Modern Pathology (2014) 27, 1296–1305

Methods in Pathology

A novel approach for HLA-A typing in formalin-fixed paraffin-embedded-derived DNA

Lisa Villabona[1], Daniel A Leon Rodriguez[1,2], Emilia K Andersson[1], Barbara Seliger[3], Tina Dalianis[1] and Giuseppe V Masucci[1]

[1]Department of Oncology-Pathology, Karolinska Institute, Radiumhemmet, P301057 Karolinska University Hospital, Stockholm, Sweden; [2]Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario, Bogotá, Colombia and [3]Institute for Medical Immunology, Halle, Germany

The aim of this study was to establish a novel approach for human leukocyte antigen (HLA)-typing from formalin-fixed paraffin-embedded-derived DNA. HLAs can be a prognostic factor in cancer and have an extensive polymorphism. This polymorphism is predominantly restricted to exons, which encode the peptide-binding domain of the protein. Formalin-fixed paraffin-embedded material is routinely collected in the clinic and therefore a great source of DNA for genetic analyses. However, its low quality due to fragmentation and nucleotide changes has often created obstacles in designing genetic assays. In this study, we amplified the most polymorphic exons of the HLA-A gene, exons 2, 3, and 4, in 16 formalin-fixed paraffin-embedded samples >10 years old. These tissue samples belonged to patients already HLA-typed by peripheral blood samples at the routine laboratory. Acquired amplification products were used for sequencing, which provided enough information to establish an HLA allele. The same method was applied to DNA extracted from peripheral blood from a healthy volunteer with known HLA type. Of the samples, 14/16 (88%) were successfully typed, in one sample only one of the alleles could be determined, and in one sample no allele could be determined. The amplification of the most polymorphic exons of HLA-A was a successful alternative when DNA quality prevented positive results with previously described methods. The method is usable when an HLA type is needed but the patients are deceased and/or no whole blood samples can be collected. It has thus potential to be used in several fields such as the clinic, research, and forensic science.
Modern Pathology (2014) **27,** 1296–1305; doi:10.1038/modpathol.2013.210; published online 7 February 2014

Keywords: formalin-fixed paraffin-embedded tissue; HLA typing; ovarian cancer sequencing

The highly polymorphic human leukocyte antigen (HLA) locus is located on chromosome 6 and encodes peptide-presenting proteins, which perform the first step in adaptive immunity and respond against foreign pathogens or abnormalities detected in the organism.[1–3] HLA class I molecules, HLA-A, -B, and -C, generally present intracellular peptides, whereas HLA class II molecules, HLA-DR, -DP, and -DQ usually present extracellular peptides.[2] The main feature of HLA genes is their extensive polymorphism. The variation is restricted predominantly to the exons, which code the peptide-binding domain of the protein, ie exons 2 and 3. As a consequence, this polymorphism defines the repertoire of peptides that can potentially bind to a specific HLA allotype, and thereby the hosts ability to respond to infectious agents.[4] The heterozygosity of HLA is advantageous for the survival of the species.

HLA typing is performed routinely for transplantation queries. HLA typing was originally performed by serology, with serum containing antibodies against different HLA antigens. This technique was based on the detection of expressed HLA class I and II molecules on the surface of T and B cells. However, the method requires live lymphocytes and a large panel of antisera.[5] Today DNA-based methods, of which there are basically three approaches, are preferred because of their superior resolution power. The sequence-specific primer PCR (PCR-SSP) uses group-specific primers in order to establish the HLA type. The second, sequence specific oligonucleotide probes (PCR-SSOP), uses hybridization of oligonucleotides to specific alleles. These oligonucleotides have a probe attached to

Correspondence: Dr L Villabona, MD, Department of Oncology-Pathology, Karolinska Institutet, Karolinska Universitetssjukhuset, CCK R8:01, 171 76 Stockholm, Sweden.
E-mail: Lisa.Villabona@ki.se
Received 27 May 2013; revised 1 October 2013; accepted 6 October 2013; published online 7 February 2014

emit light and thus identify the allele. Finally, sequence-based typing (SBT) uses sequencing to establish the identity of the allele. These methods are used routinely and are standardized to be used on good-quality DNA and thus require a large amplification product.

We have previously shown that a specific group of HLA alleles, HLA-A2, are not only over-represented, but also present a strong negative prognostic factor in serous ovarian cancer patients in stages III–IV,[5] only inferior to clinical staging.[6,7] A correlation with prognosis and HLA-A2 has also been reported in prostate cancer and malignant melanoma.[8] Epithelial ovarian cancer is a malignancy with high mortality and low survival time despite treatment with surgery and chemotherapy. Diagnosis occurs often in late clinical stages due to its diffuse symptoms.[9] In some cases, disease progress is so fast that, in research, samples from deceased patients must be used.

Formalin-fixed paraffin-embedded tissue samples have been collected for over 100 years.[10] They have been crucial for the histological diagnosis of malignancies, and are collected routinely for this purpose. This has created large biobanks that can potentially provide valuable DNA for retrospective studies. However, this formalin-fixed paraffin-embedded DNA has some disadvantages. DNA degradation, fragmentation ranging from <150 bp[15] up to 400 bp[12] and even base alterations can occur during e.g. the variation of fixation procedures, the prolonged time from tissue extraction to fixation, or during the storage of samples sometimes for decades.[12] In addition, DNA, protein crosslinkage and the presence of inhibiting substance remnants coming from extraction reactions and contamination can yield problems when performing PCRs or sequencing.[13]

To establish the HLA types of deceased patients from formalin-fixed paraffin-embedded DNA, we have here developed a novel SBT-based method for HLA typing. By only using the most polymorphic parts of the gene for sequencing, enough information can be obtained to determine an HLA-A type.

Materials and methods

Patient Description

Sixteen unrelated Swedish women diagnosed with relapsing or progressive ovarian cancer and receiving treatment at Karolinska University Hospital from 2004 to 2005 were asked for a blood sample. With the patients consent, HLA-A, -B, -Cw and DRB1 typing was performed by PCR-SSP by the European Federation of Immunology (EFI) certified immunological laboratory at Karolinska University Hospital. All women had undergone surgery previous to this study and their tissue material was formalin-fixed and paraffin-embedded routinely by the SWEDAC certified pathological laboratory at Karolinska University Hospital. The laboratory uses 4% formaldehyde for fixation and fixation times are between 24 and 72 h, depending on the size of the material. Slides of 10 μm thickness where cut from paraffin blocks and put in a 1.5 ml microcentrifuge tube instead of mounting them on slides. For each block, an additional slide of 4 μm thickness was mounted on a slide and stained with hematoxylin.

The samples of the patients were coded KT1-16. The additional slides were analyzed by a pathologist in order to determine that some normal tissue was present in each sample. One healthy Swedish woman was asked for a blood sample (coded as VB1) as a positive control for amplification. This woman had also been previously HLA-typed.

DNA Extraction

DNA was extracted from formalin-fixed paraffin-embedded samples according to the protocol for 'High Pure RNA extraction kit' (Roche, Molecular Biochemicals, Mannheim, Germany), but omitting the DNase treatment. The paraffin was removed by xylene and ethanol. The tissue pellet was disrupted by overnight treatment at 55 °C in 10 μl tissue lysis buffer, 16 μl 10% SDS, and 40 μl Proteinase K (20 mg/ml). On day 2, the DNA was bound to a membrane in collection tubes, washed, and then eluted. Additional treatment with 40 μl Proteinase K, 18 μl 10% SDS and 10 μl lysis buffer was performed (1 h, 55 °C) for removal of traces of proteins. The DNA was bound to the membrane in a fresh collection tube, washed, eluted followed by determination of the DNA amount and purity using the NanoDrop technology.

One blood sample was collected from a healthy volunteer with known HLA type. DNA from blood was extracted by the commercial kit Quiagen DNeasy Blood & Tissue kit (Qiagen AB, Sollentuna, Sweden) according to manufacturer's protocol and measured by NanoDrop technology. Extracted DNA was diluted with distilled water to a final concentration of 30 ng/μl.

Amplification and Primer Design

Primers were designed to amplify exons 2, 3, and 4 from HLA-A. Sequences from all HLA-A alleles were obtained from IMGT-HLA database[14] and aligned into a document. The primers were designed from the conserved regions for all alleles, except HLA-A*23:11 N. Each primer was designed using Gene Runner™ Software version 3.05 (Hasting Software Inc.). For exon 2, four primers were designed, two forward and two reverse, whereas for exon 3, three primers were designed, two forward and one reverse (Figure 1). For exon 4, five primers were designed, dividing the amplification of the exon into two parts, the first part with two forward primers and one reverse primer, while the second part had one

forward and one reverse primer (Figure 1). The distinct variants of the primers were determined according to the sequence changes between different alleles (see Table 1).

Amplification

A reaction mix was created consisting of the following: 5 μl 10 $\times$ amplification buffer, 8 μl DNTP (1.25 mM), 3 μl $MgCl_2$ (25 mM), 3 μl forward primer (10 pmol/μl), 3 μl reverse primer (10 pmol/μl), 22.8 μl distilled water, 0.2 μl Taq polymerase, and 5 μl sample DNA (150 ng) making it a total volume of 50 μl. For exon 2 amplification, the primers utilized were A2-1F or A2-2F and A2-1R or A2-2R. For exon 3, the primers were A3-A1F or A3-A2F, and A3-2R. Finally, for exon 4 the primers were A4-1F or A4-2F, and A4-2RH for the first half; A4-3FH and A4-1R for the second half. Amplification was run in an automated thermocycler (GeneAmp PCR system 9700, Applied Biosystems). The cycles consisted of an initial denaturation of 2 min at 94 °C, followed by 40 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 45 s. Water was used as a negative control, and DNA extracted from a fresh blood sample was used as a positive control. PCR products were run on a 2.0% agarose gel stained with ethidium bromide and visualized under UV light. The expected product was 292 bp for exon 2, 258 bp for exon 3, 166 bp for the first part of exon 4 and finally 201 bp for the second part of exon 4.

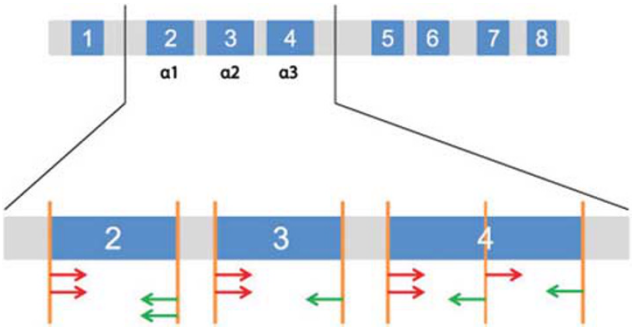

Figure 1 Overview of HLA-A gene. Primers were designed to cover all alleles by amplifying exons 2, 3, first, and second half of exon 4.

Sequencing

All amplified products were cleaned with ExoSAP-IT kit™ (Affymetrix Inc.) and sequenced with Big Dye™ (Applied Biosystems™). Each amplified product was amplified using forward and reverse primers separately, according to manufacturer's protocol.

All samples were read in an automatic sequencer (Abi Prism 3100 DNA Analyzer) in an optical 96-well plate.

Sequence Analysis

The obtained sequences were analyzed on Chromas Lite™ software version 2.01 (Technelysium Pty Ltd). Each sample had two corresponding sequences, forward and reverse, according to the primer used in the sequencing reaction. Each sequence was exported to a FASTA format. Reverse sequences were translated into complementary strands using Gene Runner™ Software version 3.05 (Hasring Software, Inc.). Using clustalw 2 tool[15], the two sequences were aligned. The remaining arms from the reverse and forward strands were added to a consensus sequence. The final sample sequence was analyzed using IMGT-HLA Blast searching tool.[14] Heterozygosity was detected by looking for double peaks in each sequence graphic. By aligning the different sequences with known HLA-A alleles registered by the IMGT organization, the HLA-A type was determined.

Results

HLA-A Exon Sequencing-Based PCR of Formalin-Fixed-Paraffin-Embedded Samples Yield Amplicons

In total, 14/16 patients could be assigned two HLA-A alleles, 1/16 patients could be assigned one HLA-A allele, while for the remaining patient no HLA-A alleles could be determined. Below further details are presented.

All samples, except the blood derived VB1, yielded a DNA concentration of >100 ng/μl. Table 2 shows the concentrations and purity measurements from the analysis with Nanodrop™. All samples had an absorbance value of >1.8 for 260/280, which is generally accepted as pure DNA,

Table 1 Primers used for each exon and their sequences

Forward primers	*A2-1F* 5′-TCTCAGCCA CTGCTCGCC-3′	*A3-1F* 5′-GGTTCTCACACCGTCCAGA-3′	*A4-1F* 5′-GCGCCCCAAAACGCATATG-3′
	A2-2F 5′-TCTCAGCCACTCCTCGTC-3′	*A3-2F* 5′-GGTTCTCACACCATCCAGA-3′	*A4-2F* 5′-CCCCCCCAAGACACATATG-3′
			A4-3FH 5′-TTCTACCCTGCGGAGATCAC-3′
Reverse primers	*A2-1R* 5′-GTCCTCGCTCTGGTTGTAG-3′	*A3-2R* 5′-CTTCCCGTTCTCCAGGTATC-3′	*A4-1R* 5′-ATCCCCTGCAGGCCTGGT-3′
	A2-2R 5′-GCCCTCGCTCTGGTTGTAG-3′		*A4-2RH* 5′-CCATCTCAGGGTGAGGGG-3′

according to manufacturer's protocol. Five samples, KT4, KT10, KT14, KT15, and KT16 had an absorbance value of <2.0 for 260/230, which may indicate the presence of contaminants, according to manufacturer's protocol (Table 2).

From the extracted DNA, products were amplified up to the size of 292 bp. For exons 2 and 3, 15/16 samples yielded products of $\sim$292 bp and 258 bp, respectively (Table 3). For exon 4, 13/16 yielded a product of $\sim$158 bp for the first half of the reaction, whereas 11/16 yielded a product of $\sim$193 bp for the second part (Table 3). The positive control VB1 yielded products with all primers (Table 3).

All products were visualized on a 2% agarose gel and examples of different amplification products, showing that all primer pairs could yield an amplification product, are illustrated for samples KT11-16 and VB1 in Figure 2.

For all samples, except KT7, KT9, and KT10, at least one amplification product was obtained for all three exons. The products of KT3 and KT9 appeared to be in a smaller quantity compared with the 100 bp ladder and other amplification products from the same reaction mix. Only one sample, KT10, did not amplify in any of the reactions.

Table 2 Results from concentration analysis by Nanodrop™

Sample ID	*Avarage conc. (ng/ul)*	*260/280*	*260/230*
VB1	62.13	1.96	2.18
KT1	525.51	1.91	2.17
KT2	606.49	1.91	2.09
KT3	838.75	1.89	2.24
KT4	957.98	1.88	1.91
KT5	1047.52	1.94	2.1
KT6	1023.86	1.95	2.16
KT7	530.63	1.92	2.2
KT8	504.5	1.92	2.16
KT9	253.69	1.88	2.1
KT10	869.28	1.9	1.72
KT11	421.59	1.91	2.08
KT12	260.56	1.95	2.07
KT13	189.5	1.95	2.18
KT14	255.54	1.97	1.77
KT15	487.49	1.82	1.90
KT16	683.19	1.87	1.95

Sequencing of the Obtained HLA-A Exon PCR-Amplified Amplicons

The obtained amplicons were thereafter sequenced. Amplicon sequences obtained for exon 2, gave clearer graphics, were more easily read and had less background interference compared to amplicon sequences obtained from exons 3 and 4 (Figure 3a–c, respectively). Nevertheless, the principal sequence which corresponded to the higher peaks in each sequence graphic was exported to a FASTA format and correlated to HLA-A alleles in all samples.[14] All samples displayed two different sequences according to the primers used in the amplification protocol. For each exon, different forward and reverse primers were used in order to amplify the different alleles. Each obtained sequence was exported to a FASTA format for further analysis. The reverse primer sequences were translated to complementary strands and then aligned with the corresponding forward primer sequence using *clustalw 2* tool (Figure 4).[15]

With these alignments, a complete sequence was obtained and used to determine the HLA-A allele with $>90\%$ accuracy (Figure 5). The final sequence,

Table 3 Overview of positive reactions and primers used in amplification

	EXON 2				*EXON 3*		*EXON 4*		
Primers/ samples	*A2-1F/1R*	*A2-2F/2R*	*A2-1F/2R*	*A2-2F/1R*	*A3-1F/2R*	*A3-2F/2R*	*A4-1F/2RH*	*A4-2F/2RH*	*A4-3FH/1R*
VB1	+	+	+	+	+	+	+	+	+
KT1	+	+	+	−	+	−	+	+	+
KT2	+	−	+	−	+	+	+	+	−
KT3	−	+	+	+	+	+	+	+	−
KT4	−	+	−	−	+	+	+	+	+
KT5	+	+	+	+	+	+	+	+	+
KT6	+	+	−	−	+	−	−	+	+
KT7	+	+	+	+	+	+	−	−	−
KT8	+	+	+	+	+	+	−	+	+
KT9	+	+	+	−	+	+	−	−	−
KT10	−	−	−	−	−	−	−	−	−
KT11	+	+	+	+	+	+	+	+	+
KT12	+	+	+	+	+	+	+	+	+
KT13	+	+	+	+	+	+	+	+	+
KT14	+	+	+	+	+	+	+	+	+
KT15	+	+	+	+	+	+	−	+	+
KT16	+	+	+	+	+	+	+	+	+
	14	15	13	10	15	13	10	13	11

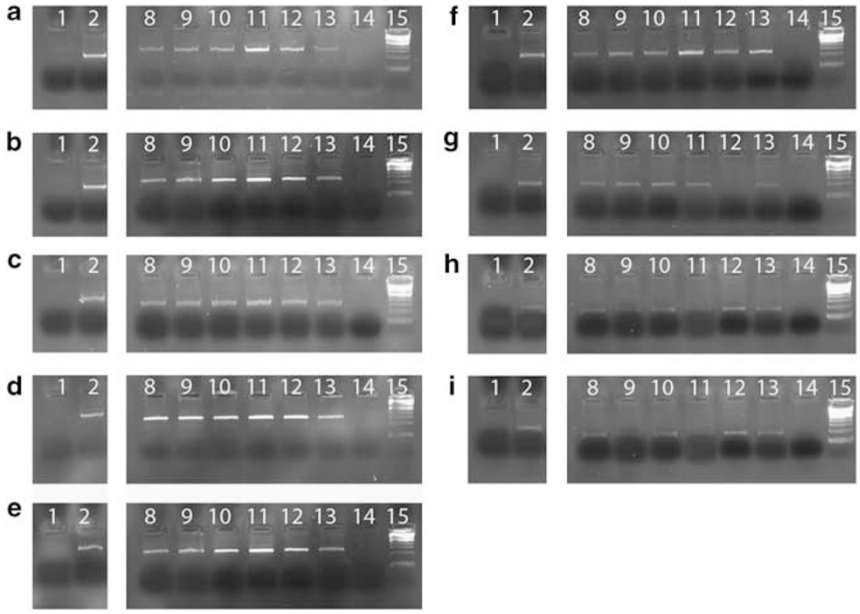

Figure 2 Formalin-fixed paraffin-embedded DNA amplification products. Lane 1, negative control (clean room); lane 2, PC (VB1); lane 8, sample KT 11; lane 9, sample KT12; lane 10, sample KT13; lane 11, sample KT14; lane 12, sample KT 15; lane 13, sample KT16; lane 14, negative control (PCR room); lane 15, 100 bp ladder. The different primer pairs used on different protocols are listed as (**a**) A2-1F and A2-1R; (**b**) A2-2F and A2-2R; (**c**) A2-1F and A2-2R, (**d**) A2-2F and A2-1R; (**e**) A3-1F and A3-2R; (**f**) A3-2F and A3-2R; (**g**) A4-1F and A4-2RH; (**h**) A4-2F and A4-2RH; (**i**) A4-3FH and A4-1R.

obtained by the consensus sequence and external arms, was then used for the blast tool in the IMGT-HLA database. The use of different amplification primers, gave more than one sequence for each sample. In almost all cases, the sequence graphics displayed double peaks that corresponded to allele changes in the IMGT-HLA database (Figure 6).

From 16 patients, 15 (88%) could be assigned an HLA-A type. From one patient, sample KT9, only one allele could be assigned. In this case the patient was HLA-A*01/26 and only HLA-A*01 was identified from the given sequence. However, two other patients, KT2 and KT13 were also HLA*A26 positive and in their case the allele was identified. From one patient, KT10, no alleles could be assigned (Table 4).

Discussion

A novel approach for HLA-A typing using DNA from formalin-fixed paraffin-embedded material, amplifying and sequencing the most polymorphic regions, exons 2–4 of the HLA-A gene, was successful in collecting enough information to establish the HLA-A alleles of a majority of patients (14/16) with formalin-fixed paraffin-embedded tissue samples. To evaluate the accuracy of this new HLA-A typing method it was compared to a standard procedure, PCR-SSP, and by using a good quality DNA control.

The low quality of formalin-fixed paraffin-embedded-derived DNA has long been known.[10,16,17] Klopfleisch *et al*[10] reviewed the effect of formalin fixation on DNA, RNA, and proteins and also discuss different alternative fixatives and extraction protocols. They conclude that when amplifying formalin-fixed paraffin-embedded DNA, one should use short products and cautiously assess mutations due to the potential detection of pseudo-mutations.

Fairley *et al*[16] focuses on using formalin-fixed paraffin-embedded cancerous tissue when identifying biomarkers. Also here the suggestions are to use short products when amplifying genes and analyzing mutations with caution. These are things we have considered when establishing this method, as we use shorter amplification products and use the sequence analysis to align with already known somatic mutations. Hewitt *et al*[17] made some very good points on the treatment of samples during the process of formalin fixation and embedding in paraffin. Even though the focus is on RNA, many conclusions can be drawn to nucleic acids in general. These points have been considered when describing the patient material in this paper, but some information is no longer available to us and thus hard to control. An example of this is the time from ischemia in the tissue until the tissue is put in a formalin bath. The most important point from this paper, however, is that quality of DNA derived from formalin-fixed paraffin-embedded material will never match the golden standard of DNA extracted from fresh frozen tissue. But this is not the goal, the goal is rather to work with what is available, and this has been our intent with this method.

The low quality of formalin-fixed paraffin-embedded derived DNA was apparent also in our study since it was not possible to HLA-A type all samples as will be discussed in more detail below.

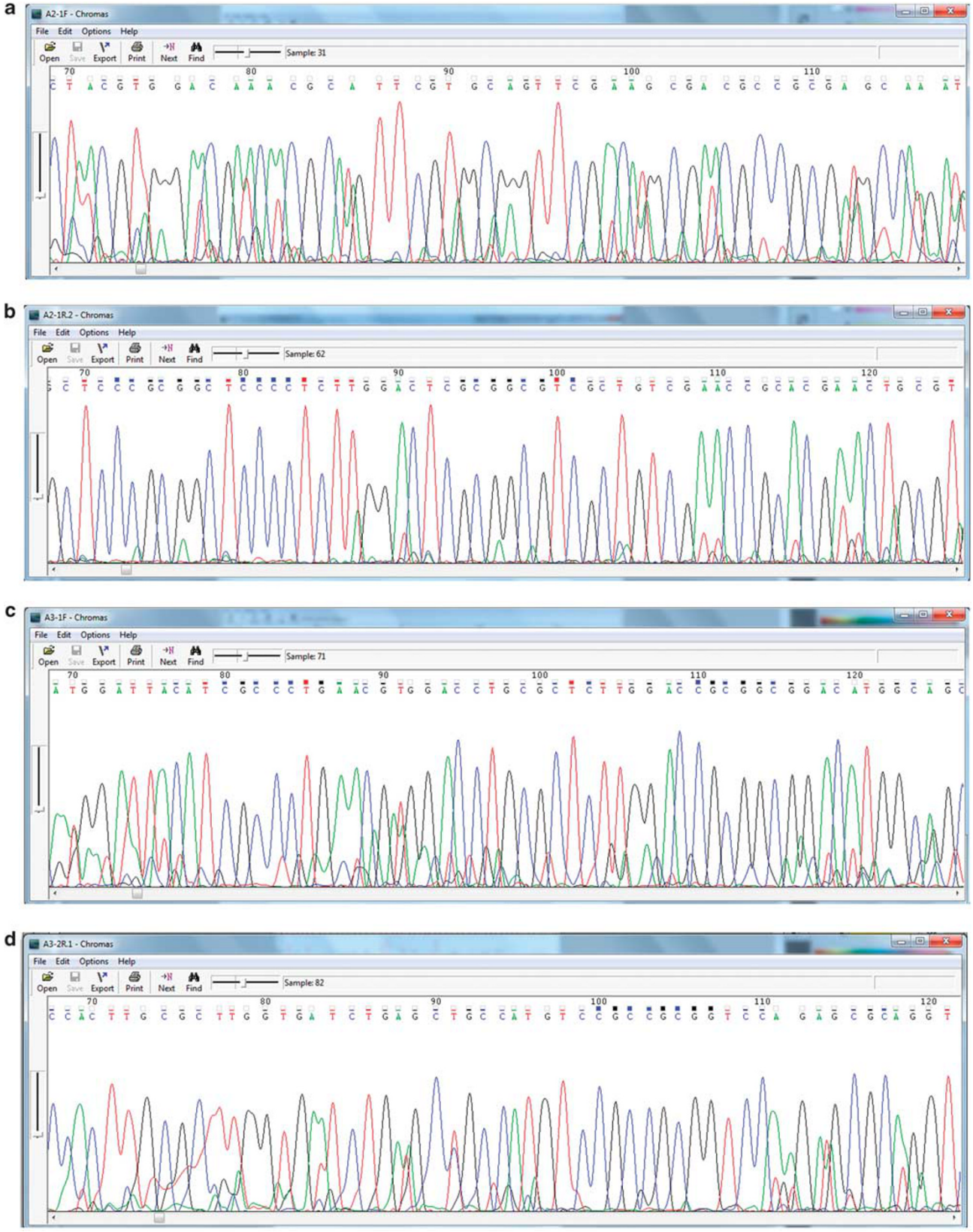

Figure 3 (**a**) Sample KT15 HLA-A exon 2 sequence graphics. Primer used was A2-1F. (**b**) Sample KT15 HLA-A exon 2 sequence graphics. Primer used was A2-1R. (**c**) Sample KT15 HLA-A exon 3 sequence graphics with background interference. Primer used was A3-1F. (**d**) Sample KT15 HLA-A exon 3 sequence graphics with background interference. Primer used was A3-1R.

```
>K16FW
CGCGTACTAGAGTATTTCTTCCATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCA
TCGCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGCGCCCCGAGGCC
AAAGGTGGGAGCCGCGGGCGCCGTGTATAGAGCAGGAGGGGCCGGAGTATTGGGACCGGG
AGACACGGAATGTGAAGGCCCAGTCACAGACTGACCGAGTGGACCTGGGGACCCTGCGCG
GCTACTACAACCAGAGCGAGGACA

>K16RV
CTCAGCCACTGCTCGCCCCCAGGCTCCCACTCCATGAGGTATTTCTTCACATCCGTGTCC
CGGCCCGGCCGCGGGGAGCCCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTC
GTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAG
CAGGAGGGGCCGGAGTATTGGGACCAGGAGACACGGAATGTGAAGGCCCAGTCACAGACG
ACCGAGGGTACCAGCACTGTACG


CLUSTAL 2.1 multiple sequence alignment


K16FW           ----------------------CGCGTACT----AGAGTATTTCTTC-CATCCGTGTCC 32
K16RV           CTCAGCCACTGCTCGCCCCCAGGCTCCCACTCCATGAGGTATTTCTTCACATCCGTGTCC 60
                                     * *  ***       ********** ***********

K16FW           CGGCCCGGCCGCGGGGAGCCCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTC 92
K16RV           CGGCCCGGCCGCGGGGAGCCCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTC 120
                ************************************************************

K16FW           GTGCGGTTCGACAGCG-CGCCCCGAGGCCAAAGGTGGGAGCCGCGGGCGCCGTGTATAGA 151
K16RV           GTGCGGTTCGACAGCGACGCCGCGAG-CCAGAGGATGGAGCCGCGGGCGCCGTGGATAGA 179
                **************** **** **** *** ***  ****************** *****

K16FW           GCAGGAGGGGCCGGAGTATTGGGACCGGGAGACACGGAATGTGAAGGCCCAGTCACAGAC 211
K16RV           GCAGGAGGGGCCGGAGTATTGGGACCAGGAGACACGGAATGTGAAGGCCCAGTCACAGAC 239
                ************************* ********************************

K16FW           TGACCGAGTGGACCTGGGGACCCTGCGCGGCTACTACAACCAGAGCGAGGACA 264
K16RV           -GACCGAGGGTACCAG----CACTGTACG------------------------ 263
                 ******* * *** *    * ***  **


>K16SeqCons
CTCAGCCACTGCTCGCCCCCAGGCTCCCACTCCATGAGGTATTTCTTCACATCCGTG
TCCCGGCCCGGCCGCGGGGAGCCCCGCTTCATCGCCGTGGGCTACGTGGACGACACG
CAGTTCGTGCGGTTCGACAGCGCGCCCCGAGGCCAAAGGTGGGAGCCGCGGGCGCCG
TGTATAGAGCAGGAGGGGCCGGAGTATTGGGACCGGGAGACACGGAATGTGAAGGCC
CAGTCACAGACTGACCGAGTGGACCTGGGGACCCTGCGCGGCTACTACAACCAGAGC
GAGGACA
```

Figure 4 Sample KT16 forward and reverse sequence analysis from exon 2, A2-2F and A2-2R primers (A2-2/2). The obtained sequence was aligned, and the complementary sequence, along with the outside arms, was analyzed using clustalw 2 tool.[15] (*) shows nucleotide identity and (-) shows gaps between both sequences.

```
>IMGTHLA:HLA01718 A*03:01:01:02N 1163 bp
        Length = 1163

  Plus Strand HSPs:

 Score = 1263 (195.6 bits), Expect = 1.3e-53, P = 1.3e-53, Group = 1
 Identities = 263/273 (96%), Positives = 263/273 (96%), Strand = Plus / Plus

Query:     20 CAGGCTCCCACTCCATGAGGTATTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGC 79
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:     71 CGGGCTCCCACTCCATGAGGTATTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGC 130

Query:     80 CCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCG-CG 138
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct:    131 CCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACG 190

Query:    139 CCCCGAGGCCAAAGGTGGGAGCCGCGGGCGCCGTGTATAGAGCAGGAGGGGCCGGAGTAT 198
              || |||| ||| |||  ||||||||||||||||||| ||||||||||||||||||||||||
Sbjct:    191 CCGCGAG-CCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGGCCGGAGTAT 249

Query:    199 TGGGACCGGGAGACACGGAATGTGAAGGCCCAGTCACAGACTGACCGAGTGGACCTGGGG 258
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:    250 TGGGACCAGGAGACACGGAATGTGAAGGCCCAGTCACAGACTGACCGAGTGGACCTGGGG 309

Query:    259 ACCCTGCGCGGCTACTACAACCAGAGCGAGGAC 291
              |||||||||||||||||||||||||||||| |
Sbjct:    310 ACCCTGCGCGGCTACTACAACCAGAGCGAGGCC 342
```

Figure 5 Result obtained from the IMGT-HLA database Blast tool. In this example, the sequence analyzed showed a nucleotide similarity to HLA*A03 of 99%. (|) shows equal nucleotides between analyzed and reference sequence.

Nevertheless, taking this into consideration along with the fact that the polymorphism in HLA class I is mainly concentrated to exons 2 and 3,[3] we designed primers for exons 2 and 3 separately. In addition, in order to avoid ambiguities, related to allele differences defined outside these exons, we

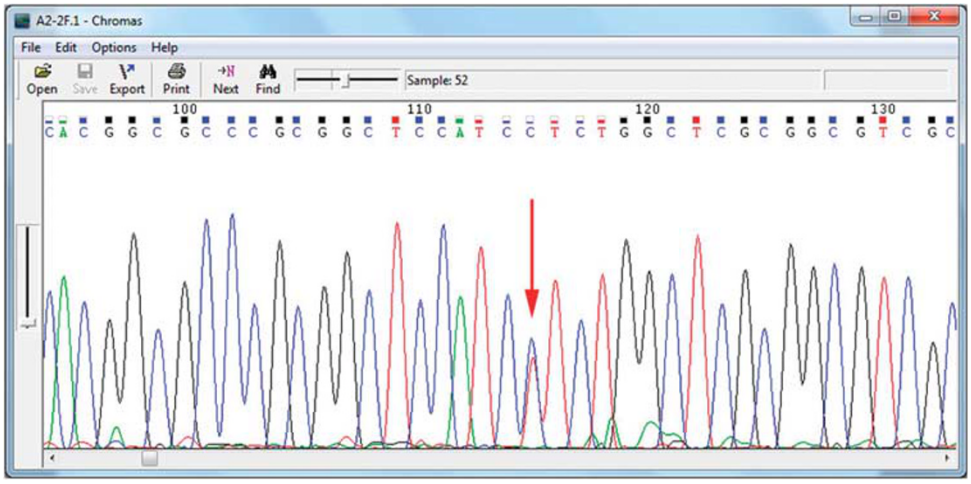

Figure 6 Sample KT15 exon 2 sequence graphic (A2-1F/1R). The arrow indicates the double peaks which correspond to heterozygocity.

Table 4 Overview of sequencing results compared with previous PCR-SSP results

	EXON 2				*EXON 3*		*EXON 4*			*HLA-A genotypes*	
Primers/ samples	*A2-1F/1R*	*A2-2F/2R*	*A2-1F/2R*	*A2-2F/1R*	*A3-1F/2R*	*A3-2F/2R*	*A4-1F/2RH*	*A4-2F/2RH*	*A4-3FH/1R*	*FFPE method*	*PCR-SSP*
VB1	NA	**A*02/24**	[a]	[a]	**A*02/24**	[a]	A*02	A*24	A*02	A*02/24	A*02/24
KT1	**A*03/25**	NA	A*25	NA	A*25	NA	A*25	A*03	NA	A*03/25	A*03/25
KT2	A*03	NA	NA	NA	A*03	A*26	A*26	A*03	NA	A*03/26	A*03/26
KT3	NA	A*02	A*01	A*02	A*02	A*01	A*02	NA	NA	A*01/02	A*01/02
KT4	NA	NA	A*03	A*02	A*02	A*03	A*02	A*03	A*03	A*02/03	A*02/03
KT5	A*03	A*02	A*03	A*02	A*02	NA	A*02	A*03	A*02	A*02/03	A*02/03
KT6	**A*01/24**	A*24	NA	NA	**A*01/24**	NA	NA	A*24	NA	A*01/24	A*01/24
KT7	A*01	A*02	A*01	A*01	A*02	A*01	NA	NA	NA	A*01/02	A*01/02
KT8	**A*01/03**	NA	**A*01/03**	NA	A*01	A*01	NA	NA	A*03	A*01/03	A*01/03
KTT9	A*01	NA	A*01	NA	A*01	A*01	NA	NA	NA	A*01/–	A*01/26
KT10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	A*01/68
KT11	A*11	A*02	A*11	A*02	A*02	A*11	A*02	A*11	A*02	A*02/11	A*02/11
KT12	A*25	A*02	A*25	A*02	A*25	A*25	A*02	NA	NA	A*02/25	A*02/25
KT13	A*26	A*02	A*26	A*02	A*02	A*26	A*02	NA	A*02	A*02/26	A*02/26
KT14	A*01	A*02	A*01	A*02	A*01	A*01	A*02	NA	A*02	A*01/02	A*01/02
KT15	**A*01/03**	NA	**A*01/03**	NA	A*01	A*01	NA	A*01	A*01	A*01/03	A*01/03
KT16	A*03	A*02	A*03	A*02	A*03	A*03	A*02	A*03	NA	A*02/03	A*02/03

Abbreviation: NA: no amplification product or poor sequencing result.
[a]VB1 sample used as positive control for amplification reactions. Results in bold correspond to genotypes established by editing FASTA formats according to sequence graphics.

also analyzed exon 4.[18] Furthermore, one must take into account that the applied method here uses already known polymorphisms in the search of an allele assessment, making base alterations and deletions more easily identified. However, this presents a problem if a patient is the carrier of an unknown allele.

For one sample, KT10, no amplification product could be obtained. For two samples, KT7 and KT9, no amplification product was obtained for exon 4. In the case of sample KT9, the information from exon 4 may have contributed to the identification of its other allele. The lack of amplification product in some samples we hypothesize could be associated with different fixation procedures that could lead to different types of damage in the DNA. Differences between fixation protocols as well as laboratory organizations also bring differences to the DNA quality.[11,12] In this particular study, the fixation protocol was not controlled, the information we have from this protocol is described above. Other problems reported with the quality of formalin-fixed paraffin-embedded-derived DNA are base alterations and base removal,[12] thus making primer hybridization difficult. This has led to discussion on the authenticity of previously reported mutations when sequencing formalin-fixed paraffin-embedded-derived DNA.[19] This might also partially explain the absence of product in these reactions.

Finally, the difference in extraction protocols can be of importance. Most studies use commercial kits and obtain different DNA length.[10,16,17] Turashvili *et al*[20] published a comparison study on different

extraction protocols when amplifying DNA from formalin-fixed paraffin-embedded material. A difference could be seen between the extraction kits, however, for almost all kits DNA amplification was successful when the product was <200 bp. Although the samples were not older than 7 days. Fifteen–year-old archival samples were also analyzed, but no amplification was successful. In our case, only one column-based commercial kit was used, all samples were >10 years old. This may however explain the absence of product in some reactions and it is possible that other extraction kits may be more useful in specific cases.

Some of the problems described above might be circumvented at an earlier stage of the procedure if DNA quality was assessed after extraction, by for example an Agilent Bioanalyzer™. In our study, we determine the fate of the samples after a Nanodrop analyses. As all samples had a 260/280 ratio of >1.8, there was no indication that any samples would not be amplifiable. Factors other than nucleic acid purity, such as high pH, can give a false high 260/280 ratio.[21] This could be an explanation as to why all samples may seem to have good purity.

The Agilent Bioanalyzer™ uses the same technique as a gel electrophoresis and separates the amplified material according to size; however, it is done in a much smaller scale, no previous amplification is needed and a fluorescent dye is introduced to the nucleic acids as they move with an electric current. Not only does this provide purity information but also the fragmentation of the sample and thus weather or not it is amplifiable, according to the manufacturer. This would give our method an earlier indication of how to proceed with a certain sample, and possibly omit the steps of gel electrophoresis and sequencing before excluding it.

Despite the above limitations, our new approach of determining HLA-A type in low-quality DNA has great potential to be expanded to other polymorphic regions of the genome. The obvious choices are the other HLA-genes for class I and II antigens, but it may also allow setting up analysis for the Killer-cell immunoglobuline-like receptors and for T-cell receptors.

Previous reports have shown HLA typing in formalin-fixed paraffin-embedded material.[22–25] Recently, eg Ota *et al*[24] showed successful HLA typing in different types of tissue after 24–96 h of formalin fixation, using a PCR-SSOP method for both HLA class I and II typing. This method was also attempted for HLA-typing our samples with no results, which possibly was due to the age of our samples, which may have accumulated more DNA damage. Furthermore, in two studies from 1997, Bateman *et al*,[22,23,25] showed that the PCR-SSP approach could generate HLA typing in formalin-fixed paraffin-embedded material. In both studies only HLA class II was determined, and in the first only skin tissue was used with conventional PCR-SSP, whereas in the second a variety of tissues were used, but only totally five samples were run with a nested PCR-SSP. These approaches were also unsuccessfully attempted on our samples, and we have no obvious explanation. However, it could be possible that the discrepancy might be explained by differences in our DNA extraction methods and fixation procedure. Finally, Lee *et al* showed in 1999,[23] that a PCR-SSOP method could be used, but needed to be confirmed with a nested PCR-SSP. This demonstrates again that the method required depends on the quality of the DNA in the samples tested.

In conclusion, our approach amplifying and sequencing the most polymorphic exons of HLA-A offers an alternative when DNA quality prevents successful results with these previously described methods. The method is usable when an HLA type is needed but the patients are deceased and/or no whole blood samples can be collected. It also has the potential to be used in several clinical disciplines, for research as well as for forensic science.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

1 Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature 1999;401:921–923.
2 Goldsby RA, Kindt TJ, Osborne BA, *et al.* Immunology, 5th edn. W. H. Freeman: New York, NY, USA, 2003; chapt 7, 161–182.
3 Little AM, Stern PL. Does HLA type predispose some individuals to cancer? Mol Med Today 1999;8: 337–342.
4 Little AM, Parham P. Polymorphism and evolution of HLA class I and II genes and molecules. Rev Immunogenet 1999;1:105–123.
5 Howell WM, Carter V, Clark B. The HLA system: immunobiology, HLA typing, antibody screening and crossmatching techniques. J Clin Pathol 2010;63: 387–390.
6 Gamzatova Z, Villabona L, Dahlgren L, *et al.* Human leucocyte antigen (HLA) A2 as a negative clinical prognostic factor in patients with advanced ovarian cancer. Gynecol Oncol 2006;103:145–150.
7 Andersson E, Villabona L, Bergfeldt K, *et al.* Correlation of HLA-A02* genotype and HLA class I antigen down-regulation with the prognosis of epithelial ovarian cancer. Cancer Immunol Immunother 2012;61:1243–1253.
8 De Petris L, Bergfeldt K, Hising C, *et al.* Correlation between HLA-A2 gene frequency, latitude, ovarian and prostate cancer mortality rates. Med Oncol 2004;21:49–52.
9 Cannistra SA. Cancer of the ovary. N Engl J Med 2004; 351:2519–2529.
10 Klopfleisch R, Weiss AT, Gruber AD. Excavation of a buried treasure—DNA, mRNA, miRNA and protein analysis in formalin fixed, paraffin embedded tissues. Histol Histopathol 2011;26:797–810.
11 Gilbert MT, Sanchez JJ, Haselkorn T, *et al.* Multiplex PCR with minisequencing as an effective

high-throughput SNP typing method for formalin-fixed tissue. Electrophoresis 2007;28:2361–2367.
12 Lehmann U, Kreipe H. Real-time PCR analysis of DNA and RNA extracted from formalin-fixed and paraffin-embedded biopsies. Methods 2001;25:409–418.
13 Farrugia A, Keyser C, Ludes B. Efficiency evaluation of a DNA extraction and purification protocol on archival formalin-fixed and paraffin-embedded tissue. Forensic Sci Int 2010;194:e25–e28.
14 Robinson J, Mistry K, McWilliam H, *et al.* The IMGT/HLA database. Nucleic Acids Res 2011;39(Database issue):D1171–D1176.
15 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
16 Fairley JA, Gilmour K, Walsh K. Making the most of pathological specimens: molecular diagnosis in formalin-fixed, paraffin embedded tissue. Curr Drug Targets 2012;13:1475–1487.
17 Hewitt SM, Lewis FA, Cao Y, *et al.* Tissue handling and specimen preparation in surgical pathology. Arch Pathol Lab Med 2008;132:1929–1935.
18 Adams SD, Barracchini KC, Chen D, *et al.* Ambiguous allele combinations in HLA Class I and Class II sequence-based typing: when precise nucleotide sequencing leads to imprecise allele identification. J Transl Med 2004;2:30–36.
19 Soussi T. Advances in carcinogenesis: a historical perspective from observational studies to tumor genome sequencing and TP53 mutation spectrum analysis. Biochim Biophys Acta 2011;1816:199–208.
20 Turashvili G, Yang W, McKinney S, *et al.* Nucleic acid quantity and quality from paraffin blocks: Defining optimal fixation, processing and DNA/RNA extraction techniques. Exp Mol Pathol 2012;92:33–43.
21 Wilfinger WW, Mackey K, Chomczynski P. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. Biotechniques 1997;22:474–476; 478-81.
22 Bateman AC, Hemmatpour SK, Theaker JM, *et al.* Nested polymerase chain reaction-based HLA class II typing for the unique identification of formalin-fixed and paraffin-embedded tissue. J Pathol 1997;181:228–234.
23 Lee AH, Bateman AC, Turner SJ, *et al.* HLA class II DRB1 and DQB1 allelic polymorphism and sclerosing lymphocytic lobulitis of the breast. J Clin Pathol 1999;52:445–449.
24 Ota M, Shimada K, Asamura H, *et al.* Highly sensitive HLA-DNA typing from formalin-fixed and paraffin-embedded tissue samples. Am J Forensic Med Pathol 2006;27:347–351.
25 Bateman AC, Hemmatpour SK, Theaker JM, *et al.* Genetic analysis of hydatidiform moles in paraffin wax embedded tissue using rapid, sequence specific PCR-based HLA class II typing. J Clin Pathol 1997;50:288–293.