Forensic Science International: Genetics 14 (2015) 168–173

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Sequence-based definition of eight short tandem repeat loci located within the HLA-region in an Austrian population

Eva-Maria Dauber *, Sabine Wenda, Elisabeth Maria Schwartz-Jungl, Barbara Glock, Wolfgang R. Mayr

Department for Blood Group Serology and Transfusion Medicine, Medical University of Vienna, Austria

ARTICLE INFO

Article history:
Received 22 April 2014
Received in revised form 23 September 2014
Accepted 13 October 2014

Keywords:
Sequence structure
Nomenclature
Short tandem repeats
Population study
HLA

ABSTRACT

Sequenced allelic ladders are a prerequisite for reliable genotyping of short tandem repeat (STR) polymorphisms and consistent results across instrument platforms. For eight STR-loci located on the short arm of chromosome 6 (6p21.3), a sequenced based nomenclature was established according to international recommendations. Publicly available reference DNA samples were sequenced enabling interested laboratories to construct their own allelic ladders. Three tetrameric (D6S2691, D6S2678, DQIV), one trimeric (D6S2906) and four dimeric repeat loci (D6S2972, D6S2792, D6S2789, D6S273) were investigated. Apart from the very complex sequence structure at the DQIV locus, three loci showed a compound and four loci a simple repeat pattern. In the flanking regions of some loci additional single nucleotide and insertion/deletion polymorphisms occurred as well as sequence polymorphisms within the repeat region of alleles with the same length. In an Austrian Caucasoid population sample (n = 293) between eight and 22 alleles were found. No significant deviation from Hardy–Weinberg expectations was observed, the power of discrimination ranged from 0.826 to 0.978. The loci cover the HLA-coding region from HLA-A to HLA-DQB1 and can be used for a better definition of HLA haplotypes for population and disease association studies, recombination point mapping, haematopoietic stem cell transplantation as well as for identity and relationship testing.

1. Introduction

Microsatellites or short tandem repeats (STR) are highly polymorphic markers dispersed over the entire genome. Their variability is based on tandemly repeated motifs 2–6 bp in length: alleles different in numbers of reiterations are thus differing in fragment length [1]. Many of them have been identified within the major histocompatibility (MHC) complex situated on the short arm of chromosome 6 (band 6p21.3) [2,3]. HLA-class I and class II loci have been used for paternity testing, but also for the analysis of forensic specimens. As HLA-class I and class II typing is complex and technically challenging, preference was given to the analysis of STRs [4,5]. They can be used for a better definition of the HLA region for population studies [6], recombination point mapping [7], susceptibility studies for HLA-associated diseases [8] and for strategies of donor selection for transplantation [9] and prediction of clinical outcome [10], as well as for identity and paternity testing [11,12]. No consensus information, however, existed on primer pairs and mapping positions; several names were used for loci and alleles, different repeat motifs were chosen in the publications. To overcome these inconsistencies, the informations on HLA-microsatellites were compiled according to the literature [13]: the public dbMHC database provides a fully accessible platform, which identifies 389 primer pairs by UniSTS numbers for amplification of 281 STR-loci, a synonym dictionary for locus names, repetition characteristics, polymorphism and mapping information [14].

Previously, 24 HLA-STRs were studied within the framework of the 13th International HLA Workshop (IHW). Fourteen laboratories reported alleles as fragment sizes, which resulted in nine different values for a given allele corresponding to the nine different groups of participants based on three different protocols/primer pairs and instrument platforms used. The results indicated that fragment sizes did not permit interlaboratory comparison of data. For the

* Corresponding author at: Department for Blood Group Serology and Transfusion Medicine, Medical University of Vienna, Waehringer Guertel 18-20, A-1090 Wien, Austria. Tel.: +43 1 40400 53200; fax: +43 1 40400 53210.
E-mail address: eva-maria.dauber@meduniwien.ac.at (E.-M. Dauber).

http://dx.doi.org/10.1016/j.fsigen.2014.10.009
1872-4973/

compilation of the consensus results a provisional nomenclature, the normalised fragment index (NFI) was used, which simply ranked alleles by size starting with 11 for the smallest allele to allow flexibility to smaller alleles not known so far. From the experience with provisional allele designations used for the integrative comparison of workshop data, the need for a standardised nomenclature based on direct sequencing was concluded [15]. Such a nomenclature was proposed following the 14th IHW to allow uniformity among research publications on the level of locus names (in D6S format), primer identification by UniSTS numbers and allele designations based on the numbers of repeated core motifs [16]. The latter guideline was based on the well-established recommendations [17,18] used in forensic genetics for many years, permitting worldwide comparison of DNA profiles obtained from STR analysis [19,20]. Sequenced allelic ladders are then to be used as external standards for genotyping of microsatellites [18,21].

This paper applies the recommended nomenclature rules to designate alleles at eight HLA-STR-loci: D6S2972, D6S2906 (C3_3_6), D6S2691 (C2_4_4, D6S2939), D6S2678 (C1_4_4, D6S2931), D6S2792 (TNFa), D6S2789 (TNFd), D6S273 and DQIV (M2_4_32) [13,22]. The main goal was to sequence representative DNA samples of the publicly available HLA reference cell lines offered by the International Histocompatibility Working Group (IHWG), to enable a convenient construction of standardised allelic ladders and to characterise further samples for a better definition of sequence and length polymorphisms. Additionally, a population study at the individual loci was carried out in an Austrian population sample to obtain the allele frequencies and the degree of polymorphism for identity and relationship testing. The loci including primer sequences have been selected from the literature according to their repeat length (if possible, preference has been given to tetrameric loci), expected degree of polymorphism and their localisation in order to cover the HLA region from HLA-A to HLA-DQB1 (Fig. 1). Three tetrameric, one trimeric and four dimeric repeat loci encompassing a range of 3 Mb have been chosen: D6S2972 is located 0.26 Mb telomeric, D6S2906 0.44 Mb centromeric of the HLA-A locus, D6S2691 and D6S2678 0.2 and 0.04 Mb telomeric to HLA-C, D6S2792, D6S2789 and D6S273 (0.21, 0.23 and 0.36 Mb centromeric of HLA-B locus, respectively), and the locus DQIV, 0.2 and 0.05 Mb centromeric to the HLA-DRB1 and DQB1 loci, respectively.

2. Materials and methods

2.1. Samples

This study was approved by the Ethics Committee of the Medical University of Vienna. Peripheral blood samples were taken from 150 Austrian Caucasoid families in search of a compatible donor for haematopoietic stem cell transplantation. The disorders involved are not HLA associated. DNA was extracted from peripheral blood lymphocytes (QIAamp® DNA Blood Mini Kit, Qiagen, Hilden, Germany) or from EDTA whole blood (salting out method or QIAamp® EZ1 Blood Kit, Qiagen). Among those 293 unrelated parents were used for population genetic investigations. Additional samples were sequenced, which were not contained in the population sample. Furthermore, 50 purified individual genomic DNA samples from the reference panels of the International Histocompatibility Working group (IHWG) were investigated [23]. Among those have been HLA homozygous samples from consanguineous families and samples of different ethnic origin (Caucasian, African Black, Hispanic, Amerindian and Asian).

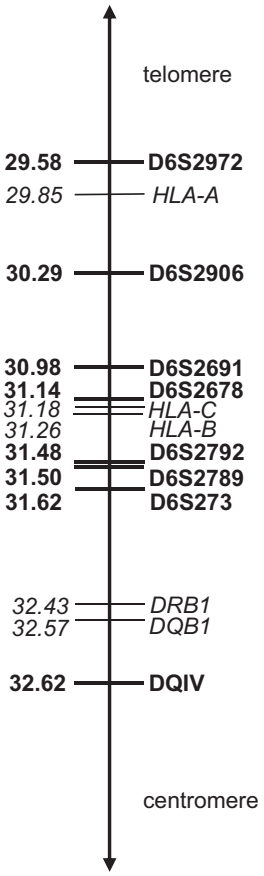

Fig. 1. Chromosomal positions of HLA-A, -B, -C, -DR, -DQ and eight HLA-STR loci.

2.2. PCR

The STR-loci were amplified with fluorescent labelled forward primers in two triplex-PCR reactions, HLA-STR Set 1 (D6S2906, D6S2691, D6S2678) and HLA-STR Set 2 (D6S2972, D6S2792, D6S273), respectively and two additional singleplex PCR reactions (D6S2789 and DQIV) on a GeneAmp 9700 PCR System, 9600 mode, (Applied Biosystems (AB) by Life technologies, Carlsbad, CA, USA). The unlabelled reverse primers for the HLA-STR Set 1 loci were modified a the 5′ end with the sequence motif [-GTGTCTT-] to achieve a complete 3′ adenosine addition of the opposite strand [24]. The primers and PCR conditions of HLA-STR Set 2 and DQIV were based on the protocol used in 13th International HLA Workshop. For DQIV the annealing temperature had to be lowered to 50 °C in a separate singleplex PCR to enable primer binding also to a polymorphic site in the middle of the primer binding region.

2.2.1. HLA-STR Set 1

2 ng template DNA, 0.2–0.8 μM each primer (forward 5′ Fluorescein labelled; 5′ pigtail modification of reverse primers): 0.4 μM D6S2906 (UniSTS:239092), 0.2 μM D6S2691 (UniSTS:239082); 0.8 μM D6S2678 (UniSTS:239073); 200 μM of each nucleotide, 1x GeneAmp PCR Buffer, 0.5 U Amplitaq Gold DNA polymerase (AB) in a reaction volume of 15 μl: 95 °C 11 min for 1 cycle; 94 °C 1 min, 60 °C 1 min, 72 °C 2 min for 30 cycles; 72 °C for 45 min, 1 cycle [25].

2.2.2. HLA-STR Set 2

2 ng template DNA, 0.125–0.4 μM each primer (0.4 μM D6S2972, forward 5′ NED® label, UniSTS:239102; 0.125 μM D6S2792, forward 5′ FAM® label, UniSTS:464285; 0.25 μM