

Differential recombination dynamics within the MHC of macaque species

Nanine de Groot · Gaby G. M. Doxiadis · Nel Otting · Annemiek J. M. de Vos-Rouweler · Ronald E. Bontrop

Received: 4 April 2014 / Accepted: 21 May 2014 / Published online: 17 June 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract A panel of 15 carefully selected microsatellites (short tandem repeats, STRs) has allowed us to study segregation and haplotype stability in various macaque species. The STRs span the major histocompatibility complex (MHC) region and map in more detail from the centromeric part of the *Mhc-A* to the *DR* region. Two large panels of Indian rhesus and Indonesian/Indochinese cynomolgus macaques have been subjected to pedigree analysis, allowing the definition of 161 and 36 different haplotypes and the physical mapping of 10 and 5 recombination sites, respectively. Although most recombination sites within the studied section of the Indian rhesus monkey MHC are situated between the *Mhc-A* and *Mhc-B* regions, the resulting recombination rate for this genomic segment is low and similar to that in humans. In contrast, in Indonesian/Indochinese macaques, two recombination sites, which appear to be absent in rhesus macaques, map between the class III and II regions. As a result, the mean recombination frequency of the core MHC, *Mhc-A* to class II, is higher in Indonesian/Indochinese cynomolgus than in Indian rhesus macaques, but as such is comparable to that in humans. The present communication demonstrates that the

dynamics of recombination ‘hot/cold spots’ in the MHC, as well as their frequencies, may differ substantially between highly related macaque species.

Keywords MHC · Primates · Recombination · Microsatellites · Evolution

Introduction

The class I and II molecules encoded by the major histocompatibility complex (MHC) play a pivotal role in activating various adaptive immune-related reactions by presenting peptides to CD8⁺ or CD4⁺T cells, respectively. The MHC of primate species is one of the most gene-dense chromosomal regions, and the genes encoding peptide-presenting molecules are often characterized by a high degree of allelic heterogeneity (polymorphism) and copy number variation (diversity). Therefore, the MHC of two relevant model species, which are used to study human biology and disease—namely, the rhesus (*Macaca mulatta*) and the Indonesian/Indochinese cynomolgus macaque (*Macaca fascicularis*)—has been thoroughly investigated (Blancher et al. 2012b; Bontrop and Watkins 2005; Doxiadis et al. 2003, 2013; Karl et al. 2013; Li et al. 2012a, b; Liu et al. 2013; Mitchell et al. 2012; Otting et al. 2005, 2012; Zhang et al. 2012). Orthologs of the human (HLA) class II genes are also present in macaques, and most of them show a high degree of allelic variation (de Groot et al. 2012). However, some differences have been observed. For example, copy number variation in *DRB* genes is far more profound in macaques than in humans, whereas human populations seem to compensate for this by displaying extensive allelic heterogeneity. Equivalents of the *HLA-A* and *HLA-B* genes are present in macaques as well. The *Mhc-C* gene emerged in hominids (human/great apes) after the split from small apes and Old World monkeys and is, therefore, not

Nanine de Groot and Gaby G. M. Doxiadis contributed equally to the article.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-014-0783-4) contains supplementary material, which is available to authorized users.

N. de Groot · G. G. M. Doxiadis (✉) · N. Otting · A. J. M. de Vos-Rouweler · R. E. Bontrop
Department of Comparative Genetics and Refinement, Biomedical Primate Research Centre, Lange Kleiweg 161, 2288 GJ Rijswijk, The Netherlands
e-mail: doxiadis@bprc.nl

R. E. Bontrop
Department of Theoretical Biology and Bioinformatics, Utrecht University, 3584 CH Utrecht, The Netherlands

present in macaques. However, this is compensated for by the fact that in macaques the *Mhc-A* and *Mhc-B* loci have been subjected to duplications and show copy number variation. On top of that, these *A* and *B* genes display transcription level polymorphisms as well. To simplify MHC-typing procedures, additional technologies to Sanger and next-generation sequencing (NGS) have been developed. Microsatellite loci or short tandem repeats (STRs)—repetitive sequence motifs of 2 to 6 bp—have been particularly useful to conduct disease association and population studies. Microsatellites have been used for molecular analysis of the HLA region, thanks to their characteristics such as a high degree of length polymorphism, chromosomal density, and Mendelian inheritance (Cullen et al. 2002; Foissac et al. 1997; Malkki et al. 2005; Martin et al. 1995a, b), and a database, *dbMHC*, is available that records all progress (Gourraud et al. 2006). Microsatellites have been shown to be excellent tools for the quick and robust *DRB*, *A*, and *B* typing of macaques (Bonhomme et al. 2007; Doxiadis et al. 2007, 2009, 2011; Mee et al. 2009; Penedo et al. 2005; Wiseman et al. 2007; Wojcechowskyj et al. 2007). To refine further biomedical studies, additional microsatellite markers have been used to analyse the extended MHC region of macaques (Aarnink et al. 2013; Mitchell et al. 2012).

In humans, microsatellite markers have been used extensively to define recombination rates within the MHC, and it is proven that meiotic recombination does not occur randomly but seems to be restricted to specific chromosomal regions. Recombination frequencies defined by microsatellite analysis revealed that there are HLA regions—‘cold spots’—where no recombination is observed. In contrast, a lack of linkage between *TAP1* and *TAP2* alleles, for example, suggested the presence of a ‘hotspot’ (Martin et al. 1995a, b). More recent studies using sperm and SNP typing confirmed high linkage disequilibrium (haplotype blocks), which are often interrupted by recombination hotspots (Cullen et al. 2002; de Bakker et al. 2006; Walsh et al. 2003). In the present study, microsatellite analyses of the MHC of Indian rhesus and Indonesian/Indochinese cynomolgus macaques allowed us to map several recombination sites and suggest the existence of differential recombination rates and localizations of recombination ‘hot/warm’ and ‘cold/low’ spots in these two species.

Materials and methods

Materials

In the present study, 321 Indian rhesus macaques and 113 cynomolgus macaques of Indonesian/Indochinese origin were analysed. The animals have been housed at BPRC for more than five generations and are part of a self-sustaining breeding

colony. Genomic DNA of the macaques was extracted from EDTA blood samples or from immortalized B cell lines, using a standard salting out procedure.

STR genotyping

In the non-duplicated parts of the core MHC (*Mhc-A - DPB1*), human microsatellites were chosen, using the *dbMHC* database (<http://www.ncbi.nlm.nih.gov/gv/mhc/main.cgi?cmd=init>). For >50 microsatellites, primers were developed according to the physical map of the rhesus MHC (Daza-Vamenta et al. 2004; Shiina et al. 2006), and the markers were tested for their suitability. Based on the degree of polymorphisms, heterozygosity, and their localization, 15 markers were chosen (Fig. 1, Table 1). The STRs were amplified by using a universal, fluorescently VIC-labeled primer adapter sequence (Applied Biosystems, Foster City, USA) together with the adapter sequence+unique 5′ primer and a unique 3′ primer (Invitrogen (Paisley, Scotland).

Most primers could be used in a multiplex PCR reaction (Suppl. Table 1; primers 1 and 11, primers 5 and 13, primers 7 and 9, primers 8 and 15, and primers 12 and 14).

The PCR reaction was performed in a 25- μ l reaction volume containing 1 unit of *Taq* polymerase (Invitrogen, Paisley, Scotland) with 0.1 μ M of the unlabeled forward primer, 1.0 μ M of the reverse primer, 1.0 μ M of the VIC-labeled forward adapter primer, 2.5 mM MgCl₂, 0.2 mM of each dNTP, 1 \times PCR buffer II (Invitrogen, Paisley, Scotland), and 10–25 ng DNA. For the multiplex PCR reactions, a second set of primers was added with the same concentrations.

The cycling parameters were a 5-min 94 °C initial denaturation step, followed by 5 cycles of 1 min at 94 °C, 45 s at 63 °C, and 45 s at 72 °C and then 25 cycles followed with 45 s at 94 °C, 30 s at 63 °C, and 45 s at 72 °C. A final extension step was performed at 72 °C for 30 min. The amplified DNA was prepared for genotyping according to the manufacturer’s guidelines and was analysed on the ABI 3130 genetic analyser (Applied Biosystems) with the GeneMapper software (Applied Biosystems).

Allele frequency analysis

Allele frequency analyses of the macaque panels (Indian rhesus macaques, $N=321$; Indonesian/Indochinese cynomolgus macaques, $N=113$) were performed with the program Cervus 3.03 including 15 STR markers. For each locus, the number of alleles is shown together with the observed (HObs) and expected (HExp) heterozygosity and with the mean number of alleles and mean Hexp (Table 1).

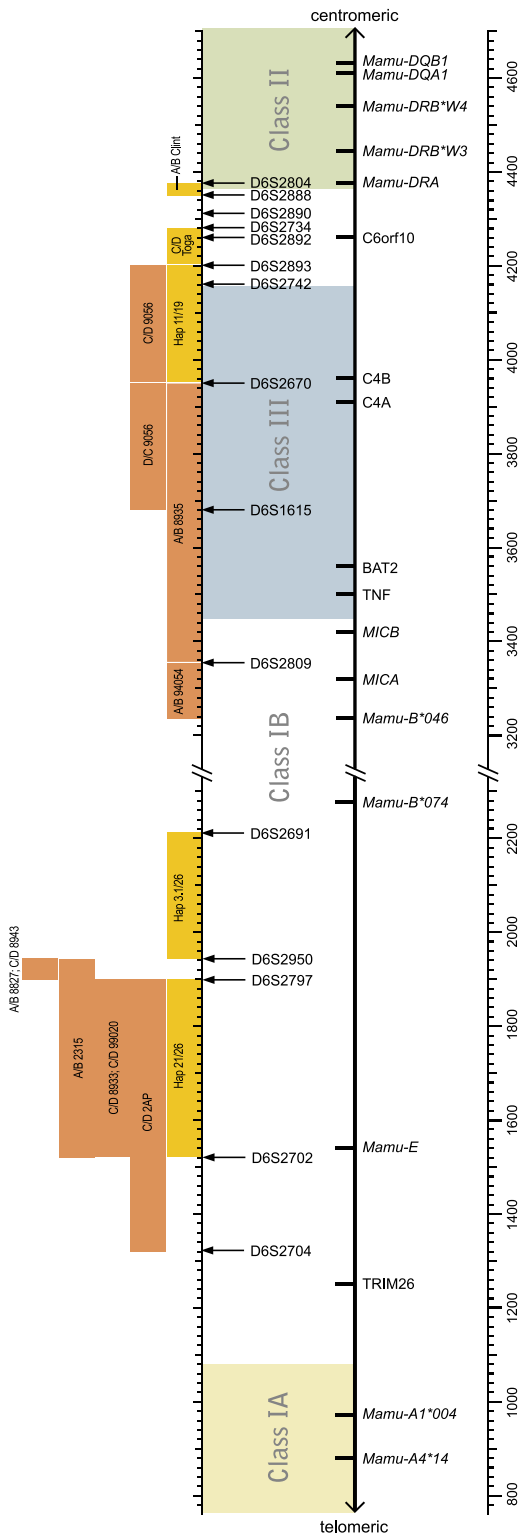


Fig. 1 Mapping of recombination sites within the MHC of Indian rhesus and Indonesian/Indochinese cynomolgus macaques. The physical map is based on the work of Daza-Vamenta et al. (2004). The *Mamu-A*, *Mamu-B*, and *Mamu-DRB* alleles are named according to rules published in the IPD-MHC NHP database (de Groot et al. 2012). Recombination sites of rhesus macaques and cynomolgus macaques are color-coded (rhesus macaque: brown; cynomolgus macaque: yellow)

Simplification of *Mamu/Mafa-A*, *-B*, and *-DRB* haplotypes by “short” names

Complete *Mamu/Mafa-A*, *-B*, and *-DRB* haplotypes have been abbreviated to short names, mainly as described previously (Doxiadis et al. 2013; Karl et al. 2013). Briefly, the short names of the *Mamu/Mafa-A* and/or *Mamu/Mafa-B* haplotypes reflect one of the ‘major’ *A* or *B* transcripts (Karl et al. 2013). *Mamu/Mafa-DRB* haplotypes are named based on different region configurations ordered by increasing numbers of genes per haplotype (Doxiadis et al. 2013) (Suppl. Table 2)

Founder haplotype definition and naming

Founder haplotypes have been defined as described earlier (Doxiadis et al. 2013). Haplotypes of Indian rhesus macaques were named according to the founder animal, with prefixes A and B indicating a founder male and prefixes C and D indicating a founder female. Haplotypes of Indonesian/Indochinese cynomolgus macaques, of which the founder animals could not always be defined unambiguously, were numbered as described earlier (Otting et al. 2012). Recombinant haplotypes were named according to the animal in which the cross-over event took place, with the prefixes A/B for a male and C/D for a female animal (e.g. C/D 2AP) (Figs. 1, 2, 3, and 4). If the recombination event could not be pinpointed to a certain animal, which is the case with some Indonesian/Indochinese cynomolgus monkeys because of unknown sires, the recombinant haplotype is named according to both donor haplotypes (e.g. Hap 11/19) (Figs. 1 and 4).

Results

Microsatellite analysis of the rhesus and cynomolgus macaque MHC region

Based on recorded microsatellite profiles in humans (Gourraud et al. 2006) (*dbMHC/NCBI*) and on the physical map of a heterozygous Indian rhesus macaque (Daza-Vamenta et al. 2004; Shiina et al. 2006), we have identified 15 different microsatellites covering the core MHC from the centromeric end of the class I *Mhc-A* to the *DRB* region of class II. Most of these microsatellites are mapping within the class III region and the adjacent section surrounding the *C6orf10* locus (Fig. 1). Although the macaque *Mhc-A* and *Mhc-B* as well as the *Mhc-DRB* regions were subjected to contraction and expansion processes during their evolutionary past (Doxiadis et al. 2008; Kulski et al. 2004; Slierendregt et al. 1994), the genomic organization of the gene-rich regions in between has remained more or less intact (Daza-Vamenta et al. 2004; Watanabe et al. 2007; Yan et al. 2011). Therefore, the

Table 1 Frequencies of microsatellites in Indian rhesus and Indonesian/Indochinese cynomolgus macaques

Locus	Rhesus macaques (N=321)			Cynomolgus macaques (N=113)		
	# of alleles	HObs	HExp	# of alleles	HObs	HExp
D6S2704	8	0.73	0.80	10	0.73	0.85
D6S2702	9	0.72	0.77	10	0.83	0.87
D6S2797	9	0.74	0.81	10	0.87	0.87
D6S2950	23	0.77	0.89	18	0.65	0.89
D6S2691	28	0.85	0.94	25	0.84	0.94
D6S2809	16	0.84	0.85	13	0.88	0.89
D6S1615	5	0.70	0.76	6	0.76	0.72
D6S2670	21	0.84	0.88	14	0.88	0.87
D6S2742	4	0.63	0.68	2	0.02	0.02
D6S2893	17	0.67	0.80	13	0.90	0.89
D6S2892	6	0.57	0.60	9	0.80	0.78
D6S2734	6	0.44	0.71	7	0.56	0.70
D6S2890	9	0.20	0.81	15	0.34	0.91
D6S2888	7	0.43	0.64	12	0.72	0.85
D6S2804	3	0.54	0.52	4	0.60	0.67
Mean #	11.4		0.76	11.2		0.78

localization of the selected microsatellites in both macaque species is expected to be similar to that in humans.

In the first instance, microsatellite profiles of the Indian rhesus and Indonesian/Indochinese cynomolgus macaque panels were assayed to determine to what extent these STRs are polymorphic and can provide relevant information. The animals possess the founder MHC (*Mamu/Mafa-A*, *Mamu/Mafa-B*, *Mamu/Mafa-DRB*, *Mamu/Mafa-DQ*, *Mamu/Mafa-DP*) haplotypes, which have been described previously

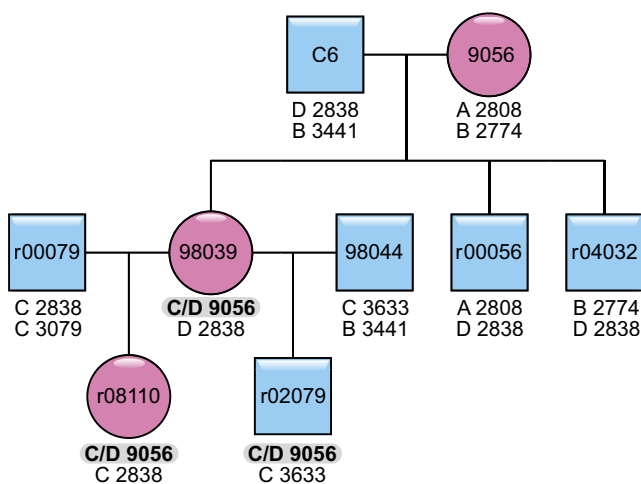


Fig. 2 Example of a recombination event, which segregates in the family. The recombination event occurred in dame 9056 and is therefore called C/D 9056. It has been observed in the F1 and F2 generation in animals 98039 and r08110. Blue squares indicate males, red circles indicate females

(Doxiadis et al. 2013; Otting et al. 2012). Most microsatellites display length polymorphism, and from 4 to 28 alleles with an expected heterozygosity (HExp) of >0.60 and a mean heterozygosity of >0.76 of all STRs, were detected (Table 1). Two markers do not fit in this schedule. Nevertheless, one of them, D6S2804, which displays a low level of polymorphism, has a HExp of >0.5 in Indian rhesus and 0.6 in Indonesian/Indochinese cynomolgus macaques and is therefore considered to be a useful marker. The second STR, D6S2742, displays polymorphism in our rhesus macaque panel but is dimorphic in Indonesian/Indochinese cynomolgus macaques, with one of the two alleles observed at a very low frequency (Table 1). The number of founder animals tested is far higher for the rhesus ($N=137$) as compared to the Indonesian/Indochinese cynomolgus macaque colony ($N\sim 30$), and one may expect the number of STR alleles to be higher in the first species. Indeed, several microsatellite markers show more allelic variation in the Indian rhesus than in the Indonesian/Indochinese cynomolgus macaque panel. However, there are two microsatellites, which display far more polymorphism in Indonesian/Indochinese cynomolgus macaques (D6S2890 and D6S2888, Table 1), and the mean number of their alleles and expected heterozygosity levels is comparable. This finding is in agreement with earlier observations that Indian rhesus macaques show on average a lower degree of polymorphism for their MHC markers than other macaque species or rhesus macaques of other origins do (Doxiadis et al. 2013).

MHC haplotype definition by microsatellite profiling

Based on STR profiles, 166 MHC founder haplotypes have been defined in the Indian rhesus macaque panel (see ‘Materials and methods’; Suppl. Table 3) (Doxiadis et al. 2013). Every STR founder haplotype profile was recorded in at least two related animals except those created by a recombination event. The profiles are listed together with their respective multi-locus MHC typing. Most founder STR profiles—namely, 161 out of 166—are unique. The few profiles, which are identical, most likely belong to related founder animals, originating from one breeding centre (Suppl. Table 3, black-bordered haplotypes). In addition, there are founder haplotypes that are identical for all their known MHC specificities, but can be distinguished based on STR patterns. Five examples are provided (Fig. 5).

Furthermore, *DRB* identical haplotypes are observed (1) to be mainly accompanied by the same *DQA1/DQB1* pair and (2) to segregate often with identical STR alleles, which map between the class III and II regions (Suppl. Table 3, colour-coded characters). The first observation confirms the tight linkage of *DR* and *DQ*, which is documented for several primate species (Begovich et al. 1992; Cullen et al. 2002; Doxiadis et al. 2000; Martin et al. 1995a, b). The second

founder hapl	C 3079	A/B 8935	D Yusa	B 1435	C/D 8933	D 499	A 2315	A/B 2315	B 2315
A	A001	A004	A004	A008	A006	A006	A001	A001	A002a
D6S2704	160	167	167	167	162	162	160	160	156
D6S2702	267	275	275	277	267	267	267	267	275
D6S2797	184	184	184	181	181	187	184	n.d.	187
D6S2950	257	253	253	252	252	268	272	257	257
D6S2691	307	274	274	303	303	307	258	290	290
B	B001a	B017a	B017a	B001a	B001a	B043a	B017a	B024a	B024a
D6S2809	138	149	149	138	138	144	151	128	128
D6S1615	122	122	122	122	122	118	118	117	117
D6S2670	176	176	191	172	172	181	178	189	189
D6S2742	201	201	203	203	203	191	201	203	201
D6S2893	199	199	199	218	218	null	215	200	200
D6S2892	213	213	213	213	213	213	214	213	213
D6S2734	209	209	209	null	null	209	202	209	209
D6S2890	201	201	185	null	null	null	209	null	null
D6S2888	146	146	146	132	132	132	146	132	132
D6S2804	239	239	251	251	251	239	239	239	239
DRB	DRB10	DRB10	DRB03a	DRB03a	DRB03a	DRB04	DRB01a	DRB04	DRB04
DQA1*	26:03	26:03	26:01	26:01	26:01	01:04	26:02	01:04	01:04
DQB1*	15:01	15:01	18:01	18:01	18:01	06:01	18:11	06:01	06:01

founder hapl	B 599	C/D 8943	A 600	A/B 94054	A 2957	A/B 8827	D 1472
A	A008	A004	A004	A008	A008	A026	A026
D6S2704	167	156	156	167	167	160	160
D6S2702	273	271	271	267	267	267	267
D6S2797	184	178	178	181	181	196	196
D6S2950	272	272	276	246	246	246	257
D6S2691	300	300	307	282	282	282	249
B	B048	B048	B024a	B047a	B047a	B047a	B012b
D6S2809	138	138	128	128	149	149	153
D6S1615	118	118	118	118	118	118	121
D6S2670	189	189	189	189	172	172	172
D6S2742	199	199	199	199	203	203	201
D6S2893	200	200	200	200	199	199	217
D6S2892	213	213	213	213	216	216	213
D6S2734	209	209	209	209	211	211	209
D6S2890	null	null	null	null	null	null	201
D6S2888	132	132	132	132	132	132	146
D6S2804	239	239	239	239	251	251	239
DRB	DRB04	DRB04	DRB04	DRB04	DRB03e	DRB03e	DRB01b
DQA1*	01:04	01:04	01:04	01:04	26:01	26:01	23:01/03
DQB1*	06:01	06:01	06:01	06:01	18:01	18:01	18:03

founder hapl	A 2808	D/C 9056	C/D 9056	B 2774	B 2775	C/D 99020	C 494	C/D 2AP	A 381
A	A019	A008	A019	A008	A012	A012	A019	A008	A008
D6S2704	167	167	167	167	158	158	156	167	167
D6S2702	275	269	275	269	275	275	277	277	277
D6S2797	184	193	184	193	184	187	187	187	193
D6S2950	280	246	280	246	269	253	253	253	234
D6S2691	290	262	290	262	279	266	266	266	274
B	B055	B045a	B055	B045a	B008	B008	B008	B008	B017a
D6S2809	149	134	149	134	136	136	136	136	136
D6S1615	117	122	117	122	118	118	118	118	118
D6S2670	193	193	193	191	178	194	194	194	191
D6S2742	201	201	201	201	203	201	201	201	199
D6S2893	215	215	217	217	199	215	215	215	null
D6S2892	213	213	205	205	205	214	214	214	213
D6S2734	209	209	202	202	211	202	202	202	209
D6S2890	185	185	194	194	201	209	209	209	null
D6S2888	146	146	146	146	146	146	146	146	132
D6S2804	251	251	239	239	239	239	239	239	251
DRB	DRB03a	DRB03a	DRB05	DRB05	DRB10	DRB01a	DRB01a	DRB01a	DRB03a
DQA1*	26:01	26:01	24:02	24:02	26:03	26:02	26:02	26:02	26:01
DQB1*	18:01	18:01	18:08	18:08	15:01	18:11	18:11	18:11	18:01

Fig. 3 Recombinant and donor MHC haplotypes of Indian rhesus macaques. Donor haplotypes are color-coded, thus showing the recombination site in the recombinant haplotype. If a microsatellite is not informative for a certain recombination, this is indicated by stripes

	Hap 11	Hap 11/19	Hap 19	A/B Clint	Hap 28	C/D Toga	Hap 24	Hap 21	Hap 21/26	Hap 26	Hap 3.1/26	Hap 3.1
A	A031b	A031b	A063a	A063a	A071a	A071a	A066	A064	A064	A069	A069	A007
D6S2704	149	149	166	166	156	156	156	156	156	153	153	160
D6S2702	277	277	271	271	277	277	281	265	265	279	279	265
D6S2797	178	178	187	187	193	193	185	178	184	184	184	184
D6S2950	269	269	null	null	253	253	277	264	253	253	253	260
D6S2691	270	270	276	276	267	267	252	349	320	320	270	270
B	B027b	B027b	B046a	B046a	B013	B013	B046c	B025	B044a	B044a	B002b	B002b
D6S2809	128	128	136	136	144	144	153	153	144	144	149	149
D6S1615	118	118	123	123	121	121	121	121	120	120	119	119
D6S2670	189	189	193	193	170	170	171	189	178	178	195	195
D6S2742	201	201	201	201	201	201	201	201	201	201	201	201
D6S2893	212	215	215	215	200	200	198	206	196	196	200	200
D6S2892	210	216	216	216	213	213	213	211	208	208	214	214
D6S2734	null	null	null	null	185	209	209	200	202	202	198	198
D6S2890	192	205	205	205	211	null	null	null	192	192	null	null
D6S2888	161	142	142	142	142	null	null	149	179	179	132	132
D6S2804	243	239	239	243	251	239	239	null	239	239	251	251
DRB	DRB19a	DRB15	DRB15	DRB19a	DRB09	DRB10	DRB10	DRB20	DRB17	DRB17	DRB04	DRB04
DQA1*	01:07	24:03	24:03	01:07	05:08	23:01	23:01	01:06	24:06	24:06	23:01	23:01
DQB1*	06:08	18:01:01	18:01:01	06:08	17:02:02	18:04	18:04	06:11	18:06	18:06	18:04	18:04

Fig. 4 Recombinant and donor MHC haplotypes of Indonesian/Indochinese cynomolgus macaques. Donor haplotypes are color-coded, thus showing the recombination site in the recombinant haplotype. If a microsatellite is not informative for a certain recombination, this is indicated by stripes

observation, however, may indicate low mutation rates of these C6orf10-surrounding markers in rhesus macaques and/or a tight linkage of the DR region not only with centromerically but also with telomerically located chromosome segments. This observation is notable for DRB04 haplotypes that are often accompanied by one specific D6S2742–D6S2804 pattern (Suppl. Table 3; D6S2742-199, D6S2893-200, D6S2892-213, D6S2734-209, D6S2890-null, D6S2888-132, D6S2804-239).

Additionally, length variation of one STR within a certain haplotype can also be observed in rhesus macaques. An example is provided by haplotype A/B 8827; as can be seen, the length of D6S2691 has changed from 282 to 278, as observed by segregation within its family (Fig. 5). Thus, a

new STR allele has been ‘born’ and its birth could be recorded.

In our Indonesian/Indochinese cynomolgus macaque panel (N=113), a total of 36 unique MHC STR profiles have been defined representing founder haplotypes. As for the Indian rhesus macaques, each STR profile has been detected in at least two related animals. All STR haplotypes have also been supplemented with relevant MHC typing (Suppl. Table 4), as published earlier (Otting et al. 2012). As in the rhesus macaque panel, we observed the appearance of a new STR allele (Suppl. Table 4, Hap 3 and 3.1), which differs in the length of marker D6S2691, and thus the potential birth of a new haplotype.

founder	A 2777	D 2849	B 2794	C 2378	A 1033	D 584	B 600	D 3837	A/B 8827	A/B 8827.1
A	A002a	A002a	A004	A004	A004	A004	A008	A008	A026	A026
D6S2704	156	156	158	158	158	156	167	167	160	160
D6S2702	271	271	273	271	275	271	275	275	267	267
D6S2797	178	178	184	178	184	178	184	184	196	196
D6S2950	272	269	257	246	268	246	272	276	246	246
D6S2691	286	282	292	245	262	245	282	286	282	278
B	B001a	B001a	B001a	B001a	B012b	B012b	B047a	B047a	B047a	B047a
D6S2809	138	138	138	138	149	149	149	149	149	149
D6S1615	121	121	122	117	117	122	117	117	118	118
D6S2670	193	193	172	181	185	191	189	176	172	172
D6S2741	201	201	203	201	201	199	201	203	203	203
D6S2893	217	null	199	215	200	200	218	200	199	199
D6S2892	211	211	216	205	213	213	216	216	216	216
D6S2734	202	202	211	null	209	209	211	211	211	211
D6S2890	null	null	null	209	null	null	null	null	null	null
D6S1888	null	null	132	146	132	132	null	null	132	132
D6S2804	239	239	239	239	239	239	251	251	251	251
DRB	DRB02	DRB02	DRB01a	DRB01a	DRB04	DRB04	DRB03f	DRB03f	DRB03e	DRB03e
DQA1*	24:01	24:01	26:01	26:01	01:04	01:04	26:02	26:02	26:01	26:01
DQB1*	18:10	18:10	18:01	18:01	06:01	06:01	18:09	18:09	18:01	18:01

Fig. 5 Examples of Mamu-A/B/DRB/DQ-identical haplotypes that can be differentiated by microsatellites. Microsatellites that show length variations are indicated by grey boxes

Mapping recombination sites within the MHC of macaques

Within the Indian rhesus macaque haplotypes analysed ($N=648$, Suppl. Table 3), we were able to define ten distinct recombination events, and the corresponding haplotypes segregate as stable entities in families (Figs. 2 and 3). According to the localization of the microsatellites, the recombination sites can be plotted on the physical map (Fig. 1). Most of the recombination events (i.e. six out of ten) took place between the *Mhc-A* and *Mhc-B* regions. More precisely, they map between the marker D6S2704, next to TRIM26, and D6S2950, which is situated in the middle between *Mamu-E* and the first *B* locus, *Mamu-B*074:01*. The other four crossing-over events are spread from the centromerical end of the *B* region to marker D6S2893, which is situated between class III and II, telomerically of C6orf10. Notably, two independent recombination events trace back to one female, 9056. One of the recombination events, C/D 9056, segregates for at least three generations, whereas the second one, D/C 9056, happened about 10 years later and has been observed in the first generation only. The second crossing-over event is nearly the complementary of the first one, and they can be mapped to both sides of the STR D6S2670, next to the *C4B* locus (Figs. 1 and 3).

A total of 225 haplotypes have been defined within our Indonesian/Indochinese cynomolgus macaque colony, and five crossing-over events have been observed (Fig. 4). Two recombinations could be designated to the animal in which the recombination took place, whereas this could not be determined for the other three recombinant haplotypes due to unknown sires. Since there are no STR length variations observed in the recombinant haplotypes in comparison to those of the ‘donor haplotypes’, it appears highly probable that all recombination events have happened within our colony, thus within the last five to six generations. Two recombination events are located between *Mhc-A* and *Mhc-B* regions as observed for most of the crossing-over events in the rhesus macaque. One of the crossing-over events, Hap 21/26, maps between markers D6S2702 and D6S2697 and the second one, Hap 3.1/26, between markers D6S2950 and D6S2691 (Figs. 1 and 4). Comparable to the rhesus macaques, one haplotype (haplotype 26) is involved twice in recombination events. However, in contrast to the observation in rhesus macaque 9056, the crossing-over events in the Indonesian/Indochinese cynomolgus haplotype 26 are not complementary and also—most probably—did not happen in the same animal. A third recombination, Hap 11/19, is identical to one in rhesus macaques, namely, C/D 9056. However, in Indonesian/Indochinese cynomolgus macaques, two recombinations, C/D Toga and A/B Clint, took place centromerically to marker D6S2893 within the region between class III and class II. More precisely, the recombination event in C/D Toga is localized between D6S2893 and D6S2734 surrounding C6orf10.

The other crossing-over event can be pinpointed between two markers, D6S2804 and D6S2888, in A/B Clint (Figs. 1 and 4). To our knowledge, no recombination has been observed in this part of the chromosome in Indian rhesus macaques.

Discussion

Profiling with 15 carefully selected STR markers in addition to typing for multiple MHC loci allowed us to define different haplotypes in great detail in large panels of Indian rhesus and Indonesian/Indochinese cynomolgus macaques. In such a way, even haplotypes that are identical for their ‘classical’ MHC alleles can be differentiated. Such a fine mapping may help to determine genes or genetic factors that are associated with susceptibility or resistance to diseases that are otherwise disguised by strong levels of linkage disequilibrium (LD) (reviewed in Trowsdale and Knight 2013).

In addition, the ‘birth’ of new haplotypes in both macaque colonies was recorded, as evidenced by the length change in one of the STRs (D6S2691). This marker represents a CCTT repeat of often more than 30 units and is therefore characterized by high levels of length variation. Although recent studies have shown that every microsatellite is unique in terms of its mutational variation, a relationship between motif length and mutagenesis has been observed (Eckert and Hile 2009). Thus, it is likely that a highly polymorphic microsatellite such as D6S2691 is prone to frequent mutagenesis. In the rhesus macaque colony, such an STR mutation was recorded; two generations after a recombination event had established a hybrid haplotype (Fig. 3, Suppl. Table 3: haplotypes A/B 8227 and A/B 8227.1). In the Indonesian/Indochinese cynomolgus macaque panel, the opposite event was observed. A new haplotype was first initiated by the birth of a new allele for marker D62691 (Suppl. Table 4: haplotypes 3 and 3.1). The subsequent recombination event happened one or two generations later (Fig. 4). Since the presence of specific sequence motifs such as $GT_{>12}$ is significantly associated with recombination hotspots in the MHC, an STR-like D6S2691 may represent the substrate for recombination events (Cullen et al. 2002). Therefore, STR typing, especially for those with high repeat numbers, seems to represent an excellent manner for characterizing and studying the evolution of different DNA segments/haplotypes of highly variable regions such as sections of the classical MHC genes.

The analysis of 2,270 haplotypes in rhesus macaques (Doxiadis et al. 2013) and the data provided in this study show that Indian rhesus macaques have a comparable recombination rate for classes *A* and *B* (~0.26%), as observed in humans (~0.21 %) (Cullen et al. 2002; Martin et al. 1995a, b). Although the low number of informative meioses observed in a previous study ($N=332$) (Otting et al. 2012) and in the present analysis may not allow statistically relevant

conclusions, the recombination frequencies of the class I region in Indonesian/Indochinese cynomolgus macaques (~0.6 %) are apparently higher than those observed in humans and rhesus macaques. The difference between both macaque species becomes even more pronounced when recombination rates between the class III and II regions are taken into account. In this section of the MHC, no recombination has been documented in the rhesus macaque, and therefore, this part of the chromosome seems to represent a recombination desert. In contrast, in Indonesian/Indochinese cynomolgus macaques, two crossing-over events have been detected in this chromosomal section. Thus, in contrast to a recombination desert, which maps between *C6orf10* and *DRA* in rhesus macaques, Indonesian/Indochinese cynomolgus monkeys apparently have one or more recombination warm/hot spots in this part of the genome. As a consequence, the mean recombination frequency of the core MHC, *Mhc-A* to class II (3.3 Mb), is also higher in our Indonesian/Indochinese cynomolgus macaque panel (1.5 % or 0.45 cM/Mb) than in Indian rhesus macaques (0.44 % or 0.13 cM/Mb), but is comparable to humans (0.49 cM/Mb) (Cullen et al. 2002).

Although at first glance the different recombination frequencies and cold/hotspots of the two related macaque species seem remarkable, comparable results have been reported in humans. Single-sperm typing of several humans has demonstrated that the distribution of recombination events may differ significantly between individuals (Cullen et al. 2002). Analyses of parent–offspring combinations of four human populations showed that linkage disequilibrium and recombination hotspots are population and haplotype specific (Ahmad et al. 2003; de Bakker et al. 2006). These results have been confirmed in a recent study comparing recombinations in the MHC of Asian, European, and African populations, which showed that >50 % of the recombination sites are unique for a single population (Lam et al. 2013). Taking into account that African and Europeans shared a common ancestor <250,000 years ago, whereas rhesus and cynomolgus macaque lineages separated ~1.3 million years ago, it seems plausible that substantial differences in recombination rates and sites are present in both species. It is noted that Mauritian cynomolgus macaques have a lower average recombination frequency (0.4–0.8 %) than the Indonesian/Indochinese cynomolgus macaques in our panel (1.5 %) (Blancher et al. 2012a). In this regard, the Mauritian cynomolgus monkeys resemble the Indian rhesus macaques. The low recombination frequencies in both populations are in accordance with a longer-ranged LD that seems to be more pronounced in bottlenecked populations, as has been documented for Mauritian cynomolgus and Indian rhesus macaques (Bonhomme et al. 2007; Doxiadis et al. 2003; Hernandez et al. 2007). However, MHC diversity after a bottle-neck may be partially restored by recombination events creating haplotypes that are composed of different MHC segments (Doxiadis et al. 2013; Karl et al.

2013; Wiseman et al. 2007). Similar observations have been made for West African chimpanzees, which experienced a selective sweep (de Groot et al. 2008). Certain forms of balancing selection may favour the selection of particular combinations of allotypes, thus actively selecting for recombinant haplotypes with novel combinations (de Bakker et al. 2006; Traherne et al. 2006). Since the Indonesian/Indochinese cynomolgus macaque colony studied was established with a relatively small number of founder animals (~30), it might be that such selection forces have been operative. At this stage, we do not understand whether this is due to pathogen-related pressure or whether this kind of selection takes place during stages of reproduction. The MHC in primates is considered to play a role in mate choice and reproduction (reviewed in Setchell and Huchard 2010; Ziegler et al. 2010). These ideas are supported by the observations that mate selection in most eukaryotic species is driven by some kind of self-/non-self-perception. Olfactory genes are localized in the extended MHC region, but many genes within the core MHC appear to play a role in reproduction, too (e.g. *POU5F1*, *BAT3*) (reviewed in Horton et al. 2004; Ziegler et al. 2010). Therefore, this panel of microsatellites may also help in understanding the role of the MHC region in reproductive biology and its success.

Acknowledgments The authors wish to thank Donna Devine for carefully editing this article and Henk van Westbroek for preparing the figures. This work was supported in part by the National Institute of Health projects (NIH/NIAID HHSN272201100013C and 5R24RR016038-05 (CFA:03.9389)).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Aarnink A, Bonhomme M, Blancher A (2013) Positive selection in the major histocompatibility complex class III region of cynomolgus macaques (*Macaca fascicularis*) of the Philippines origin. *Tissue Antigens* 81:12–18
- Ahmad T, Neville M, Marshall SE, Armuzzi A, Mulcahy-Hawes K, Crawshaw J, Sato H, Ling KL, Barnardo M, Goldthorpe S, Walton R, Bunce M, Jewell DP, Welsh KI (2003) Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Hum Mol Genet* 12:647–656
- Begovich AB, McClure GR, Suraj VC, Helmuth RC, Fildes N, Bugawan TL, Erlich HA, Klitz W (1992) Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J Immunol* 148:249–258
- Blancher A, Aarnink A, Savy N, Takahata N (2012a) Use of cumulative Poisson probability distribution as an estimator of the recombination

- rate in an expanding population: example of the *Macaca fascicularis* major histocompatibility complex. *G3 (Bethesda)* 2:123–130
- Blancher A, Aarnink A, Tanaka K, Ota M, Inoko H, Yamanaka H, Nakagawa H, Apoil PA, Shiina T (2012b) Study of cynomolgus monkey (*Macaca fascicularis*) Mhc DRB gene polymorphism in four populations. *Immunogenetics* 64:605–614
- Bonhomme M, Blancher A, Jalil MF, Crouau-Roy B (2007) Factors shaping genetic variation in the MHC of natural non-human primate populations. *Tissue Antigens* 70:398–411
- Bontrop RE, Watkins DI (2005) MHC polymorphism: AIDS susceptibility in non-human primates. *Trends Immunol* 26:227–233
- Cullen M, Perfetto SP, Klitz W, Nelson G, Carrington M (2002) High-resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am J Hum Genet* 71:759–776
- Daza-Vamenta R, Glusman G, Rowen L, Guthrie B, Geraghty DE (2004) Genetic divergence of the rhesus macaque major histocompatibility complex. *Genome Res* 14:1501–1515
- de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, Morrison J, Richardson A, Walsh EC, Gao X, Galver L, Hart J, Hafler DA, Pericak-Vance M, Todd JA, Daly MJ et al (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet* 38:1166–1172
- de Groot NG, Heijmans CM, de Groot N, Otting N, de Vos-Rouweller AJ, Remarque EJ, Bonhomme M, Doxiadis GG, Crouau-Roy B, Bontrop RE (2008) Pinpointing a selective sweep to the chimpanzee MHC class I region by comparative genomics. *Mol Ecol* 17:2074–2088
- de Groot NG, Otting N, Robinson J, Blancher A, Lafont BA, Marsh SG, O'Connor DH, Shiina T, Walter L, Watkins DI, Bontrop RE (2012) Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. *Immunogenetics* 64:615–631
- Doxiadis GG, Otting N, de Groot NG, Noort R, Bontrop RE (2000) Unprecedented polymorphism of Mhc-DRB region configurations in rhesus macaques. *J Immunol* 164:3193–3199
- Doxiadis GG, Otting N, de Groot NG, de Groot N, Rouweller AJ, Noort R, Verschoor EJ, Bontrop RE (2003) Evolutionary stability of MHC class II haplotypes in diverse rhesus macaque populations. *Immunogenetics* 55:540–551
- 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 U S A* 104:8907–8912
- Doxiadis GG, de Groot N, Bontrop RE (2008) Impact of endogenous intronic retroviruses on major histocompatibility complex class II diversity and stability. *J Virol* 82:6667–6677
- Doxiadis GG, de Groot N, Dauber EM, van Eede PH, Fae I, Faner R, Fischer G, Grubic Z, Lardy NM, Mayr W, Palou E, Swelsen W, Stingl K, Doxiadis II, Bontrop RE (2009) High resolution definition of HLA-DRB haplotypes by a simplified microsatellite typing technique. *Tissue Antigens* 74:486–493
- Doxiadis GG, de Groot N, Otting N, Blokhuis JH, Bontrop RE (2011) Genomic plasticity of the MHC class I A region in rhesus macaques: extensive haplotype diversity at the population level as revealed by microsatellites. *Immunogenetics* 63:73–83
- Doxiadis GG, de Groot N, Otting N, de Vos-Rouweller AJ, Bolijn MJ, Heijmans CM, de Groot NG, van der Wiel MK, Remarque EJ, Vangenot C, Nunes JM, Sanchez-Mazas A, Bontrop RE (2013) Haplotype diversity generated by ancient recombination-like events in the MHC of Indian rhesus macaques. *Immunogenetics* 65:569–584
- Eckert KA, Hile SE (2009) Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. *Mol Carcinog* 48:379–388
- Foissac A, Crouau-Roy B, Faure S, Thomsen M, Cambon-Thomsen A (1997) Microsatellites in the HLA region: on overview. *Tissue Antigens* 49:197–214
- Gourraud PA, Feolo M, Hoffman D, Helmsberg W, Cambon-Thomsen A (2006) The dbMHC microsatellite portal: a public resource for the storage and display of MHC microsatellite information. *Tissue Antigens* 67:395–401
- Hernandez RD, Hubisz MJ, Wheeler DA, Smith DG, Ferguson B, Rogers J, Nazareth L, Indap A, Bourquin T, McPherson J, Muzny D, Gibbs R, Nielsen R, Bustamante CD (2007) Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques. *Science* 316:240–243
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, Wain HM, Trowsdale J, Ziegler A, Beck S (2004) Gene map of the extended human MHC. *Nat Rev Genet* 5:889–899
- Karl JA, Bohn PS, Wiseman RW, Nimityongskul FA, Lank SM, Starrett GJ, O'Connor DH (2013) Major histocompatibility complex class I haplotype diversity in Chinese rhesus macaques. *G3 (Bethesda)* 3:1195–1201
- 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
- Lam TH, Shen M, Chia JM, Chan SH, Ren EC (2013) Population-specific recombination sites within the human MHC region. *Heredity (Edinb)* 111:131–138
- Li A, Wang X, Liu Y, Zhao Y, Liu B, Sui L, Zeng L, Sun Z (2012a) Preliminary observations of MHC class I A region polymorphism in three populations of Chinese-origin rhesus macaques. *Immunogenetics* 64:887–894
- Li W, Wang T, Ling F, Zhao H, Wei L, Zhuo M, Du H, Wang X (2012b) Identification of MhcMafa-DRB alleles in a cohort of cynomolgus macaques of Vietnamese origin. *Am J Primatol* 74:958–966
- Liu Y, Li A, Wang X, Sui L, Li M, Zhao Y, Liu B, Zeng L, Sun Z (2013) Mamu-B genes and their allelic repertoires in different populations of Chinese-origin rhesus macaques. *Immunogenetics* 65:273–280
- Malkki M, Single R, Carrington M, Thomson G, Petersdorf E (2005) MHC microsatellite diversity and linkage disequilibrium among common HLA-A, HLA-B, DRB1 haplotypes: implications for unrelated donor hematopoietic transplantation and disease association studies. *Tissue Antigens* 66:114–124
- Martin M, Mann D, Carrington M (1995a) Recombination rates across the HLA complex: use of microsatellites as a rapid screen for recombinant chromosomes. *Hum Mol Genet* 4:423–428
- Martin M, Mann D, Carrington M (1995b) Corrigendum: recombination rates across the HLA complex: use of microsatellites as a rapid screen for recombinant chromosomes. *Human Molecular Genetics* 4, 423-428 (1995). *Hum Mol Genet* 4:2423
- Mee ET, Badhan A, Karl JA, Wiseman RW, Cutler K, Knapp LA, Almond N, O'Connor DH, Rose NJ (2009) MHC haplotype frequencies in a UK breeding colony of Mauritian cynomolgus macaques mirror those found in a distinct population from the same geographic origin. *J Med Primatol* 38:1–14
- Mitchell JL, Mee ET, Almond NM, Cutler K, Rose NJ (2012) Characterisation of MHC haplotypes in a breeding colony of Indonesian cynomolgus macaques reveals a high level of diversity. *Immunogenetics* 64:123–129
- Otting N, Heijmans CM, Noort RC, de Groot NG, Doxiadis GG, van Rood JJ, Watkins DI, Bontrop RE (2005) Unparalleled complexity of the MHC class I region in rhesus macaques. *Proc Natl Acad Sci U S A* 102:1626–1631

- 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
- Penedo MC, Bontrop RE, Heijmans CM, Otting N, Noort R, Rouweler AJ, de Groot N, de Groot NG, Ward T, Doxiadis GG (2005) Microsatellite typing of the rhesus macaque MHC region. *Immunogenetics* 57:198–209
- Setchell JM, Huchard E (2010) The hidden benefits of sex: evidence for MHC-associated mate choice in primate societies. *Bioessays* 32: 940–948
- Shiina T, Ota M, Shimizu S, Katsuyama Y, Hashimoto N, Takasu M, Anzai T, Kulski JK, Kikkawa E, Naruse T, Kimura N, Yanagiya K, Watanabe A, Hosomichi K, Kohara S, Iwamoto C, Umehara Y, Meyer A, Wanner V, Sano K et al (2006) Rapid evolution of major histocompatibility complex class I genes in primates generates new disease alleles in humans via hitchhiking diversity. *Genetics* 173:1555–1570
- Slierendregt BL, Otting N, van Besouw N, Jonker M, Bontrop RE (1994) Expansion and contraction of rhesus macaque DRB regions by duplication and deletion. *J Immunol* 152:2298–2307
- Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, Stewart CA, Ashurst JL, Atrazhev AM, Coggill P, Palmer S, Almeida J, Sims S, Wilming LG, Rogers J, de Jong PJ, Carrington M, Elliott JF, Sawcer S, Todd JA, Trowsdale J et al (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. *PLoS Genet* 2:e9
- Trowsdale J, Knight JC (2013) Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet* 14:301–323
- Walsh EC, Mather KA, Schaffner SF, Farwell L, Daly MJ, Patterson N, Cullen M, Carrington M, Bugawan TL, Erlich H, Campbell J, Barrett J, Miller K, Thomson G, Lander ES, Rioux JD (2003) An integrated haplotype map of the human major histocompatibility complex. *Am J Hum Genet* 73:580–590
- Watanabe A, Shiina T, Shimizu S, Hosomichi K, Yanagiya K, Kita YF, Kimura T, Soeda E, Torii R, Ogasawara K, Kulski JK, Inoko H (2007) A BAC-based contig map of the cynomolgus macaque (*Macaca fascicularis*) major histocompatibility complex genomic region. *Genomics* 89:402–412
- Wiseman RW, Wojcechowskyj JA, Greene JM, Blasky AJ, Gopon T, Soma T, Friedrich TC, O'Connor SL, O'Connor DH (2007) Simian immunodeficiency virus SIVmac239 infection of major histocompatibility complex-identical cynomolgus macaques from Mauritius. *J Virol* 81:349–361
- Wojcechowskyj JA, Yant LJ, Wiseman RW, O'Connor SL, O'Connor DH (2007) Control of simian immunodeficiency virus SIVmac239 is not predicted by inheritance of Mamu-B*17-containing haplotypes. *J Virol* 81:406–410
- Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ, Du H, Chen J, Chen R, Zhang P, Huang Z, Thompson JR, Meng Y, Bai Y, Wang J, Zhuo M et al (2011) Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat Biotechnol* 29: 1019–1023
- Zhang GQ, Ni C, Ling F, Qiu W, Wang HB, Xiao Y, Guo XJ, Huang JY, Du HL, Wang JF, Zhao Sh J, Zhuo M, Wang XN (2012) Characterization of the major histocompatibility complex class I A alleles in cynomolgus macaques of Vietnamese origin. *Tissue Antigens* 80:494–501
- Ziegler A, Santos PS, Kellermann T, Uchanska-Ziegler B (2010) Self/nonsel perception, reproduction and the extended MHC. *Self Nonsel* 1:176–191