Immunogenetics (2017) 69:87–99
DOI 10.1007/s00251-016-0954-6

ORIGINAL ARTICLE

The orthologs of *HLA-DQ* and *-DP* genes display abundant levels of variability in macaque species

Nel Otting[1] · Marit K. H. van der Wiel[1] · Nanine de Groot[1] · Annemiek J. M. de Vos-Rouweler[1] · Natasja G. de Groot[1] · Gaby G. M. Doxiadis[1] · Roger W. Wiseman[3] · David H. O'Connor[3] · Ronald E. Bontrop[1,2]

Received: 23 August 2016 / Accepted: 12 October 2016 / Published online: 22 October 2016

Abstract The human major histocompatibility complex (MHC) region encodes three types of class II molecules designated HLA-DR, -DQ, and -DP. Both the *HLA-DQ* and *-DP* gene region comprise a duplicated tandem of *A* and *B* genes, whereas in macaques, only one set of genes is present per region. A substantial sequencing project on the *DQ* and *DP* genes in various macaque populations resulted in the detection of previously 304 unreported full-length alleles. Phylogenetic studies showed that humans and macaques share trans-species lineages for the *DQA1* and *DQB1* genes, whereas the *DPA1* and *DPB1* lineages in macaques appear to be species-specific. Amino acid variability plot analyses revealed that each of the four genes displays more allelic variation in macaques than is encountered in humans. Moreover, the numbers of different amino acids at certain positions in the encoded proteins are higher than in humans. This phenomenon is remarkably prominent at the contact positions of the peptide-binding sites of the deduced macaque DPβ-chains. These differences in the MHC class II DP regions of macaques and humans suggest separate evolutionary mechanisms in the generation of diversity.

Keywords Mafa · Mamu · Mane · MHC · Nonhuman primates

Introduction

The glycoproteins of the major histocompatibility complex (MHC) play an important role in immunological defense. Two main groups of classical MHC genes and gene products are distinguished: class I and class II. The class II glycoproteins are usually expressed on the surface of antigen-presenting cells. They bind peptides and display them for recognition to the $CD4^+$ T lymphocytes, thus providing a mechanism for the induction and regulation of adaptive immune responses to extracellular pathogens.

MHC class II proteins consist of two polypeptide chains, α and β, which are encoded by two distinct genes. In humans, these genes are the *HLA-DRA* and *-DRB*, *HLA-DQA* and *-DQB*, and *HLA-DPA* and *-DPB* pairs, and they are located on the short arm of chromosome 6. The *DRB*, *DQB*, *DQA*, and *DPB* genes are highly polymorphic in the human population, and hundreds of alleles are known today, whereas the *HLA-DRA* and *-DPA* genes are oligomorphic (Robinson et al. 2015). The polymorphism of MHC class II plays a role in the susceptibility or the resistance to certain diseases. In autoimmune diseases, in particular, the MHC class II genes appear to be an important predisposing factor (Thorsby and Lie 2005). Certain *HLA-DQ* alleles, for example, are associated with disorders such as type 1 diabetes, multiple sclerosis, and celiac disease. Some DQ allotypes, with aspartic acid at position β57, confer susceptibility to type 1 diabetes, whereas β chains with alanine or serine at this position are associated

Electronic supplementary material The online version of this article (doi:10.1007/s00251-016-0954-6) contains supplementary material, which is available to authorized users.

✉ Nel Otting
otting@bprc.nl

[1] Department of Comparative Genetics and Refinement, Biomedical Primate Research Centre (BPRC), Lange Kleiweg 161, 2288 GJ Rijswijk, The Netherlands

[2] Department of Biology, Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands

[3] Department of Pathology and Laboratory Medicine, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

with protection (Gough and Simmonds 2007; Jones et al. 2006; Thorsby and Lie 2005). HLA-DP may play a role in Wegener's granulomatosis (Xie et al. 2013) and beryllium disease (Amicosante et al. 2001; Dai et al. 2010; Silveira et al. 2012). The latter syndrome is a chronic allergic-lung response to the inhalation of beryllium, a lightweight alkaline earth metal. Susceptibility is associated with HLA-DP allotypes that have a glutamic acid at position β69.

Species such as the rhesus monkey (*Macaca mulatta*), the cynomolgus macaque (*Macaca fascicularis*), and the pig-tailed macaque (*Macaca nemestrina*) are important models in human biology and diseases (t Hart et al. 2015). They are used in studies on infectious diseases such as HIV (Bontrop and Watkins 2005; Breed et al. 2015; Mooij et al. 2015; Mudd et al. 2012), tuberculosis (Flynn et al. 2015), and malaria (Divis et al. 2015), and in autoimmune disorders such as multiple sclerosis (Haanstra et al. 2013). Furthermore, macaques are essential as preclinical models for the evaluation of transplantation strategies (Kean et al. 2012). The MHC of these animals has been investigated in recent years. The orthologs of the human class II genes are present in macaques and are designated, respectively, *Mamu-*, *Mafa-* and *Mane-DR*, *-DQ*, and *-DP*. Sequencing analyses have resulted in the description of about 450 class II alleles that are archived in the nonhuman primate section of the IPD-MHC database (Blancher et al. 2014; Creager et al. 2011; de Groot et al. 2004; de Groot et al. 2012; Deng et al. 2013; Karl et al. 2014; Otting et al. 2012). However, most of these studies focused only on exon 2 sequences, which encode the antigen-binding region of the proteins (Doxiadis et al. 2006; Ling et al. 2011; Sano et al. 2006). In contrast to humans, the *DPA* genes are polymorphic in the macaques, and, moreover, the polymorphism is not restricted to exon 2 (Karl et al. 2014; Otting et al. 2012).

In the present study, six cohorts of macaques, encompassing three species, were sequenced for the full-length alleles of *DQ* and *DP* genes. The animals originated from different geographical regions. The study constituted a part of an extended project, with the aim of discovering nonhuman primate MHC genes and developing typing technologies. Not only were the repertoires of alleles in the cohorts determined but the *DQA-DQB* and the *DPA-DPB* allele combinations on the chromosomes, as they segregate in families, were also explored. This large-scale approach allowed a comparison to be made regarding relevant HLA class II genes.

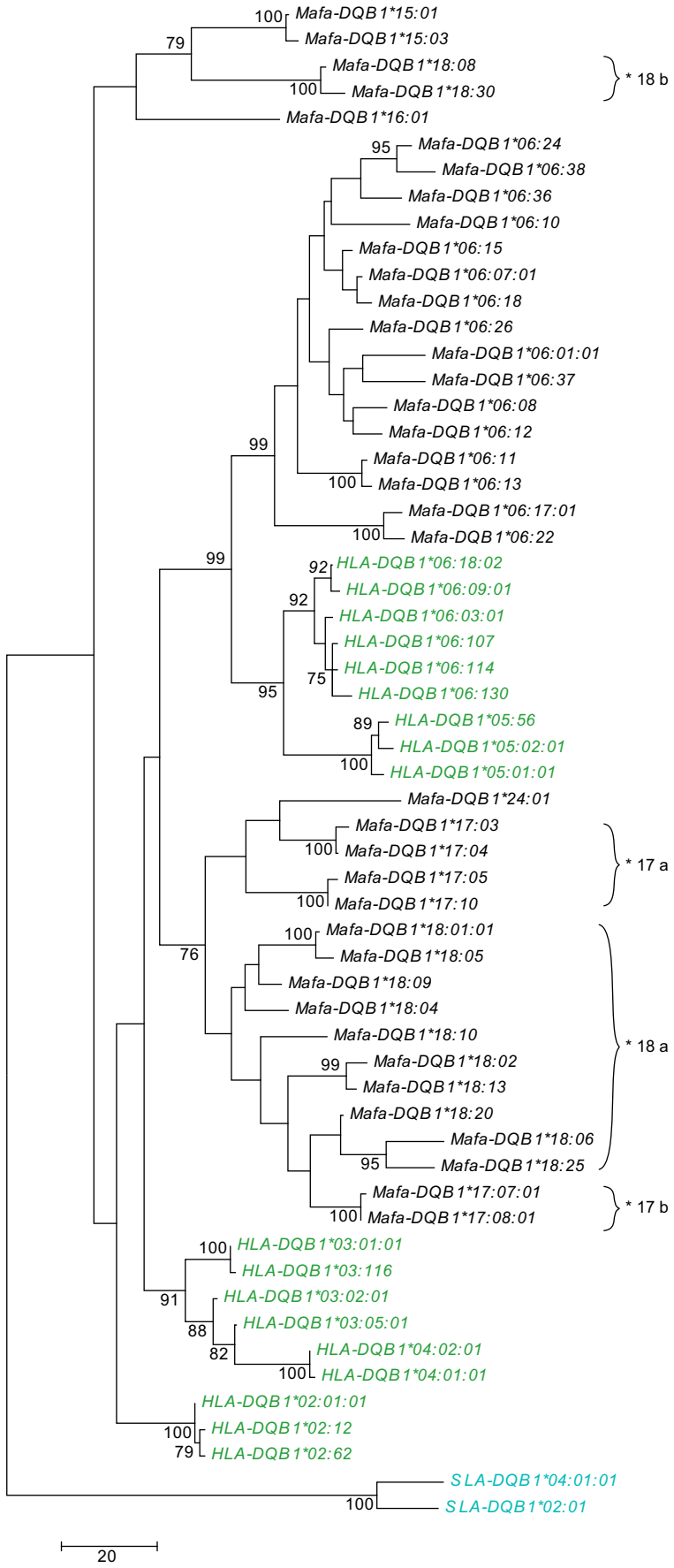
Table 1 Primers used in this study

Primer name	Sequence
DQA	
5′MHCII-DQA-F-2	CTGAGGCTGCCTTGGGAAGAA
3′MHCII-DQA-R-2	TTAGGTAGCTGGGTGGCTTACT
DQA-F	GAGGCTGCCTTGGGAAGA
DQA-R	CCTAGGTCATGTAGCAAGTCCA
DQB	
5′MHCII-DQB-F-2	CCACTACTTTTCCCTTCGTCT
3′MHCII-DQB-R-2	GGCAGGGACAAGTAGGCATT
DQB-R	CCAGTTAAAATAGTCTCAGGAGTCA
DPA	
DPA-F	CGTAGTCATCAATTAGAGACCC
DPA-R	TCCTAAGTCCTCTTCTGTTCAG
DPA-F-inex	GACAGAATGTTCSAGACCAG
DPA-fw1	CTCATCACTGTTCCTGTGCTC
DPB	
5′MHCII-DPB-F-1	GCAGCTCTTTTCATTTTGCCATCC
3′MHCII-DPB-R-2	GTCCTGGAACCAGGTGCTAACG
DPB-R738	CCTCCTGTGCATGAAGATRCCC

Table 2 The numbers of alleles detected in this study. A comparison is made to the contents of the IPD-MHC database at date 2016-01-01. The "novel" alleles and "extensions" are submitted. The "confirmed" alleles were presented as full-length sequences in the database and were also detected in the present study. Other sequences in the database were "not found" in our cohorts

	Novel	Extension	Confirmed	Not found
Mafa-DQA1	25	7	29	24
Mamu-DQA1	11	8	17	12
Mane-DQA1	22	0	0	7
Mafa-DQB1	12	17	32	41
Mamu-DQB1	9	16	21	34
Mane-DQB1	3	20	1	14
Mafa-DPA1	24	8	27	23
Mamu-DPA1	22	8	12	8
Mane-DPA1	14	0	3	1
Mafa-DPB1	24	12	23	30
Mamu-DPB1	22	6	10	23
Mane-DPB1	14	0	0	2
Σ	202	102	175	219

Fig. 1 Maximum parsimony analyses of *HLA-DQB1* and *Mafa-DQB1* ▶ sequences. The evolutionary history was inferred using the maximum parsimony method. The most parsimonious tree with length = 798 is shown. The consistency index is (0.449333), the retention index is (0.774563), and the composite index is 0.373693 (0.348037) for all sites and parsimony-informative sites. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths calculated using the average pathway method and are in the units of the number of changes over the whole sequence. The human sequences are depicted in *green*; an out-group of swine sequences is depicted in *blue*. The *DQB*17* and *DQB*18* lineages, which were originally defined based on exon 2 sequences, are both spliced in two branches, a and b


Mafa-DQB1*15:01
Mafa-DQB1*15:03
Mafa-DQB1*18:08
Mafa-DQB1*18:30
* 18 b
Mafa-DQB1*16:01
Mafa-DQB1*06:24
Mafa-DQB1*06:38
Mafa-DQB1*06:36
Mafa-DQB1*06:10
Mafa-DQB1*06:15
Mafa-DQB1*06:07:01
Mafa-DQB1*06:18
Mafa-DQB1*06:26
Mafa-DQB1*06:01:01
Mafa-DQB1*06:37
Mafa-DQB1*06:08
Mafa-DQB1*06:12
Mafa-DQB1*06:11
Mafa-DQB1*06:13
Mafa-DQB1*06:17:01
Mafa-DQB1*06:22
HLA-DQB1*06:18:02
HLA-DQB1*06:09:01
HLA-DQB1*06:03:01
HLA-DQB1*06:107
HLA-DQB1*06:114
HLA-DQB1*06:130
HLA-DQB1*05:56
HLA-DQB1*05:02:01
HLA-DQB1*05:01:01
Mafa-DQB1*24:01
Mafa-DQB1*17:03
Mafa-DQB1*17:04
Mafa-DQB1*17:05
Mafa-DQB1*17:10
* 17 a
Mafa-DQB1*18:01:01
Mafa-DQB1*18:05
Mafa-DQB1*18:09
Mafa-DQB1*18:04
Mafa-DQB1*18:10
Mafa-DQB1*18:02
Mafa-DQB1*18:13
Mafa-DQB1*18:20
Mafa-DQB1*18:06
Mafa-DQB1*18:25
* 18 a
Mafa-DQB1*17:07:01
Mafa-DQB1*17:08:01
* 17 b
HLA-DQB1*03:01:01
HLA-DQB1*03:116
HLA-DQB1*03:02:01
HLA-DQB1*03:05:01
HLA-DQB1*04:02:01
HLA-DQB1*04:01:01
HLA-DQB1*02:01:01
HLA-DQB1*02:12
HLA-DQB1*02:62
SLA-DQB1*04:01:01
SLA-DQB1*02:01
20

Materials and methods

Animals and RNA samples

For the sequencing analyses of the MHC class II genes in the cohorts of rhesus, cynomolgus, and pig-tailed macaques, we collected peripheral blood mononuclear cells (PBMCs) or RNA samples from different institutions. The animals were selected based on sire/dame/offspring combinations, though a number of unrelated animals were also present in the panels. Indian rhesus macaque samples were derived from the Oregon National Primate Research Centre (ONPRC), Beaverton, OR ($N = 61$) and from the breeding colony at the Biomedical Primate Research Centre (BPRC) in the Netherlands ($N = 69$). Samples of rhesus macaques of Chinese origin were delivered by breeding centers in Chengdu and in Yuling, both in China ($N = 78$). AlphaGenesis Incorporated, Yemassee, SC, provided RNA samples of Indonesian cynomolgus macaques ($N = 79$), whereas Cambodian/Vietnamese cynomolgus samples were supplied by the three Chinese breeding centers in Hainan, in Yuling, and in Jinggong ($N = 119$). The pig-tailed macaque RNA samples were derived from animals housed at the Johns Hopkins University, Baltimore, MD ($N = 74$).

RNA isolation, cDNA synthesis, and PCR

RNA was isolated from the PBMC samples using either the All prep DNA/RNA mini kit (Qiagen) or the TRIzol method (ThermoFisher Scientific), according to the suppliers' protocols. First-strand complementary DNA (cDNA) syntheses were performed on the RNA samples, using the Revertaid kit, as recommended by the supplier (ThermoFisher Scientific). PCR primers used for amplification of the *DQA*, *DQB*, *DPA*, and *DPB* genes were copied from other studies on cynomolgus macaques (O'Connor et al. 2007) and pig-tailed macaques (Karl et al. 2014). These primers were situated in the promotor and 3′ UTR regions of the class II genes, resulting in full-length PCR products that included the start and stop codons (Table 1). Two additional DP primers were designed based on exon sequences, resulting in PCR products that were not full-length. The amplification reactions were performed as follows: An initial step of 2 min at 94 °C was followed by 25 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and the last extension step was extended to 5 min. The PCR products were run on a 1 % agarose gel, excised from this gel, and purified using the GeneJet gel extraction kit (ThermoFisher Scientific). The PCR amplifications were always performed with the primers that were first mentioned in Table 1. In instances where no PCR product was obtained, or when only one allele was found, another round of PCR was performed with the additional primers.

Sequencing and analyses

Direct sequencing reactions on the PCR products were performed in two directions, using the BigDye terminator cycle sequencing kit, and the samples were run on a Genetic Analyzer 3130 capillary system (ThermoFisher Scientific). The sequencing primers were identical to the PCR primers. The results were, in most cases, peak-patterns containing double peaks based on two alleles. The two full-length alleles were identified by the SBT engine software (Gendx, Utrecht, the Netherlands) or by manual comparison using MacVector™ version 13.0.7 (Oxford Molecular Group). In the event that two alleles in a PCR sample could not be identified, a cloning step was introduced. These PCR products were ligated using the CloneJET PCR cloning system (ThermoFisher Scientific). After transformation into XL1-blue bacteria (Stratagene, Huissen, the Netherlands), 16 colonies were selected for culturing and plasmid isolation. Sequencing on the plasmid DNA was performed as described above.

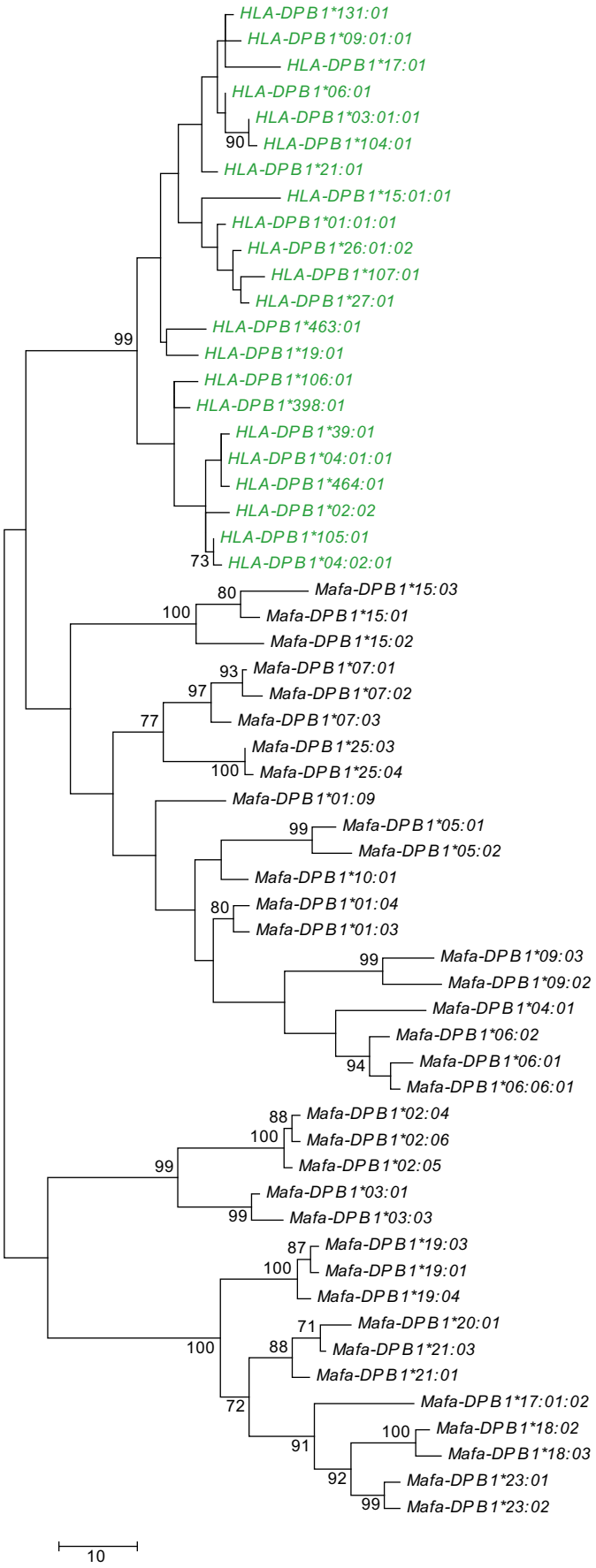
New alleles, based on at least two independent PCR reactions, were submitted to the EMBL/EBI database (www.ebi.ac.uk/ena), and for official designations, to the nonhuman primate section of the IPD-MHC database (www.ebi.ac.uk/ipd/mhc/nhp/) (de Groot et al. 2012). Phylogenetic analyses of the MHC class II sequences were conducted using MEGA7 software (Kumar et al. 2016). The evolutionary histories were inferred using the maximum parsimony analysis method.

Results

Novel alleles

All animals in the six cohorts were sequenced for the *DQ* and *DP* genes, and the obtained sequences were compared to the alleles that were already deposited in the IPD-MHC database, using the SBT engine and MacVector software. New alleles were in most cases confirmed by other animals in the sire/dame/offspring triads. In instances where an unreported sequence was observed in only one animal, a second round of PCR and sequencing was performed on the same sample to confirm this sequence. In total, 304 novel alleles were

Fig. 2 Maximum parsimony analysis of *HLA-DPB1* and *Mafa-DPB1* sequences. The evolutionary history was inferred using the maximum parsimony method. The most parsimonious tree with length = 467 is shown. The consistency index is (0.437071), the retention index is (0.832196), and the composite index is 0.393823 (0.363729) for all sites and parsimony-informative sites. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths calculated using the average pathway method and are in the units of the number of changes over the whole sequence. The *HLA-DPB1* sequences are depicted in *green*. Non-primate full-length *DPB1* sequences, to use as an out-group, were not available


HLA-DPB1*131:01
HLA-DPB1*09:01:01
HLA-DPB1*17:01
HLA-DPB1*06:01
HLA-DPB1*03:01:01
90
HLA-DPB1*104:01
HLA-DPB1*21:01
HLA-DPB1*15:01:01
HLA-DPB1*01:01:01
HLA-DPB1*26:01:02
HLA-DPB1*107:01
HLA-DPB1*27:01
HLA-DPB1*463:01
99
HLA-DPB1*19:01
HLA-DPB1*106:01
HLA-DPB1*398:01
HLA-DPB1*39:01
HLA-DPB1*04:01:01
HLA-DPB1*464:01
HLA-DPB1*02:02
HLA-DPB1*105:01
73
HLA-DPB1*04:02:01
80
Mafa-DPB1*15:03
100
Mafa-DPB1*15:01
Mafa-DPB1*15:02
93
Mafa-DPB1*07:01
97
Mafa-DPB1*07:02
77
Mafa-DPB1*07:03
Mafa-DPB1*25:03
100
Mafa-DPB1*25:04
Mafa-DPB1*01:09
99
Mafa-DPB1*05:01
Mafa-DPB1*05:02
Mafa-DPB1*10:01
80
Mafa-DPB1*01:04
Mafa-DPB1*01:03
99
Mafa-DPB1*09:03
Mafa-DPB1*09:02
Mafa-DPB1*04:01
Mafa-DPB1*06:02
94
Mafa-DPB1*06:01
Mafa-DPB1*06:06:01
88
Mafa-DPB1*02:04
100
Mafa-DPB1*02:06
99
Mafa-DPB1*02:05
Mafa-DPB1*03:01
99
Mafa-DPB1*03:03
87
Mafa-DPB1*19:03
100
Mafa-DPB1*19:01
Mafa-DPB1*19:04
71
Mafa-DPB1*20:01
100
88
Mafa-DPB1*21:03
Mafa-DPB1*21:01
72
Mafa-DPB1*17:01:02
100
Mafa-DPB1*18:02
91
Mafa-DPB1*18:03
92
Mafa-DPB1*23:01
99
Mafa-DPB1*23:02
10

identified and submitted to the databases, 102 of which were extensions of known sequences. All the submitted alleles, together with the accession numbers, a reference animal, and other relevant information, are shown in Supplementary Table 1. The alleles were designated according to the nomenclature rules for macaques and other nonhuman primates, which have been published (de Groot et al. 2012). The alleles received the locus-number "1" in their designations, since they are assumed to be orthologs of the expressed human *HLA-DQA1*, *-DQB1*, *-DPA1*, and *-DPB1* genes.

At present, the BPRC breeding colony comprises about 1100 rhesus macaques (*M. mulatta*) of Indian origin. Animals in the colony were typed for their MHC-markers for three decades, and 140 different MHC-haplotypes have been defined by segregation analyses (Doxiadis et al. 2013). The cohort of 69 Indian rhesus macaques is a sample representing haplotypes in the contemporary BPRC breeding colony. Sequencing of the class II genes in this group revealed the presence of 18 *DQA1*, 19 *DQB1*, 14 *DPA1*, and 14 *DPB1* alleles. Most of these alleles had already been published and deposited in the IPD-MHC database. Only seven novel alleles were found, and five other exon 2 sequences have been extended. Sequencing the cohort of Indian rhesus macaques obtained from ONPRC appeared to contain almost the same set of alleles as are present in the BPRC animals. Three unreported alleles were found; however, two of which were also present in Chinese rhesus macaques. Within the panel of 78 animals, derived from Chinese centers, substantially more allelic variation was encountered, and 30 *DQA1*, 37 *DQB1*, 35 *DPA1*, and 30 *DPB1* alleles were detected. From these alleles, 28 extended already published exon 2 sequences, whereas another 58, mostly *DPA1* and *DPB1*, represented unreported alleles.

In the cohort of 79 cynomolgus macaques (*M. fascicularis*) of Indonesian origin, 27 *DQA1*, 25 *DQB1*, 22 *DPA1*, and 25 *DPB1* alleles were present. Among these, 27 sequences appeared to be novel, and six were elongations of known ones. The cynomolgus monkeys, which are kept at the three Chinese breeding centers, were of Cambodian (Ca), Vietnamese (Vi), or mixed origin. Since there are no natural boundaries between these two neighboring realms, we considered them as originating from the same region (Ca/Vi). In the panel of 119 animals, as many as 51 *DQA1*, 49 *DQB1*, 47 *DPA1*, and 48 *DPB1* alleles were detected. Twelve sequences confirmed new alleles present in the Indonesian animals as well, whereas 27 were extensions of published exon 2 sequences. Furthermore, 70 unreported alleles were detected in this panel.

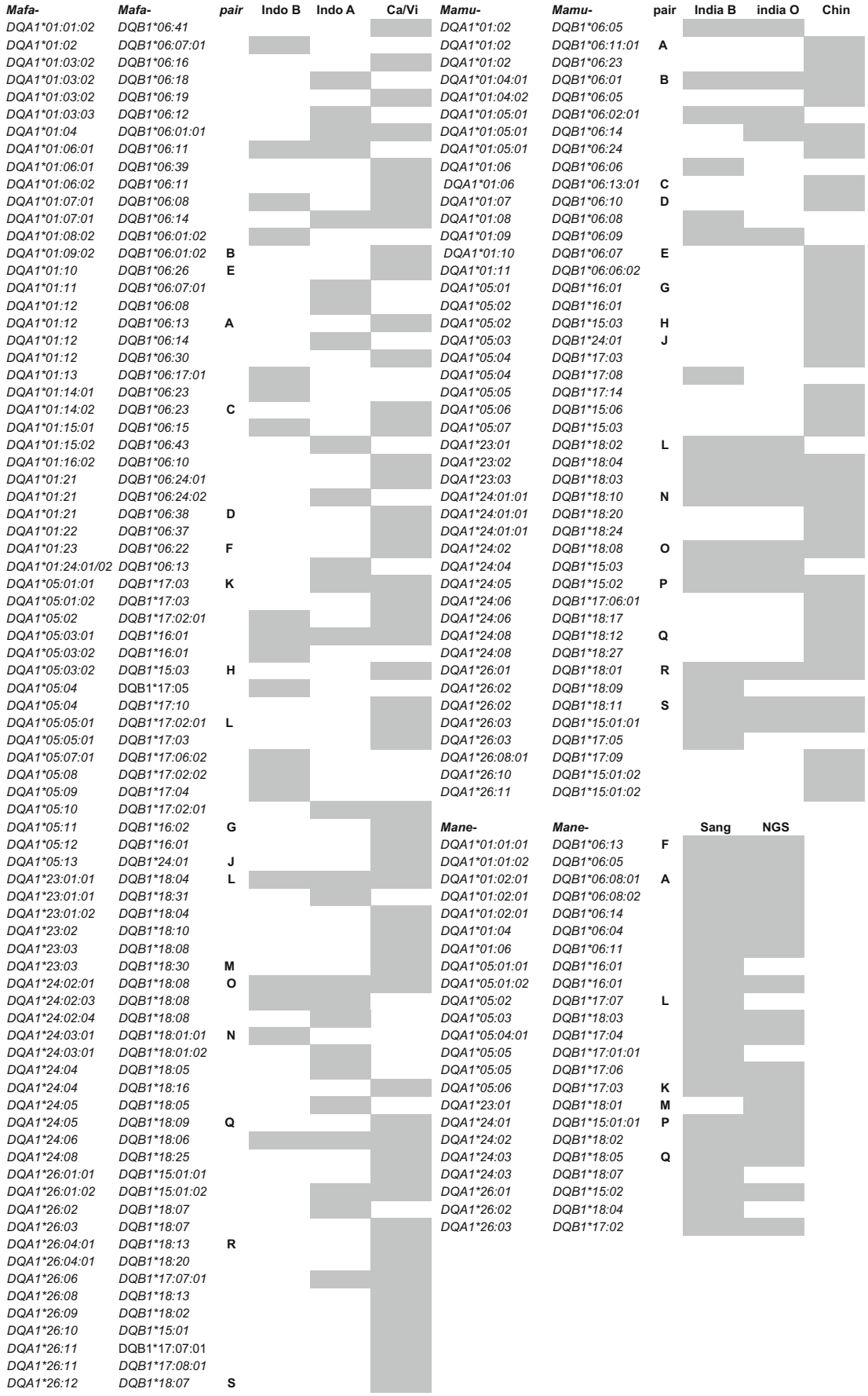
The panel of pig-tailed macaques (*M. nemestrina*) comprised 74 animals, in which 17 *DPA1*, 14 *DPB1*, 22 *DQA1*, and 24 *DQB1* alleles were found. Recently, MHC class II genes have been analyzed by next-generation sequencing (NGS) in 32 pig-tailed macaques that are also held at the Johns Hopkins University (Karl et al. 2014). The present Sanger sequencing analyses confirmed the findings in this report, and the accession numbers of both studies are provided in the Supplementary Table 1. In the IPD-MHC database, a series of *Mane-DQB1* exon 2 sequences had already been archived. Full-length sequencing elongated 20 of these deposited exon 2 sequences, and four novel *DQB1* alleles were found. In the NGS and the present Sanger sequencing studies, the *DPA1*, *DPB1*, and *DQA1* genes in the pig-tailed macaque have been sequenced for the first time, and all alleles are new. In Table 2, an overview is provided of the numbers of newly detected alleles for each gene, in comparison to the sums of those already deposited in the IPD-MHC database. In the cohorts analyzed, 202 were submitted as novel alleles and 102 as extensions of known sequences. A total of 175 alleles present in the database were confirmed by our analyses, whereas another 219—either exon 2 or full-length sequences—were not observed in these cohorts.

Allelic lineages: does the trans-species mode of evolution persist?

The different macaque species share many full-length MHC class II alleles (Doxiadis et al. 2006). For instance, out of the group of 32 submitted *Mafa-DQA1* alleles, 13 sequences are also present in the rhesus or in the pig-tailed macaques (Supplementary Table 1, last column). The high level of sharing of full-length alleles has not been reported for other primate species. Next to the fully identical ones, various alleles are observed that differ for only one or a few nucleotides, which reflect the close genetic relationship of the investigated macaque species.

In the designations of the macaque class II sequences, the first two digits following the asterisk reflect the allelic lineage. A lineage is a group of alleles that originate from one ancestral structure. Some of the macaque *DQA1* and *DQB1* lineages are shared with humans and great apes (de Groot et al. 2012). For instance, the *DQA1*01* and *DQA1*05* lineages are also present in humans, gorillas, chimpanzees, and orangutans. The *DQB1*06* lineage in macaques is shared with humans, though *DQB1*15* and *DQB1*16* are also present in, respectively, chimpanzee and orangutan (Otting et al. 2002). The definition

Fig. 3 The *DQA1-DQB1* pairs in the various macaque cohorts. The pairs ▶ are sorted on the *DQA1* lineages. Only those pairs are listed that were observed in at least two animals. The pairs A-S are identical allele sets in different macaques species. *Indo B* Indonesian cynomolgus macaques held at the BPRC, *Indo A* Indonesian cynomolgus macaques held at AlphaGenesis Inc, *Ca/Vi* cynomolgus macaques of Cambodian and/or Vietnamese origin, *India B* Indian rhesus macaques held at the BPRC, *India O* Indian rhesus macaques held at the ONPRC, *Chin* rhesus macaques of Chinese origin, *Sang* pig-tailed macaques analyzed by Sanger sequencing, *NGS* pig-tailed macaques analyzed by NG sequencing


Mafa- Mafa- pair Indo B Indo A Ca/Vi
DQA1*01:01:02 DQB1*06:41
DQA1*01:02 DQB1*06:07:01
DQA1*01:03:02 DQB1*06:16
DQA1*01:03:02 DQB1*06:18
DQA1*01:03:02 DQB1*06:19
DQA1*01:03:03 DQB1*06:12
DQA1*01:04 DQB1*06:01:01
DQA1*01:06:01 DQB1*06:11
DQA1*01:06:01 DQB1*06:39
DQA1*01:06:02 DQB1*06:11
DQA1*01:07:01 DQB1*06:08
DQA1*01:07:01 DQB1*06:14
DQA1*01:08:02 DQB1*06:01:02
DQA1*01:09:02 DQB1*06:01:02 B
DQA1*01:10 DQB1*06:26 E
DQA1*01:11 DQB1*06:07:01
DQA1*01:12 DQB1*06:08
DQA1*01:12 DQB1*06:13 A
DQA1*01:12 DQB1*06:14
DQA1*01:12 DQB1*06:30
DQA1*01:13 DQB1*06:17:01
DQA1*01:14:01 DQB1*06:23
DQA1*01:14:02 DQB1*06:23 C
DQA1*01:15:01 DQB1*06:15
DQA1*01:15:02 DQB1*06:43
DQA1*01:16:02 DQB1*06:10
DQA1*01:21 DQB1*06:24:01
DQA1*01:21 DQB1*06:24:02
DQA1*01:21 DQB1*06:38 D
DQA1*01:22 DQB1*06:37
DQA1*01:23 DQB1*06:22 F
DQA1*01:24:01/02 DQB1*06:13
DQA1*05:01:01 DQB1*17:03 K
DQA1*05:01:02 DQB1*17:03
DQA1*05:02 DQB1*17:02:01
DQA1*05:03:01 DQB1*16:01
DQA1*05:03:02 DQB1*16:01
DQA1*05:03:02 DQB1*15:03 H
DQA1*05:04 DQB1*17:05
DQA1*05:04 DQB1*17:10
DQA1*05:05:01 DQB1*17:02:01 L
DQA1*05:05:01 DQB1*17:03
DQA1*05:07:01 DQB1*17:06:02
DQA1*05:08 DQB1*17:02:02
DQA1*05:09 DQB1*17:04
DQA1*05:10 DQB1*17:02:01
DQA1*05:11 DQB1*16:02 G
DQA1*05:12 DQB1*16:01
DQA1*05:13 DQB1*24:01 J
DQA1*23:01:01 DQB1*18:04 L
DQA1*23:01:01 DQB1*18:31
DQA1*23:01:02 DQB1*18:04
DQA1*23:02 DQB1*18:10
DQA1*23:03 DQB1*18:08
DQA1*23:03 DQB1*18:30 M
DQA1*24:02:01 DQB1*18:08 O
DQA1*24:02:03 DQB1*18:08
DQA1*24:02:04 DQB1*18:08
DQA1*24:03:01 DQB1*18:01:01 N
DQA1*24:03:01 DQB1*18:01:02
DQA1*24:04 DQB1*18:05
DQA1*24:04 DQB1*18:16
DQA1*24:05 DQB1*18:05
DQA1*24:05 DQB1*18:09 Q
DQA1*24:06 DQB1*18:06
DQA1*24:08 DQB1*18:25
DQA1*26:01:01 DQB1*15:01:01
DQA1*26:01:02 DQB1*15:01:02
DQA1*26:02 DQB1*18:07
DQA1*26:03 DQB1*18:07
DQA1*26:04:01 DQB1*18:13 R
DQA1*26:04:01 DQB1*18:20
DQA1*26:06 DQB1*17:07:01
DQA1*26:08 DQB1*18:13
DQA1*26:09 DQB1*18:02
DQA1*26:10 DQB1*15:01
DQA1*26:11 DQB1*17:07:01
DQA1*26:11 DQB1*17:08:01
DQA1*26:12 DQB1*18:07 S
Mamu- Mamu- pair India B india O Chin
DQA1*01:02 DQB1*06:05
DQA1*01:02 DQB1*06:11:01 A
DQA1*01:02 DQB1*06:23
DQA1*01:04:01 DQB1*06:01 B
DQA1*01:04:02 DQB1*06:05
DQA1*01:05:01 DQB1*06:02:01
DQA1*01:05:01 DQB1*06:14
DQA1*01:05:01 DQB1*06:24
DQA1*01:06 DQB1*06:06
DQA1*01:06 DQB1*06:13:01 C
DQA1*01:07 DQB1*06:10 D
DQA1*01:08 DQB1*06:08
DQA1*01:09 DQB1*06:09
DQA1*01:10 DQB1*06:07 E
DQA1*01:11 DQB1*06:06:02
DQA1*05:01 DQB1*16:01 G
DQA1*05:02 DQB1*16:01
DQA1*05:02 DQB1*15:03 H
DQA1*05:03 DQB1*24:01 J
DQA1*05:04 DQB1*17:03
DQA1*05:04 DQB1*17:08
DQA1*05:05 DQB1*17:14
DQA1*05:06 DQB1*15:06
DQA1*05:07 DQB1*15:03
DQA1*23:01 DQB1*18:02 L
DQA1*23:02 DQB1*18:04
DQA1*23:03 DQB1*18:03
DQA1*24:01:01 DQB1*18:10 N
DQA1*24:01:01 DQB1*18:20
DQA1*24:01:01 DQB1*18:24
DQA1*24:02 DQB1*18:08 O
DQA1*24:04 DQB1*15:03
DQA1*24:05 DQB1*15:02 P
DQA1*24:06 DQB1*17:06:01
DQA1*24:06 DQB1*18:17
DQA1*24:08 DQB1*18:12 Q
DQA1*24:08 DQB1*18:27
DQA1*26:01 DQB1*18:01 R
DQA1*26:02 DQB1*18:09
DQA1*26:02 DQB1*18:11 S
DQA1*26:03 DQB1*15:01:01
DQA1*26:03 DQB1*17:05
DQA1*26:08:01 DQB1*17:09
DQA1*26:10 DQB1*15:01:02
DQA1*26:11 DQB1*15:01:02
Mane- Mane- Sang NGS
DQA1*01:01:01 DQB1*06:13 F
DQA1*01:01:02 DQB1*06:05
DQA1*01:02:01 DQB1*06:08:01 A
DQA1*01:02:01 DQB1*06:08:02
DQA1*01:02:01 DQB1*06:14
DQA1*01:04 DQB1*06:04
DQA1*01:06 DQB1*06:11
DQA1*05:01:01 DQB1*16:01
DQA1*05:01:02 DQB1*16:01
DQA1*05:02 DQB1*17:07 L
DQA1*05:03 DQB1*18:03
DQA1*05:04:01 DQB1*17:04
DQA1*05:05 DQB1*17:01:01
DQA1*05:05 DQB1*17:06
DQA1*05:06 DQB1*17:03 K
DQA1*23:01 DQB1*18:01 M
DQA1*24:01 DQB1*15:01:01 P
DQA1*24:02 DQB1*18:02
DQA1*24:03 DQB1*18:05 Q
DQA1*24:03 DQB1*18:07
DQA1*26:01 DQB1*15:02
DQA1*26:02 DQB1*18:04
DQA1*26:03 DQB1*17:02

of the lineages above, which supports the trans-species mode of evolution, was based on exon 2 studies. To investigate whether this feature extends to the full-length sequences, phylogenetic analyses were performed. Available full-length *HLA-DQ* and *-DP* alleles were downloaded from the IPD-IMGT/HLA database (www.ebi.ac.uk/ipd/imgt/hla/) and aligned using MEGA7 software, respectively, with subsets of *Mafa-DQ* and *-DP* sequences. The resulting phylogenetic trees confirm the sharing of the *DQA1*01*, *DQA1*05* (Supplementary Figure 1), and *DQB1*06* (Fig. 1) lineages. However, intermingling of human and macaque alleles in these clades was not observed. The phylogenetic analyses show that the *DQB1*17* and *DQB1*18* lineages in macaques split into two clades (a and b), alluding to their complex evolution (Fig. 1). The *DQB1*17b* sequences cluster within the main *DQB1*18a* branch, whereas a separate *DQB1*18b* cluster is formed by only four alleles, found in the three macaque species. Thus, subdivision into lineages, which was initially based on motifs in exon 2, does not remain intact for the full-length *DQB1*17* and **18* sequences. Renaming of the *DQB1 *17b* and **18b* alleles is needed in the future.

The situation in the *MHC-DP* region is markedly different in humans and macaques. In humans, the variation of the *DPB1* gene is generated, to a minor extent, by point mutations, but the exchange of small sequence motifs by recombination played a more prominent role in the generation of allelic polymorphism (Doxiadis et al. 2001; Gyllensten et al. 1996). The rapid evolution of *HLA-DPB1* makes it difficult to divide the alleles into distinct lineages, and, as a consequence, each newly detected allele is sequentially named with its own lineage number, at present 572 in total. It was presumed that this was the case for the *DPB1* genes in macaques as well, and alleles were numbered sequentially in order of their detection. However, phylogenetic analyses with available full-length *Mafa-DPB1* alleles show the inter-species division in lineages (Fig. 2). For that reason, the macaque *DPB* alleles were recently grouped into 20 lineages and renamed (de Groot et al. 2012). With the novel alleles discovered in this study, this number has been extended to 25 lineages. Furthermore, the phylogenetic tree shows that different *HLA-DPB1* lineages cluster together in one clade, separated from macaque lineages.

The polymorphism of the *HLA-DPA1* gene is limited, and about 21 alleles are available in the IPD-IMGT/HLA database, divided into four allelic lineages. The *DPA1* gene in macaques is polymorphic, and 121 alleles were detected in the present populations. Moreover, the variation is not limited to exon 2 that encodes the antigen-binding part of the protein. A previously published study indicated that the *DPA1*04* lineage is shared with humans and great apes (Otting and Bontrop 1995), whereas the other lineages *DPA1*02*, **06* up to **11* appear to be more specific for Old World monkeys. The present phylogenetic analyses on full-length *DPA1* alleles of humans and macaques do not support the trans-species mode of evolution (Supplementary Figure 2).

DQ haplotypes

Based on the sire/dame/offspring triplets in the cohorts studied, it was possible to determine the combinations of alleles, segregating on one chromosome, or haplotypes. Unrelated animals were also included in the panels, and in many cases, the *DQ* pairs were deducible by comparison to other animals. Only the *DQ* pairs that were observed in at least two animals are listed (Fig. 3). Sometimes, a combination of alleles was present in another species of macaque (Fig. 3, A-S) but the names of the shared alleles differ, since they are given in order of discovery.

The results of two independent and previously published studies are included in the overall analysis. The first study was performed on cynomolgus macaques from the breeding colony kept at the BPRC (Otting et al. 2012). These animals are mainly of Indonesian origin (Fig. 3, indo B). In addition, the NGS results involving the pig-tailed macaques are included (Fig. 3, NGS) (Karl et al. 2014). The sample of 69 Indian rhesus macaques held at the BPRC reflects, in fact, the colony of about 1100 animals that are the offspring of 137 founder animals, representing 140 different MHC haplotypes (Doxiadis et al. 2013).

The two cohorts of Indonesian cynomolgus macaques, Indo B and Indo A, appear to have their own smaller repertoires of *DQ* haplotypes, whereas in the cohort of Indochinese origin (Ca/Vi), the collection is more extended. The smaller repertoires in the two Indonesian groups may be the result of founder effects, due to by the starting of breeding colonies. Another cause may be the separation of populations on the islands in this large archipelago. However, it is possible that more similarities are observed when these cohorts are extended. The repertoires in the two groups of Indian rhesus macaques are similar, but limited in comparison with the animals derived from the Chinese breeding centers. This is in line with the hypothesis that the Indian rhesus macaques have experienced a severe bottleneck (Hernandez et al. 2007; Smith and McDonough 2005). The pig-tailed macaques analyzed in this study, and those investigated by NGS, are derived from the same breeding center, resulting in DQ series that are almost identical.

The *DQA1/B1* combinations in macaques were further investigated and compared to those in humans. *HLA-DQ1* haplotypes were downloaded from the HLA section of the allele frequencies website (http://www.allelefrequencies.net) and summarized (Gonzalez-Galarza et al. 2015). The data are based on almost 26,700 haplotypes, detected in 18 populations around the world. An overview of the *DQA1* and *DQB1* combinations, in humans and in macaques as well, is represented (Fig. 4).

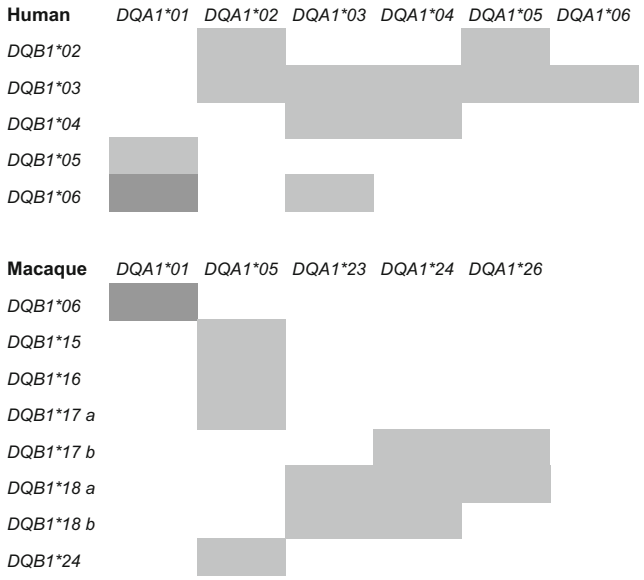


Fig. 4 The HLA and macaque *DQA1/DQB1* lineage combinations. The *DQA1*01/DQB1*06* combinations are observed both in humans and in macaques, as indicated by *darker shading*

In macaques, the alleles of the *DQA1*01* lineage are exclusively linked to sequences of the *DQB1*06* cluster. In humans, next to pairing to *DQB1*06*, combinations with *DQB1*05* sequences are observed, a lineage that is absent in macaques. The conserved linkage of *DQA1*01* and *DQB1*06* lineages predates the separation of Old World monkeys and hominoids (Doxiadis et al. 2001). For the other *DQA* lineages, promiscuous pairing with *DQB* alleles of multiple lineages is evident (Fig. 4). The *DQB1*18* lineage in macaques is split into two clades: the extended **18a* clade and the smaller **18b* group (Fig. 1). The alleles in both clades are found in combination with *DQA1*23*, **24*, or **26* lineages. The situation is different in the other fragmented lineage, *DQB1*17*. One group of alleles, **17b*—which clusters close to *DQB1*18*—is seen with *DQA1*24* and **26*, whereas alleles in the other *DQB1*17a* group pair only with *DQA1*05* alleles. This confirms our observation that the first group of *DQB1*17b* shares an evolutionary descent with the *DQB1*18a* lineage.

The pairing of specific alleles is not exclusive, and one allele can be found in combination with more than one partner. For example, *DQA1*01*:*12* is found in combination with four different *DQB1*06* alleles. Also, *DQB1* sequences are also observed that are found in combination with more than one *DQA1* allele. The same observation is made with regard to the human system (data not shown).

Although the number of *HLA-DQ* haplotypes are a hundredfold greater than those detected in the macaque cohorts, the general picture remains the same (Fig. 4). Some *DQA1* lineages pair exclusively with particular *DQB1* lineages and vice versa. Other lineages are less strict, and pairing with more than one lineage is observed. This is probably a reflection of the extent to which the gene products can form dimers.

DP haplotypes

As described for the *DQ* loci above, the *DPA1/DPB1* combinations in the cohorts were determined and listed (Supplementary Figure 3). A strong linkage between *DQ* and *DP* pairs appears to be absent due to a hotspot of recombination mapping between these two regions. The *DP* haplotype repertoires are more extended in the Chinese rhesus macaques and in the cynomolgus macaques of Cambodian/Vietnamese origin, as was observed for the *DQ* haplotypes.

The *DPA1*04* lineage alleles pair only with the *DPB1*02* or **03* alleles. Furthermore, *DPA1*07* alleles are found in combination with *DPB1*19*, **20*, **21*, and **23* sequences, with one exception: in the pig-tailed macaque, *DPB1*23*:*01* is found in combination with *DPA1*12*:*01*, a newly detected lineage not observed in the rhesus and cynomolgus macaques. *DPA1*10* alleles are always connected to *DPB1*18* sequences, whereas the *DPA1*11*:*01*/*DPB1*16*:*01* combination is observed both in the cynomolgus and in the rhesus macaque. The *DPA1*02* lineage is a large group of alleles (Supplementary Figure 2), and pairing to various *DPB1* lineages is observed in this cluster.

From the HLA section of the allele frequencies website, 13,200 *DPA1*/*DPB1* haplotypes were gathered from inhabitants of 14 different regions. The *DP* haplotypes in humans show another picture in comparison to the abovementioned macaque combinations. For the *HLA-DPA1* gene, restricted polymorphism is observed, whereas each new *HLA-DPB1* allele received a unique lineage number. At present, 572 *HLA-DPB1* alleles/lineages are known. However, only 34 of these are linked to particular *HLA-DPA1* alleles and are listed in the database. The division into lineages is more or less dependent on the available alleles, initially based on exon 2 similarities. The present study shows that in the *Mafa-DPA1*02* lineage, more variation is present than in the four *HLA-DPA1* lineages together (Supplementary Figure 2).

Variability comparison to HLA

Next to class II lineages and haplotypes, the comparison of the deduced proteins to the human equivalents may shed light on their evolution and the selective forces operating on them. For this purpose, HLA-DQ and -DP protein sequences were downloaded from the IPD-IMGT/HLA database. Full-length human and macaque sequences (about 260 amino acids) were used to construct variability plots. Because of their low numbers, sequences of the pig-tailed macaques were not included. The variability plots of the rhesus macaque appeared to be virtually identical to those of the cynomolgus macaques, and only the plots of the latter are included (Fig. 5). Although

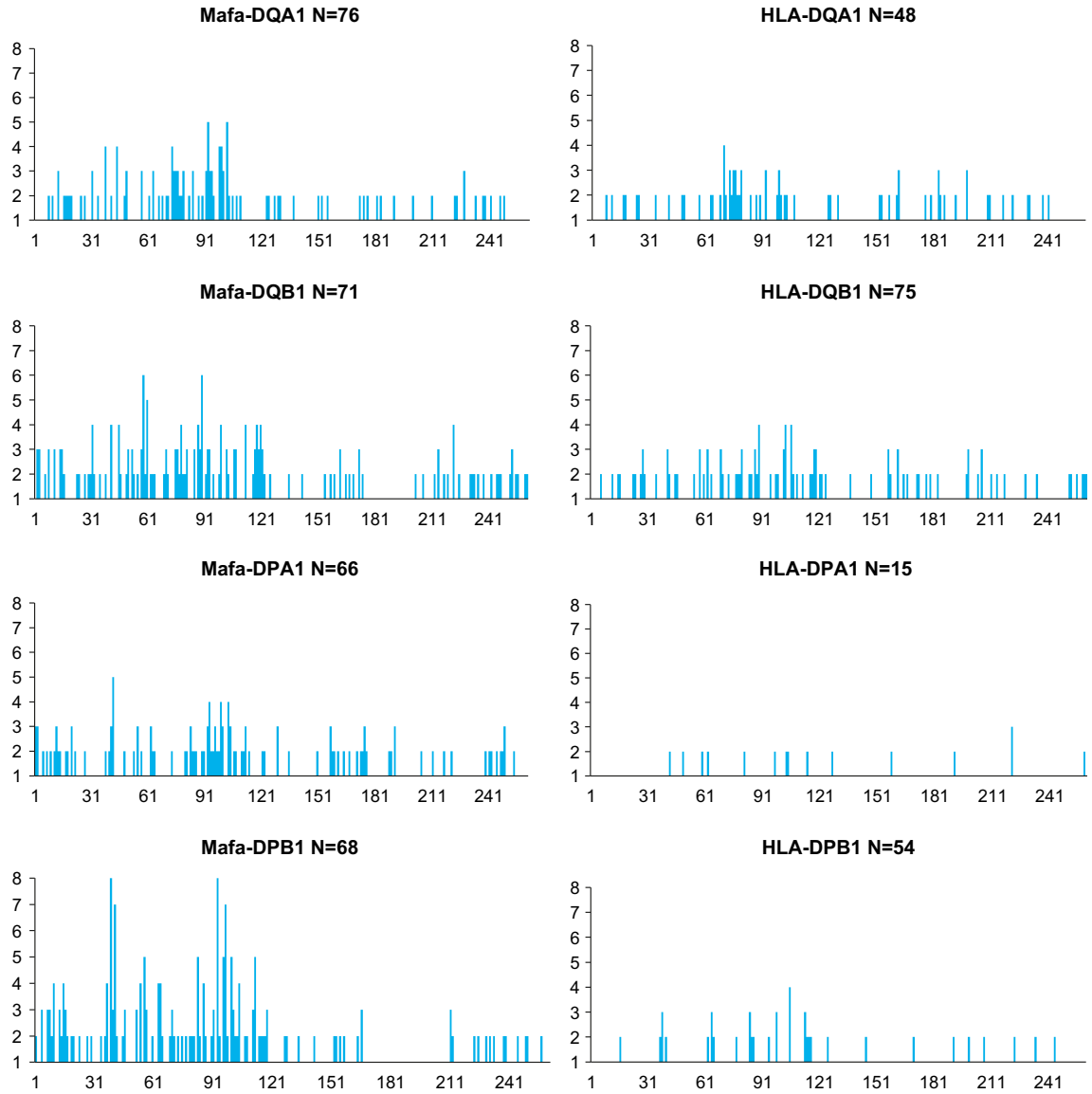


Fig. 5 Variability plots of macaque and human class II proteins. Shown on the *Y-axis* is the number of different amino acids found at a position on the protein, as depicted on the *X-axis*. The α1 and β1 domains of the proteins, encoded by exon 2, are situated between positions 30 and 120. *N* represents the number of alleles of which full-length sequences were available

the *Mafa-DQA1* and *-DQB1* plots show, as in humans, a high degree of amino acid variation, the genes in the macaque appear to be more polymorphic. Whereas in humans, only a maximum of four amino acids per position is observed; in macaques, six amino acids may be present. This suggests a higher degree of freedom to accumulate variation. In DP proteins, this type of amino acid variation is even more profound. The plots display the low variability of the *HLA-DPA1* gene and the polymorphism of its macaque ortholog over the entire length of the encoded gene. The macaque *DPB* gene, in particular, is far more polymorphic than its human counterpart.

The question is whether the MHC class II proteins of humans and macaques differ at positions that are relevant for antigen binding. To answer this question, the amino acids involved in the actual binding of antigens were investigated (Bondinas et al. 2007; Dai et al. 2010; Schneider and Stephens 1990). Since the antigen contact residues are encoded by exon 2 of the genes, most available protein sequences, downloaded from the IPD-IMGT/HLA database, were included in the analyses. The antigen-binding amino acids were aligned for all the loci and subjected to analyses by WebLogo (weblogo.berkeley.edu) (Crooks et al. 2004). The amino acids forming the antigen-binding pockets in HLA-DQA1 and -DQB1 are also observed in the macaque orthologous (Fig. 6). Mafa-DQA1 and -DQB1 have more degrees of freedom, however, since more amino acids are visible at the bases of the graphs. This observation is reinforced by the fact that a total of 886 HLA-DQB sequences were plotted versus 71 Mafa-DQB1.

Fig. 6 Weblogos of the antigen-binding amino acids. The amino acids are placed in the columns in the same order as they occur in the primary structure of the protein (Bondinas et al. 2007; Dai et al. 2010). The *height of a letter* in a column is proportional to the frequency of the corresponding amino acid at that position. The *colors of the letters* reflect the chemical properties of the encoded amino acids: *green* is polar, *purple* is neutral, *blue* is basic, *red* is acidic, and *black* is hydrophobic

Remarkably, the number of *HLA-DQB1* alleles in the databases is almost ten times that of the *HLA-DQA1* alleles, whereas in macaques, these numbers are comparable.

The situation for DP proteins is markedly different, in particular for the DPB proteins. Nine out of 15 antigen-binding residues are non-variable in humans, whereas profound variation is permitted at these positions in the macaques. This is even more striking when the numbers of allotypes are compared. In humans, the number of DPB1 allotypes (578) is almost nine times higher than in the cynomolgus macaques (68).

Discussion

The sequencing results of this study have considerably expanded the number of full-length MHC class II alleles of three macaque species in the public databases. At the start of these studies in 2014, the IPD-MHC database contained about 450 macaque DQ and DP sequences, and 203 of these were full-length alleles. In this study, we submitted 304 sequences, of which 268 represented full-length genes. This means that the sum of full-length MHC class II alleles for the three investigated macaque has been extended by 130 %.

Based on the sire/dame/offspring triads in cohorts, it was possible to determine the segregation of *DQ* and the *DP* tandems. Each cohort has its own repertoire of *DQ* and *DP* haplotypes, though intra- and inter-species sharing of allele-pairs is also observed. In the cynomolgus macaques of Indochinese origin, and in the Chinese rhesus macaques, which were derived from breeding centers in China, the number of different alleles encountered was substantially higher in comparison with the animals that are held in western breeding centers. This most likely reflects the sampling results.

The phylogenetic analyses and the variability studies on proteins reveal that the *DQA1* and *DQB1* genes in humans and macaques are subject to the same evolutionary forces. In macaque species, however, these genes are more polymorphic. Although far more human samples have been explored, it seems likely that more alleles will eventually be detected in macaques. Another observation is that DQA and DQB display more variation in the amino acids occupying the binding sites of the proteins. A logical explanation could be that Homo sapiens is a relatively young species and thus has had less time to accumulate variation. Noteworthy is the number of *HLA-*

DQB1 alleles in the database that exceeds more than ten times the sum of *HLA-DQA1* sequences. In macaques, these numbers are comparable. It is possible that the *HLA-DQB1* gene is more polymorphic.

The situation is strikingly different in the *DP* region. Despite their common ancestry, on the *DP* genes in humans and macaques, different mechanisms of generating polymorphism and selection may be at work. In the case of *HLA-DPB1*, a huge number of alleles have been discovered, which seem to have been generated by exchange of small sequence motifs. However, considering the overall result of the variation, the diversity in amino acids of the antigen-binding sites is low. In macaques, the allelic variation of *DPA1* and *DPB1* genes is much higher than in humans. There is little evidence for evolution of the *DPB1* locus by recombination, and variability seems to have been generated by point mutations. The degree of freedom observed for the contact residues of the antigen-binding site is high in the DPβ chain. These differences in the generation polymorphism may eventually result in the specialization of functions. This may be another example of the plasticity of the MHC system (van der Wiel et al. 2013).

Acknowledgments The authors wish to thank Francisca van Hassel for preparing the figures and Donna Devine for editing this article.

This work was supported by the National Institute of Health projects (NIH/NIAID HHSN272201100013C).

References

Amicosante M et al (2001) Beryllium binding to HLA-DP molecule carrying the marker of susceptibility to berylliosis glutamate beta 69. Hum Immunol 62:686–693

Blancher A, Aarnink A, Yamada Y, Tanaka K, Yamanaka H, Shiina T (2014) Study of MHC class II region polymorphism in the Filipino cynomolgus macaque population. Immunogenetics 66:219–230. doi:10.1007/s00251-014-0764-7

Bondinas GP, Moustakas AK, Papadopoulos GK (2007) The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. Immunogenetics 59:539–553. doi:10.1007/s00251-007-0224-8

Bontrop RE, Watkins DI (2005) MHC polymorphism: AIDS susceptibility in non-human primates. Trends Immunol 26:227–233. doi:10.1016/j.it.2005.02.003

Breed MW et al (2015) Elite control, gut CD4 T cell sparing, and enhanced mucosal t cell responses in Macaca nemestrina infected by a Simian immunodeficiency virus lacking a gp41 trafficking motif. J Virol 89:10156–10175. doi:10.1128/JVI.01134-15

Creager HM et al (2011) Characterization of full-length MHC class II sequences in Indonesian and Vietnamese cynomolgus macaques. Immunogenetics 63:611–618. doi:10.1007/s00251-011-0537-5

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190. doi:10.1101/gr.849004

Dai S et al (2010) Crystal structure of HLA-DP2 and implications for chronic beryllium disease. Proc Natl Acad Sci U S A 107:7425–7430. doi:10.1073/pnas.1001772107

de Groot N, Doxiadis GG, De Groot NG, Otting N, Heijmans C, Rouweler AJ, Bontrop RE (2004) Genetic makeup of the DR region in rhesus macaques: gene content, transcripts, and pseudogenes. J Immunol 172:6152–6157

de Groot NG et al (2012) Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. Immunogenetics 64:615–631. doi:10.1007/s00251-012-0617-1

Deng Q et al (2013) Identification of Mamu-DPA1, Mamu-DQA1, and Mamu-DRA alleles in a cohort of Chinese rhesus macaques. Immunogenetics 65:901–904. doi:10.1007/s00251-013-0736-3

Divis PC et al (2015) Admixture in humans of two divergent Plasmodium knowlesi populations associated with different macaque host species. PLoS Pathog 11:e1004888. doi:10.1371/journal.ppat.1004888

Doxiadis GG et al (2013) Haplotype diversity generated by ancient recombination-like events in the MHC of Indian rhesus macaques. Immunogenetics 65:569–584. doi:10.1007/s00251-013-0707-8

Doxiadis GG, Otting N, de Groot NG, Bontrop RE (2001) Differential evolutionary MHC class II strategies in humans and rhesus macaques: relevance for biomedical studies. Immunological Rev 183:76–85

Doxiadis GG, Rouweler AJ, de Groot NG, Louwerse A, Otting N, Verschoor EJ, Bontrop RE (2006) Extensive sharing of MHC class II alleles between rhesus and cynomolgus macaques. Immunogenetics 58:259–268. doi:10.1007/s00251-006-0083-8

Flynn JL, Gideon HP, Mattila JT, Lin PL (2015) Immunology studies in non-human primate models of tuberculosis. Immunological Rev 264:60–73. doi:10.1111/imr.12258

Gonzalez-Galarza FF et al (2015) Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res 43:D784–D788. doi:10.1093/nar/gku1166

Gough SC, Simmonds MJ (2007) The HLA region and autoimmune disease: associations and mechanisms of action. Curr Genomics 8:453–465. doi:10.2174/138920207783591690

Gyllensten U, Bergstrom T, Josefsson A, Sundvall M, Erlich HA (1996) Rapid allelic diversification and intensified selection at antigen recognition sites of the Mhc class II DPB1 locus during hominoid evolution. Tissue Antigens 47:212–221

Haanstra KG et al (2013) Induction of experimental autoimmune encephalomyelitis with recombinant human myelin oligodendrocyte glycoprotein in incomplete Freund's adjuvant in three non-human primate species. J Neuroimmune Pharmacol 8:1251–1264. doi:10.1007/s11481-013-9487-z

Hernandez RD et al (2007) Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques. Science 316:240–243. doi:10.1126/science.1140462

Jones EY, Fugger L, Strominger JL, Siebold C (2006) MHC class II proteins and disease: a structural perspective. Nat Rev Immunol 6:271–282. doi:10.1038/nri1805

Karl JA, Heimbruch KE, Vriezen CE, Mironczuk CJ, Dudley DM, Wiseman RW, O'Connor DH (2014) Survey of major histocompatibility complex class II diversity in pig-tailed macaques. Immunogenetics 66:613–623. doi:10.1007/s00251-014-0797-y

Kean LS, Singh K, Blazar BR, Larsen CP (2012) Nonhuman primate transplant models finally evolve: detailed immunogenetic analysis creates new models and strengthens the old. Am J Transplant 12:812–819. doi:10.1111/j.1600-6143.2011.03873.x

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. doi:10.1093/molbev/msw054

Ling F et al (2011) Characterization of the major histocompatibility complex class II DOB, DPB1, and DQB1 alleles in cynomolgus macaques of Vietnamese origin. Immunogenetics 63:155–166. doi:10.1007/s00251-010-0498-0

Mooij P et al (2015) Synthetic long peptide booster immunization in rhesus macaques primed with replication-competent NYVAC-C-KC induces a balanced CD4/CD8 T-cell and antibody response

against the conserved regions of HIV-1. J Gen Virol 96:1478–1483. doi:10.1099/vir.0.000074

Mudd PA et al (2012) Vaccine-induced CD8+ T cells control AIDS virus replication. Nature 491:129–133. doi:10.1038/nature11443

O'Connor SL et al (2007) Comprehensive characterization of MHC class II haplotypes in Mauritian cynomolgus macaques. Immunogenetics 59:449–462. doi:10.1007/s00251-007-0209-7

Otting N, Bontrop RE (1995) Evolution of the major histocompatibility complex DPA1 locus in primates. Hum Immunol 42:184–187

Otting N, de Groot N, de Vos-Rouweler AJ, Louwerse A, Doxiadis GG, Bontrop RE (2012) Multilocus definition of MHC haplotypes in pedigreed cynomolgus macaques (Macaca fascicularis). Immunogenetics 64:755–765. doi:10.1007/s00251-012-0632-2

Otting N, de Groot NG, Doxiadis GG, Bontrop RE (2002) Extensive Mhc-DQB variation in humans and non-human primate species. Immunogenetics 54:230–239. doi:10.1007/s00251-002-0461-9

Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG (2015) The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 43:D423–D431. doi:10.1093/nar/gku1161

Sano K et al (2006) Novel cynomolgus macaque MHC-DPB1 polymorphisms in three South-East Asian populations. Tissue Antigens 67:297–306. doi:10.1111/j.1399-0039.2006.00577.x

Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18:6097–6100

Silveira LJ et al (2012) Chronic beryllium disease, HLA-DPB1, and the DP peptide binding groove. J Immunol 189:4014–4023. doi:10.4049/jimmunol.1200798

Smith DG, McDonough J (2005) Mitochondrial DNA variation in Chinese and Indian rhesus macaques (Macaca mulatta). Am J Primatol 65:1–25. doi:10.1002/ajp.20094

t Hart BA, Bogers WM, Haanstra KG, Verreck FA, Kocken CH (2015) The translational value of non-human primates in preclinical research on infection and immunopathology. Eur J Pharmacol 759:69–83. doi:10.1016/j.ejphar.2015.03.023

Thorsby E, Lie BA (2005) HLA associated genetic predisposition to autoimmune diseases: genes involved and possible mechanisms. Transpl Immunol 14:175–182. doi:10.1016/j.trim.2005.03.021

van der Wiel MK, Otting N, de Groot NG, Doxiadis GG, Bontrop RE (2013) The repertoire of MHC class I genes in the common marmoset: evidence for functional plasticity. Immunogenetics 65:841–849. doi:10.1007/s00251-013-0732-7

Xie G et al (2013) Association of granulomatosis with polyangiitis (Wegener's) with HLA-DPB1*04 and SEMA6A gene variants: evidence from genome-wide analysis. Arthritis Rheum 65:2457–2468. doi:10.1002/art.38036