The ψ angle of the asparagine at the pY+2 position seems to be largely responsible for this deviation from the type-I β-turn (Table 2). The larger value of the ψ angle moves the pY+3 residue slightly away from Grb2 SH2, making room for the side-chain of Met193^{pcp}. In another words, this bulky side-chain lifted the peptide away from Grb2 SH2. The Grb2 SH2/peptide complex structures reported thus far have relatively small residues at the pY+3 position. Loss of the hydrogen bond may be compensated by the hydrophobic interaction between the methionine of the peptide and Grb2 SH2.

The φ angle of Met192^{pcp} at the pY+1 position in our structure also deviates from that of the other Grb2 SH-bound peptides with type-I β-turns although the difference is smaller than that of the ψ angle discussed above. These 2 angles are complementary for maintaining the hallmark hydrogen bond, and the change in the φ angle of Met192^{pcp} compensates for deviation of the ψ angle of Asn193^{pcp} to some extent, keeping the 2 the main-chain oxygens of pTyr191^{pcp} relatively close to the main-chain nitrogen of Met194^{pcp}. The peptide may transiently adopt a type-I β-turn conformation in solution before binding to Grb2 SH2.

It is tempting to speculate that other peptides containing a residue with a large side chain may also adopt the twisted U-shape conformation, rather than the canonical type-I β-turn. Interestingly, when an epidermal growth factor (EGF)-derived peptide, which has a relatively large glutamine residue at the pY+3 position, is bound to Grb2 SH2 (PDB ID: 1ZFP), it adopts a conformation between the type-I β-turn and the twisted U-shape found in our structure [22]; both its ψ angle value (19.8°) and the O–N distance (3.37 Å) are intermediate between those of the type I β-turn and the twisted U-shape (Table 2).

In this study, the CD28-derived phosphopeptide binds to Grb2 SH2 in the twisted U-shape conformation. This phosphopeptide also binds to the SH2 domains of phosphatidylinositol 3-kinase (PI3K), whose consensus binding motif is pY-X-X-M. The crystal structure of the amino-terminal SH2 domain of PI3K (PI3K N SH2) containing a phosphopeptide derived from c-Kit [23] offers good insight into the interaction between the CD28-derived peptide and PI3K SH2 (Fig. 5A). The sequence of the c-Kit-derived peptide is TNE(pY)MDMKPGV, and this peptide bound to the SH2 domain in an extended conformation. A molecular model of the CD28-derived peptide bound to PI3K N SH2 can be made by simply replacing the sequence of the c-Kit peptide with that of the CD28-derived peptide, SD(pY)MNMT. The model would preserve most of the key protein-peptide interactions from the pY to pY+3 potions, with no obvious unfavorable interactions (Fig 5B). Therefore, one can expect that the CD28-derived peptide changes conformation in a receptor-dependent manner. The β-turn (or twisted U-shape) conformation of the Grb2 SH2-bound peptide positions the pY+2 residue close to the protein, making this residue highly conserved, whereas the extended conformation of the PI3K N SH2-bound peptide exposes the pY+2 residue to solvent, and it has no strong interactions with the protein. Instead, the residues in the pY+1 and pY+3 positions strongly interact with the protein. CD28 exploits the differences between the molecular recognition of pY by the PIK3 and Grb2 SH2 domains to enable binding to both proteins via a single pY site.

Acknowledgments

We thank Dr. Kentaro Tomii for helpful discussions. We also thank the technical staff at the Photon Factory, High Energy Accelerator Research Organization for maintenance of the beamline.

Author Contributions

Conceived and designed the experiments: NI RA MO. Performed the experiments: KH TI JT NI. Analyzed the data: NI HM. Contributed reagents/materials/analysis tools: HM. Wrote the paper: NI KH MO.

References

1. Lowenstein EJ, Daly RJ, Batzer AG, Li W, Margolis B, Lammers R, et al. (1992) The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. Cell 70: 431–442.
2. Rozakis-Adcock M, McGlade J, Mbamalu G, Pelicci G, Daly R, et al. (1992) Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. Nature 360: 689–692.
3. McNemar C, Snow ME, Windsor WT, Prongay A, Mui P, et al. (1997) Thermodynamic and structural analysis of phosphotyrosine polypeptide binding to Grb2-SH2. Biochemistry 36: 10006–10014.
4. Ogura K, Tsuchiya S, Terasawa H, Yuzawa S, Hatanaka H, et al. (1999) Solution structure of the SH2 domain of Grb2 complexed with the Shc-derived phosphotyrosine-containing peptide. J Mol Biol 289: 439–445.
5. Rahuel J, Gay B, Erdmann D, Strauss A, Garcia-Echeverria C, et al. (1996) Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. Nat Struct Biol 3: 586–589.
6. Thornton KH, Mueller WT, McConnell P, Zhu G, Saltiel AR, et al. (1996) Nuclear magnetic resonance solution structure of the growth factor receptor-bound protein 2 Src homology 2 domain. Biochemistry 35: 11852–11864.
7. Nioche P, Liu WQ, Broutin I, Charbonnier F, Latreille MT, et al. (2002) Crystal structures of the SH2 domain of Grb2: highlight on the binding of a new high-affinity inhibitor. J Mol Biol 315: 1167–1177.
8. Mueller DL, Jenkins MK, Schwartz RH (1989) Clonal expansion versus functional clonal inactivation: a costimulatory signalling pathway determines the outcome of T cell antigen receptor occupancy. Annu Rev Immunol 7: 445–480.
9. Schneider H, Cai YC, Prasad KV, Shoelson SE, Rudd CE (1995) T cell antigen CD28 binds to the GRB-2/SOS complex, regulators of p21ras. Eur J Immunol 25: 1044–1050.
10. Rudd CE, Schneider H (2003) Unifying concepts in CD28, ICOS and CTLA4 co-receptor signalling. Nat Rev Immunol 3: 544–556.
11. Rudolph R, Lilie H (1996) In vitro folding of inclusion body proteins. FASEB J 10: 49–56.
12. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67: 235–242.
13. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658–674.
14. Benfield AP, Teresk MG, Plake HR, DeLorbe JE, Millspaugh LE, et al. (2006) Ligand preorganization may be accompanied by entropic penalties in protein-ligand interactions. Angew Chem Int Ed Engl 118: 6984–6989.
15. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67: 355–367.
16. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66: 486–501.
17. Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 24: 946–950.
18. DeLorbe JE, Clements JH, Whiddon BB, Martin SF (2010) Thermodynamic and structural effects of macrocyclic constraints in protein-ligand interactions. ACS Med Chem Lett 1: 448–452.
19. Ettmayer P, France D, Gounarides J, Jarosinski M, Martin MS, et al. (1999) Structural and conformational requirements for high-affinity binding to the SH2 domain of Grb2. J Med Chem 42: 971–980.
20. Das S, Raychaudhuri M, Sen U, Mukhopadhyay D (2011) Functional implications of the conformational switch in AICD peptide upon binding to Grb2-SH2 domain. J Mol Biol 414: 217–230.
21. Suga M, Inubushi C, Okabe N (1998) *O*-Phospo-L-tyrosine. Acta Cryst C54: 83–85.
22. Rahuel J, Garcia-Echeverria C, Furet P, Strauss A, Caravatti G, et al. (1998) Structural basis for the high affinity of amino-aromatic SH2 phosphopeptide ligands. J Mol Biol 279: 1013–1022.
23. Nolte RT, Eck MJ, Schlessinger J, Shoelson SE, Harrison SC (1996) Crystal structure of the PI3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. Nat Struct Biol 3: 364–374.

Immunogenetics (2014) 66:161–170
DOI 10.1007/s00251-014-0760-y

ORIGINAL PAPER

Divergence and diversity of *ULBP2* genes in rhesus and cynomolgus macaques

Taeko K. Naruse · Hirofumi Akari · Tetsuro Matano · Akinori Kimura

Received: 10 November 2013 / Accepted: 13 January 2014 / Published online: 28 January 2014

Abstract Non-human primates such as rhesus macaque and cynomolgus macaque are important animals for medical research fields and they are classified as Old World monkey, in which genome structure is characterized by gene duplications. In the present study, we investigated polymorphisms in two genes for ULBP2 molecules that are ligands for NKG2D. A total of 15 and 11 *ULBP2.1* alleles and 11 and 10 *ULBP2.2* alleles were identified in rhesus macaques and cynomolgus macaques, respectively. Nucleotide sequences of exons for extra cellular domain were highly polymorphic and more than 70 % were non-synonymous variations in both *ULBP2.1* and *ULBP2.2*. In addition, phylogenetic analyses revealed that the *ULBP2.2* was diverged from a branch of *ULBP2.1* along with *ULBP2*s of higher primates. Moreover, when 3D structural models were constructed for the rhesus ULBP2 molecules, residues at presumed contact sites with NKG2D were polymorphic in *ULBP2.1* and *ULBP2.2* in the rhesus macaque and cynomolgus macaque, respectively. These observations suggest that amino acid replacements at the interaction sites with NKG2D might shape a specific nature of ULBP2 molecules in the Old World monkeys.

Keywords Rhesus macaque · Cynomolgus macaque · *ULBP2/RAET1H* · NKG2D · Polymorphisms

Electronic supplementary material The online version of this article (doi:10.1007/s00251-014-0760-y) contains supplementary material, which is available to authorized users.

T. K. Naruse · A. Kimura (✉)
Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University (TMDU), 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan
e-mail: akitis@mri.tmd.ac.jp

H. Akari
Primate Research Institute, Kyoto University, Inuyama, Japan

T. Matano
AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan

T. Matano
The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Introduction

Natural-killer group 2 member D (NKG2D), a C-type lectin molecule, is an activating receptor expressing on the surface of NK, $\gamma\delta^+$ and $CD8^+$ T cells, which plays an important role in the immune system (Wu et al., 1999; Raulet 2003). In humans, several MHC class I-like molecules are known as ligands for NKG2D, including MHC class I chain-related (MIC) and UL-16 binding protein (ULBP)/retinoic acid early transcript 1 (RAET1) (Bauer et al. 1999; Cosman et al. 2001; Chalupny et al. 2003; Bacon et al. 2004). These ligands are usually stress-inducible, and their recognition by NKG2D leads to the activation of NK cells, resulting in the killing of virus-infected cells and tumor cells (Pende et al. 2002; Eagle et al. 2006, Pappworth et al. 2007; Ward et al. 2007).

The human ULBP/RAET1 molecules are encoded by the *ULBP/RAET1* gene family located on the 6q24.2, which is composed of 10 members including six functional genes, *ULBP1*, *2*, *3*, *4*, *5*, and *6*, corresponding to *RAET1I*, *H*, *N*, *E*, *G*, and *L*, respectively (Radosavljevic et al. 2001; Chalupny et al. 2003; Eagle et al. 2009a, b; Eagle et al. 2009b). In addition, several sequence variations in each *ULBP* have been identified (Romphruk et al. 2009; Antoun et al. 2010). Although it is evident that the cell surface expression of the ligand molecules on target cells is differentially regulated (Eagle et al. 2006), genetic variations or polymorphisms in the coding region might also have a functional impact.

In the medical field, non-human primates including rhesus and cynomolgus macaques are used as animal models in the

immunological studies for infectious diseases, autoimmune diseases, organ transplantation, and development of vaccines. These macaques are members of the Old World monkey and it has been reported that the genetic diversity in the rhesus macaque is quite unique, i.e., more than 60 % of the rhesus macaque-specific expansions are found in the protein coding sequences (Gibbs et al. 2007). To fully evaluate the results of immunological experiments using macaque models, it is essential to characterize the genetic diversity of immune-related molecules, which may shape the basis of individual differences in the immune response against foreign antigens and/or pathogens. It has been reported that the copy numbers of genes in the major histocompatibility complex (MHC) loci in the Old World monkey are higher than those in humans (Kulski et al. 2004; Gibbs et al. 2007; Otting et al. 2007). In addition, the extent of genetic diversity in MHC differed, in part, depending on the geographic area, and we have reported that the diversity of MHC class I genes in the rhesus and cynomolgus macaques is considerably different depending on habitat (Naruse et al. 2010, Saito et al. 2012). In our previous study, we have demonstrated that *ULBP4* is more polymorphic in the Old World monkey than in humans (Naruse et al. 2011). It also was revealed that each member of the *ULBP/RAET1* gene family, except for *ULBP6*, had been duplicated in the rhesus genome (Naruse et al. 2011).

Recent reports have indicated that the expression of *ULBP2* is upregulated in HIV infection (Richard et al. 2013, Matusali et al. 2013). Because the innate immune system may be involved in the response to environmental pathogens, it is important to investigate the polymorphisms in the ligands of NK receptors in the experimental animal models for developing HIV vaccine. Here, we report the *ULBP2* polymorphisms focusing on the divergence and diversity in the Old World monkey.

Materials and methods

Animals

A total of 37 rhesus macaques and 24 cynomolgus macaques, previously analyzed for the polymorphisms in MHC class I genes (Naruse et al. 2010, Saito et al. 2012) were the subjects. They were maintained in the breeding colonies in Japan. The founders of the rhesus macaque colonies were captured in Myanmar and Laos, whereas the founders of cynomolgus macaque colonies were captured in Indonesia, Malaysia, and the Philippines. All care including blood sampling of animals were in accordance with the guidelines for the Care and Use of Laboratory Animals published by the National Institutes of Health (NIH publication 85-23, revised 1985) and the study protocol was subjected to prior approval by the local animal protection authority.

DNA extraction and sequencing analysis

Genomic DNAs of B lymphoblastoid cell lines from rhesus macaques and whole blood samples of cynomolgus macaques were prepared, as previously reported (Naruse et al. 2010, Saito et al. 2012). Amplification of *ULBP2* from macaques was done by polymerase chain reaction (PCR) with specific primer pairs designed for the region spanning from intron 1 to intron 3 of rhesus *ULBP2*, LOC694466 (designated as *ULBP2.1*) and LOC694600 (designated as *ULBP2.2*), using FastStart Taq DNA polymerase (Roche, Mannheim, Germany). Primer sequences are as follows: UL2.1NF (5′-AGGGGCTAACTAGGGGTCTTTC) and UL2.1NR (5′ ACCGTTTCTGATCTCATTCCA) for *ULBP2.1*, and UL2.2NF (5′-GAGGGCTAACTAGGGGTCTCT) and UL2.2NR (5′-ACCATTTCTGATCTCATTCCAGA) for *ULBP2.2*. The PCR program was composed of following steps: denaturation at 95 °C for 4 min; 30 cycles of 95 °C for 30 s, 56 °C for 30 s, 72 °C for 45 s; and additional extension at 72 °C for 7 min. The PCR products, about 1,400 bp for *ULBP2.1* and about 1,080 bp for *ULBP2.2*, were cloned into pSTBlue-1 AccepTer vector (Novagen, WI, USA) according to the manufacturer's instructions and transformed into Nova Blue Single™ competent cells (Merck Biosciences Japan, Tokyo, Japan). Ten to 20 independent transformed colonies were picked up for each sample and subjected to sequencing on both strands by using a BigDye Terminator cycling system and an ABI 3730 automated sequence analyzer (Applied Biosystems, CA, USA).

Data analysis

Nucleotide sequences from cloned DNAs were aligned using the Genetyx software package (version 8.0, Genetyx Corp., Japan). When at least three clones from independent PCR or from different subjects showed identical sequences, the sequences were submitted to the DNA Data Bank of Japan (DDBJ). A neighbor-joining tree was constructed by Kimura's two-parameter method for a phylogenetic analysis of *ULBP2* sequences from exon 2 to exon 3, excluding intron 2 sequences, by using the Genetyx software. Bootstrap values were based on 5,000 replications. The *ULBP2* and *ULBP6* sequences from human (GenBank accession numbers AL583835 and AL355497, respectively), and *ULBP2* sequences from chimpanzee (NC006473), western gorilla (NC018430), rhesus (NC007861), and another member of Old World monkey, olive baboon (NC018155) were included in the phylogenetic analysis. The *ULBP1* (LOC694341), *ULBP3* (LOC694525), *ULBP4* (LOC695031), and *ULBP5* (LOC694265) sequences from rhesus macaque, and *ULBP1* (NM025218), *ULBP3* (AL355497), *ULBP4* (AL355312), and *ULBP5* (AL583835) sequences from human were also included in the analysis.

Table 1 Alleles of *ULBP2.1* and *ULBP2.2* in rhesus and cynomolgus macaques

Gene	Species	Allele name	Accession no	ID of reference animal	Clone name
ULBP2.1	*Macaca mulatta*	*Mamu-ULBP2.1*1*	NC007861[a]	Not found in the subjects of this study	
		*Mamu-ULBP2.1*2*	AB826205	R491	UL2.1NR491F-9
		*Mamu-ULBP2.1*3*	AB826206	R312, R314, R496	UL2.1NR314F-2
		*Mamu-ULBP2.1*4*	AB826207	R277, R316, R350, R396, R429, R434, R437, R455, R465, R473, R492, R495	UL2-1R227-13 F
		*Mamu-ULBP2.1*5*	AB826208	R325, R333, R337, R384, R434, R491	UL2-1R434-2 F
		*Mamu-ULBP2.1*6*	AB826209	R350	UL2.1NR350F-13
		*Mamu-ULBP2.1*7*	AB826210	R227, R234, R283, R314, R320, R321, R328, R337, R346, R384,R396, R446, R455, R465, R490, R496	UL2-1R227-7 F
		*Mamu-ULBP2.1*8*	AB826211	R495	UL2.1NR495F-8
		*Mamu-ULBP2.1*9*	AB826212	R321, R333, R360	UL2.1NR321F-8
		*Mamu-ULBP2.1*10*	AB826213	R316, R342, R408	UL2-1R408-12 F
		*Mamu-ULBP2.1*11*	AB826214	R346	UL2.1NR346F-20
		*Mamu-ULBP2.1*12*	AB826215	R342	UL2.1NR342F-14
		*Mamu-ULBP2.1*13*	AB826216	R325, R346, R360, R361, R379, R408, R429, R430, R437, R439, R446, R473, R490	UL2.1NR439F-11
		*Mamu-ULBP2.1*14*	AB826217	R453	UL2-1R453-1 F
		*Mamu-ULBP2.1*15*	AB826204	R234, R312, R361	UL2.1NR234F-7
	Macaca fascicularis	*Mafa-ULBP2.1*1*	NC007861[a]	M04, C09	2.1-2UL2-1 M04-5 F
		*Mafa-ULBP2.1*2*	AB826219	M05, C10, C11	UL2.1NFM05-12
		*Mafa-ULBP2.1*3*	AB826220	M03, C07	UL2.1NFM03-8
		*Mafa-ULBP2.1*4*	AB826221	P01, P02, P03, M01, C01, C03, C04, C05, C07, C08	2.1-1UL2-1 M01-10 F
		*Mafa-ULBP2.1*5*	AB826222	P02, C06	UL2-1P02-2 F
		*Mafa-ULBP2.1*6*	AB826223	M02, C05	2.1-6UL2-1 M02-17 F
		*Mafa-ULBP2.1*7*	AB826224	M03, M04, C06, C08, C09	UL2-1 M03-1 F
		*Mafa-ULBP2.1*8*	AB826225	P04, P05, M01, M05, M06, C02, C12, C13	2.1-3UL2-1 M01-12 F
		*Mafa-ULBP2.1*9*	AB826226	P04, M06, C10, C11, C12, C13	2.1-4UL2-1 M06-10 F
		*Mafa-ULBP2.1*10*	AB826228	M02, C04	UL2-1 M02-20 F
		*Mafa-ULBP2.1*11*	AB826218	P01, C02	UL2NP01-F-2
ULBP2.2	*Macaca mulatta*	*Mamu-ULBP2.2*1*	NC007861[b]	R283, R316, R320, R321, R325, R328, R333, R337, R342, R346, R360, R379, R384, R396, R408, R429, R430, R437, R439, R446, R453, R473, R490, R495	UL2-2R396-3 F
		*Mamu-ULBP2.2*2*	AB827340	R491	UL2.2NR491F-5
		*Mamu-ULBP2.2*3*	AB827341	R314, R321	UL2-2R314-7 F
		*Mamu-ULBP2.2*4*	AB827342	R350	UL2.2NR350F-3
		*Mamu-ULBP2.2*5*	AB827343	R234, R320	UL2-2R361-8 F
		*Mamu-ULBP2.2*6*	AB827344	R325, R333, R337, R384, R491, R492	UL2-2R325-12 F
		*Mamu-ULBP2.2*7*	AB827345	R237, R312, R453	UL2-2R237-5 F
		*Mamu-ULBP2.2*8*	AB827346	R228, R314, R396, R492, R495	UL2-2R383-3 F
		*Mamu-ULBP2.2*9*	AB827347	R496	UL2-2R496-12 F
		*Mamu-ULBP2.2*10*	AB827339	R234, R312, R328, R439,R446, R490, R496	R234UL2.2NF-16
		*Mamu-ULBP2.2*11*	AB827348	R367, R430	UL2-2R367-12 F

Table 1 (continued)

Gene	Species	Allele name	Accession no	ID of reference animal	Clone name
	Macaca fascicularis	*Mafa-ULBP2.2*1*	NC007861[b]	M05, C10, C11	UL2-2FM05-2
		*Mafa-ULBP2.2*2*	AB827350	M03, M04, C06, C08	UL2-2FM03-1
		*Mafa-ULBP2.2*3*	AB827351	P03, C08	P03UL2-2-6 F
		*Mafa-ULBP2.2*4*	AB827352	P01, C03, C04, C05	UL2-2P01-1 F
		*Mafa-ULBP2.2*5*	AB827353	P01,P03, P04, M01, M04, M06, C01, C02, C09, C11, C13	UL2-2P01-7 F
		*Mafa-ULBP2.2*6*	AB827354	P04, P05, M05, M06, C12, C13	UL2-2P04-19 F
		*Mafa-ULBP2.2*7*	AB827355	M02, C05,	UL2-2FM02-2
		*Mafa-ULBP2.2*8*	AB827356	M03, C07	UL2-2FM03-11
		*Mafa-ULBP2.2*9*	AB828102	P02, M01, C01, C03, C06, C07	UL2-2FM01-1
		*Mafa-ULBP2.2*10*	AB827349	M02, C04	UL2-2FM02-11

[a] Identical to LOC694466

[b] Identical to LOC694600

Structure model analysis

Three-dimensional (3D) structure models of ULBP2 molecules were created for amino acid positions from 1 to191, by using a molecular visualization software RasTop2.2 (http://sourceforge.net/projects/rastop/), by referring the human ULPB3 molecule in complex with NKG2D (Radaev et al. 2001) from the Molecular Modeling Database (MMCB No.18231). Polymorphic sites were mapped on the 3D structure models by using the Cn3D 4.1 program (http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml).

Results

Identification of alleles for a ULBP2 gene, *ULBP2.1*

There are two orthologous genes for *ULBP2*, LOC694466 and LOC694600, in the rhesus macaque genome. In the present study, we designated LOC694466 and LOC694600 as *ULBP2.1* and *ULBP 2.2*, respectively, and we designed primer pairs to separately amplify the *ULBP2.1* and *ULBP 2.2*. As expected, PCR products from each gene could be obtained and distinguished by their lengths, although minor length differences due to single nucleotide repeat number polymorphisms in an A stretch were found in the intron 2 sequences.

We obtained nucleotide sequences for the region from exon 2 to exon 3 of *ULBP2.1* from 37 rhesus macaques and 24 cynomolgus macaques by sequencing the cloned PCR products of 1,370–1,395 bp. The *ULBP2.1* sequences from the rhesus macaques were classified into 15 different alleles (Table 1), designated as *Mamu-ULBP2.1*1* to *-ULBP2.1*15*. The LOC4964466 sequences were given with the allele name of *Mamu-ULBP2.1*1*, although it was not found in the analyzed subjects of current study. In the cynomolgus macaques, 11 different alleles, *Mafa-ULBP2.1*01* to *-ULBP2.1*11*, were identified (Table 1). The nucleotide sequences of *Mafa-ULBP2.1*1* were identical to those of *Mamu-ULBP2.1*1* reported for rhesus macaque LOC694466.

Fig. 1 Phylogenetic tree of *ULBP2.1* and *ULBP2.2* alleles and related *ULBP2*. The tree was constructed using neighbor-joining method with bootstrap values of 5,000 replications. The values are indicated as percentages and those values less than 50 % are not shown. The sequences of human *ULBP2* (AY026825), human *ULBP5* (AL583835), human *ULBP6* (AL355497), rhesus *ULBP5* (LOC694265), chimpanzee *ULBP2* (NC006473), western gorilla *ULBP2* (NC018430), and olive baboon *ULBP2* (NC018155) were included in the analysis. The *underlined alleles* indicated with *triangles* and *stars* carried polymorphisms on the α helix structure and contact sites with NKG2D, respectively

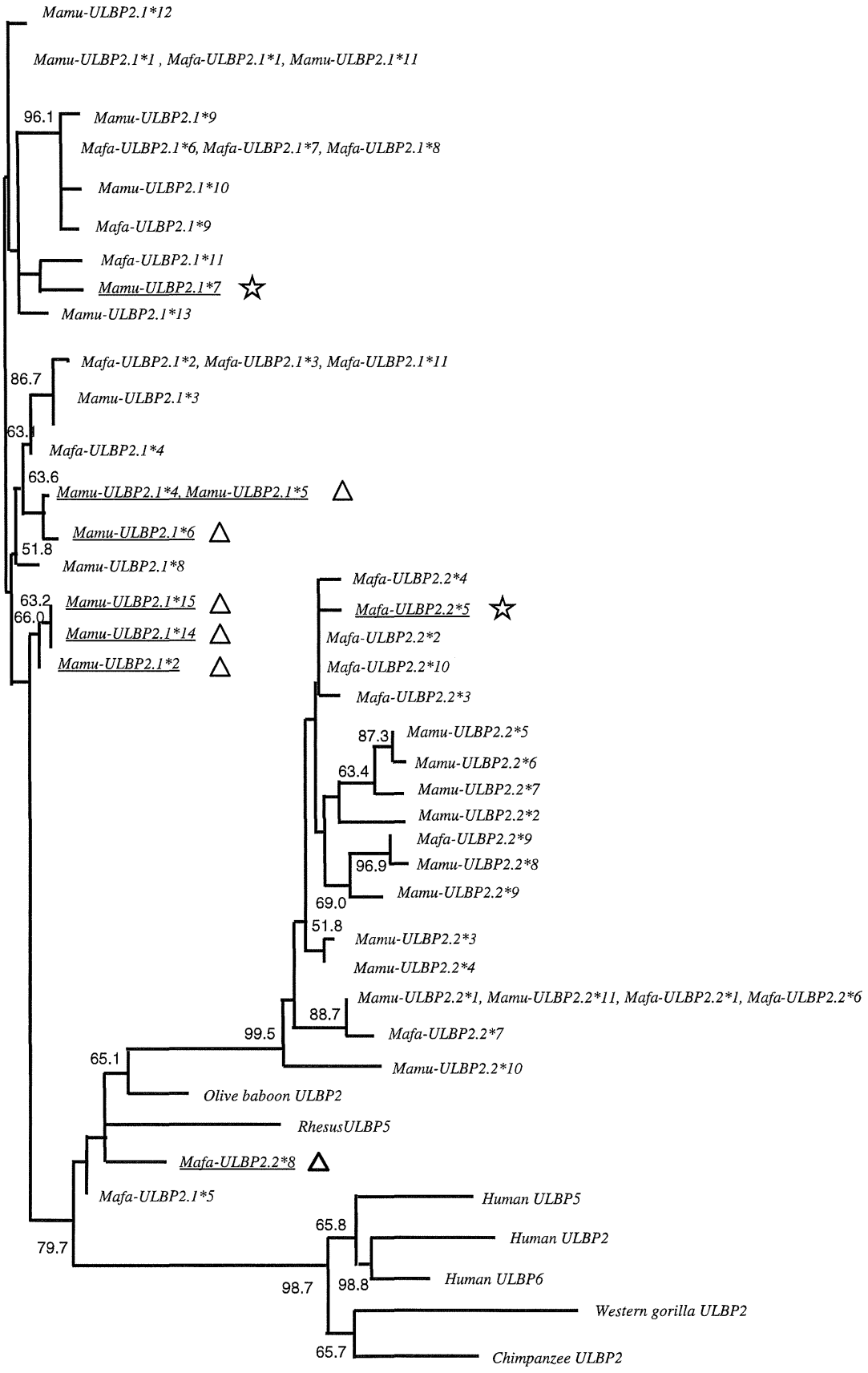
Mamu-ULBP2.1*12
Mamu-ULBP2.1*1 , Mafa-ULBP2.1*1, Mamu-ULBP2.1*11
96.1
Mamu-ULBP2.1*9
Mafa-ULBP2.1*6, Mafa-ULBP2.1*7, Mafa-ULBP2.1*8
Mamu-ULBP2.1*10
Mafa-ULBP2.1*9
Mafa-ULBP2.1*11
Mamu-ULBP2.1*7
Mamu-ULBP2.1*13
Mafa-ULBP2.1*2, Mafa-ULBP2.1*3, Mafa-ULBP2.1*11
86.7
Mamu-ULBP2.1*3
63.1
Mafa-ULBP2.1*4
63.6
Mamu-ULBP2.1*4, Mamu-ULBP2.1*5
Mamu-ULBP2.1*6
51.8
Mamu-ULBP2.1*8
63.2
Mamu-ULBP2.1*15
66.0
Mamu-ULBP2.1*14
Mamu-ULBP2.1*2
Mafa-ULBP2.2*4
Mafa-ULBP2.2*5
Mafa-ULBP2.2*2
Mafa-ULBP2.2*10
Mafa-ULBP2.2*3
87.3
Mamu-ULBP2.2*5
Mamu-ULBP2.2*6
63.4
Mamu-ULBP2.2*7
Mamu-ULBP2.2*2
Mafa-ULBP2.2*9
96.9
Mamu-ULBP2.2*8
Mamu-ULBP2.2*9
69.0
51.8
Mamu-ULBP2.2*3
Mamu-ULBP2.2*4
Mamu-ULBP2.2*1, Mamu-ULBP2.2*11, Mafa-ULBP2.2*1, Mafa-ULBP2.2*6
88.7
Mafa-ULBP2.2*7
99.5
Mamu-ULBP2.2*10
65.1
Olive baboon ULBP2
RhesusULBP5
Mafa-ULBP2.2*8
Mafa-ULBP2.1*5
Human ULBP5
65.8
Human ULBP2
79.7
Human ULBP6
98.7
98.8
Western gorilla ULBP2
65.7
Chimpanzee ULBP2

Table 2 Number of polymorphic sites and non-synonymous sites in ULBP2 genes in macaques and human

Locus	Number of alleles	Exon 2		Intron 2	Exon 3	
		Polymorphic sites	Non-synonymous sites (%)	Polymorphic sites	Polymorphic sites	Non-synonymous sites (%)
Mamu ULBP2.1	15	10	10 (100)	14	7	4 (57.1)
Mafa ULBP2.1	11	8	7 (87.5)	12	6	3 (50.0)
Mamu ULBP2.2	11	7	6 (85.7)	9	14	9 (64.3)
Mafa ULBP2.2	10	17	7 (41.2)	13	12	9 (75.0)
Human ULBP2	NC	17	14 (82.3)	0	15	11 (74.0)
Human ULBP6	NC	14	13 (92.9)	1	17	11 (64.7)

NC not counted because the polymorphisms in the public databases are indicated for each site and not for full sequence

Identification of alleles for another ULBP2 gene, *ULBP2.2*

PCR products for *ULBP2.2* could be obtained from genomic DNAs of both rhesus and cynomolgus macaques. Sequencing data from the cloned PCR products of 1,080–1,085 bp were classified into 11 different alleles, *Mamu-ULBP2.2*1* to *-ULBP2.2*11*, from the rhesus macaques and 10 alleles, *Mafa-ULBP2.2*1* to *-ULBP2.2*10*, from the cynomolgus

a

```
Mamu-ULBP2.1*1     1:PRWSQTGGDD PHSLCYDITI IPKFRPGPQW CAVQGQVDKK TFLHYDCGNK TVTSVSPLGK KLNVTKAWKA QNPVLREVVD MLTEQLLDIQ LENYTPREPL 100
Mamu-ULBP2.1*2      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*3      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------E ----------
Mamu-ULBP2.1*4      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*5      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*6      :********** L--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*7      :********** ---------- --------R- ---------- ---------- ---------- ---------- --T------- ---------- ----------
Mamu-ULBP2.1*8      :********** ---------- ---------- ---------- ---------- ------T--- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*9      :********** L--------- ---------- ---------- ---------- ----I-T--- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*10     :********** L--------- ---------- Y--------- ---------- ------T--- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*11     :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*12     :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mamu-ULBP2.1*13     :********** ---------- ---------- ---------- ---------- ---P------ ---------- ---------- ---------- ----------
Mamu-ULBP2.1*14     :********** ---------- ---------- ---------- ---------- ---------- Q--------- ---------- ---------- ----------
Mamu-ULBP2.1*15     :********** ---------- ---------- ---------- ---------- ---------- Q--------- ---------- ---------- ----------
Mafa-ULBP2.1*1      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*2      :********** ---------- ---------- ----D----- ---------- ---------- ---------- ---------- ---------E ----------
Mafa-ULBP2.1*3      :********** ---------- ---------- ----D----- ---------- ---------- ---------- ---------- ---------E ----------
Mafa-ULBP2.1*4      :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*5      :********** L--------- --------R- ---------- ---------- ---P------ ---------- ---------- ---------- ----------
Mafa-ULBP2.1*6      :********** L--------- ---------- ---------- ---------- ------T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*7      :********** L--------- ---------- ---------- ---------- ------T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*8      :********** L--------- ---------- ---------- ---------- ------T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*9      :********** L-------I- ---------- ---------- ---------- ------T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*10     :********** ---------- --------R- ----D----- ---------- ------T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.1*11     :********** ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------E ----------
Mamu-ULBP2.2*1      :SL--RA-RA- L-----E--- -------RR- ---------- ---------- I--P------ --S------- ---------- ---------- ----------
Mamu-ULBP2.2*2      :********** L-----E--- -------RR- ---------- ---------- I--P----V- -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*3      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*4      :********** L-----E--- -------RR- ---------- ---------- I--------- -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*5      :********** L-----E--- -------RR- ---------- ---------- I--PI----- -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*6      :********** L-----E--- -------RR- ---------- ---------- I--PI---V- -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*7      :********** L-----E--- -------RR- ---------- ---------- I--PI----- -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*8      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*9      :********** L-----E--- -------RR- ---------- ---------- ---P------ -FS------- ---------- ---------- ----------
Mamu-ULBP2.2*10     :********** L-----E--- -------RR- ---------- ------Y--- I--P------ -LS------- ---------- ---------- ----------
Mamu-ULBP2.2*11     :********** L-----E--- -------RR- ---------- ---------- I--P------ -LS------- ---------- ---------- ----------
Mafa-ULBP2.2*1      :********** L-----E--- -------RR- ---------- ---------- I--P------ -LS------- ---------- ---------- ----------
Mafa-ULBP2.2*2      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mafa-ULBP2.2*3      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mafa-ULBP2.2*4      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mafa-ULBP2.2*5      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mafa-ULBP2.2*6      :********** L-----E--- -------RR- ---------- ---------- I--P------ --S------- ---------- ---------- ----------
Mafa-ULBP2.2*7      :********** L-----E--- -------RR- ---------- ---------- I--P------ --S------- ---------- ---------- ----------
Mafa-ULBP2.2*8      :********** ---------- -------PR- ---------- ---------- ---P--T--- ---------- ---------- ---------- ----------
Mafa-ULBP2.2*9      :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Mafa-ULBP2.2*10     :********** L-----E--- -------RR- ---------- ---------- I--P------ -FS------- ---------- ---------- ----------
Human ULBP2         :SG--RA-RA- ------D--V -------PR- --------E- ---------- ---P------ -----T---- ---------- I-----R--- ------K---
Human ULBP6         :FG--RARR-- ------D--V -------PR- --------E- ---------- ---P------ -----M---- ---------- I--------- ------K---
```

Fig. 2 Alignment of deduced amino acid sequences of α1 and α2 domains of ULBP2-related molecules from rhesus and cynomolgus macaques and human. Amino acid sequences deduced from nucleotide sequences for *Mamu-ULBP2.1*1* and alleles of *Mamu-ULBP2.1*, *Mafa-ULBP2.1*, *Mamu-ULBP2.2,* and *Mafa-ULBP2.2*, were aligned with human *ULBP2* (AY026825), and *ULBP6* (AL355497). *Numbers* represent the amino acid positions in mature protein. Sequences for the predicted α helix structure are *underlined* and *contact sites* with NKG2D are *boxed* in the *Mamu-ULBP2.1*1* sequences. *Dashes* indicate identity to the *Mamu-ULBP2.1*1* sequences. *Asterisks* represent not sequenced regions. Amino acid sequences are shown **a** from 1 to 100 and **b** from 101 to 191

b

```
Mamu-ULBP2.1*1   101:TLQARMSCEQ KAEGHSSGSW QFGFDGQVFL LFDSENRMWT TVHPGARKMK EKWENDKDVT MSFHYISMGD CTRWLGDFLM DMDSTLEPSA G 191
Mamu-ULBP2.1*2      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- G--------- -
Mamu-ULBP2.1*3      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*4      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ------E--- ---------- -
Mamu-ULBP2.1*5      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ------E--- ---------- -
Mamu-ULBP2.1*6      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ------E--- ---------- -
Mamu-ULBP2.1*7      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*8      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*9      :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*10     :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*11     :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*12     :---------- ---------- ---------- ---------- --------I- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*13     :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.1*14     :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- G--------- -
Mamu-ULBP2.1*15     :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- G--------- -
Mafa-ULBP2.1*1      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*2      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*3      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*4      :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*5      :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*6      :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*7      :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*8      :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*9      :---------- -T-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*10     :---------- -----R---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mafa-ULBP2.1*11     :---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -
Mamu-ULBP2.2*1      :-----V---- -----R---- ---------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mamu-ULBP2.2*2      :----QV---- ---------- ---------- ---------- M--------- ---Q------ ---------- -----K---- G--------- -
Mamu-ULBP2.2*3      :-----V---- ---------- ---------- ---------- ---------- ---------- ---R------ -----K---- G--------- -
Mamu-ULBP2.2*4      :-----V---- ---------- ---------- ---------- ---------- ---------- ---R------ -----K---- G--------- -
Mamu-ULBP2.2*5      :-----V---- ---------- --S------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mamu-ULBP2.2*6      :-----V---- ---------- --S------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mamu-ULBP2.2*7      :-----V---- -----R---- ---------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mamu-ULBP2.2*8      :----QV---- ---------- ---------- ---------- ---------- ------R--- ---------- -----K---T GM-------- -
Mamu-ULBP2.2*9      :----QV---- ---------- ---------- ---------- ---------- ------R--- ---------- -----K---T GT-------- -
Mamu-ULBP2.2*10     :----QV---- ---------- ---------- ---------- M--------- ---Q------ ---------- -----K---T G--------- -
Mamu-ULBP2.2*11     :-----V---- -----R---- ---------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mafa-ULBP2.2*1      :-----V---- -----R---- ---------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mafa-ULBP2.2*2      :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- -----K---- G--------- -
Mafa-ULBP2.2*3      :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- -----K---- G--------- -
Mafa-ULBP2.2*4      :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- -----K---- G--------- -
Mafa-ULBP2.2*5      :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- -----K---- G--------- -
Mafa-ULBP2.2*6      :-----V---- -----R---- ---------- ---------- ---------- ---------- ---------- -----K---T GT-------- -
Mafa-ULBP2.2*7      :-----V---- -----R---- ---------- ---------- ---------- ---Q------ ---------- -P---K---T GT-------- -
Mafa-ULBP2.2*8      :-----V---- ---------- ---------- ---------- ---------- -------N-- ---------- -----K---T ---------- -
Mafa-ULBP2.2*9      :----QV---- ---------- ---------- ---------- ---------- ------R--- ---------- -----K---T G--------- -
Mafa-ULBP2.2*10     :-----V---- ---------- ---------- ---------- ---------- ---------- ---------- -----K---- G--------- -
Human ULBP2         :---------- ---------- --S----I-- -----K---- ---------- -------V-A -----F---- -IG--E---- G--------- -
Human ULBP6         :---------- ---------- --SI---T-- -----K---- ---------- ---------A ---------- -IG--E---- G--------- -
```

Fig. 2 continued.

macaques (Table 1). In this study, we found a repeat number polymorphism in the A stretch in intron 2 in both *ULBP2.1* and *ULBP2.2* from both rhesus and cynomolgus macaques. Polymorphisms in exons 2 and 3 and intron 2 including the repeat polymorphism were included in the allele designation. Nucleotide sequences of *Mamu-ULBP2.2*1* and *Mafa-ULBP2.2*1* were identical to those of LOC694600. That the identical *ULBP2* alleles were shared in part by both rhesus and cynomolgus macaques was consistent with a cross-breeding between these macaques as suggested by the studies of diversity in the MHC class I genes (Saito et al. 2012).

Divergence of *ULBP2* genes in the higher primates

Nucleotide sequence homologies among the alleles of *ULBP2.1* and *ULBP2.2* in rhesus and cynomolgus macaques were 94.3 and 98.7 %, respectively, suggesting that *ULBP2.2* is less diverged than *ULBP2.1*. To figure out the evolutional divergence and diversity of *ULBP2* genes in the higher primates, we conducted a neighbor-joining analysis by using nucleotide sequences of exons 2 and 3 from *ULBP2.1* and *ULBP2.2* in macaques along with sequences of corresponding region from *ULBP2* and/or *ULBP6* reported for human, chimpanzee, gorilla, and another Old World monkey, olive baboon. *ULBP5* sequences from human and rhesus were also included in the analysis. As shown in Fig. 1, alleles of *ULBP2.1* and those of *ULBP2.2* in macaques were separately clustered, but both rhesus and cynomolgus alleles were found in the same cluster. Quite interestingly, *ULBP2.2* of macaques including olive baboon *ULBP2* and rhesus *ULBP5* appeared to be diverged from an ancestral *ULBP2.1*. In addition, *ULBP2* sequences from human, gorilla and chimpanzee as well as *ULBP5* and *ULBP6* sequences from human were clustered as a branch of *ULBP2.1* (Fig. 1). Furthermore, when other *ULBP* genes, *ULBP1*, *ULBP3*, and *ULBP4*, were included in the phylogenetic analysis, these genes were also clustered as

another branch of *ULBP2.1*, implying that these genes were diverged from the ancestral *ULBP2* in the primates (Supplementary Figure S1).

Diversity of *ULBP2* genes in the Old World Monkey

As for the diversity of *ULBP2* in macaques, non-synonymous substitutions were found at 14 sites in *Mamu-ULBP2.1*, 10 sites in *Mafa-ULBP2.1*, 15 sites in *Mamu-ULBP2.2*, and 16 sites in *Mafa-ULBP2.2* (Table 2). Amino acid sequences were deduced from the nucleotide sequences and alignment of the ULBP2 alleles showed that the ULBP2/RAET1H molecules in rhesus and cynomolgus macaques were homologous by more than 90 % to the human ULBP2 molecule in the α1 and α2 domains (Fig. 2). Among the polymorphic amino acid residues, six and eight residues in the ULBP2.1 and ULBP2.2 molecules, respectively, were observed in both rhesus and cynomolgus macaques (Fig. 2).

To investigate a possible role of the polymorphic residues, we created 3D structure models for ULBP2 molecules by referring the crystallographic data for human ULBP3 in complex with NKG2D, where ULBP2 residues at positions 22, 23, 70, 73, 77, 80, 83, 94, 165, 169, 170, and 172 composed of interacting surface with NKG2D (Supplementary Figure S2). It was found that three polymorphic residues at positions 73, 177, and 181 of rhesus ULBP2.1 were on the upper surface of α helix structures and pointed to the NKG2D receptor (Fig. 3a). Interestingly, the residue at position 73 could be a contacts site with Ser195 of NKG2D receptor, as deduced from the equivalent structure of human ULBP3. In contrast, none of the polymorphic residues were mapped on the surface of α helix structures in cynomolgus ULBP2.1 (Supplementary Figure S3).

On the other hand, two polymorphic residues at positions 158 and 181 of cynomolgus ULBP2.2 were on the upper surface of the α helix structure, while another polymorphic residue at position 172 was a possible interface site with Glu183 and Met184 of NKG2D in the equivalent human ULBP3, although it was not pointed up on the α helix of ULBP (Fig. 3b). In clear contrast and quite interestingly, none of the polymorphic residues were mapped on the surface of α helix in rhesus ULBP2.2 (Supplementary Figure S4). It should be noted here that the residue at position 20 is polymorphic in human ULBP2, while the residues at positions 80 and 172 are polymorphic in human ULBP6, indicating that both ULBP2 and ULBP6 might carry the allelic differences in the interaction with NKG2D in humans (Supplementary Figure S5).

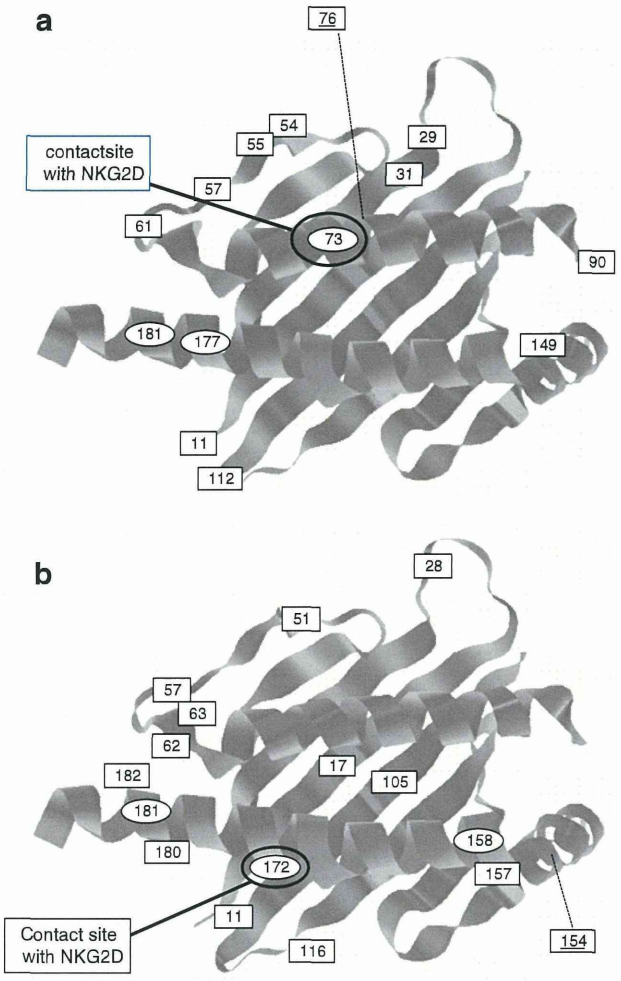

Fig. 3 Mapping of polymorphic sites on the 3D-structure model of macaque ULBP2 molecule. Polymorphic sites were mapped on the 3D structure model of ULBP2/RAET1H. Positions of polymorphic amino acid resides in the rhesus ULBP2.1 molecule (**a**) and cynomolgus ULBP2.2 molecule (**b**). Residues on the upper side of α helix are indicated by *circles*, while those on the outer side or on the β seat are indicated by *squares*. The residues mapped behind or beneath the α helix are underlined and represented by *dotted lines*. Possible contact sites with NKG2D (Radaev et al. 2001) are indicated

Discussion

In this study, we investigated the polymorphic nature of *ULBP2/RAET1H* in the Old World monkey. We previously reported that each member of the *ULBP/RAET1* gene family, except for *ULBP6/RAET1L*, was duplicated in the rhesus genome (Naruse et al., 2011). As expected, we obtained *ULBP2.1* and *ULBP2.2* sequences from both rhesus and cynomolgus macaques. On the other hand, any orthologous genes to human *ULBP6* were not detected in the macaques, even though *ULBP6* showed 96 % homology to *ULBP2* in humans (Radosavljevic et al. 2001). It was considered that *ULBP2.1* or *ULBP2.2* might be orthologous to human *ULBP6*. However, our phylogenetic analysis indicated that both human *ULBP2* and *ULBP6* were clustered with *ULBP2*s from chimpanzee and gorilla, as a branch of *ULBP2.1*. In addition, intron 2 sequences of ULBP2.2 in

macaques were relatively well conserved among *ULBP2*s from olive baboon, chimpanzee, western gorilla, and human as well as in human *ULBP6* than exon sequences. Furthermore, a Blast search showed that there was no *ULBP6*-like sequence in the genomes of chimpanzee and gorilla. These observations suggested that human *ULBP2* and *ULBP6* were diverged from an ancestral *ULBP2* after the diversification of human and other higher primates. The phylogenetic analysis also indicated that *ULBP2.2* might be diverged from *ULBP2.1* and the clustering of *ULBP2.1* alleles and *ULBP2.2* alleles did not depend on the species, supporting a trans-species evolution. Our observations were consistent with that ULBP/REAT molecule of placental mammals was originally diverged and duplicated in each species after an emigration from the *MHC* region (Kondo et al. 2010).

On the other hand, *Mafa-ULBP2.1*5* was placed at the diverging point of *ULBP2.2* in the phylogenetic tree (Fig. 1 and Supplementary Figure S1). As shown in Fig. 2, among the *ULBP2.1* alleles, only the *Mafa-ULBP2.1*5* has a replacement of Arg with Val at the position 105, which is a common feature of *Mamu-ULBP2.2* and *Mafa-ULBP2.2*. In addition, *Mafa-ULBP2.1*5* carries *ULBP2.2*-like sequences at the positions of 29 and 54. These characteristic features may eventually position the *Mafa-ULBP2.1*5* at the diverging point of *ULBP2.2*. Alternatively, gene conversion-like events from *ULBP2.2* had occurred in *ULBP2.1* to generate the *Mafa-ULBP2.1*5*.

In the present study, we denoted 15 and 11 *ULBP2.1* alleles and 11 and 10 *ULBP2.2* alleles in rhesus and cynomolgus macaques, respectively, of which more than 70 % of polymorphisms were non-synonymous. Induced expression of human ULBP2 molecule is involved in the recognition of virus-infected cell by NKG2D (Ward et al. 2009), although the functional significance of the polymorphisms in the extracellular domain of ULBP2 molecules remains to be deciphered. We demonstrated that several polymorphisms of ULBP2 molecules were located at the presumed contact sites with NKG2D or on the upper surface of α helix, which might have functional impacts. On the other hand, it has been reported that the sequence identities are less than 60 % among the ULBP molecules, while those between the ULBP and MIC molecules are about 25 % (Cosman et al. 2001). The interface residues appeared to be less conserved than the overall sequence among the ULBP molecules, and it was predicted that NKG2D could recognize the diversities of ULBP1 and ULBP2 molecules through induced-fit mechanisms in a similar manner as that of ULBP3 (Radaev et al. 2001). However, other previous studies revealed that the structural differences among the α2 domains of ULBP and MIC molecules affect the binding affinity to NKG2D and UL16, respectively (Wittenbrink et al. 2009; Spreu et al. 2006), suggesting that the ULBP polymorphisms demonstrated in this study might influence the efficacy of recognition by NKG2D. The functional impact of the polymorphisms should be investigated in future studies to decipher the evolutional and biological significance of the ULBP2 polymorphisms in the Old World monkeys.

Acknowledgments We thank Yukiko Ueda, Yasuko Saida, and Nana Ohkubo for their technical assistances. This work was supported in part by research grants from the Ministry of Health, Labor and Welfare, Japan and a Grant-in-Aid for scientific research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan. This work was also supported by a supporting program for women researchers from the Tokyo Medical and Dental University.

References

Antoun A, Jobson S, Cook M, O'Callaghan CA, Moss P, Briggs DC (2010) Single nucleotide polymorphism analysis of the NKG2D ligand cluster on the long arm of chromosome 6: extensive polymorphisms and evidence of diversity between human populations. Hum Immunol 71:610–620

Bacon L, Eagle RA, Meyer M, Easom N, Young NT, Trowsdale J (2004) Two human ULBP/RAET1 molecules with transmembrane region are ligands for NKG2D. J Immunol 173:1078–1084

Bauer S, Groh V, Wu J, Steinle A, Phillips JH, Lanier LL, Spies T (1999) Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. Science 285:727–729

Chalupny NJ, Sutherland CL, Lawrence WA, Rein-Weston A, Cosman D (2003) ULBP 4 is a novel ligand for human NKG2D. Biochem Biophys Res Commun 305:129–135

Cosman D, Mullberg J, Sutherland CL, Chin W, Armitage R, Fanslow R, Kubin M, Chalupny NJ (2001) ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. Immunity 14:123–133

Eagle RA, Traherne JA, Ashiru O, Wills MR, Trowsdale J (2006) Regulation of NKG2D ligand gene expression. Hum Immunol 67: 1159–1169

Eagle RA, Flack G, Warford A, Martinez-Borra J, Jafferji I, Traherne JA, Ohashi M, Boyle LH, Barrow AD, Caillat-Zucman S, Young NT, Trowsdale J (2009a) Celllar expression, trafficking, and function of two isoforms of human ULBP5/RAET1G. PLoS ONE 4:e4503

Eagle RA, Traherne JA, Hair JR, Jafferji I, Trowsdale J (2009b) ULBP6/RAET1L is an additional human NKG2D ligand. Eur J Immunol 39: 3207–3216

Gibbs RA, Rogers J, Katze MG et al (2007) Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234

Kondo M, Maruoka T, Otsuka N, Kasamatsu J, Fugo K, Hanzawa N, Kasahara M (2010) Comparative genomic analysis of mammalian NKG2D ligand family genes provides insights into their origin and evolution. Immunogenetics 62:441–450

Kulski JK, Anzai T, Shiina T, Inoko H (2004) Rhesus macaque class I duplicon structures, organization, and evolution within the alpha block of the major histocompatibility complex. Mol Biol Evol 21: 2079–2091

Matusali G, Tchigjou HK, Pontrelli G, Bernardi S, D'Ettorre G, Vullo V, Buonomini AR, Andreoni M, Santoni A, Cerboni C, Doria M (2013) Soluble ligands for the NKG2D receptor are released during HIV-1 infection and impair NKG2D expression and cytotoxicity of NK cells. FASEB J 27:2440–2450

Naruse TK, Chen Z, Yanagida R, Yamashita T, Saito Y, Mori K, Akari H, Yasutomi Y, Miyazawa M, Matano T, Kimura A (2010) Diversity of MHC class I genes in Burmese-origin rhesus macaque. Immunogenetics 62:601–611