

Supplementary Information dataset 1. Sequences aligned for phylogene4c analysis shown in Figure 1 and Figure 5.

Sequences aligned to compute the tree shown in figure 1

MHCII Human DQA1 classII ENSG00000196735; Mouse classII DaA ENSMUSG00000036594; **MHCI** Human HLA A ENSG00000206503; Dog ENSCAFG00000000500; Cow ENSBTAG00000002069; Rabbit ENSOCUG00000011016; Opossum ENSMODG00000003914 ; **MR1** Human ENSG00000153029; Mouse ENSMUSG00000026471; Hyrax ENSPCAG00000001327; Cow ENSBTAG00000009924; Elephant ENSLAFG00000004447; Microbat ENSMLUG00000015463; Opossum ENSMODG00000006527 ; **EPCR** Anolis ENSACAG00000009038; Chicken ENSGALG00000006547; Opossum ENSMODP00000002258; Dog ENSCAFP00000011762; Human ENSG00000101000 ; **HFE** Mouse ENSMUSG00000006611; Human ENSG0000010704 ; **CD1** Chicken CD1_1 ENSGALG00000012494; Chicken CD1_2 ENSGALG00000027865; Anolis CD1 ENSACA G00000011395; Tasmanian devil CD1E ENSSHAG00000007630;; Elephant CD1C ENSLAFG0000001869; Elephant CD1D ENSLAFG00000013716; Elephant CD1E ENSLAFG00000011444; Human CD1A ENSG00000158477; Human CD1B ENSG00000158485; Human CD1C ENSG00000158481; Human CD1D ENSG00000158473; Human CD1E ENSG00000158488; Hyrax CD1B ENSPCAG00000013046; Hyrax CD1C ENSPCAG00000002671; Hyrax CD1D ENSPCAG0000001006; Cat CD1A ENSFCAG00000000610 ; Cat CD1A8 ENSCAFG000000025281; Cat CD1B ENSFCAG00000008648; Cat CD1C ENSFCAG00000001141; Cat CD1D XP_011289534 / gnl|GNOMON|20234746.m; Cat CD1E ENSFCAG00000008651 ; Cow CD1B5 ENSBTAG00000038502 ; Cow CD1B3 ENSBTAG00000022893; Cow CD1D ENSBTAG00000024960 ; Cow CD1D2 ENSBTAG00000047988; Cow CD1D3 ENSBTAG00000039993; Cow CD1D4 ENSBTAG00000047776 ; Cow CD1E ENSBTAG00000009421; Dog CD1A ENSCAFG000000023677 ; Dog CD1A2 ENSCAFG000000023735; Dog CD1A2 ENSCAFG000000025281; Dog CD1B ENSCAFP00000016818; Dog CD1C ENSCAFG00000024030; Dog CD1D ENSCAFG000000030554; Dog CD1E ENSCAFP00000016826; Microbat CD1like ENSMLUG00000010834; Microbat CD1like ENSMLUG000000027037; Microbat CD1like ENSMLUG00000025040; Microbat CD1like ENSMLUG00000015108; Microbat CD1like ENSMLUG00000028075; Microbat CD1like ENSMLUG00000024540; Microbat CD1like ENSMLUG00000016762; Microbat CD1like ENSMLUG00000023829; Microbat CD1like ENSMLUG00000023623; Microbat CD1like ENSMLUG00000028472; Microbat CD1like ENSMLUG00000006796; Microbat CD1like ENSMLUG00000027281; Microbat CD1like ENSMLUG00000025703; Microbat CD1like ENSMLUG00000022794; Microbat CD1like ENSMLUG00000016762; Microbat CD1like ENSMLUG00000013158; Microbat CD1like ENSMLUG00000022211; Microbat CD1like ENSMLUG00000007789; Microbat CD1like19 ENSMLUG00000029770; Mouse CD1D1 ENSMUSG00000028076; Mouse CD1D2 ENSMUSG00000041750; Rabbit CD1A ENSOCUG00000009212; Rabbit CD1A2 ENSOCUG00000005421 ; Rabbit CD1B ENSOCUG00000011270; Rabbit CD1D ENSOCUG00000006549 ; Rabbit CD1E ENSOCUG00000013956

MHX sequences aligned to compute the tree shown in figure 5

As several MHX sequences were not annotated, protein sequences (instead of gene ID) are provided below.

```
>CD1dcow
VGCWLFLLFWGLPQIWRSFETPHLSFPPFRGLQISSFANRSWTRTDGLAWLGELQPYTLRN
ESNTIGFLKLPWIRSGFTSDQQWEQLQHTFLVYRSSTFRDIWEFVKMLPSDYPLEIQVSGGC
ELLPRNISESFLRAAFQGGKDVLSFQGMSWVSAPDAPPLIEEVIKVLNQDQGTKETVHWLL
HDICPELVRLGLQTKGSELEKQVKPEAWLSSGSPSPGPRLLLLVCHVSGFYPKPVVWMMR
GEQEEPTGRQGVMPNADSTWYLRVTLDDVAAGEAAGLSCQVKHSSLGDQDIIILYWDGNRV
SRGLIVALVLLVFLVLLVFGGLVFWFRKHRRYQDIP
>CD1dHUMAN
MGCLLFLLLWALLQAWGSAEVPQRLFPLRCLQISSFANSWTRTDGLAWLGELQTHSWSN
DSTVRSCLKPWSQGTFSDDQWETLQHFVYRSSTFRDVKFAKMLRSLYPLELQVSAGC
EVHPGNASNNFPHVAFQGGKDIILSFQGTSWEPQEAFLWNLAIQVLNQDKWTRRETQVLL
NGTCTPQFVSGLLESGKSELKQVKPKAWLSRGPSPGPRLLLLVCHVSGFYPKPVVWMMR
GEQEQGTQDQDILPNADETWYLRATLDDVVAAGEAAGLSCRKHSLEGGQDIIIVLYWGGSYT
SMGLIALAVLACLFLLLIVGFTSRFRKQTSYQGVV
> ENSG00000206503|HLA-A|6
MAVMAPRTEVLLLSGALALTTWAGSHSMRYFFTSVSRPGRGEPFISVGYVDDTQFVRF
DSDAASQRMEPRAPWIEQEGPEYWDQETRNKVAQSQTDRVDLGTLRGYYNQSEAGSHTIQ
IMYGCDVSGDGRFLRGYRQDAYDGKDYIALNEDLRSWTAADMAAQITKRKWEAAHEAEQL
RAYLDGTCVEWLRRLRYLENGKETLQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLT
WQRDGEDQTDQDELVEVTRPAGDGTQKWAAVVPSGEEQRYTCHVQHEGLPKPLTLRWEL
SSQPTPIVIGIAGLVLLGAVITGAVVAAMWRRKSSDRKGGSYTQAASDQAQSDVSL
TACKV*
> ENSG00000234745|HLA-B|6
MLVMAPRTEVLLLSAALALTTWAGSHSMRYFFTSVSRPGRGEPFISVGYVDDTQFVRF
DSDAASPREPRAPWIEQEGPEYWDRTQIYKAQAQTDRESLRNLRGYYNQSEAGSHTLQ
SMYGCDVGPDRLLRGRHDQYAYDGKDYIALNEDLRSWTAADTAAQITQRKWEAAEAER
RAYLEGEVCEWLRRLRYLENGKDLERADPPKTHVTHHPISDHEATLRCWALGFYPAEITLT
WQRDGEDQTDQDELVEVTRPAGDRTQKWAAVVPSGEEQRYTCHVQHEGLPKPLTLRWEP
SSQSTPIVIGIAGLVAVLVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV
TA*
> ENSG00000204525|HLA-C|6
MRVMAPRTEVLLLSGGLALTTWACSHSMRYFDTAVSRPGRGEPFISVGYVDDTQFVRF
DSDAASPRGEPAPWVEQEGPEYWDRETQYKQRQAQADRVSLRNLRGYYNQSEAGSHTLQ
RMSGCDLGPDRLLRGRYDQSAVDGKDYIALNEDLRSWTAADTAAQITQRKLEAARAAEQ
RAYLEGTVEWLRRLRYLENGKETLQRAEPPKTHVTHHPISDHEATLRCWALGFYPAEITLT
WQRDGEDQTDQDELVEVTRPAGDGTQKWAAVVPSGEEQRYTCHVQHEGLQEPPLTLRWEP
SSQPTPIVIGIAGLVAVLVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV
LITCKA*
>MHCicow
MRVMRPTEVLLLSGVLVLTETLAGSHSLRYFTGVSRLGEPFISVGYVDDTQFVRF
DSDAADPRTEPRVPWMEQEGPEYWDRETRNLKDAQTFRVNLNLLGYYNQSEAGSHTLQ
WMYGCDVEPDGRFLRGYRQDAYDGRDYIALNEDLRSWTAADTAAQITKRKWEAAGEAER
RNYLEGRVCEWLRRLRYLENGKDTLLRADPPKAHVTRHPSSDREVTLRCWALGFYPEEISLT
WQHNGEDQTDQDELVEVTRPSGDGTQKWAALVPSGEEQRYTCHVQHEGLREPLTLRWEP
PQTSFLTMGIIVGLVLLVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV
```

>HosaMR1

MGELMAFLPLIIVLMVKHSDSRTHSLRYFRLGVSDPIHGVPFISVGVVDSHPITTYDS
VTRQKEPRAPWMAENLAPDHWERYTQLLRGWQOMFKVELKRLQRHYNHSGSHTYQRMIGC
ELLEDGSTTGFLOQAYDGDQDFLIFNKDTLWSLAVDNVAHTIKQAWEANQHHELLYQKNWLE
EECIAWLKRFLYEGKDTLQRTPEPLVRVNRKETFPGVTALFCKAHGFYPPETIYMTWMKNG
EEIVQEIDYGDILPSGDGTQAWASIELDPQSSNLYSCHVEHCGVHMVLQVQSEETIPL
VMKAVSGSIVLVIIVLAGVGLVWRRRPREQNGAIYLPDPDR

>MR1cow

MMLLLPLIIVLMMKLSDARTHSLRYFRLGISEPGYGIPEFISAGYVDSHPITMYNSVSQ
KEPRALWMEENLAPDHWERYTQLLRGWQOAFKVELKQLQHHYNHSGFHTYQRMIGCELLE
DGSITGFLQYAYDGDQDFLIFNKDTLWSMAMDNVADIIRRVWEANRHELQYQKNWLEECI
AWLKRFLYEGKDALQRTPEPPKVRVNRKETFPGITTLICRAYGFYPPETIINWMKNGEEIF
QDQDYGGLPSGDGTQVWVSVELDPQNGDIYSCHVEHGGVHMVLQGFQSEETILLVVKVA
VGFIVLAIALAGVGLAWRKRPRGKKNVICLSTPEH

>MHXrabbit

MAELASLKLKLEKPPWPGGTHSLHYHYLALSEPGLPEFLAVGYVDDQPFIRYDSRTGRAE
LRAAWMAPVDYQWYETETQKQKAWKQVQVEMWTVMGYHNQSSGTHTAQRMFGCEVQEDG
SSSSFWQFGFDGQDHLSDLVEFSLWSVAEPVAVRTRKRWETERCYAEYDKAYLEGLCLTS
LHRYLELGSQSLTRKEPPMVHITKMTQDGTTLRCWALGFYPPQDISLNWLDGEELETET
EHVETRPSGDGTQVWVAVQVPAKAEWYSCHVRHPGLNHTLTVAVESPSNGLIAMAIS
TVLITIMLAVTVIWSKRCLQAWNKSQVQTPAPGAGFQKEEQYKQLC

>MHXpika

GTHSLHYHYLALSEPGLPEFLAVGYVDDQPFIRYDSRTGKAEPQAAWMDPVAAQYWET
ETQKQKAWKQVQVEMWTVMGYHNQSSGACTGIHSAQRMFGCEINEDGSSRSFWQFGFDG
HDHSLDMETLWSVAEPVAIRTKSWWEMEHYAEYDKAYLEGLCLTSLHRYLELGGQSL
TREPPMVQITKHTQDRATLRCWALGFYPRDILLSWLDGKELTLETEHVETRPSGDGTQ
QVWVAIQVPAGNEAQYSCHVQHSGLNYTLTMAW

>MHXsquirrel

ERIQAPPHLFLVLSFYSQFLILLTEYFLGTHSLHYHYLTLSEPGPLPKFLAVGYVDDQ
FIRYDSRCRKAEPQAPWMLVDSQYWETETQKQRAWKQVQVEMWTVMGYHNQSGAYTGI
HSAQRMFGCEIQEDGRSSSFQFGFDGQDHLSDLETLWSVAEPVAVRTRKRWEMERCY
AEYDKAYLESCLISLHRYLELGSQSLTRIEPPMVQVTKHTQDGTMLRCWALGFYPPDI
SLSWLGRKELVLETEHVETRPSGDGTQVWVAVQVPAKAEWYTCHVQHPGLKQRLTVA
WDSSSIGNGNGGILLGGIMVLVFGVAVIVILKKKYLQGGEDNLNSQSPERTEAAQTL C

>MHXcow

MAELANLKLKLEKSWPRGTHSLRYHYLDLSEDPPLPKFQAVGYVDDQPFIRYDSRVDAE
PQAPWIMPNAKYWEKETKKQRKWAELHQVETWRMMGYHNHSSVLTGMHSTHRMFGCEIQ
EDGHSNSFWQFGFDGQDHLSDMETLWSVANPVALLTRRMETEHCYAEYDKAYLEGPC
LASLRRYLELGGQRFTRREPPTVWVKHSAQDGGTTLRCWALGFYPPQDISLWLDGKEL
NSKLEYVEARPSGDGTQVWMAVVPAGDETQYCHVQHSGLNHTLTVWEMPSPHGLTVA
MVISLILILLVFGGVSLLTKCLQVLGGGKGG

>MHXferret

LFNRPQFLLLLMANTLVRLSYVFTGTHSLHYHYLALSEPGLDLPQFLAVGYVDDQPFH
YDSRVDRAPQAPWMAVDAQYWETETQKQRAWAKVQVETWTVMGYHNQSTAFSPYAG
MHSTQRMFGCEIREDGHTSFWQFGFDGQDHLSDLETLNWVSAKPAAMRTKSWWETERC
YAEYDKAYLEGLCLVSLRRYLELGGQSLTRKEPPVQVTRHTAQDGGITLKCWARGFYPR
DISLWLDGEEELVLETEYVETRPSGDGTQVWAAVQVPAKAEWYTCHVQHSGLNHTFT
VTWEAPPHQWITSAVVIMPLIFLLTAGVLIFIKQYSQGRWEKVKVPSQAKPCREKGS
PILPAERF

>MHXcat

CRSQICIGALLFHIFTGTHSLHYHYLALSEPGLDLPQFLAVGYVDDQPFHYDSRVDRAP
PQAPWMAAVDQYWETETQKQRAWAKVQVETWTVMGYHNQSSGTHSTQRTFGCEIRE
DGCTSSFWQFGFDGQDHLSDLETLWSVAKPAVAVQTKRWETERCYAEYDKAYLEGLCL
TSLHRYLELGSQSLTRKEPPKQVTRHTAEDGSI TLRCWARGFYPRDISLWLDGEEELV
QETEVETRPSGDGTQVWAAVQVPAKAEWYTCHVQHSGLNHTLTVAVAPSHQWITSAK
VISIIFTITFLVAGVLTFIKQYSQGRSPPYFLPAGGEEPHNSQNPMSVE

>MHXPanda

MMLLLQLLGLVPERAAESPWGLPLVADGPASLAKVMAELANLKLKGNKPWPGGTHSLHYH
YLAVSEPGPLPQFLAVGYVDDQPFHYDSCVDRAPQAPWMAVDAQYWETETQKQRTW
AKVQVETWTVMGYHNRSRGMHSTQRMFGCEIREDGRTSSFWQFGFDGQDHLSDLETL
WVSAKPAVAVRTRKRWETERCYAEYDKAYLEGLCLISLHRYLELGGWLSLTRKEPPVQVTR
HTAQDGGITLKCWARGFYPRDISLWLDGEEELVLETEYVETRPSGDGTQVWAAVQVPA
RKEDWYCHVQHSGLNHTLTVAVAPSRQWKSAVVSTIFTITFLLMAGVRIFIKQYSQ
GGEDPNSQSPATVEATQLFC

>MHXmicrobat

GTHSLRYHYLALSEPGLPQFLAVGYVNDQPFIRYDSRVGRAEPQAPWIASVDARYWET
ETQKQKAWAEVQVEMWTVMGYYSQGGASGGTHSAQRMFGCEIQEDGRSSSFQFGFDGR
DHLSDLETLWSVAEPMAVAVRTRKRWESERCYAEYDKAYLEGLCLASLRRYLELGGPSLT
RRKQAQEAQVQVTRHTIQGRGTLRCWARGFYPPQDISLWLDGEEELGLETEHM
ETRPSGDGTQVWAAVWLQAGEEAGYTCRVQHSGLNHTLTVAWG

>MHXelephant

THVSTGTHSLRYHYLALSEPGLPQFLAVGYVDDQPFIRYDSRVGRAQSQAPWMLVDT
QYWERETQQRGWKQVQVEMWTVMGYHNLSGASGKHSQRMFGCEIQEDGRSSNFQF
GFNGQDHLSDLDLTLWSVAHPVALRTRKRWETERCYAEYDKAYLQGLCLTSLRRYLEL
GQSLTRTEPPKQVQVWHIIGQDGRVTLKWCWALGFYPPQDISLWLDGEEELALETEHVET
RPSGDGTQVWAAVQVPAEESGYTCHVQHPGLNHTLMTYESPSHPRLIADLVIFISILVTI
SPVAAVVIWIKKPSG

>MHXarmadillo

TPWGLPLVADNPSMAKEVMAQLASLKLKPKWPGGTHSLHYHYLALSEPGLDLPFLAV
GYVDDQPFIRYDSRLGRAEPQAPWMAVVEARYWETETQKQVWEKQVQVEMWTVMGYHNQ
SSGIHSAQRMFGCEIQEDGHTSSFWQFGFDGQDHLSDLETLWSVAEPVAVRTRKRWEM
ERCYAEYDKAYLEGLCLTSLNRYLELGGQNLTRREPQVQVTRHLAQDGGAMLRCAWSF
YPKDITLWLDGEEELALETEYVETRPSGDGTQVWAAVHVRKGTAEQYCHVQHPGLN
HTLTVAVEPPACCGPITVTVAVLAIILLVAGVLMCAKWLARMKPSYEPAAAGESP

>MHXsloth

GTHSLHYHYLALSEPGLDLPFLAVGYVDDQPFIRYDSRLGRAEPQAPWMAVVEARYWET

ETQKQRAWEKVQVEMWTVMGYHNQSSGMSGLHSTQRMFGCEIQEDGSTSIFWQFGFDGQ
DHLSLDLETLSSWVSAEPAAVLTKRRWEMEPCYAEYDKAYLESCLISLHRYLQLGGHNL
QQEPPKVQVTKHLAQDGCAMLKCWALGFYPKDITLSSWLGGEELETEQVETRPSGDGT
YQTWAVMHVCKGVAETQYTCHVQHPSLNSTLIVA
>MHXsloth2
GTHSLHYHYLTLSEPGSDFFDFLAVGYVDDQPFIQYDSRLGRAKQPALWVAPIEAQYWEM
ETQKQRGWEKVQVEIWTMMRYHNSACPLHSIQRSFGCEIQEDGSTSIFWQFGFDGQDH
SLDLETLSSWVSAEPAAIQTKPWLERERCYAEYDKAYLEGLCLTSLHRYLQLGGQNLTRQ
>MHXplatypus
MLRISQLKVQVSRWLWGSVQFHFLFSWKGSQSLHYFFLTMSSSSLEVPNFVSVGYVNDL
IWDYDRTHLAEARAPWMSAMKAEYWERENQNQRAWEKVQRAEQAIMRNLNHSGGSHSIQ
RMFGCEIHEDNSTSVFHHFGYDGEDHLTLDLDTLWISANPLAQDTKKWSEQEVCAQYN
KAYLQGLCLTSLRYLELGRGRLNRKVPVVRVTRHLAHNGEPLRCWAHGFYPRDIRLS
WWWGQELTQETEHVGSRPGDGTYSWGAVEVPSGEEHRYTCRVEHLGLERPLTVTWE
PIDPSPIFTGAVVVLALSIVAAVTWRRRKRKRPVSDSGSPSHLDCKPHIKLSDLPILC