,GRE4,QUA_CAP16)))))	-0,8367	0,213	0,637
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	K3	(Out2,((STL4,HIP10,ONO5,CAB16),(SUS3,(HYD_HEM_KIA10,ASI3,ZEB4,GRE4,QUA_CAP16))))	-0,8367	0,213	0,637
J	(Out2,((STL4,HIP10,ONO5,CAB16),((HYD_HEM_KIA10,ASI3,SUS3),(ZEB4,GRE4,QUA_CAP16))))	K4	(Out2,((STL4,HIP10,ONO5,CAB16),(SUS3,((HYD_HEM_KIA10,ASI3),GRE4,(ZEB4,QUA_CAP16)))))	-0,8367	0,213	0,637

Among caballines and New-World equids:
All topologies receive rather similar support, suggesting that both (i) the exact branching order within caballine horses and New World equids and (ii) the exact position of the root cannot be deciphered thanks to the current HVR-1 dataset

L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**L1**	(Out2,(STL4,(Others40,(CAB16,(HIP10,ONO5)))))	-1,1997	0,252	0,259
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**M**	(Out2,(Others40,(CAB16,(STL4,(HIP10,ONO5)))))	-0,0279	0,359	0,846
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**N**	(Out2,((HIP10,ONO5),(STL4,(CAB16,Others40))))	-0,0093	0,432	0,815
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**N1**	(Out2,((HIP10,ONO5),(CAB16,(STL4,Others40))))	-1,2209	0,252	0,253
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**N2**	(Out2,((HIP10,ONO5),(Others40,(STL4,CAB16))))	-1,2209	0,252	0,253
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**O**	(Out2,((STL4,(HIP10,ONO5)),(Others40,CAB16)))	-0,0247	0,341	0,833
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**P**	(Out2,(CAB16,(Others40,(STL4,(HIP10,ONO5)))))	-0,0506	0,304	0,752
L	(Out2,(STL4,((HIP10,ONO5),(CAB16,Others40))))	**Q**	(Out2,((HIP10,ONO5),STL4,(CAB16,Others40)))	-0,0093	0,427	0,819

Among caballines and New-World equids: (merged dataset: HVR-1 – excluding positions 15518-15577 because of alignment uncertainty -, and *cyt b* sequences)
Topology R showing hippidions nested within paraphyletic *Equus* appears most likely than the alternative topology, though the difference of likelihood is only marginally significant (SH-test p-value = 0.065)

R	(*C.simum,R.unicornis*,((*E.b.boehmi*1,*E.asinus*,CH1069,*E.capensis*_ACAD226, *E.b.boehmi*2,*E.capensis*_ACAD236,*E.hydruntinus*_CH561,*E.hemionus*_CH28,Sussemione_ACAD2302,Sussemione_ACAD2305,Sussemione_ACAD2303,*E.grevyi*_Oraye),(*E.caballus*,(*H.saldiasi*,*H.principale*))))	**S**	(*C.simum,R.unicornis*,(((*E.b.boehmi*1,*E.asinus*,CH1069,*E.capensis*_ACAD226, *E.b.boehmi*2,*E.capensis*_ACAD236,*E.hydruntinus*_CH561,*E.hemionus*_CH28,Sussemione_ACAD2302, Sussemione_ACAD2305,Sussemione_ACAD2303,*E.grevyi*_Oraye),*E.caballus*),(*H.saldiasi,H.principale*)))	-1,2145	0,105	0,065

Sequence accession numbers for HVR-1 and *cyt b* respectively, except for sequences obtained from complete mitochondrial genomes (asterisk, *). *C. simum*: Y07726*; *R. unicornis*: X97336*; *E.b.boehmi*1: AF220916-AY534349; *E.asinus*: X97337*; *E.b.boehmi*2: AF220918-AY534349; *E. caballus*: X79547*; *H. saldiasi*: EU030679-AY152860; *H. principale*: DQ007562-AY152860

Among caballines and New-World equids: (merged dataset: HVR-1 – including positions 15518-15577 -, HVR-2 and *cyt b* sequences)
Topology T showing hippidions nested within paraphyletic *Equus* appears most likely than the alternative topologies, though the difference of likelihood is only marginally significant (p-values = 0.083)

T	((*C.simum,R.unicornis*),(*E.asinus*,(*E.caballus,H.saldiasi*)))	**U**	((*C.simum,R.unicornis*),(*E.caballus*,(*E.asinus,H.saldiasi*)))	-4,5715	0,083	0,083
T	((*C.simum,R.unicornis*),(*E.asinus*,(*E.caballus,H.saldiasi*)))	**V**	((*C.simum,R.unicornis*),(*H.saldiasi*,(*E.asinus,E.caballus*)))	-4,5715	0,083	0,083

Sequence accession numbers for HVR-1, HVR-2 and cyt b, respectively, except for sequences obtained from complete mitochondrial genomes (asterisk, *). *H.saldiasi* (DQ007562, DQ007615 and AY152859) ; *E.caballus* (X79547*); *E.asinus* (X97337*); *R.unicornis* (X97336*); *C.simum* (Y07726*).

Among caballines and New-World equids: (merged dataset: HVR-1 – excluding positions 15518-15577 because of alignment uncertainty -, HVR-2 and *cyt b* sequences)
Topology T showing hippidions nested within paraphyletic *Equus* appears most likely than the alternative topologies, though the difference of likelihood is only not significant (p-values > 0.072)

T	((*R.unicornis,C.simum*),(*E.asinus*,(*E.caballus,H.saldiasi*)))	**U**	((*R.unicornis,C.simum*),(*E.caballus*,(*E.asinus,H.saldiasi*)))	-4,4405	0,072	0,072
T	((*R.unicornis,C.simum*),(*E.asinus*,(*E.caballus,H.saldiasi*)))	**V**	((*R.unicornis,C.simum*),(*H.saldiasi*,(*E.asinus,E.caballus*)))	-4,4444	0,073	0,073

Sequence accession numbers for HVR-1, HVR-2 and cyt b, respectively, except for sequences obtained from complete mitochondrial genomes (asterisk, *). *H. saldiasi* (DQ007562, DQ007615 and AY152859); Complete mitochondrial genomes: *E. caballus* (X79547*); *E. asinus* (X97337*); *R. unicornis* (X97336*); *C. simum* (Y07726*)

Topologies (presented in suppl fig S3) were tested and the difference in likelihood and p-values (unilateral) of KH- and SH-tests were determined according to the RELL procedure implemented in PAUP* 4.0.b10. For HVR-1 datasets or HVR-1/cyt *b* merged datasets, phylogenetic relationships were assessed excluding positions 15518-15577 because of alignment uncertainty and a K81uf+Γ+I model of molecular evolution (selected according to the AIC criterion of Modeltest). For HVR-1/HVR-2/cyt *b* merged datasets, a K81uf+Γ model of molecular evolution was selected according to the AIC criterion of Modeltest. The principal phylogenetic nodes are reported according to the names presented on suppl table T3. ASI: *E. asinus*; CAB: *E. caballus*; CAP: *E. capensis*; GRE: *E. grevyi*; HEM: *E. hemionus*; HIP: *Hippidion saldiasi/principale*; HYD: *E. hydruntinus*; KIA: *E. kiang*; ONO: *Hippidion devillei* (Peruvian hippidions); QUA: *E. quagga* (all Plains zebras, including *E. quagga quagga*); SUS: Sussemiones (SW Siberia); STL: New World Stilt Legged horses); ZEB: *E. hartmannae*. The number of sequences considered for each group is reported.