

Study of cynomolgus monkey (*Macaca fascicularis*) *Mhc DRB* gene polymorphism in four populations

Antoine Blancher · Alice Aarnink · Keiko Tanaka ·
Masao Ota · Hidetoshi Inoko · Hisashi Yamanaka ·
Hiroshi Nakagawa · Pol-André Apoil · Takashi Shiina

Received: 18 January 2012 / Accepted: 12 March 2012 / Published online: 4 April 2012
© Springer-Verlag 2012

Abstract The cynomolgus macaque (*Macaca fascicularis*) is currently used as an animal model in various fields of immunology especially in the development of innovative vaccines for the prevention and treatment of infectious diseases. The polymorphism of the major histocompatibility complex (MHC) influences the development of adaptive immune responses, and it is crucial to characterize the polymorphism of cynomolgus MHC genes. Among all macaque species, the cynomolgus macaque has the most diversified geographical area encompassing continental and insular populations. By the study of a large sample of animals from the Philippines ($N=359$), we have characterized 20 *DRB* haplotypes. The *DRB* genotyping was

performed by denaturing gradient gel electrophoresis (DGGE) sequencing of exon 2 and was confirmed by polymerase chain reaction–sequence-specific oligonucleotide. The *DRB* and *DRA* cDNA of 126 animals were characterized by cloning and sequencing. By means of DGGE sequencing, we characterized the polymorphism of genomic *DRB* exon 2 in three other cynomolgus macaque population samples (Java, Vietnam, and Mauritius), and we discuss about the origin of the founders of the Mauritian and the Filipino cynomolgus macaque populations.

Keywords MHC · *DRB* · Cynomolgus macaque · Polymorphism

Electronic supplementary material The online version of this article (doi:10.1007/s00251-012-0613-5) contains supplementary material, which is available to authorized users.

A. Blancher (✉) · A. Aarnink · P.-A. Apoil
Laboratoire d'Immunogénétique moléculaire, EA 3034,
Faculté de Médecine Purpan, Université Paul Sabatier, Toulouse 3,
CHU de Toulouse, 1 avenue Jean Poulhes, TSA, 50032, 31059,
Toulouse cedex 9, France
e-mail: blancher.a@chu-toulouse.fr

K. Tanaka · H. Inoko · T. Shiina
Department of Molecular Life Science,
Division of Basic Medical Science and Molecular Medicine,
Tokai University School of Medicine, 143 Shimokasuya,
Isehara, Kanagawa 259-1143, Japan

M. Ota
Department of Legal Medicine, Shinshu University School
of Medicine, Matsumoto, Nagano 390-8621, Japan

H. Yamanaka · H. Nakagawa
Ina Research Inc, 2148-188 Nishiminowa,
Ina, Nagano 399-4501, Japan

Introduction

Among all macaque species, *Macaca fascicularis fascicularis* (*Mafa*) has the most diversified geographical area with continental and insular populations which are clearly distinguished by mitochondrial and nuclear gene polymorphisms (Tosi et al. 2002; Smith et al. 2007; Blancher et al. 2008; Shiina et al. 2010). The Philippine cynomolgus macaque population which belongs to the “insular population” was isolated from the continental population long time ago. The founding of the macaque Philippine population is probably as old as 110,000 years before present and resulted from the settlement of animals arriving from the Sunda shelf thanks to the low level of the sea during the last strong Pleistocene glaciation (Blancher et al. 2008). Contrary to what we found in the *Mafa* Vietnam population, the *Mafa* Philippine population is free of introgression from the rhesus monkey (Bonhomme et al. 2009; Stevison and Kohn 2009). Therefore, the *Mafa* Philippine population is an isolated population that had

experienced a founding effect followed by a probable rapid initial expansion of the population which reached a plateau following, by and large, a classical logistic model. This population constitutes an excellent model to observe the consequences of such a demographic trajectory on the polymorphism of the major histocompatibility complex (MHC).

Previously, we reported on the *DRB* polymorphism in the Philippine and Mauritian populations and more recently on the *DRA* polymorphism in various cynomolgus populations (Blancher et al. 2006; Aarnink et al. 2010). The *DR* region of animals originating from Indochina and the Indonesian islands was recently described (Doxiadis et al. 2010; Creager et al. 2011; Doxiadis et al. 2012).

We present here the characterization of 20 *DRA-DRB* haplotypes of the *Mafa* Philippine population by the study of a large animal sample. Moreover, by means of denaturing gradient gel electrophoresis (DGGE) sequencing, we characterized the polymorphism of genomic *DRB* exon 2 in animals originated from Indonesia, Vietnam, and Mauritius and discuss about the origin of the Mauritian and the Filipino population founders.

Materials and methods

Animals

We studied 1,369 unrelated cynomolgus monkeys from the Mauritius Island ($N=750$, from Noveprim Co. and Le Tamarinier Co.), Philippine archipelago (257 from Sicombrec Co. and 102 from INA Research Philippines, Inc.), Java ($N=138$, Tsukuba Primate Center), and Vietnam ($N=122$, Nafovanny). The blood collection and animal studies were conducted in accordance with the guidelines for animal experiments specific to each location.

DNA extraction

Genomic DNA was extracted from peripheral blood using either QIAamp Blood Mini Kit (Qiagen, Courtaboeuf, France) or a standard phenol–chloroform method.

RNA extraction

Total RNA was extracted from buffy coats by means of an RNeasy Mini Kit (Qiagen, Courtaboeuf, France) or by using the TRIzol reagent (Invitrogen, CA, USA).

Characterization of *DRB* alleles by denaturing gradient gel electrophoresis and sequencing (DGGE sequencing)

Genomic *DRB* exon 2 sequences were determined by DGGE sequencing as described previously (Blancher et al. 2006).

The recurrent association of certain alleles characterized by DGGE sequencing in various animals suggested that these bands could correspond to *DRB* genes inherited as blocks corresponding to *DRB* putative haplotypes.

Characterization of *DRB* alleles by polymerase chain reaction–sequence-specific oligonucleotide

Briefly, a *DRB* exon 2-amplified fragment from genomic DNA was obtained by using the Taq QIAGEN kit (Qiagen, Courtaboeuf, France), with primers DRBP1 (sense) and DRBP2 (antisense) previously described by Blancher et al. (2006). The cycling parameters were 2 min at 94 °C, followed by 32 cycles of 15 s at 94 °C denaturation step, 30 s at 60 °C annealing step, and 30 s at 72 °C extension step and 7 min at 72 °C. Each PCR product will be screened with 49 probes labeled at the 3' with digoxigenin [Table 1 in the Electronic supplementary material (ESM)]. For that, the PCR products were blotted on Hybond N+ membranes (Amersham) and fixed by exposure to UV (5 min). A prehybridization was carried for 1 h at 54 °C in a hybridization buffer containing 3 M tetramethylammonium chloride (TMAC), 50 mM Tris–HCl pH=8, 2 mM ethylenediaminetetraacetic acid (EDTA), 5× Denhardt's reagent, 0.1 % sodium dodecyl sulfate (SDS), and 10% salmon sperm. For each of the 49 probes, the hybridization was performed overnight at 54 °C with a solution of digoxigenin (DIG)-labeled probes (2 pm/ml) in the hybridization buffer. A first washing was performed in 20 mM sodium phosphate, pH 7.4/0.3 M NaCl/2 mM EDTA/0.1 % SDS at room temperature for 10 min. A second washing was performed at a temperature varying as a function of the probe (see Table 1 in the ESM) in a buffer containing 3 M TMAC, 50 mM Tris–HCl pH=8, 2 mM EDTA pH=8, and 0.1 % SDS. After two washings in appropriate buffers, the membranes were incubated with alkaline phosphatase-conjugated anti-DIG antibodies (Roche, Paris, France). After two washings in appropriate buffers, the detection was carried out using a chemiluminescent substrate for alkaline phosphatase (CDP-Star: Disodium 4-chloro-3-(methoxy)spiro {1,2-dioxetane-3,2'-(5'-chloro)tricyclo [3.3.1.^{13,7}]decan}-1-4-yl)phenyl phosphate, Roche, Paris, France) and exposure of the membrane to X-ray film.

Characterization of *DRB* cDNA sequences

For 24 animals from Sicombrec Co., the characterization of *DRB* cDNA sequences was performed as previously described by Blancher et al. (2006) with the pair of primers DRB5SALP (sense) and DRB3BAMP (antisense) described by Lekutis and Letvin (1995). This pair of primer was referred as “set 1.”

For 102 animals from INA Research Philippines Inc, *DRB* cDNA sequences were performed as previously described by Lekutis and Letvin (1995) and O'Connor et al. (2007) with the

pairs of primers DRB5SALP (sense) and DRB3BAMP (antisense) and 5'MHCII-DRB-F-2 (sense) and 3'MHCII-DRB-R-2 (antisense). These pairs of primers were referred as “set 1” and “set 2,” respectively. The cDNA was synthesized by oligo d(T) primer using the ReverTra Ace for reverse transcriptase reaction (Toyobo, Japan) and was used for PCR amplification. In brief, the 20- μ l amplification reaction contained 10 ng of cDNA, 0.4 units of KOD FX polymerase (Toyobo, Japan), 1 \times PCR buffer, 2 mM of each dNTP, and 0.5 μ M of each primer. The cycling parameters were as follows: an initial denaturation of 94 °C/2 min followed by 30 cycles of 98 °C/10 s and 68 °C/1 min. PCR reactions were performed by using the thermal cycler GeneAmp PCR system 9700 (Applied Biosystems, CA, USA). Reverse transcription PCR (RT-PCR) products of the 102 cynomolgus macaques were cloned into the pGEM-T Easy vector and TArget vector with the TA cloning kit according to the protocol provided by the manufacturer (Promega, Madison, WI, USA or Toyobo, Japan) and sequenced by using the ABI3130 genetic analyzer (Applied Biosystems, CA, USA) in accordance with the protocol of BigDye terminator method. To avoid PCR and sequencing artifacts generated by polymerase errors, 24 clones per individual were sequenced. The nucleotide sequences of all individuals were also determined by direct sequencing of the RT-PCR products using PCR primers as sequencing primers.

Characterization of *DRA* cDNA sequences

For 24 animals from Sicombrec Co., the characterization of *DRA* cDNA sequences was performed as previously described by Aamink et al. (2010). For 102 animals from INA Research Philippines Inc, *DRA* cDNA sequences were performed as previously described by O'Connor et al. (2007) with the pair of primers 5'MHCII-DRA-F-2 (sense) and 3'MHCII-DRA-R-2 (antisense). The protocol of sequencing of the RT-PCR product of the 102 cynomolgus macaques was the same with that of *DRB* cDNA sequencing.

Sequence analysis

All new *cDRA* and *cDRB* sequences were deposited in the GenBank database (accession references are given in Tables) and submitted to the Immuno Polymorphism Database of Major Histocompatibility Complex for Non-Human Primates (IPD-MHC NHP database) (Robinson et al. 2005) for allele nomenclature. Cynomolgus monkey allele names were assigned to *DRA* and to *DRB* sequences by the IPD-MHC NHP database following classical rules (Klein et al. 1990). Multiple sequence alignments were obtained by using CLUSTALW 1.83 (Thompson et al. 1994). Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

Results

Characterization of Philippine macaque *DR* haplotypes

By means of DGGE sequencing, we characterized 32 *DRB* alleles in a sample of 359 animals originated from the Philippines. The *DRB* transcripts of 126 animals out of the 359 of the Philippine population sample were amplified by RT-PCR and characterized by cloning sequencing. The *DRB* cDNA were amplified by two sets of primers referred to as “set 1” and “set 2” (see “Materials and methods” section). Most of the *DRB* alleles defined by their genomic exon 2 sequences (by DGGE sequencing) were transcribed (see Fig. 1 for details). In total, we characterized 30 cDNA *DRB* alleles, 8 of which were described here for the first time (Table 2 in ESM). It has to be noted that the primer “set 1” failed to amplify the *Mafa-DRB*W36:01* and *Mafa-DRB*W53:01* transcripts (*DR* haplotype #11 and #15, respectively). Moreover, the *Mafa-DRB*W1:04* allele (*DR* haplotype #10) was not detected either by DGGE sequencing or by the amplification of cDNA with the primer “set 1.” The transcript of this allele was amplified by the pair of primers “set 2.” The presence of this allele at the genomic level was confirmed by means of PCR–sequence-specific oligonucleotide (PCR–SSO) and PCR–SSP (data not shown).

By means of PCR–SSO with a set of 48 probes specific for *DRB* alleles or groups of alleles, we checked the presence of the various alleles defined by DGGE sequencing and/or by the cDNA studies. Taking into account the results obtained by the three methods used here to characterize the *DRB* gene polymorphism (DGGE sequencing, PCR–SSO, and cDNA cloning and sequencing), we deduced the *DRB* alleles present in each animal.

By the study of the association of *DRB* alleles in the genome of the animals, we confirmed the presence of a high-frequency haplotype (around 30 %). By means of heterozygous animals having this haplotype (*DRB* haplotype #7) and another one, it was possible to define ten additional *DRB* haplotypes. The characterization of nine rarer haplotypes was obtained by the study of heterozygous animals having one of these ten haplotypes but not the haplotype #7. In total, we succeeded to characterize 20 *DRB* haplotypes among which 8 were not previously reported (Blancher et al. 2006). The characterization of the *DRB* haplotypes #13 a, b, and c requires the cDNA sequencing (Fig. 1 for details).

By PCR–SSO, we defined 18 PCR–SSO profiles associated with the *DRB* haplotypes (the haplotypes #13a, #13b, and #13c gave identical profiles of hybridization; Fig. 2). By cloning and sequencing the *DRA* cDNA of the 126 animals used for the cDNA *DRB* study, we characterized 13 *Mafa-DRA* alleles, 2 of which are reported here for the first time (Table 2 in ESM). By the combination of the *DRB* and *DRA* genotyping of the 126 animals, we reconstructed 20 *DRA-DRB* haplotypes (Fig. 1).

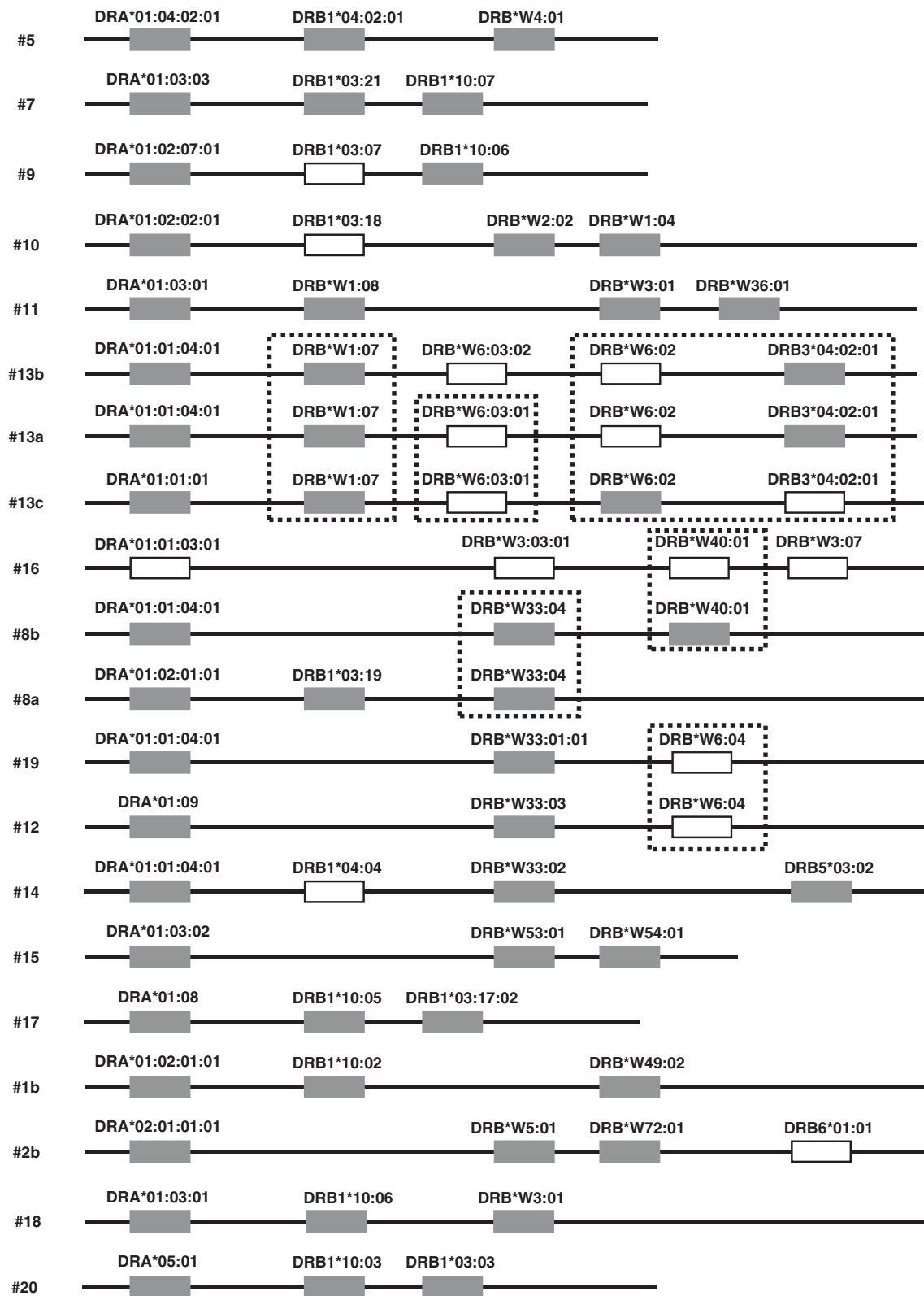


Fig. 1 Characterization of Filipino DR haplotypes. Haplotypes are defined by number on the left. The DRA and DRB genes are represented by rectangles. Gray rectangle indicates sequence found as cDNA. The order of DRB genes on each haplotype is arbitrary. Haplotypes #13a and #13b differ by sequences DRB*W6:03:01 and DRB*W6:03:02, which differ by

one nucleotide on exon 2. Haplotype #13c encompasses the same DR exon 2 that #13a haplotype but does not lead to the expression of the same DRB genes. Haplotype #16 was studied only at the genomic level because mRNA from animals having this haplotype is not available. Alleles shared by various haplotypes are shown in dotted lines

B, C, and D (Fig. 3). For each haplotype, the association with the expression of DR beta protein of the various clusters is given in Fig. 3.

The DR beta proteins deduced from the Philippines were aligned with the cynomolgus macaque DR beta proteins available in the international ImMunoGeneTics (IMGT) database. The Simpson variability index was computerized with the alignment of 111 aligned *Mafa* DR beta proteins (Fig. 4 and Table 3 in the ESM). Remarkably, out of 24 amino acid

positions in contact with the peptides, 18 have variability indexes higher than 0.45. We optimized a sinusoidal curve with a slightly variable period in order to fit the index distribution along the sequence (dotted curve in Fig. 4). From the period of the theoretical curve which fitted the best with the observed curve, we deduced that there are between 3.14 and 3.21 residues between two consecutive picks of variation. This suggests that the alpha helix could be more coiled than the classic model which predicts 3.6 amino acids per helix turn.

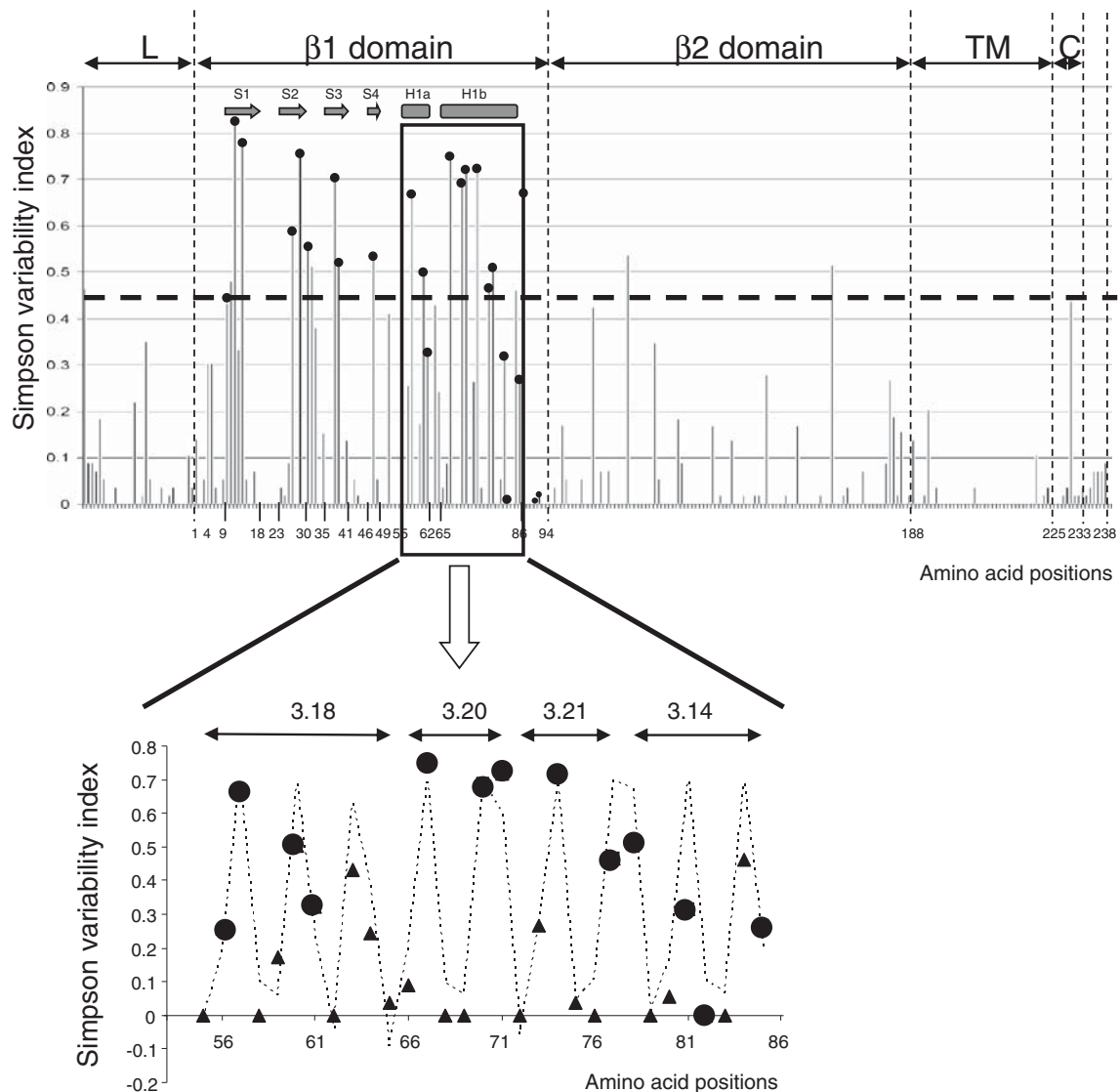


Fig. 4 Variability of *Mafa*-DR beta protein in cynomolgus macaque. The Simpson variability index (in ordinate) of each amino acid of *Mafa*-DR beta proteins was deduced of the analysis from 111 full-length transcripts (available in the IMGT database; the list of sequences taken into account is given in Table 3 of the ESM) by means of Protein Variability Server (Diez-Rivero and Reche 2009; Garcia-Boronat et al. 2008). The amino acid numeration (in abscissa) is based on the IMGT numeration. The Simpson variability indexes of amino acids in contact with peptides were noted with a black point. Each *Mafa*-DR beta protein

domain (*L* leader peptide; $\beta 1$ and $\beta 2$ domains; *TM* transmembrane domain, and *C* cytoplasmic domain) was delimited by a vertical dotted line. Secondary structure elements are indicated by gray arrows for β -strands (*S1*, *S2*, *S3*, and *S4*) and gray cylinders for α -helices (*H1a* and *H1b*). At the bottom of the figure, by a simulation curve (dotted curve), we deduced that α -helices have 3.18, 3.20, 3.21, or 3.14 amino acid residues per turn. The Simpson variability indexes of amino acids in contact with peptide were noted with a black point, while the Simpson variability indexes of other amino acids were noted with a black triangle

Study of polymorphism of *DRB* alleles in three other cynomolgus macaque populations

We genotyped 750 Mauritian cynomolgus macaques by using the PCR–SSO method. For 60 animals, the PCR–SSO genotyping was compared to that obtained by DGGE sequencing. The nine profiles of hybridization (Fig. 1 in the ESM) corresponded to the nine *DRB* haplotypes defined by DGGE sequencing. Among them, two haplotypes (#4c and #4d) were very rare. From the results obtained here, we redefined the *DRB* haplotypes of the Mauritian macaque population (see Fig. 2 in the ESM and their commentaries).

By using DGGE sequencing method, we genotyped 260 animals originated from Vietnam ($N=122$) or Java ($N=138$), and characterized 64 and 60 *Mafa-DRB* exon 2 sequences, respectively (Table 4 in the ESM). Putative *DRB* haplotypes were deduced in Vietnam (13 haplotypes) and Java (19 haplotypes) from the recurrence of allele associations in the animals of the two *Mafa* population samples (Table 5 in the ESM).

Comparison of *DRB* polymorphism in four cynomolgus macaque populations

For each cynomolgus macaque population (The Philippines, Mauritian, Java, and Vietnam), we determined the total

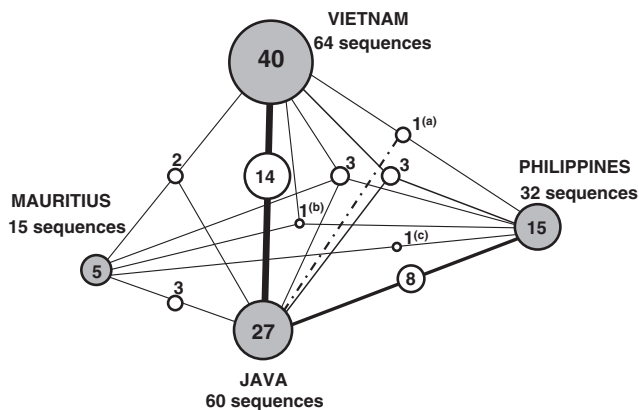


Fig. 5 *Mafa-DRB* exon 2 alleles obtained by DGGE sequencing in four cynomolgus macaque populations. The total number of exon 2 *Mafa-DRB* alleles found in each population sample is indicated under the name of that population. The numbers of *Mafa-DRB* exon alleles found exclusively in each population sample are indicated in the gray circles. The numbers of *Mafa-DRB* exon 2 alleles present in several population samples are indicated in white circles connected to gray circles by continuous lines indicating allele sharing between population samples. (a), (b), (c) These three *Mafa-DRB* exon 2 alleles were not characterized in our Javanese macaque sample while they were observed in the Mauritian and/or Philippine macaque samples. Only the *Mafa-DRB**W40:01^(a) allele was found in the Indonesian cynomolgus macaque population by other authors (HM580017 and DQ156971). This is indicated by a discontinuous line connecting this allele to the Java population

number of exon 2 *Mafa-DRB* alleles found by DGGE sequencing (Fig. 5 and Table 4 in the ESM). In the Vietnam and the Java macaque population samples, which were of similar sizes, the numbers of *DRB* alleles characterized by DGGE sequencing were very close ($N=64$ and $N=60$ in Vietnam and Java, respectively), but the number of population sample-specific alleles was much greater in Vietnam ($N=40$) than in Java ($N=27$). Indeed, 32 % of *DRB* exon 2 alleles characterized in the Java sample, and only 15% of those characterized in the Vietnam sample, were present in the Mauritian and/or Philippine population samples.

The variety of *DRB* alleles found in the Mauritian and the Philippine macaque population samples is lesser than those observed in the Java and Vietnam samples. There were 5 and 15 *Mafa-DRB* population sample-specific alleles in the Mauritian and the Philippine macaque samples, respectively (Fig. 5).

One Mauritian haplotype (#1a) was found identical in Vietnam and another one (#2a) in Java and Vietnam (Table 6 in the ESM). The most frequent Filipino *DRB* haplotype (*DRB* haplotype #7) was also found in the Java *Mafa* population sample (Table 6 in the ESM). As for all other Mauritian and Filipino haplotypes, only some of the *DRB* alleles of each haplotype were found in the two other populations. The Filipino *DRB* haplotype #17 was constituted of two sequences which were absent in all other *Mafa* population samples.

Discussion

The present study is focused on the Philippine cynomolgus population. It is based on a large sample of animals ($N=359$) from two different breeding companies, Sicombrec Co. ($N=257$) and INA Research Philippines Inc ($N=102$). It is important to note that all the animals studied here were F1 animals born in captivity from breeding animals originated from the Mindanao island of the Philippine archipelago. Therefore, we have not explored the MHC polymorphism of the macaque populations of other islands of the Philippine archipelago.

We describe here 20 Filipino *DRB* haplotypes of which 12 among them were previously defined (Blancher et al. 2006). We characterized the *DRA* and *DRB* transcripts associated with each haplotype and described two *DRA* and eight *DRB* full-length transcripts never reported so far. Our previous study (Blancher et al. 2006) was based on DGGE sequencing of exon 2 amplified from genomic DNA and on cDNA amplified by one set of primers. In the present study, we have characterized the amplified exon 2 by DGGE sequencing and also by PCR–SSO by using 48 probes described here for the first time. The PCR–SSO genotyping described here allows fast and easy characterization of *DRB* haplotypes in the population of the Philippines, although the definition of some

haplotypes (#13a, #13b, and #13c) requires the characterization of the *DRB* cDNA and DGGE sequencing (Figs. 1 and 2). The PCR–SSO technique is also adapted for the genotyping of cynomolgus macaque originated from Mauritius (Fig. 1 of the ESM).

In total, the number of *DRB* genes per Filipino haplotype varies from 2 to 4. However, this number is underestimated because the primers we have used here to amplify genomic *DRB* exon 2 fail to amplify *DRB6* alleles of most haplotypes. Indeed, only the allele *DRB6* of haplotype #2 (Filipino haplotype #2b and Mauritian haplotype #2a) was amplified and characterized in this study. By using primers different from those we have used here, Doxiadis et al. (2010) characterized the exon 2 of the *DRB6* gene of 41 out of 54 haplotypes that they have characterized in cynomolgus monkey originated from Indonesia or Indochina.

The study of the *DRA* and *DRB* transcripts confirmed, by and large, the results that we reported previously in the Philippine *Mafa* population (Blancher et al. 2006; Aarnink et al. 2010). The number of *DRB* genes transcribed varies from 1 to 3 per haplotype (Fig. 1). We have used here two different primer pairs to amplify the *DRB* mRNA; the first is identical to that we have used in our previous study, and the second was described by O'Connor et al. (2007). It is important to note that some cDNA can be amplified only with one of the two sets of primers (see “Results” section for details). This led us to conclude that the priming regions targeted by the two sets of primers are polymorphic in macaques from the Philippines. Therefore, failure to characterize the cDNA of a given allele characterized at the genomic level does not imply that this allele is not transcribed but could signify that neither of the two primer sets match with this particular cDNA sequence.

It is important to note that the precise definition of *DRB* alleles requires the characterization of the full-length cDNA sequence because two alleles identical at the exon 2 can differ outside exon 2. For example, *Mafa-DRB*W33:01:01* of the haplotype (#19) differs from *Mafa-DRB*W33:04* (haplotype #8a and 8b) by only one nucleotide in exon 1. These observations were recently reported by Doxiadis et al. (2012).

The study of the variability of amino acid residues deduced from *Mafa-DRB* transcripts confirms that only the amino acid positions the most variable correspond to amino acids in contact with the peptide (Reche and Reinherz 2003) (Fig. 4). At the alpha-helical β 1 domain of DR beta protein, the variability of amino acids varies cyclically. We simulate this cyclic variation with a sinusoidal function with a variable period (from 3.14 to 3.21 amino acid residues). The oscillation period of the variability index is slightly less than the period deduced from the classic alpha helix (3.6 amino acid residues per turn of the helix). By restricting the comparison of DR beta proteins expressed by Filipino animals to the amino acid positions in contact with the peptide, we defined four clusters (Fig. 3).

Out of the 15 *DRB* haplotypes expressing two *DRB* genes, 12 are associated with the expression of DR beta proteins belonging to two different clusters and three with the expression of two DR beta proteins belonging to the same cluster. Remarkably, these three haplotypes were present at low frequencies in the animal sample studied here (see legend in Fig. 3). This observation suggests that the haplotypes associated with the expression of DR beta proteins belonging to different clusters have been favored during the evolution of the Filipino *Mafa* population. The selective advantage of haplotypes composed of *DRB* genes encoding proteins belonging to different clusters resides probably in the correlative increase of the variety of peptides presented by the *DRB* gene products associated with such haplotypes.

We have also studied here the polymorphism of genomic *DRB* exon 2 by means of DGGE sequencing in animals originated from Java and Vietnam. In the case of the Mauritian animal, the polymorphism of exon 2 was characterized by means of PCR–SSO in a large sample of animals ($N=750$) (Fig. 1 in the ESM). The results obtained here allowed us to redefine the *DRB* haplotypes in the Mauritian population (Fig. 2 in the ESM). We compared the variety of exon 2 *DRB* alleles evidenced by DGGE sequencing in each of the four macaque populations (Philippines, Java, Vietnam, and Mauritius). In the populations of Java and Vietnam, the levels of the *DRB* gene polymorphism are comparable and are much higher than in the Philippine and Mauritian populations.

The Indonesian origin of Mauritian population founders and the Sunda shelf origin of the Philippine population are illustrated by the large allele sharing between these two populations and the Javanese population sample studied here (Tosi and Coke 2007; Blancher et al. 2008; Bonhomme et al. 2008). However, it is important to note that almost one half of the Mauritian alleles are found neither in Java nor in Vietnam. These alleles could have been imported with animals originated from Sumatra as evoked by Tosi and Coke (2007), or these alleles are present at low frequencies in the Java population and our Javanese animal sample was too small to give us the opportunity to find them. It is also possible that we failed to characterize these alleles by DGGE sequencing. By evidence, it is very unlikely that *DRB* alleles differing from those found in the Javanese ancestral population have emerged in Mauritius only 400 years after the founding of the macaque population.

One half of the *DRB* Filipino alleles are found also in the Javanese population sample studied here. The 15 alleles which are found only in the Filipino population could result from a local differentiation or could be part of the allele stock of the Sunda shelf ancestral population. To resolve this alternative, more animals of the Sunda shelf *Mafa* populations (Java, Borneo, Sumatra, and Malaysian peninsula) have to be studied. In these future studies, DGGE sequencing could be replaced

by a more exhaustive method such as high throughput pyrosequencing of amplified *DRB* exon 2.

The recurrent associations of exon 2 allele in the Javanese and Vietnamese animal samples allowed us to propose 19 and 13 putative *DRB* haplotypes, respectively (Table 5 in the ESM). It is remarkable that 14 out of these combinations have been previously reported by Doxiadis et al. (2010 and 2012) in animals originating from Indochina or Indonesia (Table 6 in the ESM). Five *DRB* haplotypes are observed in at least two of the four *Mafa* populations studied here. Notably, the haplotype present at very high frequency (around 30 %) in the Philippines (haplotype #7) is also present in the Java population (Table 6 in the ESM).

On the other hand, among the 30 *DRB* haplotypes described in the rhesus monkey (Doxiadis et al. 2007), 9 are similar to those described in *Mafa* (this study and Doxiadis et al. 2010, 2012) (Table 7 in the ESM). The persistence of these haplotypes over time in different populations of rhesus and cynomolgus macaques could have been favored by selection. This haplotype sharing can result from either inheritance of haplotypes of the common ancestral species or introgressive hybridization between the two species (Bonhomme et al. 2009; Stevison and Kohn 2009). It is interesting to remind that modern human MHC alleles inherited by introgressive hybridization with old human species (Neanderthal or Denisovan) could have been maintained by selection (Abi-Rached et al. 2011).

In conclusion, the macaques represent an exceptional model of a radiation in Asia of species from a common ancestor species and offer the opportunity to apprehend the impact of the selection on the evolution of the MHC genes and haplotypes across the speciation, introgression, and fragmentation of each species in highly differentiated populations.

Acknowledgments We are pleased to thank all the technicians of the Toulouse Laboratory of Immunogenetics for their excellent technical assistance: Béatrice Atlan, Audrey Dauba, Stéphanie Despiau-Schiavinato, and Sylvie Hébrard. We thank Kouji Otabe, Kengo Sakamoto, Masaaki Kurata, and Mamoru Nomura for helping with the cynomolgus macaque blood collection. This work was supported by Grant-in-Aid for Scientific Research (B) (21300155) from Japan Society for the Promotion of Science (JSPS) and by a grant from the French Ministry of Research (contract EA3034) and a grant from University Paul Sabatier Toulouse 3.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aarnink A, Estrade L, Apoil PA, Kita YF, Saitou N, Shiina T, Blancher A (2010) Study of cynomolgus monkey (*Macaca fascicularis*) *DRA* polymorphism in four populations. *Immunogenetics* 62:123–136
- Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, Babrzadeh F, Gharizadeh B, Luo M, Plummer FA, Kimani J, Carrington M, Middleton D, Rajalingam R, Beksac M, Marsh SG, Maiers M, Guethlein LA, Tavoularis S, Little AM, Green RE, Norman PJ, Parham P (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science* 334:89–94
- Blancher A, Tisseyre P, Dutaur M, Apoil PA, Maurer C, Quesniaux V, Raulf F, Bigaud M, Abbal M (2006) Study of cynomolgus monkey (*Macaca fascicularis*) *MhcDRB* (*Mafa-DRB*) polymorphism in two populations. *Immunogenetics* 58:269–282
- Blancher A, Bonhomme M, Crouau-Roy B, Terao K, Kitano T, Saitou N (2008) Mitochondrial DNA sequence phylogeny of 4 populations of the widely distributed cynomolgus macaque (*Macaca fascicularis fascicularis*). *J Hered* 99:254–264
- Bonhomme M, Blancher A, Cuartero S, Chikhi L, Crouau-Roy B (2008) Origin and number of founders in an introduced insular primate: estimation from nuclear genetic data. *Mol Ecol* 17:1009–1019
- Bonhomme M, Cuartero S, Blancher A, Crouau-Roy B (2009) Assessing natural introgression in 2 biomedical model species, the rhesus macaque (*Macaca mulatta*) and the long-tailed macaque (*Macaca fascicularis*). *J Hered* 100:158–169
- Creager HM, Becker EA, Sandman KK, Karl JA, Lank SM, Bimber BN, Wiseman RW, Hughes AL, O'Connor SL, O'Connor DH (2011) Characterization of full-length MHC class II sequences in Indonesian and Vietnamese cynomolgus macaques. *Immunogenetics* 63:611–618
- Diez-Rivero CM, Reche PA (2009) Discovery of conserved epitopes through sequence variability analysis. In: Flower DR, Davies M, Ranganathan S (eds) *Bioinformatics for immunomics*, vol 3. Springer, New York, pp 95–101
- Doxiadis GG, de Groot N, Claas FH, Doxiadis II, van Rood JJ, Bontrop RE (2007) A highly divergent microsatellite facilitating fast and accurate *DRB* haplotyping in humans and rhesus macaques. *Proc Natl Acad Sci USA* 104:8907–8912
- Doxiadis GG, de Groot N, de Groot NG, Rotmans G, de Vos-Rouweler AJ, Bontrop RE (2010) Extensive *DRB* region diversity in cynomolgus macaques: recombination as a driving force. *Immunogenetics* 62:137–147
- Doxiadis GG, de Vos-Rouweler AJ, de Groot N, Otting N, Bontrop RE (2012) *DR* haplotype diversity of the cynomolgus macaque as defined by its transcriptome. *Immunogenetics* 64:31–37
- Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA (2008) PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res* 36:W35–W41
- Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, Watkins DI (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics* 31:217–219
- Lekutis C, Letvin NL (1995) Biochemical and molecular characterization of rhesus monkey major histocompatibility complex class II *DR*. *Hum Immunol* 43:72–80
- O'Connor SL, Blasky AJ, Pendley CJ, Becker EA, Wiseman RW, Karl JA, Hughes AL, O'Connor DH (2007) Comprehensive characterization of MHC class II haplotypes in Mauritian cynomolgus macaques. *Immunogenetics* 59:449–462
- Reche P, Reinherz E (2003) Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *J Mol Biol* 331:623–641
- Robinson J, Waller MJ, Stoehr P, Marsh SG (2005) IPD—the Immuno Polymorphism Database. *Nucleic Acids Res* 33:D523–D526
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425

- Shiina T, Tanaka K, Katsuyama Y, Otabe K, Sakamoto K, Kurata M, Nomura M, Yamanaka H, Nakagawa H, Inoko H, Ota M (2010) Mitochondrial DNA diversity among three subpopulations of cynomolgus macaques (*Macaca fascicularis*) originating from the Indochinese region. *Exp Anim* 59:567–578
- Smith DG, McDonough JW, George DA (2007) Mitochondrial DNA variation within and among regional populations of longtail macaques (*Macaca fascicularis*) in relation to other species of the *fascicularis* group of macaques. *Am J Primatol* 69:182–198
- Stevison LS, Kohn M (2009) Divergence population genetic analysis of hybridization between rhesus and cynomolgus macaques. *Mol Ecol* 18:2457–2475
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Tosi AJ, Coke CS (2007) Comparative phylogenetics offer new insights into the biogeographic history of *Macaca fascicularis* and the origin of the Mauritian macaques. *Mol Phylogenet Evol* 42:498–504
- Tosi AJ, Morales JC, Melnick DJ (2002) Y-chromosome and mitochondrial markers in *Macaca fascicularis* indicate introgression with indochinese *M. mulatta* and a biogeographic barrier in the isthmus of Kra. *Int J Primatol* 23:161–178
- Zuckermandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Vogel HJ (ed) *Evolving genes and proteins*. Academic, New York