차세대염기서열분석법을 이용한 HLA-A, -B 그리고 -DR 형별 분석법 개발

서동희 · 이정민 · 박미옥 · 이현주 · 문서윤 · 오미진 · 김소영 · 이상현 · 형기은 · 허혜진 · 조대연

랩지노믹스의학연구소

Development of HLA-A, -B and -DR Typing Method Using Next-Generation Sequencing

Dong Hee Seo, Jeong Min Lee, Mi Ok Park, Hyun Ju Lee, Seo Yoon Moon, Mijin Oh, So Young Kim, Sang-Heon Lee, Ki-Eun Hyeong, Hae-Jin Hu, Dae-Yeon Cho

LabGenomics Clinical Laboratories, Seongnam, Korea

Background: Research on next-generation sequencing (NGS)-based HLA typing is active. To resolve the phase ambiguity and long turn-around-time of conventional high resolution HLA typing, this study developed a NGS-based high resolution HLA typing method that can handle large-scale samples within an efficient testing time.

Methods: For HLA NGS, the condition of nucleic acid extraction, library construction, PCR mechanism, and HLA typing with bioinformatics were developed. To confirm the accuracy of the NGS-based HLA typing method, the results of 192 samples HLA typed by SSOP and 28 samples typed by SBT compared to NGS-based HLA-A, -B and -DR typing.

Results: DNA library construction through two-step PCR, NGS sequencing with MiSeq (Illumina Inc., San Diego, USA), and the data analysis platform were established. NGS-based HLA typing results were compatible with known HLA types from 220 blood samples.

Conclusion: The NSG-based HLA typing method could handle large volume samples with high-throughput. Therefore, it would be useful for HLA typing of bone marrow donation volunteers. (Korean J Blood Transfus 2018;29:310-319)

Key words: HLA, NGS, HLA typing

서 론

사람 백혈구 항원(Human Leukocyte Antigen:

HLA)은 6번 염색체의 단완에 위치한 HLA 유전자에 의해 생성되며, 구조 및 기능에 따라 class I과 class II 유전자로 구분된다. HLA class I 유전

Received on April 23, 2018. Revised on September 17, 2018. Accepted on November 7, 2018

Correspondence to: Dong Hee Seo

LabGenomics Clinical Laboratories, 700 Daewangpangyo-ro, Bundang-gu, Seongnam 13488, Korea

Tel: 82-31-628-0730, Fax: 82-31-628-0701, E-mail: seo2023@nate.com, ORCID: <http://orcid.org/0000-0003-0828-3773>

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright ©2018 The Korean Society of Blood Transfusion

자는 HLA-A, HLA-B, HLA-C 등의 항원 생성에 관여하고 이들 class I 항원은 거의 모든 세포에 표현되는 반면에, class II 유전자는 HLA-DR, HLA-DQ, HLA-DP 등의 항원 생성에 관여하며 class II 항원은 주로 B림프구 및 단핵구에 표현된다[1].

HLA는 조혈모세포 및 장기의 이식시 거부반응을 유발하는 면역항체를 생성하여 조직적합항원으로도 알려져 있다. 골수나 장기 이식시 성공율을 높이기 위해서는 환자와 공여자의 HLA 형별이 일치되어야 하고, 만약 HLA가 불일치할 경우는 이식된 조혈모세포가 환자로부터 이물질로 간주되어 거부 반응이 일어나거나 또는 반대로 공여자의 T-림프구가 환자를 공격하는 이식편대숙주 질환이 발생하여 환자가 사망할 수도 있다[2]. 두 경우 모두 조혈모세포나 장기 이식을 실패로 만드는 요인이기 때문에 장기 및 조혈모세포 이식 전에 HLA형별 및 교차시험 검사는 필수적인 검사로 시행되고 있다.

HLA검사방법은 검사의 해상도에 따라 혈청학적 수준의 형별 결과를 보이는 저해상도, 대립유전자 수준의 고해상도 및 대립유전자 수준에서 종종 그룹 형별을 보이는 중해상도로 분류할 수 있다. 저해상도 분석인 혈청학적 검사법은 숙련된 기술과 경험이 요구되며, DNA검사법들과 비교했을 때 불일치하는 결과가 보고되면서, DNA 검사법이 점차 혈청학적 검사법을 대체하고 있다[3].

중해상도 분석을 위해 이용되는 SSOP (sequence specific oligonucleotide probe) 방법은 중합효소연쇄반응(Polymerase Chain Reaction; PCR) 후에 스트립에 고정된 프로브와의 교잡반응 양상을 토대로 HLA 형을 정하는 방법으로, SSOP방법은 혈청학적 검사법에서 흔히 보이는 교차반응 및 약한 반응으로 인한 모호한 결과의 가능성이 적어 결과 판독이 다소 용이하다는 점과 많은 양

의 검체를 동시에 처리할 수 있는 장점이 있다. 하지만, PCR 후에 교잡 과정을 거치므로 검사 시간이 다소 걸린다는 점과 양성 판정이 어려운 경우가 간혹 발생하는 단점이 있다[4,5].

중해상도 분석을 위해 이용되는 PCR-SSP (Sequence Specific Primer) 기법은 PCR 시발체의 염기서열차이를 이용하여 유전자를 선택적으로 증폭하며, PCR 시발체 세트의 수를 조절하여 원하는 수준의 형별을 판정할 수 있다[6]. 고해상도 방법으로 이용되는 SBT (sequence-based typing)법은 HLA 형별 분석을 위한 가장 우수한 방법으로 HLA 대립유전자 수의 급격한 증가에 기여한 검사법이다[7]. 하지만, 대립유전자의 변이 부분이 시스(cis)형인지 트랜스(trans)형인지를 구별하지 못하는 위상 모호성(phase ambiguity) 문제가 있고 시간과 비용이 많이 드는 단점이 있다[8].

저해상도HLA형별 분석에 기반한 이식 성적이 좋지 않다는 보고가 많아지면서, 고해상도 분석 필요성에 대한 공감대가 형성되고 있고, 비혈연 조혈모세포이식의 경우에는 혈청학적 수준에서 HLA 형별이 일치하여도 고해상도 수준에서 불일치하면 거부반응, 급성 이식편대숙주병의 발생과 사망률이 높아진다는 연구결과가 보고되어 있다[9]. 따라서, HLA 고해상도 분석법인 SBT검사법의 이용이 요구되고 있지만, SBT 분석법은 높은 검사 비용, 적은 검체 처리량의 단점이 있어서, 많은 검체의 HLA형별 분석이 요구되는 국내 조혈모세포기증희망자의 검사에는 적용에 한계가 있다.

최근 차세대염기서열분석법(Next Generation Sequencing: NGS)을 이용한 HLA 형별 분석에 대한 연구가 활발히 진행되고 있으며, 기존 고해상도 방법인 SBT법에 비해 모호성이 많이 줄어든 결과가 보고되고 있다[10-13]. 이에 HLA 고해상도 분석법의 내재적 한계인 위상 모호성의 문제를

해결하고, 짧은 시간에 대량의 검체가 처리 가능한 NGS기반 고해상도 HLA 형별 검사법을 자체 기술로 개발하고자 본 연구를 실시하였다.

대상 및 방법

NGS 기반 HLA 형별 검사법 개발을 위해, MiSeq (Illumina Inc., San Diego, USA) 기기 활용에 최적화된 핵산 정량 조건 확립, 라이브러리 제작, PCR 체계 확립, 그리고 생물정보학을 이용한 HLA 형별 분석을 실시하였다. 개발한 NGS 기반 HLA 형별 검사의 정확성을 알아보기 위해, 본 기관에서 WAKFlow HLA Tying kit (Wakunaga Pharmaceutical Co., Osaka, Japan)를 사용한 SSOP 검사법으로 HLA 형별을 알고 있는 192개 검체와 SBT검사법으로 HLA 형별을 알고 있는 28개 검체에 대해 NGS 기반으로 검사한 HLA 형별 결과를 비교해 보았다. 이들 대상 검체는 익명으로 관리되는 건강인 검체로 본 기관의 잔여 혈액 검체를 이용하였으며, 기관생명윤리위원회의 심의를 거쳤다(2015-R05).

1. DNA 추출 및 정량

전혈 혈액에서 NucleoMag Blood 시약(MACHEREY-NAGEL, Düren, Germany)과 Hamilton Star 자동화 장비(Hamilton, Reno, USA)를 이용하여 DNA를 추출하였다. 추출된 핵산의 농도는 Quant-iT PicoGreen dsDNA Assay 시약(Thermo Fisher Scientific, Eugene, USA)과 Quantus fluorometer (Promega, Madison, USA)를 이용하여 측정하였다.

2. 시발체 디자인

시발체는 HLA Class I의 HLA-A와 B의 엑손 2, 3, 4, 그리고 HLA Class II의 HLA-DRB1의 엑손

2의 상이한 HLA 형별을 모두 증폭할 수 있고, 비특이 증폭을 초래하지 않는 시발체들로 디자인하였다. PCR은 두 단계로 구분되며, 첫 단계 PCR에 사용된 시발체는 표적 특이적인 시퀀스(18~23 bp)와 일루미나 시퀀싱에 이용되는 paired end 어답터 시퀀스의 일부(33~34 bp)가 결합된 형태이다. 이때, PCR I의 7쌍 시발체들은 세 개의 그룹으로 나누어 여러 표적이 한 번에 증폭될 수 있도록 하였다. 그룹 1은 HLA-A 유전자의 4번 엑손, B의 2번과 4번 엑손을 증폭하고, 그룹 2는 HLA-A의 3번 엑손과 DRB1의 2번 엑손을, 그룹 3은 HLA-A의 2번 엑손과 B의 3번 엑손을 증폭하도록 디자인되었다(Fig. 1). PCR II에 사용된 시발체는 PCR I의 각 시발체에 연결된 paired end 어답터 시퀀스와 상보적인 결합을 하는 시퀀스, 각 개별 검체를 식별할 수 있는 바코드 시퀀스(10 bp), 그리고 일루미나 시퀀싱 flow cell에 심겨진 시퀀스와 상보적으로 결합을 할 수 있는 시퀀스(~21 bp)로 구성하였다.

3. 라이브러리 제작 및 NGS

HLA NGS 분석을 위해 두 단계의 DNA 라이브러리 제작이 필요하였다. 첫 단계는 유전자 특이적인 PCR 산물을 증폭하는 과정으로 PCR I이고, 두 번째 단계는 검체 식별 바코드가 부착된 일루미나 시퀀싱이 가능한 형태로 앰플리콘(amplicon)을 만드는 과정으로 PCR II였다. 두 단계의 PCR은 모두 변성, 결합, 신장의 기본적인 PCR 순서로 진행되었다. PCR I에서는 gDNA (그룹별 ~20 ng 사용), PCR I 시발체쌍, 그리고 Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific, Carlsbad, USA)를 이용하여, 유전자 특이적인 PCR 산물이 생산되도록 C1000 증폭기기(Bio-Rad, Hercules, USA)를 이용하여 다중 중합효소연쇄반응을 실시하였다. PCR II에서는 세 그룹의

Fig. 1. Products of multiplex PCR for HLA gene targeted. Group 1 is the PCR products of HLA-A exon 4, HLA-B exon 2 and 4. Group 2 is that of HLA-A exon 3 and HLA-DRB1 exon 2. Group 3 is for HLA-A exon 2 and HLA-B exon 3.

PCR I 증폭산물을 희석하여 특정 비율로 합친 후, 하나의 PCR II 시발체쌍(공통 forward 시발체와 reverse 바코드 시발체), 그리고 Taq polymerase를 이용하여 최종 라이브러리를 완성하였다(Fig. 2). 완성된 라이브러리들을 동일한 농도(nM)로 합친 후 TopQXSEP MagBead (Celegence, Seoul, Korea)를 이용하여 정제함으로써 시퀀싱의 품질을 저해하는 라이브러리 내에 존재하는 시발체와 시발체 다이머를 제거하였다. 각 검체 별 라이브러리 및 통합된 라이브러리의 농도는 Quant-iT PicoGreen dsDNA Assay 시약(Thermo Fisher Scientific, Carlsbad, USA)과 Quantus fluorometer (Promega)를 이용하여 측정하였으며, NGS 시퀀싱은 Miseq v2 (2×250 cycles) 시약과 MiSeq/

MiSeqDx (Illumina) 기기를 이용하여 실행하였다(Fig. 2).

4. NGS 데이터 분석

NGS 데이터를 이용한 HLA 형별 분석은 본 기관에서 자체 개발한 PEERE 라는 프로그램을 이용하였다. PEERE 프로그램은 참조 서열을 준비하는 단계와 테스트 샘플을 이용하여 HLA 형별 분석을 수행하는 두 단계로 구성되어 있다. 첫 단계에 필요한 참조 서열은 IMGT/HLA 데이터베이스 (Release 3.26.0, 2016-10-14 기준)에서 다운로드하였고(<https://support.illumina.com/downloads/imgt-hla-3-26-reference.html>), 이 참조 서열을 기반으로 엑손(exon) 영역이 아닌 서열을 제거하고 7개 표적 엑손 별 컨센서스 엑손 서열을 생성하였다. 샘플서열을 이용하여 형별 분석을 수행하는 단계는 다음의 세부 과정으로 구성되어 있다. 샘플 서열은 정방향과 역방향 서열로 구성된 두 개의 fastq 파일로 되어 있는데, 이 파일들이 병합(merge) 과정을 거치면서 하나의 fastq 파일로 전환된다. 이 fastq 파일은 본 연구에 사용된 시발체 서열에 기초하여 7개의 표적 엑손 각각에 대한 서열 군으로 분류되는데, 이 과정에서 한 개의 fastq 파일이 7개의 fastq 파일로 나뉘어진다. 이후 각 표적 엑손 서열은 개별적으로 분석된다. 즉, 각 서열들은 첫 번째 단계에서 준비한 참조서열의 컨센서스 엑손 서열에 정렬(aligned)되고, 트리밍(trimming) 과정을 통해 인트론(intron) 영역이 제거된다. 인트론이 제거된 각각의 서열들은 일치하는 대립 유전자의 유형과 빈도를 확인하기 위해 참조 서열에 정렬된다. 각 엑손에서 주요 빈도를 보이는 HLA 형별들을 취하여, 후보 대립 유전자 리스트에 넣는다. 이들 리스트에 있는 후보들 중, 표적 엑손에 동일한 조합으로 일관성 있게 나타나는 단 한 쌍을 최종 대립형으로 결정하였다.

Fig. 2. HLA NGS workflow. HLA NGS workflow started with multiplex PCR (PCR I) of HLA-A, -B, and -DRB1. After PCR I, the amplicons were diluted, pooled and bar-coded (PCR II). The bar-coded libraries were pooled in equimolar concentrations and bead-purified. The resulting pooled library was denatured, diluted, and sequenced on MiSeq/MiSeqDx instruments using the paired-end 2×250 cycle MiSeq Reagent Kit. The sequence data were exported and analyzed using HLA NGS analysis software (PEERE) with 3.19.0 IMGT/HLA database serving as the reference.

5. SBT 검사

국내외 HLA 형별 검사용 외부정도관리 28개

검체에 대해 SBT를 실시하였다. SBT는 SeCore Sequencing Kits (One Lambda Inc., Los Angeles, USA)을 사용하였으며, 제조사에 의해 제공된 방

법에 따라 증폭 및 정제된 PCR 산물은 BigDye terminator cycle sequencing v3.1 kits (Life Technologies, Carlsbad, USA)와 3730xl DNA analyzer (Hitachi High-Technologies Co, Tokyo, Japan)를 사용하여 시퀀싱 되었다. HLA 형별 데이터 분석은 uTYPE Dx HLA Sequence analysis software (One Lambda Inc., Los Angeles, USA)를 사용하였다.

결 과

1. HLA NGS 검사체계 구축

조혈모세포 기증희망자의 HLA NGS 검사를 위한 DNA 추출은 검체 당 하나의 튜브 내에서 추출의 모든 과정을 수행하여 검체 간 오염과 바뀔을 최소화 하였으며, 자동추출 장비를 이용하여 다량의 검체를 동시에 처리함으로써 핵산 추출 시간을 단축하였다. Multiplex PCR 방법을 이용하여 HLA-A, -B, -DRB1 유전자좌를 증폭하였다(PCR I). 일 회 시퀀싱에서 샘플 처리량을 최대화(96샘플/1회검사)하여 검사의 효율성을 높이기 위해, PCR II 과정에서는 바코드 시퀀스(10 bp 길이의 96개 바코드 시퀀스)와 Illumina MiSeq 시퀀싱을 위한 어답터 시퀀스를 수반하는 시발체 세트를 이용하여 검체마다 고유의 바코드를 부여하였고, 생성된 라이브러리를 MiSeq/MiSeqDx로 시퀀싱 하였다. 본 기관에서 개발한 NGS 기반 HLA 형별 분석 시스템(HLAveiw로 명명)은 96개의 검

체 기준으로 gDNA 추출에서 NGS 데이터 분석 및 HLA 형별 결과 보고까지 4 단계로 구분되며, 총 검사소요일은 4일이다(Fig. 2).

NGS로 생성된 데이터 분석은, 자체 개발된 HLA 형별 분석 소프트웨어인 PEERE 분석 소프트웨어를 이용하였다. PEERE 분석 소프트웨어는 HLA 형별 분석의 모호성 문제를 해결하고, 6 자리(6 digit) 이상 수준의 고해상도 분석이 가능하도록 개발된 소프트웨어로, 시퀀서에서 생산된 FASTQ 파일을 입력하면 자동으로 분석이 진행되며, 최종 결과는 엑셀 형식으로 얻어지게 된다. 평균 100개 샘플을 한 번에 분석하는 경우, 소요 시간은 약 30분 정도이다.

2. HLA NGS 검사체계 검증

총 220개의 검체에 대해 HLAview 검사를 실시하였고, 시행된 검사에 대한 NGS 데이터 매트릭스를 Table 1에 요약하였다. 검체당 평균 raw 데이터 생산량은 4.91 Mb, 리드수는 10,242개, 품질 점수(quality score, Q30)는 94.01%, 평균 리드 길이는 241 bp였다. 검체 당 생성된 총 리드 중 69%가 표적 영역에 매핑 되었으며, 표적 영역인 HLA-A, -B, -DRB1에 대한 평균 coverage depth는 각각 667x, 714x, 1679x였다.

PEERE 분석 소프트웨어를 포함한 HLAview 검사의 분석 성능을 검증하기 위해, 첫 번째, SSOP 방법으로 HLA 형별을 결정된 총 192개의 검체에 대해 PEERE 소프트웨어에 의해 분석된

Table 1. Summary of the HLA sequencing metrics from the MiSeq run (n=220)

Average raw data (Mb)	Average read numbers	Average read length (bp)	Average quality score (Q30) (%)	Average mapped read number	Average coverage depth (X)		
					HLA-A	HLA-B	HLA-DRB1
4.91	10,242	241	94.01	6,597	667	714	1,679

결과를 비교하였다. 분석에는 HLA-A, -B, DRB1 유전자와 각각에 대해 2개의 대립유전자를 나타내는 총 146개의 대립유전자형(HLA-A: 39종, HLA-B: 69종, HLA-DRB1: 40종)이 포함되었다. 총 192개 검체로부터 충분한 품질의 NGS 데이터가 생산되었고, 분석 결과 4 자리 수준에서 모호성을 보이지 않고 100% 일치하였다(data not shown). 두 번째로, HLA 형별이 알려진 28개 국내·외 외부정도관리 검체를 대상으로 HLAview와 SBT 검사를 동시에 시행한 후, 검사 결과를 비교하고 위상 모호성 해소율을 평가하였다. 두 방법에 의해 얻어진 결과는 6자리 수준에서 모두 일치하였다. SBT 결과 중 2개의 조합 결과를 보인 HLA-B 유전자 1개 유전자좌와 HLA-DRB1 유전자 3개의 유전자좌에서 관찰된 4개의 형별은, 본 기관의 PEERE 소프트웨어 분석에 의해 명확한 대립유전자 유형으로 구별되었다(Table 2). 따라서 NGS를 활용한 HLA 형별 분석이 Sanger 기반 형별 분석에서 관찰되는 위상 모호성을 해소시킴으로써 고해상도 형별 분석을 향상시킬 수 있음을 확인하였다.

PCR 증폭에 의한 짧은 가닥의 앰플리콘을 이용하여 신뢰할 수 있는 HLA 형별 분석 결과를 얻기 위해서는 각 유전자좌의 2개 대립유전자 각각에서 유래한 리드의 비율(allele balance)이 중요하게 고려되어야 한다[14,15]. 이를 조사하기 위해 총 220개 검체로부터 얻어진 NGS 데이터를

활용하여 세 개의 유전자좌에 대해 동형접합 유전형을 제외하는 2개의 대립유전자간 비율을 측정하였다. 각 유전자좌별 2개의 대립유전자 비율의 범위는 HLA-A는 0.26~0.74, HLA-B는 0.41~0.58, HLA-DRB1는 0.27~0.73로 나타났다. 이는 검사에 사용된 각 엑손 표적 시발체들이 상이한 형별을 모두 증폭함과 동시에 대립유전자간 증폭이 균형 있게 되었음을 보여준다.

고 찰

차세대염기서열분석법은 동시에 병렬적으로 여러 클론의 핵산을 시퀀싱할 수 있는 방법으로 2010년경에 상용화되어, 현재 암유전체 검사, 비침습적 산전선별검사, 미생물검사 등에 사용되고 있다[16-18]. 현재 전 세계적으로 많이 사용되고 있는 NGS 기기는 Ion Torrent사의 Personal Genome Machine 시리즈와 Illumina사의 MiSeq 시스템이다. 본 기관에서는 MiSeq 기기를 이용하여 HLA NGS 형별 분석법을 개발하였다. MiSeq 기기는 flow cell을 이용하며, 브릿지 증폭을 통해 형광으로 시퀀싱 데이터를 만들어낸다. NGS 분석 시간이 빠르고 높은 데이터 무결성이 MiSeq 기기의 특징이다[19]. NGS 검사 체계는 보통 핵산 추출, 라이브러리 제작, 표적 증폭, 시퀀싱 순으로 이루어지는 검사부분과 생물정보학으로 시퀀싱 데이터를 분석하는 정보분석 부분으로 나눌

Table 2. Resolved SBT combination results by the HLA NGS method

Sample	Locus	SBT combination 1		SBT combination 2		NGS	
UCLA-N04	DRB1	DRB1*04:03:01	DRB1*08:03:02	DRB1*04:03:01	DRB1*08:14	DRB1*04:03:01	DRB1*08:03:02
UCLA-N08	B	B*15:01:01G	B*39:06:02	B*15:04:01G	B*39:01:01G	B*15:01:01G	B*39:06:02
UCLA-N11	DRB1	DRB1*08:02:01	DRB1*08:03:02	DRB1*08:01:01	DRB1*08:30:01	DRB1*08:02:01	DRB1*08:03:02
UCLA-N12	DRB1	DRB1*04:02:01	DRB1*14:17	DRB1*04:04:01	DRB1*13:01:01G	DRB1*04:04:01	DRB1*13:01:01G

수 있다[20]. 본 기관에서 자체적으로 개발한 HLA NGS 부분은 검사부분의 라이브러리 제작과 정보분석 부분이다.

본 연구에서는 조혈모세포기증 희망자의 HLA 형별 분석을 위해 HLA-A, B, 그리고 DR에 대한 대립유전자만을 표적으로 하였다. 2017년 12월 기준으로 알려진 HLA-A 대립유전자 4,081개, B 대립유전자 4,950개, DRB1 대립유전자 2,146 개 모두를 증폭할 수 있도록 PCR I의 시발체를 디자인하였다. 총 HLA 표적 유전자가 7개로, 7개의 시발체 쌍을 디자인하였으며, 개별 PCR으로는 7회의 PCR 수행이 요구되지만, 시간과 노동력을 줄이기 위해 3그룹으로 다중 PCR 조건을 확립하였다. PCR I의 증폭산물의 크기가 500 bp 정도였어, 일반적인 HLA NGS 분석에서 시행되는 단편화(fragmentation) 과정 없이 바로 PCR II로 진행되었다[21]. PCR II에 사용된 시발체는 PCR I의 각 시발체에 연결된 paired end 어답터 시퀀스와 상보적인 결합을 하는 시퀀스, 바코드 시퀀스, 그리고 일루미나 시퀀싱 flow cell에 심겨진 시퀀스와 상보적으로 결합을 할 수 있는 시퀀스로 이루어져 NGS 시퀀싱이 가능하도록 만들었다. PCR II에 사용되는 바코드 시퀀스는 최대 96개의 개별 검체를 식별할 수 있도록 다양한 시퀀스 조합으로 제작되었다. PCR II 과정에서 고가의 NGS 라이브러리 준비 키트를 사용하지 않고 PCR을 통하여 어답터와 바코드를 붙이는 방법을 사용하였다. 96종류의 NGS-MiSeq 식별 바코드가 연결된 퓨전 시발체(fusion primer)를 이용하여, HLA 표적 PCR 산물을 대상으로 퓨전 PCR 조건을 확립하였다. PCR I의 세 그룹의 PCR 산물을 혼합하여 한 번에 퓨전 PCR을 수행하였고, 비특이적인 증폭 없이 7개의 산물이 모두 증폭됨을 확인하였다. 본 기관에서 연구 개발한 HLA NGS 형별 분석을 위한 라이브러리 제작 방법은 독자적인 기

술력을 인정받아 2017년 9월에 ‘차세대염기서열 분석기술 기반의 고효율성, 고해상도 조직적합성 형별 분석 방법 및 키트’로 국내에 특허 등록(제 10-1782806호)되었다.

본 기관에서는 차세대염기서열분석 파이프라인(PEERE)을 자체적으로 구축하여 HLA 대립유전자 분석을 수행하였다. PEERE 소프트웨어는 사용자가 파이프라인을 실행할 때 자동으로 IMGT/HLA 데이터베이스가 업데이트되도록 구현되어, 최신 버전의 참조 서열을 바탕으로 형별 분석이 이루어져 새로운 형별 식별이 용이하다는 장점이 있다. PEERE 알고리즘은 시퀀싱 오류로 나타난 형과 유사형을 구별할 수 있고, 두 형별이 비대칭적으로 증폭되어 마치 동형접합자처럼 나타나는 오류를 식별할 수 있게 설계되었다. 해외에서는 Assign ATF, HLA Twin 그리고 Sequence Pilot 등이 HLA NGS 데이터 분석 소프트웨어로 개발되어 있는데, 이들은 특정 시퀀싱 플랫폼에서 생산된 데이터 분석 전용이거나, 혹은 상용 소프트웨어이다[22]. 즉, Assign ATF 프로그램은 Roche 454 시퀀싱 플랫폼에서 생산된 엑손 시퀀싱 데이터에 적용하기 위해 개발된 소프트웨어이다. HLA Twin은 Omixon사에서 개발한 상용 소프트웨어로 consensus genotyping과 statistical genotyping이라는 두 개의 독립적인 알고리즘을 적용해서 HLA 형별 분석을 수행한다. 또 다른 상용 소프트웨어로 Sequence Pilot은 독일의 JSI Medical Systems사에서, 주로 임상 실험실에서의 사용을 고려하여 개발한 프로그램이다[22].

본 기관에서 개발한 NGS 기반 HLA 형별 분석 시스템(HLAveiw)은 96개 검체 기준으로 gDNA 추출, HLA 유전자 라이브러리 제작, 시퀀싱, NGS 데이터 분석 및 HLA 형별 보고까지 4단계로 구분되며, 총 4일이 소요된다. 본 기관에서 SBT법으로 96개의 검체를 처리하려면 약 2주 정

도가 소요된다. 또 자체 개발한 HLAVIEW 시스템은 기존 SSOP와 SBT HLA 형별 검사 결과와 일치하는 결과를 보여 기존 검사와 HLA 형별 결정력이 동등하거나 우수함을 보였다. NSG 기반 HLA 형별 검사법은 많은 검체를 짧은 시간에 처리가 가능하여 조혈모세포기증 희망자 HLA 검사 등에 유용할 것으로 기대된다.

감사의 글

본 연구개발은 중소기업청의 중소기업기술개발지원사업(S2091203)의 지원에 의해 이루어짐.

요 약

배경: 최근 차세대염기서열분석법(Next Generation Sequencing: NGS)을 이용한 HLA 형별 분석에 대한 연구가 활발히 진행되고 있다. 이에 HLA 고해상도 분석법의 내재적 한계인 위상 모호성의 문제를 해결하고, 대량 검체 처리가 가능한 NGS 기반 고해상도 HLA 형별 검사법을, 자체 기술로 개발하고자 본 연구를 실시하였다.

방법: HLA NGS를 위한 핵산 추출 조건, 라이브러리 제작 및 PCR 체계 확립, 그리고 생물정보학을 이용한 HLA 형별 분석법을 개발하였다. 본 기관에서 개발한 NGS 기반 HLA 형별 검사의 정확성을 알아보기 위해 SSOP법으로 HLA 형별을 알고 있는 192개 검체와 SBT법으로 HLA 형별을 알고 있는 28개 검체에 대해 NGS 기반으로 검사한 HLA-A, -B 그리고 -DR 형별 결과를 비교해 보았다.

결과: 두 단계의 PCR을 통한 DNA 라이브러리 제작과 MiSeq (Illumina Inc., San Diego, USA) 기기를 이용한 NGS 시퀀싱 그리고 데이터 분석 시스템을 구축하였다. 기존에 HLA 형별을 알고 있

는 220개 혈액 검체에 대해 NGS 기반 HLA 형별 검사 결과가 모두 일치함을 확인하였다.

결론: NSG 기반 HLA 형별 검사법은 많은 검체를 효율적인 시간 내에 처리가 가능하여 조혈모세포기증 희망자 HLA 검사 등에 유용할 것으로 기대된다.

References

1. Guy K, Krajewski AS, Dewar AE. Expression of MHC class II antigens in human B-cell leukemia and non-Hodgkin's lymphoma. *Br J Cancer* 1986;53:161-73
2. Kim HJ. HLA mismatch allogeneic blood and marrow transplantation. *Korean J Med* 2014; 86:1-13
3. Kim MH, Choi SE, Oh HB. Human leukocyte antigen typing proficiency surveys in Korea, 2005-2006. *Korean J Lab Med* 2007;27:442-50
4. Lee KW, Park MH. A comparative study of HLA-DR6 by serology and PCR-SSOP typing. *Korean J Clin Pathol* 1994;14:109-22
5. Cao K, Chopek M, Fernández-Viña MA. High and intermediate resolution DNA typing system for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). *Rev Immunogenet* 1999;1:177-208
6. Cha CH, Kim MH, Chung HJ, Choi SE, Oh HB. Evaluation of BioSewoom™ HLA-A,-B,-C PCR/SSP kit. *Korean J Lab Med* 2007;27: 360-8
7. In JW, Roh EY, Oh SH, Shin S, Park KU, Song EY. Allele and haplotype frequencies of Human leukocyte antigen-A, -B,-C,-DRB1, and -DQB1 from sequence-based DNA typing data in Koreans. *Ann Lab Med* 2015;35:429-35
8. Park YJ, Kim HS, Yoon CE, Kwon OJ, Kim YS. Resolution of ambiguous HLA genotyping in

- Korean by multi-group-specific sequence-based typing. *Yonsei Med J* 2014;55:1005-13
9. Petersdorf EW, Gooley TA, Anasetti C, Martin PJ, Smith AJ, Mickelson EM, et al. Optimizing outcome after unrelated marrow transplantation by comprehensive matching of HLA class I and II alleles in the donor and recipient. *Blood* 1998;92:3515-20
10. Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, et al. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc Natl Acad Sci U S A* 2012;109:8676-81
11. Lange V, Böhme I, Hofmann J, Lang K, Sauter J, Shöne B, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics* 2014;15:63
12. Gabriel C, Fürst D, Faé I, Wenda S, Zollikofer C, Mytilineos J, et al. HLA typing by next-generation sequencing-getting closer to reality. *Tissue Antigens* 2014;83:65-75
13. Gandhi MJ, Ferriola D, Huang Y, Duke JL, Monos D. Targeted next-generation sequencing for human leukocyte antigen typing in a clinical laboratory: metrics of relevance and considerations for its successful implementation. *Arch Pathol Lab Med* 2017;141:806-12
14. Profaizer T, Coonrod EM, Delgado JC, Kumánovics A. Report on the effects of fragment size, indexing, and read length on HLA sequencing on the Illumina MiSeq. *Hum Immunol* 2015;76:897-902
15. Pirooznia M, Kramer M, Parla J, Goes FS, Potash JB, McCombie WR, et al. Validation and assessment of variant calling pipelines for next-generation sequencing. *Hum Genomics* 2014;8:14
16. Kim JH, Park WY, Kim NKD, Jang SJ, Chun SM, Sung CO, et al. Good laboratory standards for clinical next-generation sequencing cancer panel tests. *J Pathol Transl Med* 2017;51:191-204
17. Seo DH, Cho DY, Kim JH, Kim SY, Cho SE, Oh M. Clinical study of non-invasive prenatal testing using next-generation sequencing. *J Lab Med Qual Assur* 2015;37:214-18
18. Gargis AS, Kalman L, Lubin IM. Assuring the quality of next-generation sequencing in clinical microbiology and public health laboratories. *J Clin Microbiol* 2016;54:2857-65
19. Na HS. Performance comparison of benchtop next-generation sequencing systems. *J Bacteriol Virol* 2014;44:208-13
20. Yohe S, Thyagarajan B. Review of clinical next-generation sequencing. *Arch Pathol Lab Med* 2017;141:1544-57
21. Yin Y, Lan JH, Nguyen D, Valenzuela N, Takemura P, Bolon YT, et al. Application of high-throughput next-generation sequencing for HLA typing on buccal extracted DNA: results from over 10,000 donor recruitment samples. *PLoS One* 2016;11:e0165810
22. Kunkel M, Duke J, Ferriola D, Lind C, Monos A. Molecular methods for human leukocyte antigen typing: current practices and future directions. In: Detrick B, Hamilton RG, Schmitz JL. *Manual of molecular and clinical laboratory immunology*. 8th ed. Washington, DC: ASM, 2016