A highly divergent microsatellite facilitating fast and accurate *DRB* haplotyping in humans and rhesus macaques

Gaby G. M. Doxiadis[†‡], Nanine de Groot[†], Frans H. J. Claas[§], Ilias I. N. Doxiadis[§], Jon J. van Rood[‡§], and Ronald E. Bontrop[†]

[†]Department of Comparative Genetics and Refinement, Biomedical Primate Research Centre, P.O. Box 3306, 2280 GH, Rijswijk, The Netherlands; and [§]Department of Immunohematology and Blood Transfusion, Leiden University Medical Centre, E3-Q, P.O. Box 9600, 2300 RC, Leiden, The Netherlands

Contributed by Jon J. van Rood, March 30, 2007 (sent for review October 27, 2006)

The *DRB* region of the MHC in primate species is known to display abundant region configuration polymorphism with regard to the number and content of genes present per haplotype. Furthermore, depending on the species studied, the different *DRB* genes themselves may display varying degrees of allelic polymorphism. Because of this combination of diversity (differential gene number) and polymorphism (allelic variation), molecular typing methods for the primate *DRB* region are cumbersome. All intact *DRB* genes present in humans and rhesus macaques appear to possess, however, a complex and highly divergent microsatellite. Microsatellite analysis of a sizeable panel of outbred rhesus macaques, covering most of the known *Mamu-DRB* haplotypes, resulted in the definition of unique genotyping patterns that appear to be specific for a given haplotype. Subsequent examination of a representative panel of human cells illustrated that this approach also facilitates high-resolution *HLA-DRB* typing in an easy, quick, and reproducible fashion. The genetic composition of this complex microsatellite is shown to be in concordance with the phylogenetic relationships of various *HLA-DRB* and *Mamu-DRB* exon 2 gene/lineage sequences. Moreover, its length variability segregates with allelic variation of the respective gene. This simple protocol may find application in a variety of research avenues such as transplantation biology, disease association studies, molecular ecology, paternity testing, and forensic medicine.

biological science | MHC

Gene products of the MHC play a key role in generating adaptive immune responses. Most vertebrate species express two distinct classes of MHC cell surface markers that execute differential biological functions. MHC class I glycoproteins are involved in the presentation of peptides of intracellular origin to cytotoxic T cells, possibly resulting in the lysis of infected or malignant structures. MHC class II molecules bind peptides from extracellular sources, and subsequent interaction with T helper cells may result, for instance, in antibody production. During vertebrate evolution, the MHC class I and II region genes have been subject to many expansion and contraction processes, and as a consequence a given (sub)region may harbor multiple, highly related genes. A next layer of complexity is the phenomenon that some of these MHC genes display unprecedented levels of polymorphism. The compound evolutionary history of MHC genes, in combination with their abundant levels of polymorphism, represents a challenge for the research community, because one has to come up with a simple and understandable nomenclature system. For example, in the human population five *HLA-DRB* region configurations have been defined. Each configuration contains a *DRA* gene in conjunction with a unique combination of polymorphic genes, designated *DRB1* to *DRB9* (1), which arose from a intricate series of duplication and crossing-over events during primate evolution (2, 3). Some of these genes are functional, whereas others have various types of deleterious mutations and are considered to represent pseudogenes (4, 5). The *HLA-DRB1* gene is, at least in part, responsible for the fact that each of the human region configurations displays plentiful allelic heterogeneity (6).

The *Mamu-DRB* region in rhesus macaques (*Macaca mulatta*) is also highly plastic (7, 8), and at least 30 region configurations have been documented that, as found in humans, display marked differences with regard to number and content of genes present per haplotype (7, 9, 10). Most of the *Mamu-DRB* alleles belong to loci/lineages that are shared with humans, and these alleles have been named accordingly (6, 11, 12). For example, *HLA-DRB1*0301* and *Mamu-DRB1*0306* define highly similar alleles encoded by apparently orthologous genes in different primate species. This high degree of genetic resemblance is also reflected by similarities at the functional level, such as peptide selection and T cell activation (13). Different primate species may also share similar pseudogenes. *HLA-DRB6* orthologues have been found, for instance, in chimpanzees (14–16), gorillas (17), and different macaque species (17–19). Additionally, particular lineage members that may be regarded as functional in one species remain untranscribed in another species (20). Some pseudogenes located in the primate MHC region have remained stable entities over long evolutionary time spans. Recently it was shown that intact gene segments of pseudogenes can be recruited again for immune functions by exon shuffling (21). Although some *Mamu-DRB* loci/lineages display moderate levels of polymorphism, within the region configurations themselves only marginally low levels of allelic polymorphism are observed (7, 22).

Allelic polymorphism ensures that different allotypes select distinct peptides for T cell activation, preventing one particular pathogen from sweeping through an entire population. Particular HLA class II alleles are indeed associated with either susceptibility or resistance to contract particular diseases (23–26). Apart from its genuine biological function, in transplantation protocols HLA-DR matching may also influence graft survival positively (27–29).

Because of the complexity of the *HLA-DRB* region, routine typing is performed by using different techniques (or a combination thereof) such as sequence-specific oligonucleotide typing (PCR-SSOP) for lineage-specific low resolution analysis (30) or sequence-specific priming (PCR-SSP) with unique primer sets for character-

Author contributions: J.J.v.R. and R.E.B. designed research; N.d.G. performed research; F.H.J.C. and I.I.N.D. contributed new reagents/analytic tools; G.G.M.D. analyzed data; and G.G.M.D. and R.E.B. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviation: STR, short tandem repeat.

Data deposition: The sequences reported in this paper have been deposited in the EMBL Nucleotide Sequence Database (accession nos. AM490200–AM490205).

[‡]To whom correspondence may be addressed. E-mail: doxiadis@bprc.nl or vanrood@europdonor.nl.

This article contains supporting information online at www.pnas.org/cgi/content/full/0702964104/DC1.

© 2007 by The National Academy of Sciences of the USA

Table 1. *Mamu-DRB* haplotypes defined by exon 2 sequencing and D6S2878 genotyping

Hap	First *DRB* locus	STR	Second *DRB* locus	STR	Third *DRB* locus	STR	Fourth *DRB* locus	STR	Fifth *DRB* locus	STR	Sixth *DRB* locus	STR
1a	*DRB1*0406*	205, 207, 211	*DRB5*0301*	169								
1b	*DRB1*0411*	223	*DRB5*0304*	153								
2	*DRB3*0403*	192	*DRB*W305*	227								
3†	*DRB1*0412*	199	*DRB*W3701*	235								
4†	*DRB1*1010*	205	*DRB6*0121*	216								
5†	*DRB3*0408*	203	*DRB*W404*	293								
6†	*DRB3*0411*	190	*DRB*W314*	219								
7‡	*DRB4*0102*	291	*DRB5*0306*	173								
8‡	*DRB1*0321*	181	*DRB1*0322/3*	209								
9‡	*DRB1*0321*	181	*DRB1*0322/3*	209	*DRB1*1003*§	175?						
10‡	*DRB1*0309*	229	*DRB*W2507*	181	*DRB6*0109*	?						
11a	*DRB1*0303*	195	*DRB1*1007*	197	*DRB6*0103*	(202–208)						
11b	*DRB1*0312*	195	*DRB1*1007*	201	*DRB6*0103*	?						
11c	*DRB1*0306*	226	*DRB1*1007*	197	*DRB6*0103*	?						
11d	*DRB1*0306*	226–230	*DRB1*1003*	199–213	*DRB6*0107*	182						
12	*DRB1*0309*	233–247	*DRB6*0101*	?	*DRB*W201*	263 (261, 267)						
13	*DRB6*0112*	214	*DRB*W2501*	214, 218?	*DRB*W2002*	218	*DRB6*0106*?	205?				
14	*DRB6*0114*	218	*DRB*W303*	193	*DRB*W401*	187						
15a	*DRB1*0403*	229	*DRB6*0107*	184	*DRB*W501*	?	*DRB6*?*	208				
15b‡	*DRB1*0403*	229	*DRB6*0107*	183	*DRB*W502*	?	*DRB6*0120*	208				
16	*DRB1*0404*	225	*DRB6*0102*	164	*DRB*W307*	195	*DRB*W702*	188				
17	*DRB1*0701*	197	*DRB3*0405*	237	*DRB5*0303*	169	*DRB6*0123*	184				
18a†	*DRB1*0306*	214	*DRB1*07032*	211	*DRB6*0124*	170	*DRB*W2603*	177	*DRB6*0120*	?		
18b‡	*DRB1*0306*	214	*DRB1*07033*	207	*DRB6*0124*	168	*DRB*W2603*	177	*DRB6*0120*	208		
19a	*DRB1*0318*	181	*DRB6*0105*	190, 192	*DRB*W604*	201, 205	*DRB*W603*	275, 277	*DRB6*0104* (+ALU)	222 (220)		
19b	*DRB1*0318*	181	*DRB6*0105*	194	*DRB*W604*	197	*DRB*W611*	275	*DRB6*01?*	222		
20†	*DRB3*0403*	182	*DRB*W402*	185	*DRB*W2701*	227	*DRB6*0116*	206	*DRB6*0107*	182		
21a	*DRB6*0111*	178	*DRB*W606*	211, 213, 197	*DRB*W2104*	211, 213	*DRB*W2603*	177, 187	*DRB6*0122*	236		
21b‡	*DRB6*0108*	176	*DRB*W606*	213	*DRB*W2104*	213	*DRB*W2604*	177	*DRB6*0122*	232		
22	*DRB1*0310*	241	*DRB6*0118*	210	*DRB*W101*	219, 223	*DRB*W602*	279 (285)	*DRB*W609*	219 (223)	*DRB6*0106*	192

Data in parentheses are STR length observed in some (one to three) animals. Question marks indicate STRs not detected or rarely detected but presence of gene ascertained by sequencing. "205?" and "175?" indicate STRs detected but not confirmed by sequencing. Hap, haplotype.
†Animals with Chinese origin.
‡Animals with Burmese origin.
§*DRB1*1003* belong to haplotypes 8 or 9.

izing various *HLA-DRB* genes and their alleles. Because rhesus macaques possess complex *DRB* region configurations, laborious methods such as exon 2 or full-length *DRB* cloning and subsequent sequencing are inevitable for accurate typing.

In close proximity to exon 2, responsible for most sequence variation, maps a complex dinucleotide repeat, D6S2878, that is present in the majority of the *HLA-DRB* genes (31–33). This microsatellite, located in the beginning of intron 2, displays a complex profile concerning repeat length variability as well as sequence variation (34). This repeat is also attractive from an evolutionary point of view, because it is present in apes and Old and New World monkeys as well (33–36). Therefore, D6S2878 is considered to represent a promising marker for the development of a *DR* haplotyping protocol in different primate species, provided that a single set of informative primers can be developed. Because the *DR* regions in humans (high allelic variability in combination with a low number of region configurations) and rhesus macaques (low allelic variability and a high number of region configurations) represent each other's extremes, they were selected as candidate systems for setting up a fast and accurate high-resolution haplotyping protocol.

Results

***Mamu-DRB* Region Configuration Definition by Microsatellite Markers.** The rhesus macaques studied are part of a large self-sustaining colony that has been thoroughly pedigreed based on segregation of serologically and molecularly defined markers (9, 20). The current panel covers 22 different region configurations characterized by unique combinations from two to six distinct *Mamu-DRB* genes (Table 1). Limited allelic polymorphism is detected within region configurations 1, 11, 15, 18, 19, and 21. To denote allelic variation observed, for instance, in region configuration 1, the relevant haplotypes have been designated 1a and 1b (Table 1). The gene frequencies of the different region configurations as encountered in our population of animals are provided as well, in concert with the number of animals tested. Some haplotypes are frequently observed whereas others appear to be rare [supporting information (SI) Table 3].

Nearly all *Mamu-DRB* loci/lineages possess the D6S2878 microsatellite, and the relevant DNA segment could be amplified by means of the unique primer set developed for this protocol (Table 1). Amplification failures have only been observed for a few *DRB6*/*DRBW* pseudogenes for which the corresponding amplicons could be scarcely detected, most probably because of primer inconsistencies. The overall results illustrate, however, that the lengths of the amplified short tandem repeat (STR) products are highly variable and range from 153 to 293 bp (Table 1). Additionally, various STR markers seem to be predictive for the presence of a particular *Mamu-DRB* allele, and, moreover, family studies demonstrated that they segregate in a Mendelian manner. Subsequent sequencing of the DNA segment ranging from exon 2 to intron 2 was conducted to unequivocally link each D6S2878 allele to an individual *DRB* gene/allele. This approach proved that diverse STR lengths could distinguish even highly similar alleles, differing for only one or two nucleotides. An instance is provided by the *Mamu-DRB1*07032* and *Mamu-DRB1*07033* alleles, which are part of the two haplotypes belonging to configuration 18 (Table 1).

On average, most of the STRs linked to an individual allele appear to be rather conservative in composition and length. In the case of the frequently observed haplotype 11a (SI Table 3), for

Table 2. *HLA-DRB* haplotypes defined by sequencing and D6S2878 genotyping

Haplotype	First *DRB* locus	STR	Second *DRB* locus	STR	Third *DRB* locus	STR
DR8						
DR8a	*DRB1*0801(03)*	170 (172)				
DR8b	*DRB1*080302*	176				
DR1						
DR1a/DR103	*DRB1*010101/0103*	154 (156)	*DRB6*010101*	136		
DR1b	*DRB1*010201*	160	*DRB6*0101*	136		
DR10	*DRB1*100101*	161	*DRB6*0101*	136		
DR52						
DR17a	*DRB1*030101*	172	*DRB3*010101(2)*	180		
DR17b	*DRB1*0301*	178–184	*DRB3*0202(01*	208 (206, 210)		
DR11a	*DRB1*110101*	182–192	*DRB3*020201*	206–218		
DR11b	*DRB1*110201*	186	*DRB3*020201*	214		
DR11c	*DRB1*110401*	188	*DRB3*020201*	208 (210)		
DR12a	*DRB1*120101*	200	*DRB3*010102*	180		
DR12b	*DRB1*1201*	206 (208)	*DRB3*02(0201)*	210, 212		
DR12c	*DRB1*120201*	216	*DRB3*030101*	186		
DR13a	*DRB1*1301(01)*	194	*DRB3*0101(02)*	180		
DR13b	*DRB1*1301(01)*	194, 196	*DRB3*02(0201)*	204–214		
DR13c	*DRB1*1302(01)*	188, 206–218	*DRB3*0301/0101*	186		
DR14a	*DRB1*140101*	192	*DRB3*0211*	212		
DR14b	*DRB1*140101*	186, 188, 196	*DRB3*020201*	210		
DR14c	*DRB1*140101*	198	*DRB3*0201*	214		
DR53						
DR4a	*DRB1*0401(01)*	180, 182, 184	*DRB4*(010101)*	178, 180 (176, 182/4)	*DRB7*010101*	135
DR4b	*DRB1*0402*	194	*DRB4*010101*	176	*DRB7*010101*	135
DR4c	*DRB1*0404*	192 (196)	*DRB4*010101*	178	*DRB7*010101*	135
DR4d	*DRB1*040701*	184	*DRB4*010101*	178	*DRB7*010101*	135
DR4e	*DRB1*0408*	180	*DRB4*010101*	178	*DRB7*010101*	135
DR7	*DRB1*070101*	149	*DRB4*010101*	170, 176, 180	*DRB7*010101/2*	135
DR9	*DRB1*09(0102)*	170 (196)	*DRB4*010101*	180	*DRB7*010101*	135
DR51						
DR15a	*DRB1*150101*	186	*DRB5*0101(01)*	180	*DRB6*(0201)*	154
DR15b	*DRB1*150101*	186, 190 (184, 188)	*DRB5*0101(01)*	184 (190)	*DRB6*(0201)*	152
DR15c	*DRB1*150201*	202–208	*DRB5*0102*	220	*DRB6*(0201)*	174
DR16	*DRB1*160101*	178	*DRB5*?*	184	*DRB6*0202*	144

Data in parentheses indicate STR lengths detected only once or twice.

example, no length variability for both *DRB1* gene-associated D6S2878 alleles is observed, which also holds true for the STR that is linked to *DRB5*0301* as seen in haplotype 1a (Table 1). The *DRB1*0406*-linked STR of the latter configuration, however, may slightly vary in length. Such differences, as can be seen for example in haplotype 12, are reproducible and do segregate in families with a particular haplotype as defined by serological methods (37). Because each region configuration is composed of an exclusive combination of different *Mamu-DRB* genes, the most essential conclusion that can be drawn is that the combination of STR markers appears to be unique for a given region configuration/haplotype.

***HLA-DRB* Haplotype Definition by Microsatellite Analysis.** Genotyping for the highly divergent (GT)x(GA)y microsatellite allows speedy and accurate *DR* haplotyping in rhesus macaques, a species known to possess a high number of region configurations; these display, however, low levels of allelic variation. Next, it was investigated whether the same approach would work in a species that possesses a relatively low number of region configurations that parade abundant levels of allelic polymorphism. Therefore, in the first instance, a thoroughly characterized panel of 160 unrelated human samples of Caucasoid origin was chosen to conduct such an analysis. This selection also comprised 64 homozygous typing cells, facilitating the simple definition of haplotypes. Furthermore, the panel included the five known *HLA-DRB* region configurations, designated DR8, DR1, DR51, DR52, and DR53 (38), but also covered the 13 most common DR serotypes present in the Caucasoid population, differing for their *DRB1* alleles (Table 2). To denote allelic variation as observed within some serotypes, the different *DRB* haplotypes have been designated a, b, c, and so on (Table 2). The number of samples tested per haplotype, as well as the gene frequencies of the serotypes (39), have also been summarized (SI Table 3).

Again, a specially designed generic primer pair allowed amplification of the relevant intron 2 segment for virtually all *HLA-DRB* genes/alleles. Amplification failed only for one particular *HLA-DRB5* allele present on DR16 haplotypes. This may be due to a mutated primer site. Subsequent extensive sequencing of exon 2–intron 2 DNA segments verified for each of the *HLA-DRB* alleles the unique linkage to its adjacent D6S2878 allele. The lengths of the respective repeats are highly polymorphic in the human population as well and range from 135 to 220 bp (Table 2). Moreover, STR lengths are highly predictive for the presence of individual *HLA-DRB* alleles, as was also shown earlier for rhesus monkeys. Again, differential STR lengths can make the distinction between highly similar *-DRB* alleles, differing for a few nucleotides: for example in the case of *DRB1*0801*(*03*) and *DRB1*080302* (Table 2).

The DR53 region configuration contains a pseudogene, named *HLA-DRB7*, which is absent in macaques. This pseudogene displays no allelic variation, and indeed the associated D6S2878 is invariant in length independent of the adjacent *-DRB1* allele. Thus, the *DRB7*-associated STR typifies all of the haplotypes belonging to the DR53 region configuration (Table 2). The same holds true for the *DRB6* allele and its adjacent STR, which are characteristic for the DR1 region configuration, covering the DR1 and DR10 serotypes. As observed in rhesus macaques, some *HLA-DRB* alleles

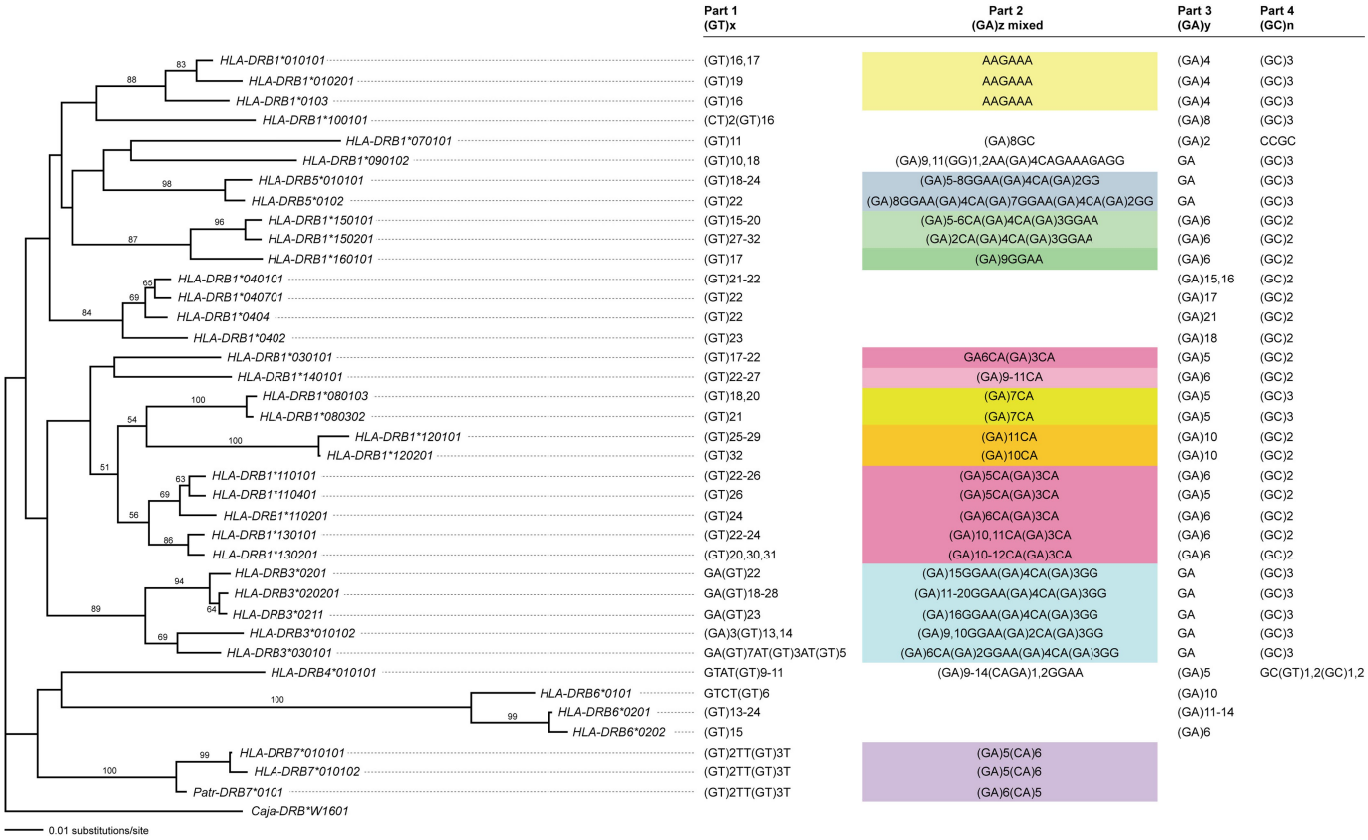

Fig. 1. Phylogenetic analysis of *HLA-DRB* exon 2 with D6S2878 sequences superimposed. The tree was rooted by including the *Caja-DRB*W1601* allele, and bootstrap values are shown. Identical or similar colors indicate identity or similarities of the (GA)z mixed part of D6S2878.

appear to be associated with multiple D6S2878 variants. A case is provided by the *HLA-DRB1*140101* allele observed within the DR14 haplotypes, but some of the repeats grouping in the DR11a and DR13c family of haplotypes also show length differences (Table 2). Family analyses are needed to demonstrate their segregation. Most of the human repeats, however, appear to be conservative in composition and length and are linked to an individual allele. Within this panel of 30 different *HLA-DRB* haplotypes, all but two can be readily dissected based on by their D6S2878 profiles. Thus, complex *DR* region configuration/haplotype information, as present in at least two populations of primate species, is readily obtained and defined based on D6S2878 genotyping after a simple amplification protocol conducted with one primer set.

Genetic Stability and Evolutionary History of the STR-DRB Complex. The D6S2878 microsatellite has a composite character in primates (33–36), and phylogenetic comparisons of different *DRB1* sequences obtained from humans and chimpanzees indicated that the ancestral structure most likely must have been a (GT)x(GA)y dinucleotide repeat. The *HLA-DRB1*-associated microsatellite comprises three sections exhibiting different evolutionary stabilities. The 5′(GT)x repeat represents the longest segment and evolves most rapidly, which is a known feature for long, uninterrupted dinucleotide repeats. The middle section or the (GA)z part is shorter and interrupted; its constellation appears to correlate well with different lineages/loci, and its length seems to segregate with specific *DRB1* alleles. The length of the 3′(GA)y part appears to be specific for a certain *DRB* lineage/locus (34). All *HLA-DRB* exon 2 sequences described in this study have been subjected to phylogenetic analyses, on which the D6S2878 sequences have been superimposed (Fig. 1). As can be seen, this article extends the knowledge on this particular microsatellite but also underscores its compound character (32, 34). Additionally, a short dinucleotide $(GC)_{1–3}$ part could be observed at the 3′ end of the microsatellite of all *HLA-DRB* genes except for those belonging to the *DRB6* and *DRB7* pseudogenes. The variation seen in this newly recognized section of the microsatellite seems to be prognostic for certain *HLA-DRB* lineages or loci, respectively. The 5′(GT)x part is indeed the most polymorphic, which may, especially in the case of fairly long (GT)x repeats, evolve faster than the mutation rate operative on exon 2 itself (Fig. 1). This phenomenon, already described for DQCAR (40), could explain why, for instance, the *HLA-DRB1*1302* allele is associated with a STR displaying length variation (Table 2).

The ancient *HLA-DRB6* and *HLA-DRB7* pseudogenes appear to miss the middle, interrupted (GA)z or the (GA)y part, respectively. The *HLA-DRB6*0101* gene harbors a genetically stable D6S2878 showing no length differences at all. This is most probably because of its composition represented by a short and interrupted 5′(GT) and 3′(GA) part (41, 42). For *HLA-DRB6*0201* the opposite is true, and the long and uninterrupted (GT)x and (GA)y parts are indicative for unstable STR lengths (Table 2 and Fig. 1) (43).

The *Mamu-DRB* exon 2 sequences have also been subjected to phylogenetic analyses, and the genetic composition of the D6S2878 sequences has been superimposed (SI Fig. 2). Like in humans, D6S2878 has a compound character. The newly described fourth 3′(GC) dinucleotide part is also present in rhesus macaques with repeat lengths ranging from 1 to 5. Comparable to the human situation, *Mamu-DRB6* alleles form a distinct clade in the phylogenetic tree. Because the *DRB6* locus is thought to predate the divergence of Old World monkeys, great apes, and homonoids, it is not surprising that the same ancestral (GT)x(GA)y structure was detected in rhesus monkeys just as in humans. In contrast to the human microsatellite, some *Mamu-DRB* lineages/loci are charac-

terized by a multipart (GT)x and/or a (GA)z middle segment. In general, the rhesus macaque STR appears to be even more complex than its human equivalent. This multifaceted composition seems to correlate with the high number of *DRB* region configuration/haplotype diversity observed in rhesus macaques. Despite the complex microsatellite patterns observed in the rhesus macaque, the composition and length of repeats associated with known *Mamu-DRB* alleles are, as expected, highly similar. The rhesus macaque *DRB* region has been subject to several rounds of duplication and contraction processes. For that reason, it is difficult to understand which highly related genes located on different region configurations/haplotypes represent separate loci or whether these sequences have an allelic affiliation. This microsatellite will be helpful in sorting out such genetic relationships.

Discussion

Microsatellite D6S2878 is considered to represent a promising marker for the development of a quick and accurate *DRB* haplotyping protocol in primate species, on the condition that one single set of informative primers can be developed. Indeed, a specifically designed primer pair allowed amplification of the relevant intron 2 segment for virtually all *HLA-DRB* and *Mamu-DRB* genes/alleles. For both species, amplification artifacts have rarely been observed and such types of problems are easily overcome by designation of specific primers that can be added to the same reaction mixture. The D6S2878 STR was proven to be highly variable in length, not only in humans but also in the rhesus macaque, thus verifying this microsatellite as a useful marker for *DRB* typing of both species. Phylogenetic analyses of human as well as rhesus macaque exon 2 sequences have been performed and compared with microsatellite composition. For humans, the evolutionary relationships of exon 2 and the adjacent microsatellite seem to segregate closely together (Fig. 1). In rhesus macaques, the microsatellite composition was far more variable than in humans, and a comparison of microsatellite and exon 2 phylogeny does not always seem to match the microsatellite composition (SI Fig. 2). One explanation for these results may be that rhesus macaque *DRB* loci/lineages are much older than their human equivalents, and this can be the reason for the higher diversity of the adjacent microsatellite as well. Furthermore, preliminary results of intron sequences illustrate that some of the *Mamu-DRB* sequences that are considered alleles of a given locus probably represent monomorphic loci themselves. However, truly allelic variants manifest the reliability of the comparison of exon 2 and the D6S2878 marker not only in humans but also in rhesus macaques.

Microsatellite typing was in the first instance performed on large human and rhesus macaque panels for which the typing information was known. The most essential conclusion to be drawn (Tables 1 and 2) is that in both species the combination of STR markers appears to be unique for a given haplotype. Within the rhesus macaque panel of 31 haplotypes all could be defined unambiguously; within the human panel of 30 haplotypes all but two could be thus defined. As stated earlier, ambiguities can easily be solved by development of additional primer pairs. As a control, a blind test was performed with 47 human and 26 rhesus monkey samples. More than 90% of the samples were scored correctly for their *Mhc-DRB* haplotypes. Some of the samples were not scored properly because they contained allotypes that were not present in the original test panel. Thus, this D6S2878 typing protocol provides a highly reliable method for *Mamu-DRB* and *HLA-DRB* haplotyping, with its main advantage being simplicity and speed. Because differently labeled primers can be used for microsatellite typing, multiplexing is possible, and 96 samples or even more can be analyzed within one test panel. The simplicity of the test is especially useful for *Mamu-DRB* haplotyping, which is otherwise extremely time-consuming because of the unprecedented high number of different *-DRB* region configurations. For the human situation, this approach is very helpful in the analysis of large amounts of samples, as they are needed, for example, in population and/or disease association studies. Furthermore, this method may also be of use in forensic medicine as well as in paternity-testing protocols. Additionally, high-resolution *-DRB* haplotyping will simplify donor–recipient matching in organ as well as bone marrow transplantation. Because the D6S2878 STR is an old entity, it may also be used to study other populations of primate species.

The evolutionary stability of this microsatellite has been a matter of debate (32–34, 36, 44). The *HLA-DRB7*-associated D6S2878 allele is especially remarkable, because this repeat has the shortest (GT)x as well as the (GA)y part and is highly stable in length, showing no polymerase slippage at all. This is in accordance with the fact that short and/or interrupted repeats are more stable than long, uninterrupted dinucleotides (41–43). To what extent the repeat composition may have a direct or indirect influence on the low mutation rate of the adjacent exon 2 segment of the *DRB7* pseudogene is not well understood at present. It has been proposed, for example, that microsatellites near genes may increase and probably also decrease local mutation rates (45). Interestingly, exon 2 of the *DRB7* pseudogene, present on the only shared *DR* region configuration of humans and chimpanzees, is highly conserved between both species. Moreover, the D6S2878 sequence is nearly identical (Fig. 1). It has been suggested that the intron 2 segment containing the (GT)x(GA)y repeat may bind a zinc-dependent protein and forms non B-DNA structures; thus, functionality of these so-called "junk" DNA sequences cannot be ruled out and should be subjected to further analysis (44).

Materials and Methods

Samples. The 167 rhesus monkeys analyzed, housed at the Biomedical Primate Research Centre's breeding colony, originated mostly from India, but some are also of Burmese or Chinese origin. Seven of these animals are completely homozygous for their MHC region and derived from consanguineous matings; two additional animals are homozygous for their Mamu-A, Mamu-B, and Mamu-DR serotypes. Genomic DNA of human individuals or rhesus macaques was extracted from EDTA blood samples or from immortalized B cell lines using a standard salting-out procedure. Of the 160 human samples tested, 64 were *HLA-DRB* homozygous, 17 of which belong to a thoroughly characterized homozygous typing cell panel of the 14th International Histocompatibility Workshop (2005).

STR-DRB Genotyping. The relevant DNA segment in rhesus macaques was amplified with a forward primer located at the end of exon 2 (5′-*Mamu-DRB*-STR, TTC ACA GTG CAG CGG CGA GGT) and two labeled reverse primers in intron 2 (3′-*Mamu-DRB*-STR_VIC, ACA CCT GTG CCC TCA GAA CT; and 3′-*Mamu-DRB*-STR_FAM_1007, ACA TCT GTG TCC TCA GAC CT). For human samples a labeled forward primer located at the end of exon 2 (5′-*HLA-DRB*-STR_VIC, GAG AGC TTC ACA GTG CAG C) and one reverse primer in intron 2 (3′-*HLA-DRB*-STR, GAG AGG ATT CTA AAT GCT CAC) was used. The labeled primers were synthesized by Applied Biosystems (Foster City, CA), and the unlabeled primers were synthesized by Invitrogen (Paisley, Scotland). The PCR for rhesus macaques was performed in a 25-μl reaction volume containing 1 unit of *Taq* polymerase (Invitrogen) with 1.0 μM of the unlabeled forward primer, 0.6 μM of the VIC-labeled reversed primer, 0.2 μM of the FAM-labeled reversed primer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1× PCR buffer II (Invitrogen), and 100 ng of DNA.

The PCR mixture for the human STR amplification was the same as that used for rhesus macaques with 0.1 μM of the VIC-labeled forward primer and 0.1 μM of the unlabeled reverse primer. The cycling parameters for both amplifications were a 5-min 95°C initial denaturation step followed by five cycles of 1 min at 94°C, 45 s at 58°C, and 45 s at 72°C. Then the program was followed by 25 cycles of 45 s at 94°C, 30 s at 58°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 analyzed on an ABI 3130 genetic analyzer (Applied Biosystems). STR analysis was performed with the Genemapper program (Applied Biosystems), and all samples were analyzed at least twice.

PCR, Cloning, and Sequencing. Seventy-five different *Mamu-DRB* alleles and 38 *HLA-DRB* alleles were sequenced from exon 2 to intron 2 including the microsatellite with a generic 5′ DRB–exon 2 primer (CGT GTC CCC ACA GCA CGT TTC) together with the same 3′ primers as used for DRB-STR genotyping but without label. The PCRs for rhesus monkey and human *DRB* were performed in a 100-μl volume containing 4 units of *Taq* polymerase (Invitrogen) with 0.2 μM of each primer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1× PCR buffer II (Invitrogen), and 200 ng of DNA. The cycling parameters were the same as described for STR-DRB genotyping. The resulting amplicons were cloned and sequenced as described recently (19, 22). The resulting sequences were analyzed by using the Sequence Navigator program (Applied Biosystems).

Phylogenetic Analyses. Multiple sequence alignments of exon 2 of human and rhesus macaque *-DRB* sequences were created by using MacVector version 8.1.1 (Oxford Molecular Group), and phylogenetic analyses were then performed by using PAUP version 4.0b.10 (46). Pairwise distances were calculated by using Kimura-2 parameter and the neighbor-joining method for creating a phylogram. Confidence estimates of grouping were calculated according to the bootstrap method generated from 1,000 replicates.

We thank W. Verduin and J. Anholts for human DNA preparation, D. Devine for editing the manuscript, and H. v. Westbroek for artwork. This study was supported in part by National Institutes of Health Grant 1 1-R24-RR16038-01 (Catalog of Federal Assistance 93.306) and by National Institutes of Health/National Center for Research Resources Project U24-RR1814401.

1. Marsh SG (1998) *Tissue Antigens* 51:467–507.
2. Bontrop RE, Otting N, de Groot NG, Doxiadis GG (1999) *Immunol Rev* 167:339–350.
3. Klein J, Satta Y, O'Huigin C, Takahata N (1993) *Annu Rev Immunol* 11:269–295.
4. Raymond CK, Kas A, Paddock M, Qiu R, Zhou Y, Subramanian S, Chang J, Palmieri A, Haugen E, Kaul R, Olson MV (2005) *Genome Res* 15:1250–1257.
5. Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, Stewart CA, Ashurst JL, Atrazhev AM, Coggill P, Palmer S, *et al.* (2006) *PLoS Genet* 2:e9.
6. Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SG (2003) *Nucleic Acids Res* 31:311–314.
7. Doxiadis GG, Otting N, de Groot NG, Noort R, Bontrop RE (2000) *J Immunol* 164:3193–3199.
8. Slierendregt BL, Otting N, van Besouw N, Jonker M, Bontrop RE (1994) *J Immunol* 152:2298–2307.
9. Doxiadis GG, Otting N, de Groot NG, Bontrop RE (2001) *Immunol Rev* 183:76–85.
10. Khazand M, Peiberg C, Nagy M, Sauermann U (1999) *Tissue Antigens* 54:615–624.
11. Ellis SA, Bontrop RE, Antczak DF, Ballingall K, Davies CJ, Kaufman J, Kennedy LJ, Robinson J, Smith DM, Stear MJ, *et al.* (2006) *Immunogenetics* 57:953–958.
12. Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, Watkins DI (1990) *Immunogenetics* 31:217–219.
13. Geluk A, Elferink DG, Slierendregt BL, van Meijgaarden KE, de Vries RR, Ottenhoff TH, Bontrop RE (1993) *J Exp Med* 177:979–987.
14. Bergstrom TF, Erlandsson R, Engkvist H, Josefsson A, Erlich HA, Gyllensten U (1999) *Immunol Rev* 167:351–365.
15. Mayer WE, O'Huigin C, Klein J (1993) *Proc Natl Acad Sci USA* 90:10720–10724.
16. Moreno-Pelayo MA, Fernandez-Soria VM, Paz-Artal E, Ferre-Lopez S, Rosal M, Morales P, Varela P, Arnaiz-Villena A (1999) *Immunogenetics* 49:843–850.
17. Kenter M, Otting N, de Weers M, Anholts J, Reiter C, Jonker M, Bontrop RE (1993) *Hum Immunol* 36:205–218.
18. Doxiadis GG, Otting N, de Groot NG, de Groot N, Rouweler AJ, Noort R, Verschoor EJ, Bontjer I, Bontrop RE (2003) *Immunogenetics* 55:540–551.
19. Doxiadis GG, Rouweler AJ, de Groot NG, Louwerse A, Otting N, Verschoor EJ, Bontrop RE (2006) *Immunogenetics* 58:259–268.
20. de Groot N, Doxiadis GG, De Groot NG, Otting N, Heijmans C, Rouweler AJ, Bontrop RE (2004) *J Immunol* 172:6152–6157.
21. Doxiadis GG, van der Wiel MK, Brok HP, de Groot NG, Otting N, 't Hart BA, van Rood JJ, Bontrop RE (2006) *Proc Natl Acad Sci USA* 103:5864–5868.
22. Penedo MC, Bontrop RE, Heijmans CM, Otting N, Noort R, Rouweler AJ, de Groot N, de Groot NG, Ward T, Doxiadis GG (2005) *Immunogenetics* 57:198–209.
23. Jones EY, Fugger L, Strominger JL, Siebold C (2006) *Nat Rev Immunol* 6:271–282.
24. Nepom GT, Erlich H (1991) *Annu Rev Immunol* 9:493–525.
25. Tisch R, McDevitt H (1996) *Cell* 85:291–297.
26. Undlien DE, Lie BA, Thorsby E (2001) *Trends Genet* 17:93–100.
27. Dankers MK, Roelen DL, Nagelkerke NJ, de Lange P, Persijn GG, Doxiadis II, Claas FH (2004) *Hum Immunol* 65:13–19.
28. Heemskerk MH, de Paus RA, Lurvink EG, Koning F, Mulder A, Willemze R, van Rood JJ, Falkenburg JH (2001) *Proc Natl Acad Sci USA* 98:6806–6811.
29. Opelz G, Wujciak T, Dohler B, Scherer S, Mytilineos J (1999) *Rev Immunogenet* 1:334–342.
30. Verduyn W, Doxiadis II, Anholts J, Drabbels JJ, Naipal A, D'Amaro J, Persijn GG, Giphart MJ, Schreuder GM (1993) *Hum Immunol* 37:59–67.
31. Andersson G, Larhammar D, Widmark E, Servenius B, Peterson PA, Rask L (1987) *J Biol Chem* 262:8748–8758.
32. Epplen C, Santos EJ, Guerreiro JF, van Helden P, Epplen JT (1997) *Hum Genet* 99:399–406.
33. Riess O, Kammerbauer C, Roewer L, Steimle V, Andreas A, Albert E, Nagai T, Epplen JT (1990) *Immunogenetics* 32:110–116.
34. Bergstrom TF, Engkvist H, Erlandsson R, Josefsson A, Mack SJ, Erlich HA, Gyllensten U (1999) *Am J Hum Genet* 64:1709–1718.
35. Kriener K, O'Huigin C, Tichy H, Klein J (2000) *Immunogenetics* 51:169–178.
36. Trtkova K, Mayer WE, O'Huigin C, Klein J (1995) *Mol Phylogenet Evol* 4:408–419.
37. Bontrop RE, Otting N, Slierendregt BL, Lanchbury JS (1995) *Immunol Rev* 143:33–62.
38. Schreuder GM, Hurley CK, Marsh SG, Lau M, Fernandez-Vina M, Noreen HJ, Setterholm M, Maiers M (2005) *Tissue Antigens* 65:1–55.
39. Marsh SG, Parham P, Barber LD (2000) *The HLA FactsBook* (Academic, London).
40. Lin L, Jin L, Lin X, Voros A, Underhill P, Mignot E (1998) *Tissue Antigens* 52:9–18.
41. Petes TD, Greenwell PW, Dominska M (1997) *Genetics* 146:491–498.
42. Wierdl M, Dominska M, Petes TD (1997) *Genetics* 146:769–779.
43. Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E (1996) *Proc Natl Acad Sci USA* 93:15285–15288.
44. Maueler W, Bassili G, Arnold R, Renkawitz R, Epplen JT (1999) *Gene* 226:9–23.
45. Vowles EJ, Amos W (2004) *PLoS Biol* 2:E199.
46. Swafford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)* (Sinauer, Sunderland, MA), Version 4.