4.3 Cloning of CD28 from horse and zoo animal species

CD28, a member of the immunoglobulin supergene family, is a heavily glycosylated homodimeric glycoprotein expressed on the surface of almost all $CD4^+$ $CD8^+$ thymocytes, mature $CD4^+$ T cells, and $CD11b^-$ $CD8^+$ T cells in humans (Aruffo and Seed, 1987; Linsley and Ledbetter, 1993). The physiologic ligands for CD28 are B7-1 (CD80) and B7-2 (CD86), which are expressed on APC such as macrophages, dendritic cells, and B cells (Thompson, 1995; Lu et al., 1997).

In horses, a unique T cell surface molecule was suggested to be the orthologue of human CD28. Although this molecule has similar biochemical characteristics to CD28, the number of equine PBL expressing it differed considerably from human lymphocytes and lacking of *in vitro* expressed equine CD28 made it difficult to designate this antigen (Kydd. et al., 1994; Byrne et al., 1997; Lunn et al., 1998). Unfortunately, none of anti-human CD28 mAbs analyzed at this study was defined positive. The importance of CD28 as a co-stimulatory molecule and the inability of any of mAbs directed to human CD28 to stain equine leukocytes increased the interest to analyze this molecule in horses and different zoo animal species.

4.3.1 RT-PCR of CD28

To analyze CD28, members of the order *Perissodactyla* including three members of family *Equidae*: Somali wild ass, Hartmann's mountain zebra, and Grevy's zebra, and a member of family *Rhinocerotidae*, the Greater one-horned rhinoceros were included next to horses here. Asian elephant representing order *Proboscidae*, family *Elephantidae* and members of order *Artiodactyla*, family *Bovidae*, represented by European bison and African Buffalo and one member of family *Giraffidae*, Nubian Giraffe were also analyzed. As described in sections 3.2.5-3.2.8, total cellular RNA was extracted from PHA activated PBMC of these animal species, reverse transcribed to cDNA using a hexamer random primer followed by two consecutive PCRs using different sets of primers (Table 5) to amplify CD28 ORF. Specific primers were delineated from human and other animals GenBank published sequences (Table 4). To isolate CD28 ORF, V1R1 (nested primer) covering the complete ORF was used. V1R1 amplified CD28 as a product of about 663 bp (Fig. 73b). A different approach was used to amplify CD28 ORF from European bison where an additional amplification was necessary. Here, V4R4 product was amplified using V1R4 heterologous primer pair in a semi-nested PCR.

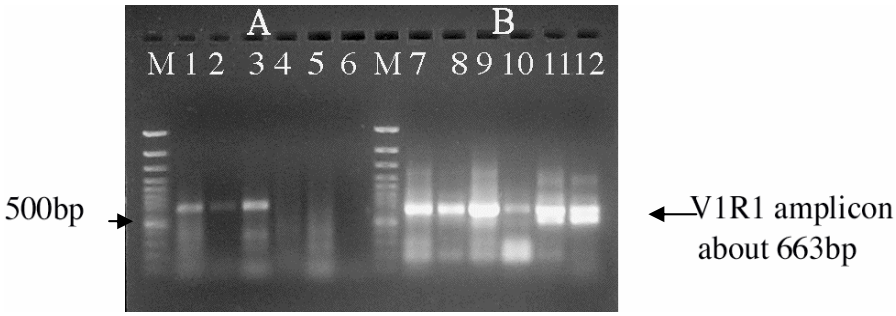


Fig.73 RT-PCR of CD28 ORF cDNA from different animal species.
M= 100 bp page ruler DNA ladder. Lanes 1 to 6 and 7 to 12 are amplicons from domestic horse (1, 7), Somali wild ass (2, 8), Hartmann's mountain zebra (3, 9), Asian elephant (4, 10), Affrican Buffalo (5, 11), Nubian Giraffe (6, 12), respectively. Amplification using the most outer primers V4R4 (Fig. 73A) produced either a smear or a band of a size a bit above 663bp. Nested PCR to isolate CD28 ORF using V1R1 primer (Fig. 73B) revealed an amplicon of 663 bp approximate size. CD28 ORFs represented as bands of about 663 bp were excised from agarose gel and purified before ligation into cloning vector. A second band of slightly smaller size than 663 bp appeared in some instances which could represent a splice variant of CD28.

4.3.2 Sequence analysis of CD28

Gel purified CD28 cDNA was ligated into a PCR cloning vector before transformation of competent bacteria. Sequence was then amplified using primers specific to SP6 and T7 in a sequencing reaction (3.2.10-3.2.14). Multiple sequence alignment of CD28 from a set of 13 sequences including human (*Homo sapiens*), Genbank accession No.: NM006139; Cat (*Felis catus*), Genbank accession No.: AB025316; Dog (*Canis familiaris*), Genbank accession No.: AF259962 and Mouse (*Mus musculus*), Genbank accession No.: NM007642, was performed using MacVector software based on CLUSTAL W (v1.4) analysis.

The CD28 precursor protein, with typical features of an integral membrane protein, was shown to contain ~220 aa residues in the different animal species. The observed slight differences in the number of the coding amino acid residues was due to insertion or deletion of nucleotide triplets in CD28 cDNA at various places (Fig. 74). Removal of a predicted N-terminal signal sequence (~18 aa) should result in mature proteins of ~202 aa residues (Fig. 75) comprising of an extra-cellular domain encoded by ~134 aa residues (Fig. 75, red arrow) with homology to an IgV-like domain, a hydrophobic membrane-spanning domain encoded by ~27 aa residues (Fig. 75, yellow arrow), and a cytoplasmic tail encoded by ~41 aa residues (Fig. 75-blue arrow) (Aruffo and Seed, 1987).

The majority of mismatches detected in the CD28 nucleotide sequences resulted in silent mutations without effect on CD28 sequences (Fig. 75). 18 mismatches between equids and human aa residues were detected in the extra cellular domain (Fig. 75) which could argue for the inability of six human CD28 specific clones (Appendix, tables 16 and 17) to stain horse leukocytes. Besides, the CD28 extra-cellular domain was highly conserved in the tested

members of family *equidae* and there was a unique position for all species where variability was detected (Fig. 75, red colored residue, aa 145). The most divergent sequence among the analyzed zoo animals was from the Asian elephant (Fig. 75, green colored residues).

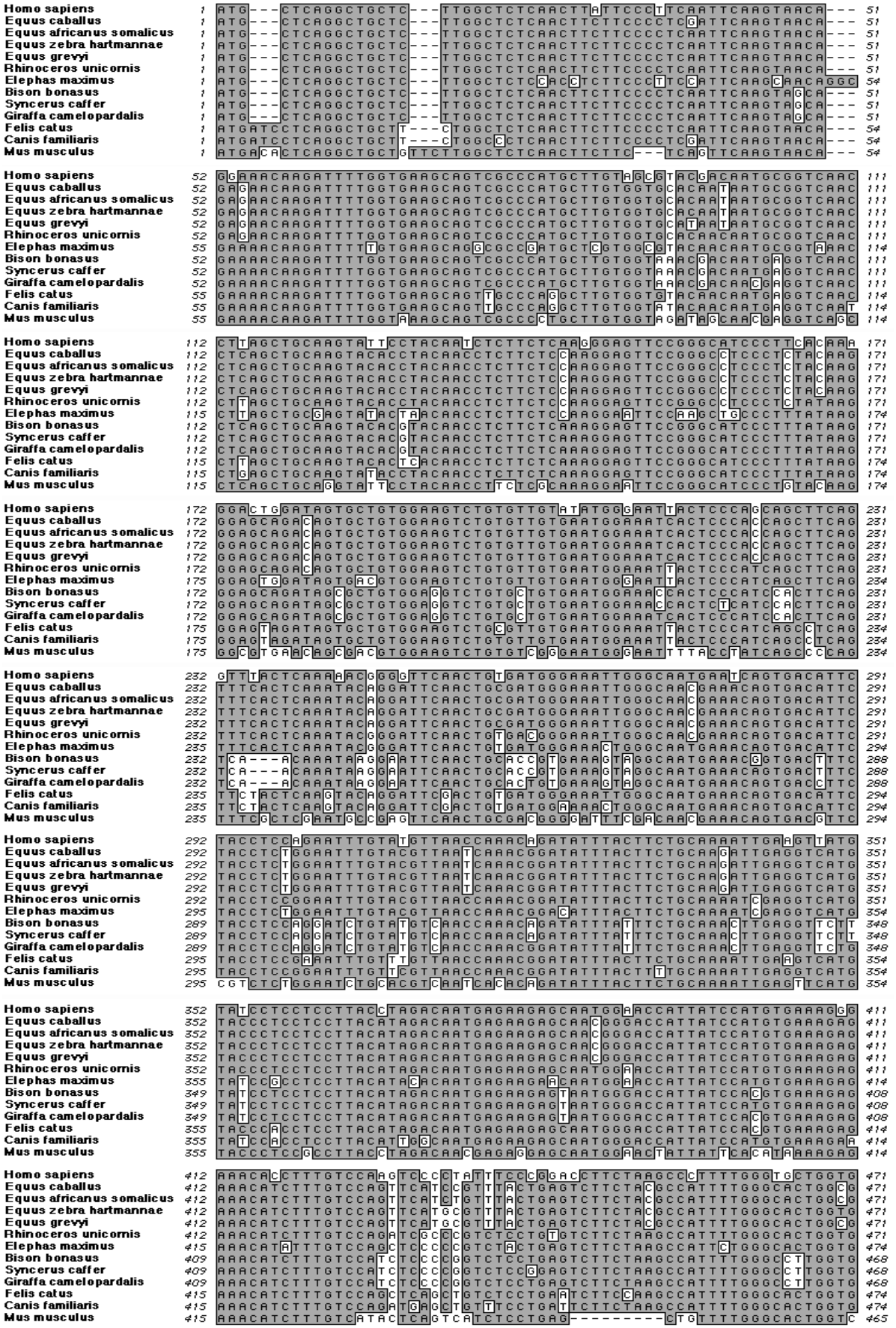
The "MYPPPY" motif in the extra cellular domain, as represented by a yellow box (Fig. 75, aa 117-122), is critical for B7.1 and B7.2 ligands binding (Peach et al., 1994; Wang et al., 2002; Lühder et al., 2003). "MYPPPY" motif was conserved in the analyzed species except for members of order *Artiodactyla* where the first methionine (M) was replaced by leucine (L). The "classical" TCR signaling-dependent co-stimulatory mAbs recognize an epitope close to the binding site for CD28 natural ligands. Binding of the "classical" mAbs is critically dependent on the aa side chain at position 116 adjacent to the "MYPPPY" motif (Peach et al., 1994; Lühder et al., 2003). Conservation of the "MYPPPY" motif and the adjacent aa residue in human and the analyzed wild life animals indicated the ability of an Ab specific to this area to *in vitro* cross-link and activate lymphocytes of these species in the presence of an antigen/TCR complex as described (Lühder et al., 2003; Beyersdorf et al., 2005).

"Superagonistic" anti-CD28 mAbs are capable of inducing the full activation of primary resting T cells in the absence of TCR ligation both *in vitro* and *in vivo* (Lühder et al., 2003). The amino acid residues 78-82 at the C"D motif (Fig. 75, blue box), form the epitope critical for binding "superagonistic" mAbs (Lühder et al., 2003). The C"D motif was not conserved between human and the analyzed animal species (Fig. 75).

The cytoplasmic domain (Fig. 75, blue arrow), was highly conserved in all species. The "YMNM" motif (Fig. 75, aa 191-194), potential for binding signaling proteins (Songyang et al., 1993; June et al., 1994), was 100% conserved in human and the analyzed animal species.

Paired CD28 aa sequence analysis revealed conservation between animal species with long stretches of conserved and semi-conserved amino acid motifs. The percentage of identical residues among different species was 68-99% (Table. 15). Identity between mouse and human was 68%, while identity between horse and human CD28 was 81%. CD28 identity among members of order *Perisodactyla* was significantly higher (93-99%).

CD28 cDNA sequence alignment:



CD28 cDNA sequence alignment (continued):

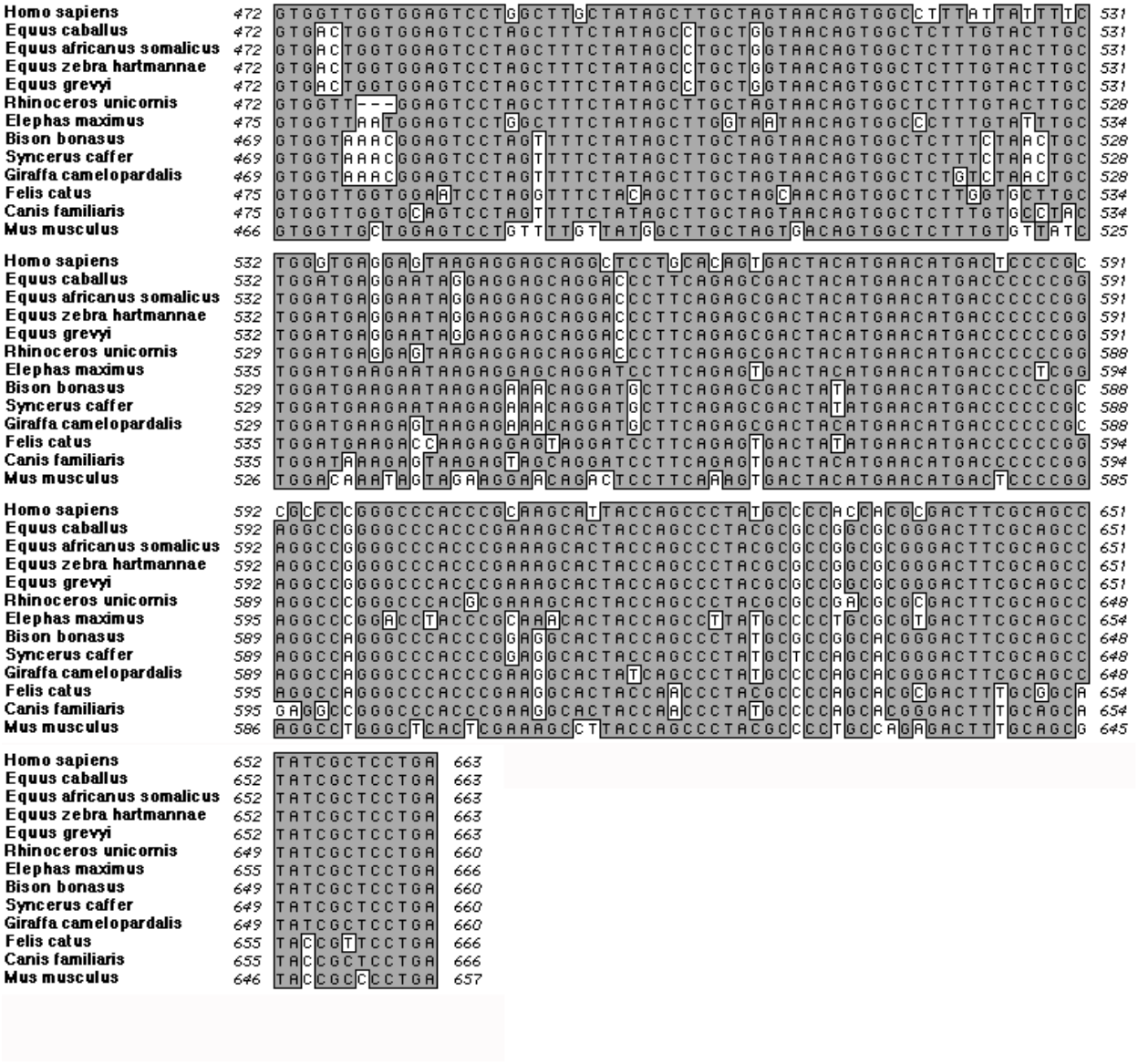


Fig. 74 CD28 cDNA multiple sequence alignment of different animal species and humans.
Analysis of CD28 ORF nucleotide residues in human (*Homo sapiens*), horse (*Equus caballus*), Somali wild ass (*Equus africanus somalicus*), Hartmann's mountain zebra (*Equus zebra hartmannae*), Grevy's zebra (*Equus grevyi*), Greater one-horned rhinoceros (*Rhinoceros unicornis*), Asian elephant (*Elephas maximus*), European bison (*Bison bonasus*), African Buffalo (*Syncerus caffer*), Nubian Giraffe (*Giraffa camelopardalis*), Cat (*Felis catus*), Dog (*Canis familiaris*), and Mouse (*Mus musculus*) respectively. Identical residues are grey coloured, while white coloured residues represent sequence mismatches. "-" represent gaps (deletion or insertion mutation of one or more nucleotides) leading to slight differences in ORFs sizes: Horse, Somali wild ass, Hartmann's zebra, and Grevy's zebra ORFs have a size of 663bp. Greater one-horned rhinoceros, European bison and African Buffalo have ORFs of 660 bp. Elephant has a 666bp CD28 ORF. First 22 and last 23 bases in all sequences represent V1 and R1 primer sequences. cDNA alignment revealed a broad homology between different species and long stretches of conserved motifs. Mouse significantly differs while cluster of changes according to zoological groups were already visible.

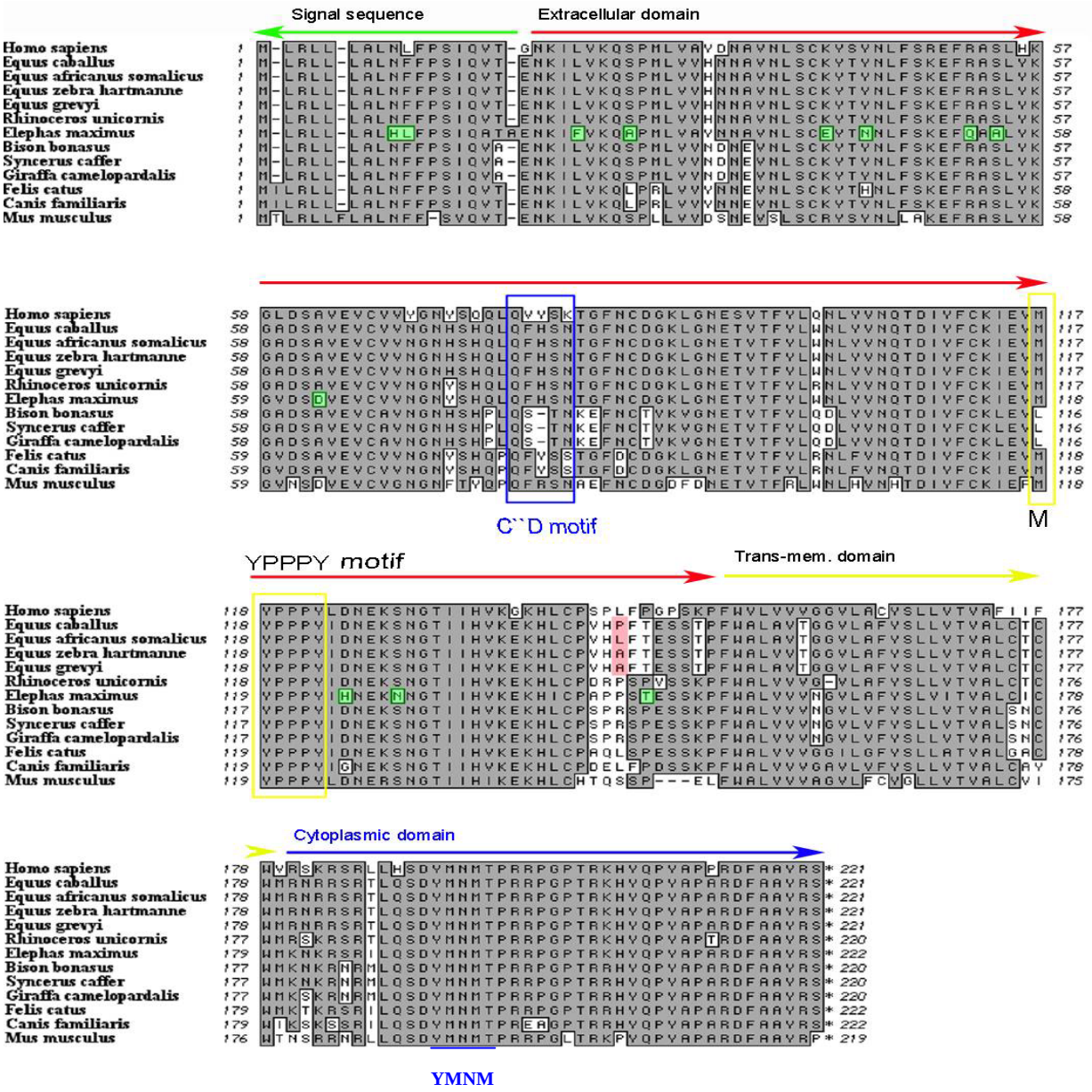


Fig. 75 CD28 amino acid multiple sequence alignment of different animal species and humans.
Amino acids sequence alignment of human, horse, Somali wild ass, Hartmann's mountain zebra, Grevy's zebra, Greater one-horned rhinoceros, Asian elephant, European bison, Affrican Buffalo, Nubian Giraffe, Cat, Dog and Mouse revealed the broad homology between different species. Dark grey represent identical residues while white color represents residues mismatches. "-" represent gaps. "*" represents C terminus of CD28 protein. A N-terminal signal sequence ~18 aa residues was predicted (green arrow). Extra-cellular domain encoded by ~134 aa residues (red arrow). A hydrophobic membrane-spanning domain encoded by ~27 aa residues (yellow arrow), and a cytoplasmic tail encoded by ~41 aa residues (blue arrow) are displayed. The region corresponding to the "MYPPPY" motif (yellow boxes, aa 117-122), which is critical for B7 ligand binding, is highly conserved in the analyzed species except for members of order *Artiodactyla*. The aa residue 116 "V", which is critical for binding of "classical" bioactive anti-CD28 mAbs, is conserved between human and other species. The C"D motif (blue box, aa 78-82) which is critical for binding of "superagonistic" mAbs, is not conserved between human and other species. The "YMNM" motif (aa 191-194), potential for binding signaling proteins, was 100% conserved in humans and the analyzed animal species. Sequence variability between different equids (red color residues, aa 145) is located at a non-conserved amino acid motif (aa 143-147). Mismatches in elephant amino acid residues are represented as green colored residues.

Table 15: CD28 amino acid sequences pair wise analysis percentages.

Human	Buffalo	Bison	Giraffe	Elephant	Horse	Wild ass	Hartmann's zebra	Grevy's zebra	Rhino	Mouse	***Species***
100	**77**	**77**	**77**	**78**	**81**	**82**	**82**	**81**	**84**	**69**	*Homo sapiens*
	100	**100**	**99**	**79**	**79**	**77**	**84**	**84**	**85**	**70**	*Syncerus caffer*
		100	**99**	**79**	**84**	**84**	**84**	**84**	**85**	**70**	*Bison bonasus*
			100	**78**	**83**	**83**	**84**	**83**	**82**	**71**	*Giraffa camelopardalis*
				100	**86**	**85**	**86**	**85**	**86**	**68**	*Elephas maximus*
					100	**99**	**99**	**99**	**93**	**73**	*Equus caballus*
						100	**99**	**99**	**93**	**73**	*Equus africanus somalicus*
							100	**99**	**93**	**74**	*Equus zebra hartmannae*
								100	**93**	**73**	*Equus grevyi*
									100	**74**	*Rhinoceros unicornis*
										100	*Mus musculus*

Legend to table 15:
Green coloured numbers represent CD28 aa identity between different animal species and human (69-84%). Blue coloured numbers represent amino acid residues identity within order *perisodactyla* (93-99%). Red coloured numbers represent CD28 aa identity between Elephant and other species (68-86%).
Brown coloured numbers represent CD28 aa identity within order *Artiodactyla* (99-100%).

The relation of CD28 amino acid sequences from different species was analyzed by drawing a phylogenetic dendogram (Fig. 76). The phylogenetic tree was built using the Neighbour joining (NJ) method, proportionally estimated distances ("P"-correction), and 1000 bootstrap replicates, as implemented in MEGA3® software package. Mouse (*Mus musculus*) was used as the out-group species. Members of family *Equidae* formed a monophyletic cluster originating from a common ancestor, where *Equus zebra hartmannae* was clearly distinct within. Likewise other species were grouped according to their zoological nomenclature. This phylogenetic tree underlined the results obtained from pair-wise analysis of CD28 aa residues sequence (Table 15).

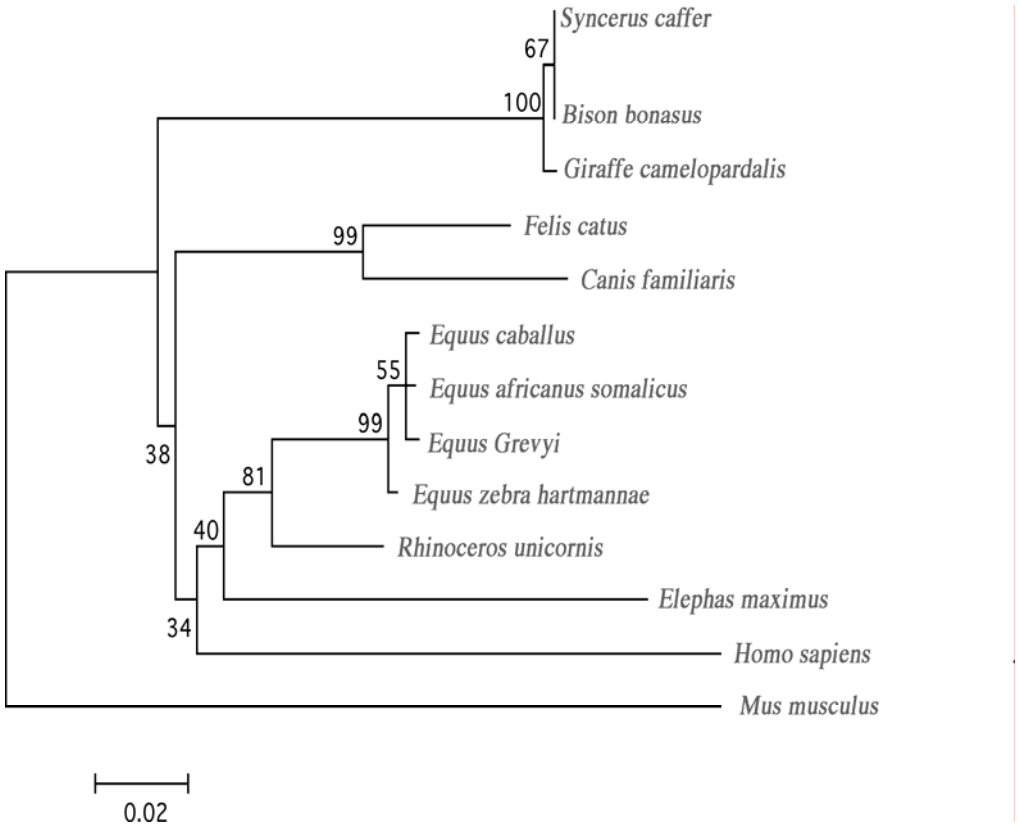


Fig. 76 Phylogenetic analysis of CD28 amino acid residues sequence.
The phylogenetic tree was estimated from CD28 amino acid residues sequences (Fig. 75) by the Neighbour joining (NJ) method, using proportionally estimated distances, and 1000 bootstrap replicates, as implemented in MEGA3 software package. Values of internal branches indicate the percentage of bootstrap replicates in which the branch was found. Horizontal branches were drawn to scale. Bar, 0.02 aa substitutions per site. Mouse (*Mus musculus*) resulted as the out-group species.