

Pengembangan Metode *In-House HLA-Typing* Gen HLA Kelas I (HLA A, HLA B, dan HLA C) Menggunakan *Next Generation Sequencing Illumina MiSeq*

Rika Yuliwulandari,^{1,5} Kinasih Prayuni,^{2,5} Kenconoviyati,³ R. W. Susilowati,³ Abdul Salam M. Sofro⁴

¹Departemen Farmakologi, Fakultas Kedokteran, Universitas YARSI Jakarta, ²Departemen Anatomi, Fakultas Kedokteran Universitas YARSI Jakarta, ³Department of Histology, Fakultas Kedokteran, Universitas YARSI Jakarta, ⁴Sekolah Pasca Sarjana Universitas YARSI Jakarta, ⁵Genomic Medicine Research Group, Lembaga Penelitian Terpadu Universitas YARSI, Jakarta

Abstrak

Human leucocyte antigen (HLA) adalah protein penyaji antigen yang lokus genetiknya berada di kromosom 6p21 dengan ukuran sebesar 3,8 Mb dan berasosiasi dengan lebih dari 100 penyakit berbeda yang kebanyakan merupakan penyakit autoimun. Proses *HLA-typing* menggunakan sekuensing Sanger masih memberikan ambiguitas terhadap determinasi alel, *low-throughput*, dan membutuhkan biaya besar untuk sampel dalam jumlah besar. *Next generation sequencing* (NGS) menjadi metode yang dapat mengatasi kelemahan sekuensing Sanger. MiSeq dari Illumina merupakan salah satu NGS yang digunakan untuk *HLA-typing*. MiSeq memberikan kemudahan preparasi dan fleksibilitas metode yang dapat dikembangkan sesuai dengan kebutuhan laboratorium penelitian. Penelitian dilakukan di Laboratorium Molekular Genetik, Laboratorium Terpadu Universitas YARSI pada empat orang mahasiswa Fakultas Kedokteran Universitas YARSI etnik Melayu selama periode Mei–Desember 2014. Hasil menunjukkan terdapat total 546 SNP *heterozygous*, 888 SNP *homozygous*, 25 insersi, dan 23 deleksi dari keseluruhan 11 sampel amplicon dengan *coverage* 2.106,536x dengan 2x25 siklus pembacaan. Optimasi metode *HLA-typing* dapat dikatakan berhasil dengan mengombinasikan *long-range* PCR dan pemilihan ukuran *library* 300–600 bp. [MKB. 2015;47(3):152–59]

Kata kunci: HLA Kelas I, MiSeq, next generation sequencing

Development of Class I HLA Gene *In-House HLA-Typing* Methods (HLA A, HLA B, and HLA C) using *Next Generation Sequencing Illumina MiSeq*

Abstract

Human leukocyte antigen (HLA) is a 3.8 Mb protein presenting antigen whose genetic locus is located in chromosome 6p21 area and have association with more than 100 different diseases that are mostly autoimmune diseases. *HLA-typing* process using Sanger sequencing still creates ambiguity in the determination of alleles, *low-throughput*, and costly as it requires a large quantity of sample. *Next generation sequencing* (NGS) is a method that can overcome the drawbacks of Sanger sequencing. MiSeq Illumina is one of the NGSs that are used for *HLA-typing*. This study was conducted at the Laboratory of Molecular, Universitas YARSI in a period from May to December 2014. MiSeq provides convenience and flexibility in the preparation methods that can be developed according to the needs of the research laboratory. The results showed that there were a total of 546 SNPs that were heterozygous, 888 homozygous SNP, 25 insertions and 23 deletions from the overall 11 amplicon samples with an average coverage with 2x25 read length of 2,106,536x. Our protocol generates good result as we combined long PCR amplicon and size selection method to 300–600 bp fragment. [MKB. 2015;47(3):152–59]

Key words: HLA Class I, MiSeq, next generation sequencing

Korespondensi: Rika Yuliwulandari, dr., Ph.D, Departemen Farmakologi Fakultas Kedokteran Universitas Yarsi, Jalan Letjen Suprpto Cempaka Putih, Jakarta 10510, Tlp:(021) 4206674, *mobile* 087782065501, *e-mail* rika.yuliwulandari@yarsi.ac.id

Pendahuluan

Gen *Human Leucocyte Antigen* (HLA) yang berada di kromosom 6p21 dan mempunyai sekuens genomik sebesar 3,6 Mbp¹ diketahui berperan penting pada transplantasi organ dan *haematopoietic stem cell*,^{2,3} penyakit infeksius (HIV, hepatitis, CMV),⁴ penyakit autoimun (diabetes, *rheumatoid arthritis*, *celiac*),⁵ kanker (*nasopharyngeal carcinoma*, kanker payudara),^{6,7} dan reaksi hipersensitivitas obat.⁸

Proses identifikasi alel HLA telah dilakukan dengan berbagai cara, dari metode berbasis serologis hingga metode berbasis DNA seperti RFLP, PCR-SSO, PCR-SSP, and *Sanger sequence-based typing* (SSBT).^{9,10} Saat ini sekuensing DNA merupakan metode baku terbaik untuk HLA-typing yang telah digunakan secara luas, telah teroptimasi dengan baik, dan merupakan teknik yang handal serta efisien untuk mendeteksi dan mengidentifikasi alel-alel pada 6 major lokus HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1 serta HLA-DPB1 yang memiliki tingkat polimorfisme tinggi.

Sekuensing Sanger terkenal karena akurasi dan panjang pembacaannya (800–1.000 bp) telah mendominasi selama hampir 30 tahun sebagai metode baku untuk analisis genotipe. Sekuensing sanger juga merupakan metode umum yang digunakan selama hampir 10 tahun untuk HLA-typing. Namun demikian, metode ini secara simultan melakukan sekuensing pada dua kromosom sehingga pada region HLA yang sangat polimorfik sangat sulit dilakukan. Umumnya proses sekuensing Sanger mengikutsertakan ekson 2 dan 3 dari gen HLA kelas I dan ekson 2 gen HLA kelas II, namun demikian pada beberapa kasus, alel yang berbeda berbagi sekuens yang mirip pada sekuens di daerah tersebut sehingga mengakibatkan ambiguitas pada determinasi alel.¹¹ Biaya yang besar dan *low-throughput* membuat metode sekuensing Sanger ini sangat terbatas untuk proyek HLA-typing menggunakan jumlah sampel besar.¹²

Kemajuan teknologi dengan hadirnya *next generation sequencing* (NGS) mengakibatkan revolusi di dalam bidang medis, terutama untuk menyediakan data sekuensing yang lebih besar daripada metode sekuensing Sanger. Proyek 1000 genom sudah dilaporkan menemukan varian baru termasuk varian yang langka dengan metode NGS.¹³ Namun demikian, untuk menganalisis gen HLA yang sangat polimorfik dibutuhkan prosedur tertentu dan teknologi NGS memiliki potensi untuk menganalisis seluruh bagian gen HLA yang sangat polimorfik tersebut.

Keterbatasan dalam sekuensing Sanger dapat diatasi dengan metode NGS, antara lain NGS dapat menganalisis urutan dari satu kromosom, *cost effective* untuk sampel dalam jumlah besar, dan juga *high-throughput*.¹¹

Sejauh ini beberapa metode HLA-typing yang menggunakan NGS telah dikembangkan, antara lain dengan 454 GS FLX Titanium untuk menganalisis HLA kelas I dengan akurasi yang cukup tinggi HLA-A (95,9%), HLA-B (99,4%), dan juga HLA-C (94,4%).¹⁰ Teknologi Illumina menunjukkan keakuratan genotipe dan cakupan yang lebih luas untuk tujuh ekson pertama HLA kelas I (HLA-A, -B, dan C) dan ekson 2–5 dari HLA kelas II (HLA-DRB1).¹⁴ Teknologi Roche GS Junior dan juga Ion PGM yang dikombinasikan dengan *long-range* PCR menemukan 8-digit alel HLA.¹⁵

Wang dkk.¹⁴ dan Hosomichi dkk.¹¹ telah mendemonstrasikan suatu metode amplicon sekuensing mempergunakan MiSeq. Kombinasi *long-range PCR amplicon* dengan konstruksi *library* berbasis transposase, telah mampu mendeterminasi alel dan haplotipe sekuens dari keseluruhan region gen HLA (HLA-A, -C, -B, -DRB1, -DQB1, dan -DPB1). Di Indonesia belum banyak dilakukan HLA-typing menggunakan NGS. Oleh karena itu penelitian ini bertujuan mengembangkan metode *in house HLA typing* dengan NGS MiSeq Illumina untuk jumlah sampel besar dan *high throughput*.

Metode

DNA genomik berasal dari empat orang mahasiswa Fakultas Kedokteran Universitas YARSI yang berasal dari etnik Melayu. Ekstraksi DNA dilakukan dengan menggunakan QiaQuick DNA Blood Mini Kit (Qiagen, Hilden, Germany).

Amplifikasi PCR mempergunakan primer spesifik untuk HLA kelas I (HLA A, -B, dan -C) sebagaimana tercantum pada Tabel 1. Desain primer PCR dilakukan dengan bantuan tim Illumina dengan posisi primer *forward* berada pada ekson 1 dan posisi primer *reverse* berada pada ekson 7 dan sebagian ekson 8 berdasarkan sekuens acuan *Homo sapiens chromosome 6, GRCh38 Primary Assembly, NCBI Reference Sequence NC_000006.12*. Amplifikasi PCR dikerjakan mempergunakan KAPA Taq Extra HotStart Ready Mix (KAPA Biosystem, Salt River, Cape Town, South Africa) untuk 12 amplicon dari 4 individu. Reaksi PCR dengan total volume 25 dengan 60 ng *template* DNA, 1x KAPA Taq Ready Mix, 0,5 μ M primer *forward* dan *reverse*. Profil

PCR menggunakan kondisi suhu 94°C selama 3 menit diikuti sebanyak 35 siklus dari 94°C selama 30 detik, 60°C selama 20 detik, dan 72°C selama 3 menit 20 detik, diakhiri dengan suhu 72°C selama 10 menit, dan *fase hold* pada 4°C. Hasil amplifikasi berupa satu *band* tunggal pada gel agarosa 1% (Invitrogen, Invitrogen, Thermo Fisher Scientific, Grand Island, NY, USA) yang kemudian dipurifikasi menggunakan *QiaQuick PCR Purification Kit* (Qiagen, Hilden, Germany), kemudian amplikon dikuantifikasi dengan *Qubit Fluorometric Quantitation, dsDNA HS Assay Kit* (Life Technologies, Thermo Fisher Scientific, Grand Island, NY, USA).

Sekuensing dilakukan menggunakan MiSeq (Illumina) yang memiliki tiga tahapan, yaitu konstruksi *library* DNA, proses sekuensing, dan analisis data yang sudah terintegrasi pada mesin MiSeq. Konstruksi *Library* dilakukan menggunakan *Nextera DNA Sample Preparation Kit* (Illumina, San Diego, CA, USA) yang dimediasi oleh enzim transposom dengan konsentrasi awal DNA genomik sebesar 50 ng selama 90 menit. Secara umum proses konstruksi *library* DNA dilakukan dengan menambahkan produk amplikon PCR dengan enzim transposom yang telah didesain sedemikian rupa sehingga membawa fragmen oligonukleotida sintetik berisi adapter P5 dan P7, primer pembacaan sekuensing dan sekuens indeks spesifik (Gambar 1).

Enzim transposom secara simultan melakukan fragmentasi sekaligus menambahkan sekuens adapter pada kedua ujung amplikon DNA. Purifikasi dilakukan dengan *Zymo Purification Kit* (Zymo Research, Irvine, CA, USA) untuk menghilangkan sisa enzim transposom. Setiap sampel memiliki *dual indexed* dan secara sama

rata di-*pooling* dengan mengevaluasi konsentrasi molar dengan menghitung berdasarkan hasil *real time* PCR (Ligth Cycler, Roche Diagnostic, Mannheim, Germany) dan pengukuran dengan Qubit HS DNA (Life Technologies, Thermo Fisher Scientific, Grand Island, NY, USA). Fragmen berukuran 300–600 bp diseleksi dari *library* yang sudah terbentuk dengan menambahkan prosedur pemotongan gel dari gel agarosa 2% (Invitrogen, Thermo Fisher Scientific, Grand Island, NY, USA) yang kemudian dipurifikasi menggunakan *GeneJet Purification Kit* (Thermo Scientific, Grand Island, NY, USA).

Sampel hasil konstruksi *library* disekuensing secara multipleks pada *sequencer* MiSeq (Illumina, San Diego, CA, USA) dengan 50 *Cycle MiSeq Reagent Kit V.1* (Illumina) secara *paired-end* (2x25 *reads length*), yaitu sekuensing dari 2 arah pembacaan, secara *forward* dan *reverse* sepanjang 25 bp pada masing-masing kluster yang terbentuk. Kontrol DNA PhiX digunakan dalam proses sekuensing yang digabungkan dengan *pooling library* sebesar 5%.

Data analisis menggunakan *MiSeq Reporter* (MSR) yang dilakukan *on-board* pada MiSeq serta *Sequencing Analysis Viewer* (SAV) (Illumina, San Diego, CA, USA) untuk memvisualisasi hasil sekuensing. MSR melakukan *alignment* dengan *reference* yang digunakan sebagai acuan untuk membuat primer HLA A, -B, dan -C (*Homo sapiens chromosome 6, GRCh38 Primary Assembly, NCBI Reference Sequence: NC_000006.12*).

Hasil

Desain primer PCR gen *HLA* kelas 1 dilakukan pada daerah ekson yang sangat polimorfik,

Tabel 1 Primer yang Digunakan untuk Amplifikasi PCR Gen *HLA* Kelas I

Primer	Urutan Sekuens (5'→3')	Ukuran Fragmen (bp)
HLA A		
AF2	GCGGTGTATGGATTGGGAG	3.190
AR2	ACAAAGGGAAGGGCAGGAAC	
HLAB		
BF1	GTGTCGGGTCTTCTTCCAG	3.076
BR4	CAAAGGGGAGGCGTGAAGAA	
HLA C		
CF1	TTCACCTTCTCCCAACCTGCG	3.193
CR2	GCTAACAGGAACGCAGACAC	

Keterangan: FA2: Primer forward HLA-A; RA2: Primer reverse HLA-A; BF1: Primer forward HLA-B; BR4: Primer reverse HLA-B; CF1: Primer forward HLA-C; CR2: Primer reverse HLA-C

Tabel 2 Persebaran Indeks Sampel Amplikon Gen HLA Kelas I

No	Nama Sampel	Indeks 1 (I7)	Indeks 2 (I5)	% Reads Identified (Passing Filter)*
1	16A	GGACTCCT	TAGATCGC	2,3968
2	16B	TAGGCATG	TAGATCGC	1,8927
3	16C	GGACTCCT	TATCCTCT	2,6502
4	22A	TAGGCATG	TATCCTCT	2,2438
5	22C	CGTACTAG	CTCTCTAT	6,3357
6	40A	AGGCAGAA	CTCTCTAT	5,7932
7	40B	TCCTGAGC	CTCTCTAT	11,5576
8	40C	GGACTCCT	CTCTCTAT	10,622
9	42A	TAGGCATG	CTCTCTAT	6,5488
10	42B	GGACTCCT	AGAGTAGA	5,122
11	42C	TAGGCATG	AGAGTAGA	9,8423

Keterangan: *pembacaan indeks yang teridentifikasi pada mesin

yaitu pada ekson satu hingga ekson 7 (Gambar 2A). Produk PCR amplikon dari HLA-A, -B, dan -C adalah 3190 bp, 3076 bp, dan 3193 bp yang diperlihatkan dengan visualisasi gel agarosa 1% berupa satu pita tunggal (Gambar 2B).

Hasil konstruksi *library* DNA gen *HLA* kelas I memperlihatkan pola *smear* yang rata pada gel agarosa 2% yang kemudian dipotong pada ukuran 300–600 bp untuk mengoptimalkan proses sekuensing (Gambar 3). Hasil *real time* PCR purifikasi *library* sampel memperlihatkan

bahwa konsentrasi berada di bawah 4 nM dan juga di atas 1 nM. Illumina merekomendasikan konsentrasi *library* 4 nM untuk disekuensing, sehingga dilakukan modifikasi untuk normalisasi sampel. Sampel 22B tidak diikutsertakan dalam proses selanjutnya, yaitu proses sekuensing, karena konsentrasi sampel berada di bawah 1 nM.

Berdasarkan *sequencing analysis viewer* (SAV), nilai Q30 lebih besar dari 90%, dan data tersebut sesuai dengan spesifikasi dari Illumina (Gambar

Tabel 3 Hasil Alignment pada MiSeq Reporter (MSR)

No	Nama Sampel	Cluster Passing Filter* (ribu/mm ²)	Coverage**	SNP Heterozygous	SNP Homozygous	Inseri	Delesi	Median Panjang Basa Nukleotida
1	16A	185.235	877,4	42	60	1	3	227
2	16B	146.276	687,9	49	93	3	0	213
3	16C	204.819	919,9	14	100	2	3	212
4	22A	173.411	775,6	65	62	2	2	273
5	22C	489.648	2.182	42	85	2	3	234
6	40A	447.723	2.121,4	43	54	2	2	239
7	40B	893.216	4.234	75	101	4	1	228
8	40C	820.910	3.697,9	12	103	2	3	239
9	42A	506.117	2.403,3	101	57	2	2	273
10	42B	395.848	1.858,3	94	69	3	1	232
11	42C	760.651	3.414,2	9	104	2	3	208

Keterangan: *banyaknya kluster yang terbentuk; **Banyaknya pembacaan nukleotida selama proses sekuensing

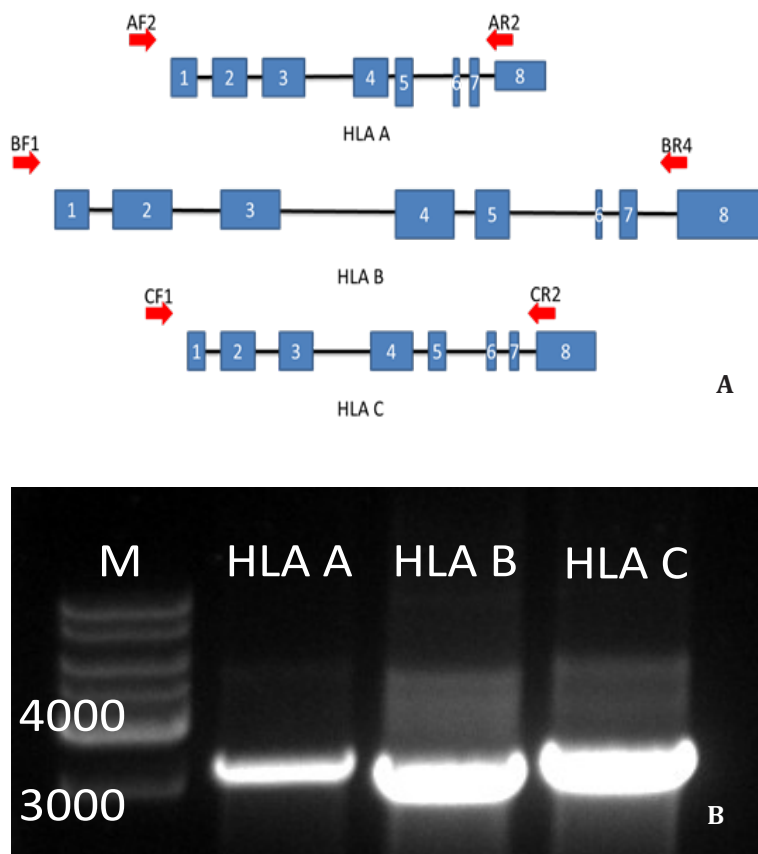


Gambar 1 Oligonukleotida Sintetik yang Ditempelkan pada DNA Insert

Keterangan: P5 dan P7: adapter spesifik Illumina; *Index 1* dan *Index 2*: sekuens indeks spesifik untuk membedakan sampel amplikon *multiplexing*; Rd1SP: primer universal sekuensing untuk pembacaan 1 (secara *forward*); Rd2SP: primer universal sekuensing untuk pembacaan 2 (secara *reverse*)

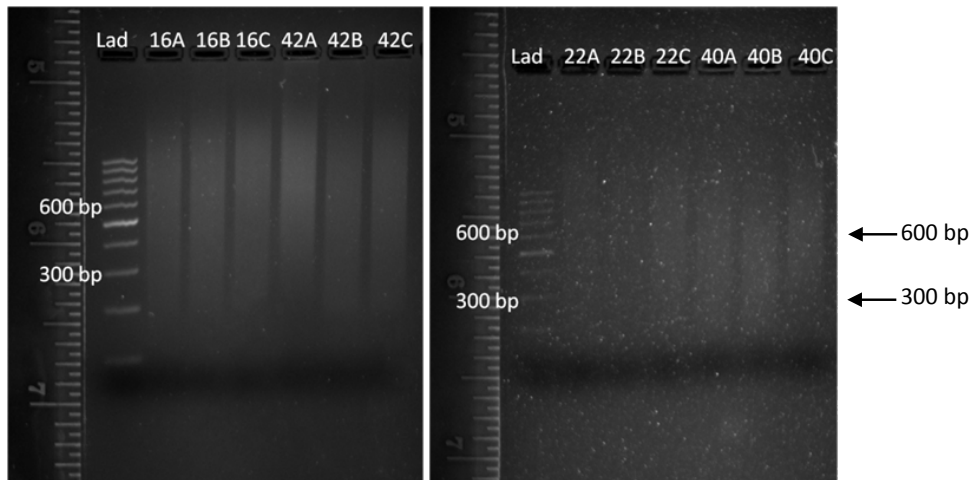
4). Nilai Q30 menandakan nilai kesalahan 1 basa di antara 1.000 basa. Kluster yang terbentuk dari proses amplifikasi pada MiSeq sebesar 864 K/mm², menandakan pembentukan kluster juga sesuai dengan spesifikasi dari Illumina. Sekitar 93,97% dari kluster tersebut dapat dianalisis secara statistik.

Persebaran indeks pada sampel amplikon terlihat pada Tabel 2 yang memuat persebaran pembacaan sampel yang disekuensing dengan kisaran persebaran yaitu 1,8% hingga 11,55%. Berdasarkan atas data *MiSeq Reporter (MSR)*, didapatkan hasil seperti tertera pada Tabel 3 dengan total 546 SNP *heterozygous*, 888 SNP



Gambar 2 Posisi Primer dan Visualisasi Gel Agarosa 1% Hasil PCR

Keterangan: A. Posisi primer *forward* dan primer *reverse* pada gen *HLA* Kelas I
 HLA-A: FA2 dan RA2
 HLA-B: BF1 dan BR4
 HLA-C: CF1 dan CR2
 B. Visualisasi gel agarosa 1% berupa satu *band* tunggal



Gambar 3 Visualisasi Agarosa 2% Sampel *Library* DNA

homozygous, 25 insersi dan 23 delesi dari keseluruhan 11 sampel amplicon dengan rata-rata *coverage* 2.106,536x.

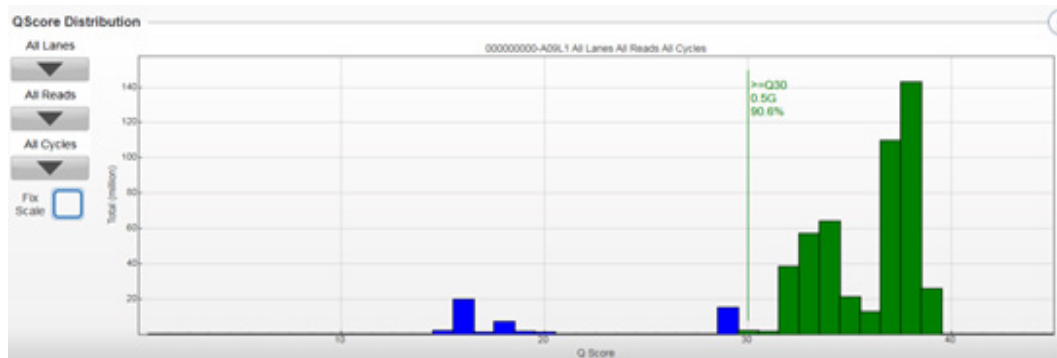
Pembahasan

Metode *next generation sequencing* atau NGS telah mulai digunakan untuk melakukan *HLA-typing*. Dibanding dengan metode sekuensing generasi pertama, sekuensing Sanger, metode NGS merupakan metode yang *high-throughput* dan *cost-effective*. Metode NGS dapat melakukan *HLA-typing* pada kesemua daerah gen *HLA* kelas I maupun kelas II dan mengurangi ambiguitas yang masih sering ditemukan pada metode terdahulu untuk dapat mendeterminasi alel yang ditemukan.^{11,12}

Salah satu NGS yang dipergunakan untuk mengembangkan *HLA-typing* adalah MiSeq Sequencer. Menurut Lange dkk.¹⁶ MiSeq tersebut

memberikan keuntungan dibanding dengan platform NGS lain yang digunakan seperti 454 GS FLX titanium dan *ion torrent system*, yaitu MiSeq memiliki fasilitas *on-board* untuk amplifikasi kluster, memiliki metode preparasi sampel yang cepat dan mudah, serta alur kerja yang fleksibel untuk dikembangkan dalam pengerjaan sampel rutin. Pengembangan metode *HLA-typing* telah dilakukan untuk mendeterminasi 162 haplotype sekuen gen *HLA* dari *homozygous cell lines*, 132 haplotype dari sampel *heterozygous* dan 60 haplotype dari trio/kuartet tiga keluarga¹¹ dan juga pengembangan dengan menggunakan *bead* dalam preparasi sampel untuk jumlah sampel besar.¹⁷

Fleksibilitas pengembangan metode tersebut mendorong untuk dilakukan optimasi dalam penelitian ini. Hasil optimasi dengan pemilihan *library* pada ukuran 300–600 bp memberikan hasil yang cukup baik dengan penggunaan 2x25 siklus sekuensing. Pemilihan *library* pada ukuran



Gambar 4 Nilai Q30 berdasarkan Visualisasi *Sequencing Analysis Viewer (SAV)*

tersebut bertujuan mengoptimalkan proses sekuensing karena daerah gen *HLA* memiliki polimorfisme dan daerah dengan kandungan GC yang tinggi. Daerah dengan kandungan GC yang tinggi dapat memengaruhi kerja enzim transposom.

Berdasarkan *data sheet Nextera DNA Library Preparation Kits* (2014),¹⁸ proses tagmentasi dengan transposom akan menghasilkan fragmen rata-rata dengan ukuran 300 bp. Namun demikian, pada penelitian tidak terdapat fragmen dengan ukuran pada kisaran 300 bp, yang terlihat adalah visualisasi gel agarosa dengan *pola smear* yang rata yang berukuran antara 100 bp hingga lebih dari 1.000 bp dibanding dengan marker 1 kb (Gambar 3). Pola tersebut mengindikasikan bahwa fragmen *library* DNA yang terbentuk tidak seragam. Fragmen dengan ukuran lebih dari 1.000 bp tidak efisien untuk proses amplifikasi pada pembentukan kluster. Gangguan pada proses pembentukan kluster akan berakibat juga pada proses sekuensing yang tidak optimal sehingga dalam penelitian dilakukan pemilihan ukuran *library* DNA sebelum 300–600 bp.

Hasil yang didapatkan NGS sangat berbeda dengan hasil sekuensing Sanger. Sekuensing Sanger menghasilkan elektroferogram berupa grafik pembacaan dari mesin yang dikonversikan menjadi urutan basa nukleotida DNA. Hasil yang diperoleh dari NGS Illumina berupa data intensitas pembacaan kluster. Data intensitas tersebut kemudian dikonversi oleh mesin menjadi data *quality score* (Q30) (Gambar 4) dan FASTQ. FASTQ merupakan *format file* berbasis teks untuk menyimpan urutan nukleotida dan sesuai dengan data *quality score*. Format FASTQ awalnya dikembangkan oleh *Wellcome Trust Sanger Institute* untuk mengabungkan data urutan FASTA dan data kualitas sekuensing, namun pada perkembangannya menjadi standar *de facto* untuk menyimpan data keluaran instrumen *high-throughput* seperti instrumen *Illumina Genome Analyzer*.¹⁹ Data FASTQ yang disejajarkan dengan sekuens acuan akan dikonversi menjadi *file* BAM (*.bam). *File* BAM merupakan versi biner terkompresi dari *file* SAM (*The Sequence Alignment MAP*) yang digunakan untuk merepresentasikan hasil penjejajaran sekuens DNA dengan sekuens acuan. *File* SAM merupakan format hasil penjejajaran dengan sekuens acuan yang mendukung pembacaan panjang dan pembacaan pendek (sampai 128 Mbp) yang diproduksi oleh platform NGS yang berbeda, seperti *Illumina/Solexa*, *AB/SOLiD*, dan *Roche/454*.²⁰

Hasil menunjukkan terdapat total 546 SNP

heterozygous, 888 SNP *homozygous*, 25 insersi, dan 23 delesi dari keseluruhan 11 sampel amplicon dengan rata-rata *coverage* 2.106,536x. Semakin tinggi nilai *coverage* maka semakin valid data yang dihasilkan. Pada daerah polimorfisme dan GC-content tinggi sebaiknya *coverage* lebih dari 1.000x untuk melihat konsistensi dari pembacaan sekuensing, walaupun belum ada ketentuan standar untuk *coverage* yang harus dicapai untuk sampel dengan polimorfisme dan juga GC-content tinggi. Hosomichi dkk.¹¹ mendapatkan kisaran *coverage* 146x hingga 6.678x dengan rata-rata *coverage* 2.281x untuk 66 amplicon menggunakan siklus 2 x 250 siklus yang menghasilkan *output* sekitar 8,5 Gb.

Penelitian ini menggunakan 2x25 siklus pada 11 amplicon, *coverage* yang dihasilkan berkisar dari 687,9x hingga yang terbesar 3.697,9x. Hal tersebut cukup baik, mengingat *output* dari 2x25 siklus hanya sebesar 5 Mb. Penggunaan siklus yang lebih besar seperti 2x150 atau 2x250 menghasilkan *output* yang lebih besar sehingga diharapkan memberikan *coverage* sekuensing yang lebih dari 1.000x sehingga data lebih menjadi lebih valid dan handal.

Secara umum metode yang dilakukan pada proses optimasi ini sudah berhasil. Namun demikian, masih diperlukannya suatu analisis data tambahan dengan perangkat lunak pihak ketiga sehingga dari data yang diperoleh dapat dideterminasi alel apa saja yang diperoleh yang masih belum dapat diketahui dari MSR. MSR hanya mampu mengetahui SNP *heterozygous*, SNP *homozygous*, dan indel serta perubahan dan posisi SNP yang mengalami perubahan jika dibanding dengan sekuens acuan.

Simpulan, penggunaan metode *long-range* berbasis PCR serta pemilihan ukuran *library* memberikan hasil yang handal, serta dapat diaplikasikan juga untuk penggunaan sampel berikutnya, namun demikian perlu dilakukan sekuensing dengan siklus yang lebih besar dari 2x 25 bp untuk mendapatkan lebih banyak *coverage* sehingga data menjadi lebih sah dan handal. Analisis data juga perlu dilakukan menggunakan perangkat lunak pihak ketiga, seperti *Omixon*, untuk mendeterminasi alel gen *HLA* kelas I.

Ucapan Terima Kasih

Ucapan terima kasih kami sampaikan kepada Dikti atas pendanaan melalui skema Hibah Desentralisasi. Kami juga menyampaikan terima kasih kepada Raden Indah Kendarsari dan M. Taufik Sukarno dari tim aplikasi PT.

Pandu Biosains sebagai perwakilan Illumina di Indonesia atas diskusi dan saran yang telah diberikan.

Daftar Pustaka

1. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet.* 2009;54(1):15-39.
2. Eapen M, Rubinstein P, Zhang M, Stevens C, Kurtzberg J, Scaradavou A, dkk. Outcomes of transplantation of unrelated donor umbilical cord blood and bone marrow in children with acute leukaemia : a comparison study. *Lancet.* 2007;369:1947-54.
3. Lee SJ, Klein J, Haagenson M, Baxter-low LA, Confer DL, Eapen M, dkk. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood.* 2007;110(13):4576-84.
4. Gao X, Bashirova A, Iversen AK, Phair J, Goedert JJ, Buchbinder S, dkk. AIDS restriction HLA allotypes target distinct intervals of HIV-1 pathogenesis. *Nature Med.* 2005;11(12):1290-2.
5. Lie BA, Thorsby E. Several genes in the extended human MHC contribute to predisposition to autoimmune diseases. *Current Opinion Immunol.* 2005;17(5):526-31.
6. Simons MJ. The origin of genetic risk for nasopharyngeal carcinoma: a commentary on "Is nasopharyngeal cancer really a Cantonese cancer?". *Chinese J Cancer.* 2010; 1:527-37.
7. De Kruijf EM, Sajat A, van Nes JGH, Natanov R, Putter H, Smit VT, dkk. HLA-E and HLA-G expression in classical HLA class I-negative tumors is of prognostic value for clinical outcome of early breast cancer patients. *J Immunol.* 2010;185(12):7452-9.
8. Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, dkk. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet.* 2002;359(9308):727-32.
9. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, dkk. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics.* 2011;12(42):41-3.
10. Dunn P. Human leucocyte antigen typing: techniques and technology, a critical appraisal. *Int J Immunogen.* 2011;38(6): 463-73.
11. Hosomichi K, Jinam TA, Mitsunaga S, Nakaoka H, Inoue I. Phase-defined complete sequencing of the HLA genes by next-generation sequencing. *BMC Genomics.* 2013;14(355):1-16.
12. Cao H, Wang Y, Zhang W, Chai X, Zhang X, Chen S, dkk. A short-read multiplex sequencing method for reliable, cost-effective and high-throughput genotyping in large-scale studies. *Hum Mutat.* 2013;34(12):1715-20.
13. 1000 Genomes Project Consortium: a map of human genome variation from population-scale sequencing. *Nature.* 2010;467:1061-73.
14. Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, dkk. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc Natl Acad Sci.* 2012;109(22):8676-81.
15. Shiina T, Suzuki S, Ozaki Y, Taira H, Kikkawa E, Shigenari A, dkk. 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(4):305-16.
16. Lange V, Böhme I, Hofmann J, Lang K, Sauter J, Schöne B, dkk. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics.* 2014;15(63):1-11.
17. Hosomichi K, Mitsunaga S, Nagasaki H, Inoue I. A bead-based normalization for uniform sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. *BMC Genomics.* 2014;15(645):1-9.
18. Illumina Inc. Nextera DNA Library Preparation Kits: Sequencing's fastest and easiest library preparation workflow, delivering libraries in 90 minutes. *Illumina Data Sheets: DNA Sequencing;* 2011.
19. Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* 2009;38(6):1767-71.
20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, dkk. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-9.