Cellular Therapy and Transplantation (CTT). Vol.5, No.4 (17), 2016
doi: 10.18620/ctt-1866-8836-2016-5-4-63-70
Submitted: 14 September 2016, accepted: 25 November 2016

Current methods and opportunities of next-generation sequencing (NGS) for HLA typing

[1]Youri A. Serov, [1]Ildar M. Barkhatov, [1]Anton S. Klimov, [2]Andrey S. Berkos
[1]Research Institute of Children Oncology, Hematology and Transplantation, The First State I. Pavlov Medical University, St. Petersburg, Russia
[2]Russian Research Institute of Hematology and Transfusiology, Federal Medical-Biological Agency, St. Petersburg, Russia

Youri A. Serov, MD, Ph. D., CHS(ABHI), The Head of Biobank Research Unit, Raisa Gorbacheva Memorial Institute of Children's Hematology, Oncology and Transplantation, The First St. Petersburg I. P. Pavlov State Medical University, Leo Tolstoy str. 6-8, 197022, St. Petersburg, Russia
Phone: +7(812) 338-6260 (office); +7(981) 973-8041 (mob.)

Summary

The region encoding human leukocyte antigens (HLA) is a 3.6-Mb segment at chromosome 6p21. A highly complex HLA system comprises important molecules involved in transplantation immunity, classical direct sequencing methods making it very complicating to resolve the genomic specificities of HLA genes, due of the complex nature of them. Current limitations of existing methods and the increasing rate of new alleles, strongly demanding for a new methodology approach for the HLA genotyping. Next-Generation Sequencing (NGS) is a method that can provide a complete solution to the HLA typing problem. Last decade, the development of (NGS) has shown the easy way for whole-genome analysis in individuals. HLA typing has also benefited from NGS technologies. Several high-throughput HLA- NGS-based typing methods were developed using different technological platforms. In this review, we summarize the possible applications and instrumentation developments of NGS in the in-depth HLA typing, with a focus on future clinical applications. However, some issues of bioinformatics and data mining remain an obstacle to more effective analysis of HLA analysis in the context of donor/recipient compatibility assessment.

Keywords

Next-Generation Sequencing (NGS), technological solutions, HLA typing, allogeneic hematopoietic stem cell transplantation.

Introduction

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) is a proven option for treatment of hematological disorders, e.g., lymphoma, leukemia, and aplastic anemias. Among other factors, the HSCT outcomes strongly depend on aggressive reactions of donor lymphocytes against the patient's cells. This complication is known as a graft-versus-host disease (GVHD) and represents a serious threat to the patients, along with the risk of relapse and concomitant infections [1]. Over last decades, a progress in the donor-recipient matching lead to an imminent increase of patients' survival post transplant, due to decreased rates of immune complications [3,20]. These advances are made, mostly, due to an improved technique of HLA gene typing [24]. E.g., sufficiently increased survival rates and better outcomes were shown for the patient-donor pairs typed by means of high-resolution allele typing for HLA-A, -B, -C and -DRB1 [27]. Analysis of HLA-DQB1 and -DPB1 loci is still under development [10].

Hence, implementation of high-resolution techniques for the classic HLA typing brings about higher clinical efficiency in allo-HSCT from relatives and unrelated donors. The aspects of optimal donor matching are connected with a sufficient allelic diversity (polymorphism) of the gene loci, mainly, due to gene recombination of the chromosome 6 regions, thus producing repeats at distinct HLA gene segments and causing some difficulties with HLA typing, even at four-digit resolution level. Therefore, novel technological achievements in immunogenetics are of special importance, and, first of all, new approaches to DNA typing [6].

Milestones of molecular HLA typing

Since the very beginning of molecular biology, individual HLA matching was oriented for gene polymorphism analysis. Forty years ago, Fred Sanger has proposed a method of DNA sequencing, by the generation of distinct DNA strands based on a complementary matrix, using DNA polymerase I Klenoff fragment, and then followed by analysis of radioactively labeled DNA fragments in polyacrylamide gel. A breakthrough in this technology and development of routine HLA typing methods occurred with the application of fluorescent dideoxynucleotides in the course of the cyclic process followed by microcapillary electrophoresis. Direct sequencing of heterozygous DNA samples has its limitations since, in the case of allelic polymorphism, the nucleic bases may be located at the same or homologous chromosome, thus precluding detection of subtle differences in the given HLA gene sequence.

Over the last 20 years, the automated Sanger technique is a prevalent approach to genome sequencing in humans, animals, bacteria, and viruses. However, a need for more rapid routine genome screening stimulated novel technologies of multiplex DNA sequencing. These modern methods are depicted as the second-generation approaches (Next-Generation Sequencing, NGS). These technological platforms are based on different strategies, with respect to unique preparations of DNA templates, their sequencing, registration, retrieval and evaluation of the nucleotide sequences by means of novel bioinformatic approaches [13]. A principal benefit of the new-generation sequencing is an opportunity of getting large databases of multiple defined gene sequences within a short time period and at relatively low cost.

A known polymorphism of HLA genes presents a special problem in search of optimal donors for hematopoietic stem cell transplantation (HSCT). E.g., as for September 2016, the International Nomenclature (HLA International ImMunoGeneTics (IMGT) [26, 29] included 15,813 HLA and related alleles, registered in the numerical nomenclature used worldwide (Fig. 1).

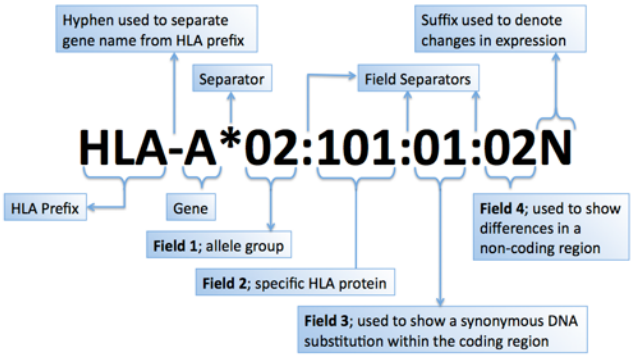

Figure 1. Current numerical indexing of HLA alleles.

In addition to the unique allele number, there are additional optional suffixes that may be added to an allele to indicate its expression status. Alleles that have been shown not to be expressed – "Null" alleles – have been given the suffix 'N'. Alleles that have been shown to be alternatively expressed may have the suffix "L", "S", "C", "A" or "Q". The suffix "L" is used to indicate an allele which has been shown to have "Low" cell surface expression when compared to normal levels. The "S" suffix is used to denote an allele specifying a protein which is expressed as a soluble, "Secreted" molecule but is not present on the cell surface. The "C" suffix is assigned to alleles that produce proteins that are present in the "Cytoplasm" and not on the cell surface. An "A" suffix indicates an "Aberrant" expression where there is some doubt as to whether a protein is actually expressed. A "Q" suffix is used when the expression of an allele is "Questionable", given that the mutation seen in the allele has been shown to affect normal expression levels in other alleles.

No wonder that the general NGS approach adapted for HLA typing proved to be a breakthrough in molecular biology applications being quite promising to the transplantation clinics and bone marrow donor registries. However, to promote the NGS implementation, we need specialized typing strategies and digital program algorithms. The sequencing costs per single run are sharply decreased with NGS approach which may be quite accessible to the tissue typing laboratories in sooner time.

Multiple NGS technological platforms are offered on the market. At the present time, four types of NGS systems are implemented in research and clinics. E.g., the 454 Junior (Roche) was the first system equipped with commercial kits for HLA-typing. Illumina platform is most widely applied in fundamental and clinical research, however, not yet registered for HLA-typing. The Ion Torrent system was offered something later, being, however, a growing point in the field. Lately, the fourth generation of NGS appears – Single Molecule Sequencing. They can do long reads with very high throughput. Advantages: Low Cost, Simultaneous sequencing of several samples (barcoding). Currently manufactured by 2 companies, PacBio and Nanopore MiniON. They are amplicon-based and "true" single molecule sequencing. It has a very small footprint and even fits in the pocket.

At the present time, a four-digit resolution approach is used for optimal matching in the donor-recipient pairs, i.e., testing of HLA-A, -B, -C, (exons 2, 3); -DRB1 (exon 2), and –DQB1 (exons 2, 3) by means of conventional PCR technique.

Pitfalls in NGS-based HLA typing

To perform high-resolution typing, one should exclude non-coding alleles from HLA-A (exons 1, 4 and introns 2, 4), HLA-B (exons 1, 4 and intron 1), and HLA-C (exons 1, 4, 7).

Phase determination of the sequence motifs becomes an increasingly difficult task, due to new HLA polymorphisms reported, thus causing an exponential increase in genotyping ambiguity with every ongoing release of the HLA database release [10]. Therefore, novel typing strategies have been arranged, e.g., group-specific PCR [15], or allelic discrimination with distinct group-specific primers [18] before subsequent sequencing. These discrimination techniques are used at HLA laboratories for a long time. However, they are labor-consuming, and appropriate commercial kits are largely lacking.

The novel NGS-based HLA typing was developed over 2000's, having been characterized by increased resolution ability and high throughput [22, 30]. In 2009, Gabriel et al. [11] and Erlich et al. [1] have independently shown an opportunity of HLA typing by means of a 454 technological platform. A double-blinded multicentric study based on exon amplification by primers' annealing was performed and reported in 2011 [13]. This cooperative work has shown that HLA sequencing with 454 system allows reliable identification of the HLA genotypes. However, the analytic approaches proved to be quite diligent, labor-intensive and require automation. Moreover, a systemic algorithm for the Class I HLA typing was performed with different PCR-based barcoding methods [9]. Later on, Lank et al. [16, 19] proposed a DNA processing protocol for HLA typing with a less complicated DNA library preparation. Finally, it was shown that, as unlike with amplicon sequencing, the shotgun sequencing approach using long PCR products, fragmenting and ligation of the multiplex identifiers (MID) is a realistic way for the entire HLA gene sequencing. [20, 27, 32].

Meanwhile, the existing protocols provide additional proofs of NGS complexity. This technology is still difficult to introduce in laboratories working in the area of histocompatibility and immunogenetics. There were some unresolved technical issues concerning labor intensity of a gene library preparation, and too long time required for processing of the databases produced. Noteworthy, the bioinformatic tools in NGS rarely generated a readable report of HLA typing.

Taking this into account, an automated approach to the high-resolutionHLA typing was developed, being adapted for the 454 GS Junior, covering 17 exons of the following genes: HLA-A, -B, -C, -DQB1, -DPB1, -DRB1, -DRB3, -DRB4 и -DRB5 (DRB3 / 4/5). However, exact data on 5"- and 3"- untranslated sequences (UTR) are not available for all the alleles. Therefore, binding of primers in these regions cannot be definitely predicted, thus missing some alleles from analysis or nucleotide shifts due to suboptimal amplification primers. Another potential problem may be due to simultaneous amplification of pseudogenes since their sequences are quite similar to those of target HLA genes.

Despite these issues, NGS has a great future in the area. An opportunity of the whole-gene analysis should ultimately lead to better understanding of HLA role for HSCT outcomes and the course of the primary disease. Moreover, the zero alleles could be better defined. So far, we cannot exploit this benefit. However, some questions will be answered, e.g., assessing a "complete" HLA allele database and overcoming existing hurdles when assembling a single complete allelic sequence.

At the initial step, a maximal length of sequences readable with Myseq was limited of 2×150 bp. However, the manufacturer (Illumina) improves its software, as well as chemical equipment. Hence, their last version (version 3) allows sequencing of ca.300 bp, in order to cover one exon within each sequenced amplicon, thus diminishing potential error rates and simplifying the work process.

Main steps of the NGS as performed with Illumina platform

1. Template generation: the complete sequences of HLA genes are amplified with complementary primers in a single reaction, using Long-Range DNA polymerase.

2. Double-stranded DNA fragmentation by means of specific fragmentase optimized by its size for the given HLA locus

3. Preparation of gene libraries: Appropriate reagents are used for the end repair, 5' phosphorylation of poly-A and poly-T ends for subsequent adaptors' ligation.

4. Adapters' ligation and their indexing.

5. Mixing of the samples and their loading to the NGS sequencer.

6. Analysis of results (Fig.1)

Following generation of a gene library, the DNA molecules are bound to the solid phase carrier, the s.c. flow well, by means of direct and reverse oligonucleotide adapters. During generation of the gene library, the complementary (adapter) sequences are fixed to the templates at the 5"- and 3"-ends, thus allowing them to bind the flow well surface covered with adapter molecules. To enhance the sequencing signal, molecular clusters are generated from one DNA molecule by means of solid-phase amplification bridge, resulting in up to 1000 tightly packed copies of the initial template in the flow well. Hence, each cluster corresponds to a single read by the end of sequencing. The adapter oligonucleotides are bound to the surface of flow well, both for the capture of sequenced templates molecules, and primers for bridge amplification. Sequencing is then carried out by means of chemical reaction "reversible terminator"' All 4 fluorescently labeled oligonucleotides are simultaneously driven through the flow well. Their labels are intended for termination of the chain synthesis. Therefore, only one nucleotide type is added for each polymerization cycle. The nonbound nucleotides are washed off. The labels and terminator are removed following CCD-assisted conversion of the pictures, in order to allow further elongation of the chain at the next sequencing cycle. To facilitate the multiplexing procedure, the adapter sequences are followed by the s.c. index sequences (DNA barcoding). These DNA barcodes are attached to the target fragments during the library design, with a unique code for given sample. Double indexing is possible by differential DNA barcoding from 5" to 3" end. In the course of four sequencing rounds, two indexing reads and two reads of target sequences are initiated. The data are accumulated for each single cluster and are used for generation of the output files in two directions of sequencing.

HLA data analysis with MiSeq platform

Miseq generates a working file in FASTQ format with two corresponding files (resp., direct and reverse reads). All the

files contain equal numbers of the read sequences presented in the same order. Moreover, all the sequences read are of equal length according to the chosen specifications. Homogeneous structure of initial data provides some benefits for the pre-analytic data handling. Upon filtration and cutting, some algorithms are required here which are more applicable than with platforms generating variable reading length (e.g., Roche 454, or Ion Torrent systems). Data quality obtained with Miseq is high enough to perform HLA typing.

To assign the HLA alleles, software is required which allows communicating with updated HLA databases (IMGT last version, renewed 4 times a year). A number of such commercial packages are currently available. To date, the solutions with an open initial code are also offered, as follows:

HLAminer: http://genomemedicine.com/content/4/12/95

seq2HLA: http://genomemedicine.com/content/4/12/102

Athlates: http://nar.oxfordjournals.org/content/41/14/e142.long

HLAforest: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0067885

PHLAT: http://www.biomedcentral.com/1471-2164/15/325

According to the guidelines of clinical trials and in order to observe acting regulations (European Federation for Immunogenetics (EFI), American Society for Histocompatibility and Immunogenetics (ASHI) standards), extensive studies of QC and QA are required for development of appropriate in-house reagents, and their testing for HLA typing bases on the Miseq platform. They should include both developments of the tests, and methods of automated data mining, and their presentation. Sufficient time is needed for design, feasibility evaluation, and implementation of NGS working protocols. The optimal choice of NGS platform and proper protocol is a difficult task for any individual setting in this rapidly developing technology [14].

Modern protocols of sequencing and data retrieval/handling differ, generally, in the aims of analysis, i.e., whole-genome evaluation (from 5'UTR to 3'UTR), or exome studies. So far, a limit of 400 bp exists for the length of sequencing, independently of the working protocol applied. This length of DNA fragment is sufficient to reading of whole exons and providing unequivocal clonal information for these exons. To sequence longer PCR products, a series of overlapping fragments is required, followed by their subsequent assembly (shotgun sequencing). The working processes under these two protocols significantly differ in labor intensity and complexity.

The output data are collected for each molecular cluster and used for generation of the output files containing results of direct and reverse reads as displayed on Fig. 2 [5]

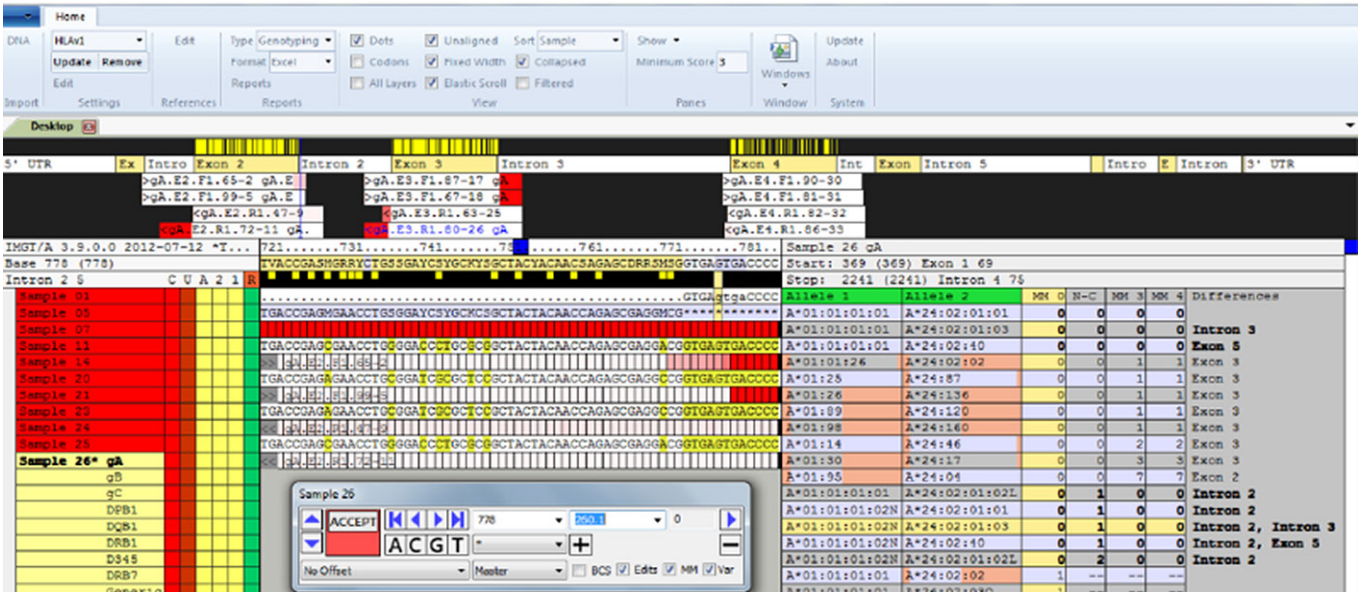

Figure 2. Interface ATF 4.7. (Conexio Genomics, Fremantle, Australia) Interim results of HLA typing.

The detailed information concerns HLA-A genotype assignment, as show in the master layer. Direct and reverse sequence reads are aligned to exons 2, 3, and 4, with partial coverage of intronic regions [33]. Read numbers and sequence direction are shown in white cells. For example, exon 4 is represented by two (alleles 1 and 2) direct and reverse sequences, with 90 coverages, and, respectively, 81 nucleotides read, and two reverse sequences with 82 and 86 nucleotides read. Low case: exon map, consensus sequence, and genotype assigned. Stepwise distribution of the sequences at exon2/intron 2 border is also shown, and an HLA-A * 01: 01: 01: 01, 24: 02: 01: 01 combination is proposed with zero mismatch conditions (MM master layer) as compared with the database. A non-expressed (A* 01: 01: 01: 02 N), and low-expressed (A * 24: 02: 01: 02 L) variants are discriminated by means of exclusion of intron 2 variants, as shown in the non-coding column (N-C). Discordances in phase layers 3 and 4 (MM3, MM4) demonstrate potential ambiguities which could occur if the SNP phase in exons 3 and 2 could not be revealed (with conventional Sanger sequencing). Other detectable alleles, A * 24: 02: 01: 03 (change in intron 3), and A * 24: 02: 40 (synonymous substitution in exon 5) cannot be excluded, and therefore, are displayed in a similar way with zero mismatches. [5]

Factors to consider in cost analysis of in-depth sequencing

Reagents for the entire process include those used for conventional pre-analytic methods (e.g., DNA extraction, quality assessment, and initial low-resolution typing step. Additional expenditures are a subject to some ambiguities, due to different prices reagents and equipment offered by distinct manufacturers. Moreover, it should be stressed that all commercial NGS systems are of closed type, thus causing broad variations in prices for the entire NGS procedure per single DNA sample, and depending on the per year capacity of the given HLA typing laboratory.

However, even taking into account maintenance costs (estimated 10% equipment cost), usage of core facilities or shared equipment, the Sanger sequencing (220 K) proves to be twice more expensive than NGS (variable, all less than Sanger), as shown in Fig.3. I.e., the sample preparation costs remain the same however the sequencing cost decreases (Fig. 4).

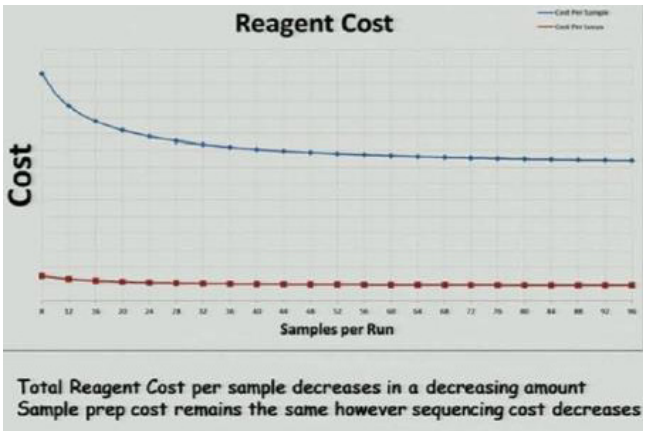

Figure 3. Total Reagent Cost per sample decreases with increased HLA-typing activity. [19]

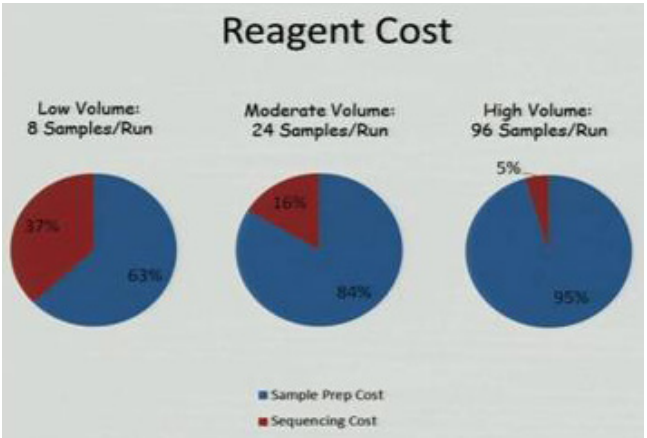

Figure 4. The more samples running, the less sequencing cost of NGS procedures [19]

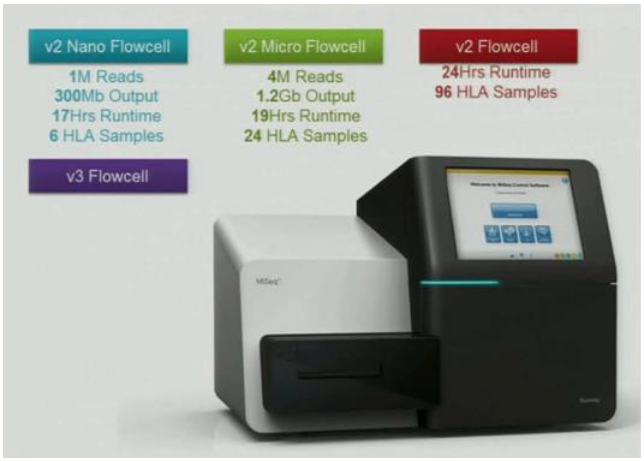

Figure 5. The MiSeq desktop sequencer is suitable for smaller volume labs, allowing access more focused applications such as targeted gene sequencing, metagenomics, amplicon sequencing etc. starting at 10 ng DNA, and HLA typing. New MiSeq reagents enable up to 15 Gb of output with 25M sequencing reads and 2×300 bp read lengths. [19].

Current progress and future directions

In 2005, Roche produced the first NGS instrument, the Genome Sequencer 20. By means of this device, 100-bp reads are possible, and in future, up to 20 Mbp could be sequenced. Over last decade, rapid progress in NGS technologies has led to revolutionary changes in general genomics and its applications for clinical sequencing or medical exomics (RNA expression analysis). However, PCR-SSO and PCR-sequencing still remain the first-line methods in HLA typing. Meanhwile, commercial kits for NGS-assisted HLA typing have been recently appeared [8]. Hence, elucidation of the whole HLA gene sequence will bring a knowledge which will be of value for medical science in sooner time. However, this knowledge will be insufficient to the total understanding of MHS and HLA interactions. One should assess the sequences and haplotypes of other HLA segments, including regulatory (non-coding) regions [12]. Further analysis should determine transcriptional factors and events regulating the entire HLA unit, including HLA genes and relevant gene structures. They all should be carefully studied, in order to get complete HLA description, i.e., HLA-omics [4]. These efforts will bring real clinical benefits for the patients. Novel methods of HLA typing will be helpful to more accurate medical examination, assignment of biologically diverse clinical groups and individualized treatment of the patients.

Conclusions

A routine tissue-typing laboratory should be flexible, providing valid and reproducible HLA data within short time period. NGS-based HLA analysis is performed with 100% reliability, and well fits the tasks of HLA typing in unrelated donors, in concordance with EFI and ASHI policies. This work process well corresponds to the working schedules for

medium- and high-capacity laboratories, thus being potentially attractive to the donor registries.

Recently introduced next-generation sequencing techniques have a facilitating potential for the high-resolution genotyping via decrease of general uncertainty, like as due to extended sequencing regions. The entire set of MHS genes could be sequenced by means of these technologies, in order to uncover the latent immune determinants relevant to autoaggresive reactions [7, 25]. In near future, the whole exome sequencing may be considered a method of decoding some genetic variants, influencing HSCT outcomes. The NGS approaches will be a quite effective and cost-effective technology when evaluating histocompatibility parameters and immunogenetic interactions.

Conflict of interests

No conflict of interests is declared.

References

1. Bentley G, Higuchi R, Hoglund B, Goodridge D, Sayer D, Trachtenberg EA, Erlich HA: High-resolution, high-throughput HLA genotyping by nextgeneration sequencing. Tissue Antigens 2009; 74: 393–403.

2. Blazar BR, Murphy WJ, Abedi M: Advances in graft-versus-host disease biology and therapy. Nat Rev Immunol 2012; 12: 443–458.

3. Brunstein CG, Eapen M, Ahn KW, Appelbaum FR, Ballen KK, Champlin RE, Cutler C, Kan F, Laughlin MJ, Soiffer RJ, Weisdorf DJ, Woolfrey A, Wagner JE: Reduced-intensity conditioning transplantation in acute leukemia: the effect of source of unrelated donor stem cells on outcomes. Blood 2012; 119: 5591–5598.

4. Danzer M, Polin H, Proll J, Hofer K, Fae I, Fischer GF, Gabriel C: Highthroughput sequence-based typing strategy for HLA-DRB1 based on real-time polymerase chain reaction. Hum Immunol 2007; 68: 915–917.

5. Danzer M, Niklas N, Stabentheiner S, Hofer K, Pröll J, Stückler C, Raml E, Polinand H, Gabriel C. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. BMC Genomics 2013, 14: 221.

6. Dunn PP: Human leucocyte antigen typing: techniques and technology, a critical appraisal. Int J Immunogenet 2011; 38: 463–473.

7. Elsner HA, Blasczyk R: Immunogenetics of HLA null alleles: implications for blood stem cell transplantation. Tissue Antigens 2004; 64: 687–695.

8. Erlich H: HLA DNA typing: past, present, and future. Tissue Antigens 2012; 80: 1–11.

9. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, Gupta N, DePristo MA, Henn MR, Lennon NJ, de Bakker PI: Next-generation sequencing for HLA typing of class I loci. BMC Genomics 2011; 12: 42.

10. Fleischhauer K, Shaw BE, Gooley T, Malkki M, Bardy P, Bignon JD, Dubois V, Horowitz MM, Madrigal JA, Morishima Y, Oudshoorn M, Ringden O, Spellman S, Velardi A, Zino E, Petersdorf EW: Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoieticcell transplantation: a retrospective study. Lancet Oncol 2012; 13: 366–374.

11. Gabriel C, Danzer M, Hackl C, Kopal G, Hufnagl P, Hofer K, Polin H, Stabentheiner S, Proll J: Rapid high-throughput human leukocyte antigen typing by massively parallel pyrosequencing for high-resolution allele identification. Hum Immunol; 2009, 70: 960–964.

12. Harismendy O, Frazer K: Method for improving sequence coverage uniformity of targeted genomic intervals amplified by LR-PCR using Illumina GA sequencing-by-synthesis technology. Biotechniques 2009;46:229–231.

13. Holcomb CL, Hoglund B, Anderson MW, Blake LA, Bohme I, Egholm M, Ferriola D, Gabriel C, Gelber SE, Goodridge D, Hawbecker S, Klein R, Ladner M, Lind C, Monos D, Pando MJ, Proll J, Sayer DC, Schmitz-Agheguian G, Simen BB, Thiele B, Trachtenberg EA, Tyan DB, Wassmuth R, White S, Erlich HA: A multi-site study using high-resolution HLA genotyping by next generation sequencing. Tissue Antigens 2011, 77: 206–217.

14. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM: Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biol 2007; 8: R143.

15. Kotsch K, Wehling J, Blasczyk R: Sequencing of HLA class II genes based on the conserved diversity of the non-coding regions: sequencing based typing of HLA-DRB genes. Tissue Antigens 1999; 53: 486–497.

16. Lank SM, Golbach BA, Creager HM, Wiseman RW, Keskin DB, Reinherz EL, Brusic V, O'Connor DH: Ultra-high resolution HLA genotyping and allele discovery by highly multiplexed cDNA amplicon pyrosequencing. BMC Genomics 2012; 13: 378.

17. Lank SM, Wiseman RW, Dudley DM, O'Connor DH: A novel single cDNA amplicon pyrosequencing method for high-throughput, cost-effective sequence-based HLA class I genotyping. Hum Immunol 2010; 71: 1011–1017.

18. Lebedeva TV, Mastromarino SA, Lee E, Ohashi M, Alosco SM, Yu N: Resolution of HLA class I sequence-based typing ambiguities by groupspecific sequencing primers. Tissue Antigens 2011, 77: 247–250.

19. Baxter-Lowe LA. Tailoring NGS for smaller volume labs. 42nd ASHI Annual Meeting. St. Louis, MO, Poster report (Sept 28, 2016).

20. Li HW, Sykes M: Emerging concepts in haematopoietic cell transplantation. Nat Rev Immunol 2012; 12: 403–416.

21. Lind C, Ferriola D, Mackiewicz K, Heron S, Rogers M, Slavich L, Walker R, Hsiao T, McLaughlin L, D'Arcy M, Gai X, Goodridge D, Sayer D, Monos D: Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. Hum Immunol 2010; 71: 1033–1042.

22. McDevitt SL, Bredeson JV, Roy SW, Lane JA, Noble JA. HAPCAD: An open-source tool to detect PCR crossovers in next-generation sequencing generated HLA data. Hum Immunol. 2016; 77 (3): 257-63. doi: 10.1016/j.humimm.2016.01.013.

23. Metzker ML: Sequencing technologies – the next generation. Nat Rev Genet 2010; 11: 31–46.

24. Petersdorf EW: Optimal HLA matching in hematopoietic cell transplantation. Curr Opin Immunol 2008; 20: 588–593.

25. Proll J, Danzer M, Stabentheiner S, Niklas N, Hackl C, Hofer K, Atzmuller S, Hufnagl P, Gully C, Hauser H, Krieger O, Gabriel C: Sequence capture and next generation resequencing of the MHC region highlights potential transplantation determinants in HLA identical haematopoietic stem cell transplantation. DNA Res 2011; 18: 201–210.

26. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 2015; 43(Database issue): D423-31. doi: 10.1093/nar/gku1161.

27. Shaw BE, Arguello R, Garcia-Sepulveda CA, Madrigal JA: The impact of HLA genotyping on survival following unrelated donor haematopoietic stem cell transplantation. Br J Haematol 2010; 150: 251–258.

28. Shiina T, Suzuki S, Ozaki Y, Taira H, Kikkawa E, Shigenari A, Oka A, UmemuraT, Joshita S, Takahashi O, Hayashi Y, Paumen M, Katsuyama Y, Mitsunaga S, Ota M, Kulski JK, Inoko H: Super high resolution for single molecule sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. Tissue Antigens 2012; 80: 305–316.

29. Marsh SG, WHO Nomenclature Committee for Factors of the HLA System. Nomenclature for factors of the HLA system, update March 2016. Human Immunology 2016; 77 (8): 706-709.

30. Voorter CE, Mulkers E, Liebelt P, Sleyster E, van den Berg-Loonen EM: Reanalysis of sequence-based HLA-A, -B and -Cw typings: how ambiguous is today's SBT typing tomorrow. Tissue Antigens 2007; 70: 383–389.

31. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin Chem 2009; 55: 641–658.

32. Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, Levinson D, Fernandez-Vina MA, Davis RW, Davis MM, Mindrinos M: Highthroughput, high-fidelity HLA genotyping with deep sequencing. Proc Natl Acad Sci USA 2012; 109: 8676–8681.

33. Weimer ET, Montgomery M, Petraroia R, Crawford J, Schmitz JL. Performance characteristics and validation of next-generation sequencing for human leucocyte antigen typing. J Mol Diagnostics 2016; 18 (5): 668-675.

Современные методы и возможности секвенирования следующего поколения (NGS) для типирования системы HLA

[1]Юрий А. Серов, [1]Ильдар М. Бархатов, [1]Антон С. Климов, [2]Андрей С. Беркос
[1]НИИ детской онкологии, гематологии и трансплантации им. Р. М. Горбачевой, Первый Санкт-Петербургский государственный медицинский университет им. И. П. Павлова, Санкт-Петебург, Россия;
[2]Российский НИИ гематологии и трансфузиологии, Санкт-Петербург, Россия

Резюме

Область генома, кодирующая лейкоциты антигенов человека (HLA) – это сегмент длиной 3,6 миллиона пар оснований на хромосоме 6(p21). Сложная по антигенному составу система HLA содержит важные молекулы, участвующие в трансплантационном иммунитете. Специфические генные последовательности генов HLA довольно сложно расшифровать с помощью классических методов прямого секвенирования в связи с их техническими ограничениями, сложным составом генов HLA, и возрастающим числом новых аллелей, что требует новых методов генотипирования. Секвенирование следующего поколения (NGS) является методом, который может обеспечить полное решение проблемы HLA-типирования. За последнее десятилетие разработка NGS обеспечила более легкий способ полногеномного анализа у человека, в том числе для типирования генов HLA. Несколько вариантов высокопроизводительных методов оценки HLA, основанные на мультиплексных аналитических технологиях NGS были разработаны на основе различных технологических платформ. В этом обзоре мы обсуждаем возможные области применения и прогресс в оборудовании для NGS в плане глубокого типирования системы HLA, с особым вниманием к будущим аспектам его клинического использования. Общеизвестный полиморфизм генов HLA и минимальные межаллельные различия представляют собой особую проблему. Поэтому некоторые вопросы, связанные с биоинформатикой и углубленной обработкой данных для более эффективного анализа HLA в контексте оценки совместимости донора и реципиента. В этом плане различными производителями предложены специализированные компьютерные программы для анализа больших баз данных, получаемых посредством технологий NGS. Наконец, эти высокопроизводительные подходы позволяют минимизировать финансовые затраты на 1 образец, особенно в крупных HLA-лабораториях, где соответствующие расходы на NGS-исследования стали теперь значительно затрат, возникающих при использовании классических технологий секвенирования по Сэнгеру.

Ключевые слова

Секвенирование следующего поколения (NGS), технологические решения, HLA-типирование, аллогенная трансплантация гемопоэтических стволовых клеток.