Original Article

Comparison of single nucleotide polymorphisms and short tandem repeats as markers for differentiating between donors and recipients in solid organ transplantation

Hyun-Ji Lee[1], Kyoung-Hwa Shin[2], Su Jeong Jeong[1], In Suk Kim[1]

[1]Department of Laboratory Medicine, and Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, Yangsan, Korea
[2]Department of Laboratory Medicine, Pusan National University Hospital, Busan, Korea

Corresponding author: Hyun-Ji Lee
Department of Laboratory Medicine, and Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, 20 Geumo-ro, Mulgeum-eup, Yangsan 50612, Korea
Tel: +82-55-360-1875
Fax: +82-55-360-1880
E-mail: hilhj1120@gmail.com

Background: To analyze transplant rejection and to distinguish between donor and recipient, it is necessary to select a marker from single nucleotide polymorphism (SNP), short tandem repeat (STR), and human leukocyte antigen (HLA) testing. SNPs are bi-allelic and the polymerase chain reaction method used for SNP testing has the advantage of lower cost than sequencing methods. In this study, we aimed to distinguish donors from recipients using a combination of existing commercialized STRs and the SNPs identified.
Methods: All selected SNPs complied with the following criterion: known and validated minor allele frequency (MAF) ≥43% in Korean and reported ethnicities from global populations (HapMap, 1000 Genomes, and the Korean Reference Genome project). The STR assays were performed for 16 tetranucleotide repeat loci.
Results: DNA from the 52 donor/recipient pairs were tested for informative markers. The median age of the recipients was 47 years. MAF in the 52 pairs was 1.0%–76.0%. The probability of informative genotypes (I) was 0.001–0.124. The summation of I was 0.680. In the 52 donor recipient pairs, the selected SNPs showed a 0.031 average probability of being informative. The probability of identity in our study was 0.122–0.348. SNP panel configuration distinguished 100% of 52 donors/recipient pairs.
Conclusions: Donors and recipients were distinguished by STR and 22 SNPs with MAF identified from SNP databases. Seventeen SNPs were able to distinguish between donors and recipients (I value=0.039).

Keywords: Organ transplantation; Single nucleotide polymorphism; Short tandem repeats; Chimerism

INTRODUCTION

Microchimerism including donor-derived cell-free DNA (ddcfDNA) can be used for predicting organ rejection after transplantation [1]. For the analysis of transplant rejection, it is necessary to distinguish between the donor and the recipient, hence, there is need to select a marker from single nucleotide polymorphisms (SNPs), short tandem repeats (STRs), or human leukocyte antigen (HLA) testing [2]. STRs are highly polymorphic and are used as markers to distinguish between individuals; however, disadvantages of STR analysis include high cost as it involves sequenc-

![KJT]

HIGHLIGHTS

- DNA from 52 organ transplantation donor-recipient pairs was used for single nucleotide polymorphism (SNP) and short tandem repeat (STR) analysis.
- Informative STR and SNP markers were identified.
- Seventeen SNPs to differentiate between Korean donors and recipients were identified.

ing and longer duration than polymerase chain reaction (PCR) analysis [3]. SNPs are bi-allelic and can be used as a combination of several useful SNP markers; they have an advantage of reduced cost compared to sequencing methods as PCR is used for SNP analysis. The development of the highly sensitive droplet digital PCR (ddPCR) method enables cheaper and faster analysis compared to sequencing [2]. ddcfDNA can be detected by ddPCR using a SNP marker to predict early rejection after transplantation [4].

Serial inspection using markers of acute injury must be performed for early detection of rejection after solid organ transplantation. Therefore, it is practical to use an inexpensive PCR-based method with performance similar to the STR analysis method. This is a study to select markers to differentiate donor and recipient to measure ddcfDNA and recipient-derived cell free DNA. In order to compare the usefulness of SNPs to STR, we selected SNPs that enabled us to distinguish donor-recipient and compared them with commercially available STR.

METHODS

The study's protocol was reviewed and approved by the Institutional Review Board of Pusan National University Yangsan Hospital (IRB No. 04-2019-006) and written informed consent was obtained from all participants.

Patients

We tested DNA of patients from January 2016 to May 2019, comprising 52 donors and 52 recipients of solid organ transplantation, with organ transplants including liver, kidney, heart, lung, and pancreas. SNP markers were selected to differentiate donor DNA from recipient DNA. The study was conducted in compliance with the principles of the Declaration of Helsinki.

Genomic DNA Extraction

Whole blood samples were collected from the donors and recipients. Extraction of genomic DNA from peripheral blood was performed using the QIAamp DNA blood mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

SNP Assays

All selected SNPs complied with the following criterion: known and validated minor allele frequency (MAF) ≥43% in Korean and all reported ethnicities (data from public databases: HapMap, 1000 Genomes, and the Korean Reference Genome project). Assuming Hardy-Weinberg equilibrium, a SNP with a MAF between 0.4 and 0.5 would be homozygous for each allele in both donor and recipient about 22% to 25% of the time. The probability of both donor and recipient having a different allele is therefore 11.5% to 12.5%, using accepted estimation models for calculation of exclusion probabilities [5]. A set of 22 assays targeting different SNPs were performed for 52 donor and recipient pairs. A TaqMan probe labeled with fluorescein amidite at the 5'-end nucleotide, and with a quencher (Black Hole Quenchers; Sigma-Aldrich, St. Louis, MO, USA) at the 3'-end nucleotide was used. SNP genotyping was performed using the QuantStudio real-time PCR (qPCR) instrument. The 20 μL PCR mixture contained Taqman Genotyping Master Mix (10 μL), Taqman Genotyping Assay (20X, 1 μL), DNase-free, RNase-free water (7.5 μL), and genomic DNA (1.5 μL). The reaction conditions were as follows: (95°C for 10 minutes)/(95°C for 15 seconds and 60°C for 1 minute) 40 cycles.

STR Assays

STR assays were performed using the AmpFLSTR Identifiler Plus PCR Amplification Kit (Applied Biosystem, Foster City, CA, USA) consisting 16 tetranucleotide repeat loci (D3S1358, vWA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, CSF1PO, TH01, D16S539, D2S1338, D19S433, TPOX, and FGA) and the Amelogenin gender-determining marker (AMEL). Each PCR reaction was performed with a final volume of 25 μL containing 10 μL Identifiler Plus Master mix, 5 μL Identifiler Plus Primer Set, and 10 μL genomic DNA (10 μL, 0.1 ng/μL) or control (10 μL, 0.1 ng/μL). The PCR reaction conditions were as follows: 11 minutes at 95°C; followed by 20 seconds at 94°C, 3 minutes at 59°C, and 10 minutes at 60°C for 29 cy-

Lee HJ et al. Comparison of SNP and STR in donors and recipients

KJT

cles; and a final hold at 4°C. A 10 μL sample contained 8.7 μL Hi-Di Formamide, 0.3 μL gene scan 500 LIZ Size Standard, and 1 μL of PCR product or an allelic ladder denatured at 95°C for 3 minutes. PCR products were analyzed on an ABI Prism 3500 DNA sequencer (Applied Biosystems, Foster City, CA, USA) using the GeneMapper software 5 (Thermo Fisher Scientific, Waltham, MA, USA). The PCR products, which were 280−312 bp in size were analyzed with the GeneScan software.

Analysis of SNP Marker Performance

We counted the number of markers showing informative genotypes. The SNP haplotypes of a minor DNA contributor that are nonshared with the major DNA represents the DNA profile that can be detected in the mixture of donor and recipient DNAs. Based on Hardy-Weinberg assumptions, the probability of informative genotypes (I) at a given SNP marker can be calculated as $I=2a^2b^2+2a^3b+2ab^3$, where a and b are the frequencies of the A and B alleles, $2a^2b^2$ is the probability that the major and minor DNA

contributors are homozygous for the opposite allele ($a^2b^2+a^2b^2$), and $2a^3b+2ab^3$ are the probabilities that the major DNA contributor is homozygous (either A or B) and the minor DNA contributor is SNP heterozygous, respectively ($[a^2[2ab]+b^2[2ab]]$). The I value for the SNP markers is reported in Table 1. If the I value is 0.5, there is a 50% chance of distinguishing the donor/recipient at the type I level of that SNP. The higher the I of the SNP, the less the number of SNPs required, and the donor/recipient can be distinguished only with the type I marker.

Analysis of STR

The probability of identity (Pi), also referred to as the matching probability, is the chance that two unrelated people selected at random will have the same genotype and was first described by George Sensabaugh in 1982. The Pi value of a single locus is determined by summing the square of the observed genotype frequencies:

$$\sum_{i=1}^{n} xi^2$$

Table 1. Allele frequencies and probability of informative genotypes in the 52 solid organ transplant donor and recipient pairs

No.	Name	Chromosome	MAF (all)	Korean MAF	SNP 1>2	MAF[1]	MAF[2]	I
1	rs12096438	1	0.433	0.489	C>T	0.279	0.260	0.031
2	rs12064796	1	0.487	0.569	A>G	0.192	0.279	0.018
3	rs6436409	2	0.46	0.566	C>G	0.462	0.115	0.030
4	rs1871686	8	0.495	0.492	A>G	0.260	0.269	0.029
5	rs13317873	3	0.481	0.537	C>T	0.413	0.212	0.053
6	rs251022	5	0.436	0.287	C>T	0.760	0.019	0.017
7	rs2523860	6	0.487	0.455	C>G	0.385	0.192	0.038
8	rs1265094	6	0.478	0.381	C>T	0.154	0.067	0.001
9	rs10228737	7	0.476	0.274	C>T	0.673	0.010	0.006
10	rs1764980	20	0.351	0.488	A>G	0.327	0.183	0.024
11	rs11103106	9	0.477	0.704	A>G	0.250	0.327	0.041
12	rs6419563	2	0.394	0.5	T>C	0.413	0.135	0.027
13	rs10734083	10	0.492	0.494	C>T	0.423	0.135	0.029
14	rs7072759	10	0.454	0.413	A>G	0.452	0.010	0.002
15	rs1467245	11	0.487	0.565	G>A	0.221	0.125	0.005
16	rs1328368	13	0.487	0.483	G/A	0.702	0.144	0.124
17	rs10460162	19	0.286	0.49	A>G	0.298	0.337	0.061
18	rs2072042	16	0.404	0.366	T>C	0.500	0.077	0.023
19	rs10164176	18	0.457	0.478	C>T	0.010	0.317	0.001
20	rs2497654	13	0.325	0.464	C>T	0.365	0.202	0.037
21	rs4361824	9	0.53	0.493	G>A	0.096	0.279	0.006
22	rs2298065	x	0.419	0.695	A>C	0.240	0.442	0.076

MAF, minor allele frequency; SNP, single nucleotide polymorphism; MAF[1], minor allele frequency of first allele on SNP; MAF[2], minor allele frequency of second allele on SNP; I, probability of informative genotypes.

where Xi is the genotype frequency.

Lower Pi values indicate more variability of the genetic marker in the measured population because there are more genotypes occurring at a lower frequency.

Informative Maker Analysis

The SNP or STR loci of informative markers were analyzed. Type I markers, or loci type I, were defined as loci or alleles presenting a homozygous genotype in the recipient and an alternative homozygous genotype in the donor. Conversely, type II markers were defined as loci or alleles presenting a homozygous genotype in the recipient and a heterozygous in the donor. Type III markers were those that did not fulfil the criteria for type I or type II markers.

RESULTS

DNA from 52 donor recipient pairs were tested for informative markers. The number of deceased donors was 28 and the number of living donors was 24. Males (n=30, 57.7%) outnumbered females in the total number of recipients (Table 2) and the median age of recipients was 47 years. The MAF of SNPs in the 52 pairs were 1.0%–76.0% (Table 1). The probability of informative genotypes (I) ranged from 0.001–0.124 and the summation of I was 0.680. In all donor recipient pairs, the selected SNP markers had an average probability of being informative of 0.031. After the exclusion of the least polymorphic markers (rs1265094, rs10228737, rs7072759, rs1467245, rs10164176), the average probability of being informative increased to 0.039.

The most selected type I informative markers were rs10460162 and rs11103106, observed in 9 donor recipient pairs (17.3%), followed by rs2523860. Four SNPs, rs10228737, rs10164176, rs7072759, and rs1265094, had no discriminatory power as type I informative markers. The most selected type II informative marker was rs12096438 observed in 17 donor recipient pairs (32.7%) (Fig. 1).

The Pi or matching probability is the chance of the same genotype in two unrelated people selected at random. The Pi of the present study ranged from 0.122 to 0.348 (Table 3). AMEL was not an optimal marker for the determination of mixed chimeras due to the sharing of alleles between donor and recipient. Except AMEL, the 15 STR loci were informative markers with a discrimination

Table 2. Baseline characteristics of 52 donors and recipients

Variable	Number
Organ	
Kidney	25
Heart	13
Pancreas	11
Liver	3
Sex	
Male	30
Female	22
Relationship	
Unrelated	
Deceased	28
Spouse	10
Family in law	3
Related	
Parent/child	8
Brother/sister	2
Uncle	1
Rejection	
AMR	2
TCMR	9
Age (yr), median (range)	47 (26–66)

AMR, antibody-mediated rejection; TCMR, T-cell−medicated rejection.
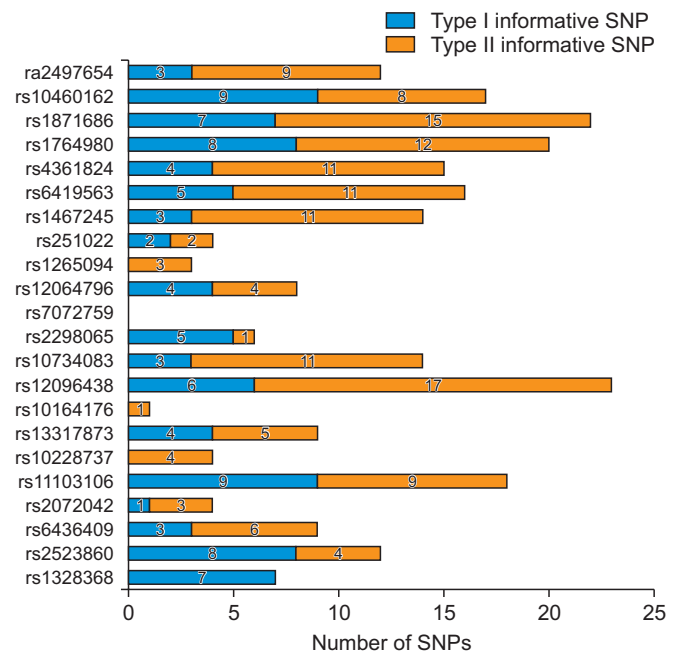
Fig. 1. Number of informative single nucleotide polymorphism (SNP) loci identified in the 52 donor and recipient pairs. The histogram shows the number of informative SNP loci identified in the donors and recipients.

Lee HJ et al. Comparison of SNP and STR in donors and recipients

KJT

power of 42.9%−92.9%. The selection rate of type I informative markers among 16 alleles was 17.1% to 56.1%, with D2S1338 observed in 27 donor recipient pairs (Fig. 2). There were no type I alleles in two mother-child pairs, one sister-sister pair, and one uncle-nephew pair.

In 52 pairs of donors and recipients, the median values of type I SNPs were 2 and the median values of type II SNPs were 3. The median value of type I STR was 6 and type II STR was 5. The sum of meaningful markers on combining type I and type II was median 4 (range, 1−9) in

SNP and median 11 (range, 2−15) in STR (Fig. 3).

DISCUSSION

The development of rejection after transplantation is a very important event leading to loss of organ transplant, and hence, early detection of rejection is important for organ transplant survival. The use of noninvasive and sensitive test methods for diagnosis of rejection have shown that an increase in amount of ddcfDNA beyond a certain level after organ transplantation can be used as a predictor of transplant rejection [5,6]. This is a preliminary study

Table 3. Loci sorted by Pi values

Allele	Observed allele type	Pi
AMEL	X,Y	0.348
CSF1PO	9,10,11,12,13	0.217
D13S317	8,9,10,11,12,13,14	0.147
D16S539	9,10,11,12,13,14	0.186
D18S51	12,13,14,15,16,17,18,19,20,22,23	0.156
D19S433	11.2,12,13,13.2,14,14.2,15,15.2,16.2	0.205
D21S11	27,28,28.2,29,29.2,30,30.2,31,31.2,32.2,33.2	0.206
D2S1338	17,18,19,20,21,22,23,24,25,26,27	0.125
D3S1358	14,15,16,17,18,19	0.221
D5S818	8,9,11,12	0.249
D7S820	7,9,10,11,12,13	0.192
D8S1179	10,11,12,13,14,15,16,17	0.141
FGA	18,19,20,21,22,23,24,25,26,27	0.122
TH01	6,7,8,9,9.3,10	0.191
TPOX	8,9,11,12	0.249
vWA	14,15,16,17,18,19,20	0.178
Total Pi		2.6478E-12

Pi, probability of identity.
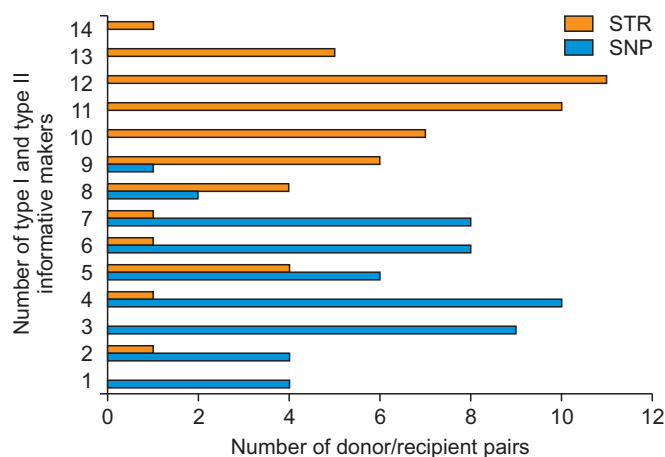
Fig. 3. Comparison of the number of informative alleles of single nucleotide polymorphism (SNP) and short tandem repeat (STR).
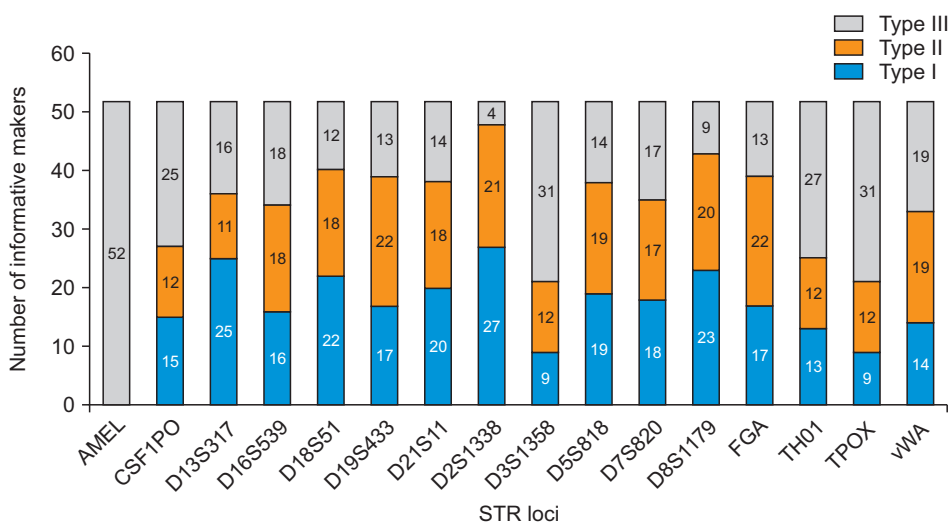
Fig. 2. Informative short tandem repeat (STR) loci in the 52 solid organ transplant donor and recipient pairs.

for a follow-up study to predict organ transplant rejection by measuring ddcfDNA in recipients after organ transplantation. Variation among races should be considered when using the SNP method [7,8]. In order to compare the usefulness of SNPs as markers, we selected SNPs to distinguish between donors and recipients in a Korean population by comparing them with STR markers that are commercially available. Of the 22 SNPs, 17 markers could be configured as the primary marker set, and the excluded SNPs included 5 SNPs. Four SNPs, rs10228737, rs10164176, rs7072759, and rs1265094, in which donor/recipient pairs could not be differentiated, as well as rs1467245, with an I value of 0.005 or less were excluded. All donor ddcfDNA were categorized in STR as well as SNP and were classified into 52 pairs (100.0%).

Several tests are widely used in forensic science to distinguish between individuals. HLA present on chromosome 6 has a wide variety and is inherited as a haplotype [9]. The probability of HLA agreement is high in kidney transplantation, and hence is not suitable for discrimination between donor and recipient post transplantation. STRs are the first choice of genetic markers in forensic science as their use yields a satisfactory result in almost all cases, mainly due to their high degree of polymorphism and consequent discrimination power [3]. STRs are multiallelic, and the probability of two test objects coinciding is almost zero if many markers are used. STRs vary depending on the marker composition but the probability of a match or the likelihood of a match is approximately 0.0001 to 0.000005 or 10^{12} power [8]. The probability of STR match is very low, and hence the large ability to distinguish between two individuals. In our study, the probability of difference between donor and beneficiary using the STR set of 16 markers was 2.6478E-12.

SNPs are genetic markers that consist DNA sequence variations resulting from single base changes in the genome sequence, and have been considered as additional informative markers to provide reliable likelihood ratios (LRs) [3]. Most SNPs are bi-allelic and exhibit a maximum heterozygosity of 0.50. The most efficient marker of the ideal heterozygosity of 0.50 has 37.5% probability of distinguishing two test subjects [10]. However, even though markers with an MAF between 0.4 and 0.5 are selected, the probability of one marker varies depending on the number of evaluations (N), so the set configuration differs among the analyzers. On an average, our markers show 0.039 probability of being informative, after the exclusion of the least polymorphic marker. In the present study, a MAF of 0.4−0.5 in the global population and a MAF of 0.4−0.5 in the Korean population were studied. However, the MAF of the study subjects was different from the documented global and Korean MAFs. In addition, some Korean markers had a MAF less than 0.4, resulting in lower I value.

There is need of a sensitive assay as the amount of ddcfDNA is small compared to the amount of recipient DNA. This can be expressed as a detection limit, i.e., if the detection limit is 1 ng, it means that substances present below 1 ng cannot be detected. Therefore, a test method with a low detection limit is essential for ddcfDNA detection. The detection limit for the test methods is 0.1−1 pg for qPCR, 1 fg for digital PCR, and 0.05 to 0.1 ng for STR. Digital PCR can detect the smallest amount of DNA [11,12]. In digital PCR, the sample is divided into many independent partitions. The distribution of target sequences in the partitions can be approximated with a Poisson's distribution. Each partition acts as an individual PCR micro reactor and the partitions containing amplified target sequences are detected by fluorescence. The ratio of positive partitions over the total number allows determination of the concentrations of the target DNA in the sample [13]. Due to this principle, the sensitivity of digital PCR is 0.001% [14]. Development of digital PCR has allowed a low detection limit and increased sensitivity, hence studies are underway to detect ddcfDNA for monitoring rejection after transplantation.

The study uses the PCR method, which is cheaper than the STR method (sequencing method), is easy to test, with a shorter test time [7]. STR can be quantitatively measured and is used to measure mixed chimerism in stem cell transplantation patients [15]. Sensitivity is important to detect mixed chimerism; thus, it is advantageous to use ddPCR, since it can detect a very small amount of ddcfDNA, and the detection amount can be quantitatively analyzed [16,17].

An obvious limitation of the current study is the small size as only 52 donor/recipient pairs were enrolled. Therefore, SNP configuration requires further validation by discovering additional markers and expanding SNP research in the Korean population.

In our study, donors and recipients were distinguished by STR and 22 SNPs with MAF identified from SNP databases. With the exception of five SNPs, 17 SNP markers were able to distinguish both donors and recipients. This finding could be further used to predict organ damage by identifying ddcfDNA in the recipient's blood after solid or-

Lee HJ et al. Comparison of SNP and STR in donors and recipients

KJT

gan transplantation.

ORCID
Hyun-Ji Lee   https://orcid.org/0000-0002-9021-5632
Kyung-Hwa Shin   https://orcid.org/0000-0002-8454-4448
Su Jeong Jeong   https://orcid.org/0000-0002-5581-3813
In-Suk Kim   https://orcid.org/0000-0002-7243-9173

Author Contributions
Conceptualization: HJL. Data curation: HJL, SJJ. Formal analysis: HJL. Funding acquisition: HJL. Investigation: SJJ. Methodology: HJL, SJJ. Project administration: HJL. Visualization: SJJ. Writing−original draft: HJL, KHS. Writing−review & editing: KHS, ISK.

REFERENCES

1. Snyder TM, Khush KK, Valantine HA, Quake SR. Universal noninvasive detection of solid organ transplant rejection. Proc Natl Acad Sci U S A 2011;108:6229-34.
2. Gielis EM, Ledeganck KJ, De Winter BY, Del Favero J, Bosmans JL, Claas FH, et al. Cell-free DNA: an upcoming biomarker in transplantation. Am J Transplant 2015;15:2541-51.
3. Pontes ML, Fondevila M, Laréu MV, Medeiros R. SNP markers as additional information to resolve complex kinship cases. Transfus Med Hemother 2015;42:385-8.
4. Beck J, Bierau S, Balzer S, Andag R, Kanzow P, Schmitz J, et al. Digital droplet PCR for rapid quantification of donor DNA in the circulation of transplant recipients as a potential universal biomarker of graft injury. Clin Chem 2013;59:1732-41.
5. Thongprayoon C, Vaitla P, Craici IM, Leeaphorn N, Hansrivijit P, Salim SA, et al. The use of donor-derived cell-free DNA for assessment of allograft rejection and injury status. J Clin Med 2020;9:1480.
6. Knight SR, Thorne A, Lo Faro ML. Donor-specific cell-free DNA as a biomarker in solid organ transplantation: a systematic review. Transplantation 2019;103:273-83.
7. Li L, Wang Y, Yang S, Xia M, Yang Y, Wang J, et al. Genome-wide screening for highly discriminative SNPs for personal identification and their assessment in world populations. Forensic Sci Int Genet 2017;28:118-27.
8. Børsting C, Morling N. Genomic applications in forensic medicine. In: Kumar D, Antonarakis S, eds. Medical and Health Genomics. Cambridge, MA: Academic Press; 2016. p.295-309.
9. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet 2009;54:15-39.
10. Castella V, Gervaix J, Hall D. DIP-STR: highly sensitive markers for the analysis of unbalanced genomic mixtures. Hum Mutat 2013;34:644-54.
11. Oldoni F, Podini D. Forensic molecular biomarkers for mixture analysis. Forensic Sci Int Genet 2019;41:107-19.
12. Robin JD, Ludlow AT, LaRanger R, Wright WE, Shay JW. Comparison of DNA quantification methods for next generation sequencing. Sci Rep 2016;6:24067.
13. Quan PL, Sauzade M, Brouzes E. dPCR: a technology review. Sensors (Basel) 2018;18:1271.
14. Huggett JF, Cowen S, Foy CA. Considerations for digital PCR as an accurate molecular diagnostic tool. Clin Chem 2015;61:79-88.
15. Lion T, Watzinger F, Preuner S, Kreyenberg H, Tilanus M, de Weger R, et al. The EuroChimerism concept for a standardized approach to chimerism analysis after allogeneic stem cell transplantation. Leukemia 2012;26:1821-8.
16. Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, et al. Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods 2013;10:1003-5.
17. George D, Czech J, John B, Yu M, Jennings LJ. Detection and quantification of chimerism by droplet digital PCR. Chimerism 2013;4:102-8.