〔ミニレビュー〕

次世代シーケンサーによる MHC 遺伝子解析手法の現状

細道 一善

金沢大学 医薬保健研究域医学系 革新ゲノム情報学分野

The impact of next-generation sequencing technologies on MHC research

Kazuyoshi HOSOMICHI

Department of Bioinformatics and Genomics, Graduate School of Medical Sciences, Kanazawa University, Takaramachi 13-1, Kanazawa, Ishikawa 920-8640, Japan

はじめに

我々が表現型を理解する上でゲノム多様性を知るとは極めて重要である。次世代シーケンサー (NGS) を駆使したゲノム配列解析技術の高速化によって、ヒトをはじめ一個体のゲノムが数日で決定できることが可能となった現在、そのゲノム多様性をより深く理解するための膨大なシーケンシング情報が日々産出されている。世界初の次世代シーケンサー GS20 が 2005 年に発表されてから、まもなく 10 年目の節目を迎えようとしている。このわずかな期間でヒトゲノムの解析が容易になったばかりか、100 万人規模でのプロジェクトが進むなど、次世代シーケンサー技術の爆発的な進歩は、研究デザイン自体を大きく変化させた。発表当初の次世代シーケンサーの短所であった読み取り精度の問題も解決され、今後、次世代シーケンサーは医療機器としての新たな役目も担っていくことから、医療を変え、我々の生活までも変える存在となっていく。2013 年 9 月イルミナ社の次世代シーケンサー MiSeq Dx システムが米国食品医薬品局 (FDA) の体外診断用医療機器の認可を受け、臨床診断に使用可能となった。このように、次世代シーケンサーが医療機器として認可を受けた意義は極めて大きい。これはクリニカルシーケンス時代がすでに始まったことを意味し、個々人のゲノム情報を基に、精密な診断方法に基づく分類とそれに基づく治療法と予防法を意味する”精密医療 (Precision medicine)”が始まったといえる。この診断のためには特定のゲノム領域にターゲットを絞り、多検体を迅速に解析するターゲットリシーケンシング法が用いられ、変異の検出や検証、スクリー

ニングのための効率的な手法の一つとなっている。本稿ではターゲットリシーケンス法の一つとして、NGS によるヒトの MHC (HLA) のシーケンス法の最近の知見と我々が開発した手法の一例を紹介する。

HLA 遺伝子のシーケンスの意義

ヒトの主要組織適合遺伝子複合体 (Human leukocyte antigen; HLA) 領域はヒト第 6 染色体短腕部 6p21.31 に位置し、免疫応答をつかさどる HLA 遺伝子を含む 3.8 Mb から構成される (Shiina ら 2009)。ヒトゲノムの中でも群を抜いて多型性に富むゲノム領域であり、HLA 遺伝子 6 座 (HLA-A, -B, -C, -DR, -DQ, -DP) に計 12,000 種類を超える膨大な HLA アレルがこれまでに同定されている。また各遺伝子座における特定の HLA 対立遺伝子が強い連鎖不平衡により、HLA-A24-B52-DR15 など特有の HLA ハプロタイプを形成している。さらにはリウマチや糖尿病など 100 種近くの生活習慣病、自己免疫疾患、がん、造血幹細胞移植に伴う GVH 病、ウイルス感染症における防御と重症化など数多くの医学的興味を有し、さらに HLA 分子による抗原提示を利用したワクチンによる感染防御の有用性とがんの免疫療法の新しい展開が期待されている。特に、ステイブンソン・ジョンソン症候群、中毒性表皮壊死融解症および薬剤性過敏症症候群などの薬剤副作用と HLA アレルはオッズ比が 1,000 を超えるもの、陽性率

連絡先：金沢大学 医薬保健研究域医学系 革新ゲノム情報学分野

〒920-8640 石川県金沢市宝町 13-1

(email: khosomic@med.kanazawa-u.ac.jp)

100% のものが多々あるなど、極めて強い関連が複数報告されており、HLA アレルが薬剤副作用に大きく寄与していることが確認されている。

HLA タイピング法の問題点

現在の HLA タイピングは蛍光ビーズ法による PCR-Sequence specific oligonucleotide (PCR-SSO) およびダイレクトシーケンスによる PCR-Sequencing based typing (PCR-SBT) により行われており、これらの技術は 10 年以上もの間置き換わっていない。PCR-SSO は高頻度アレルのみを同定する手法であり、低頻度のものを同定できない。また、塩基配列を直接決定する PCR-SBT は *HLA-A*、*B* および *-C* それぞれの遺伝子座で 8,000、18,000 および 5,000 以上の HLA アレルの組み合わせにおいて phase ambiguity に起因する曖昧なタイピング結果が得られるという問題点がある。さらにこの SBT 法は HLA クラス I 遺伝子ではエクソン 2 および 3、クラス II 遺伝子ではエクソン 2 のみで判定されており、上述の曖昧なタイピング結果の問題とあわせて解決すべき問題である。今後 HLA タイピングはより高精度なタイピング法に置き換わっていくことが予測され、次世代シーケンサーとしては Roche 454 GS や MiSeq による HLA タイピングの報告がある。初期には Roche 454 GS による HLA タイピングが報告され、シーケンサーの進歩とともに近年では MiSeq による報告例へと主流となるシーケンサーも推移してきている (表 1)。これらはいずれも解析対象のゲノム領域を PCR 増幅し、相として分けられた断片的な配列をデータベースに検索することで高精度でのタイピング精度を実証している。しかしながら、いずれの手法も既知の HLA アレル配列を参照することでエクソン間の相を特定する”みなしタイピング”であることには変わらない。HLA タイピングを診断指針にまで実用するには、将来的に HLA タイピングをより高解像度、高精度なものへ改良する必要がある。

HLA 遺伝子の多検体アンプリコンシーケンス

我々は上述の既存の HLA タイピング法の問題点を解決するために、次世代シーケンサーによるタイピング法を確立している (Hosomichi ら 2013; Hosomichi ら 2014)。既存の手法と異なる点は *HLA-A*、*-C*、*-B*、*-DRB1*、*-DQB1* および *-DPB1* 各遺伝子座の上流数 kb を含む遺伝子領域を単一の PCR で増幅させていること、シーケンスデータの解析についてはアセンブルの段階で既知の HLA 遺伝子配列データベース (IMGT/

HLA データベース) 参照せずに行い、最終的に相を完全に分けて決定された各 HLA 遺伝子の 2 つのハプロタイプのシーケンスが新規か既知か、既知であれば HLA アレル番号を得るということのみに用いていることある。ワークフローとしてはまず *HLA-A*、*-C*、*-B*、*-DRB1*、*-DPB1* および *-DQB1* 各遺伝子座の遺伝子を含む約 3 ~ 13kb の領域を PCR 増幅し、得られた PCR 増副産物を Nextera DNA Sample Preparation Kit (illumina) を用いて DNA ライブラリーを調整、MiSeq(illumina) により 150bp ペアエンドでシーケンスするという流れである (図 1)。Nextera DNA Sample Preparation Kit は Nextera transposome を用いることでタグメンテーションと呼ばれる DNA 切断とタグ配列の挿入を同時に行うことにより反応時間を含め 90 分と短時間で DNA ライブラリーを調整することが可能である (図 2)。また、2 種類の Index を用いることで

図 1 Nextera と MiSeq による HLA 遺伝子のシーケンスの概要

図 2 タグメンテーションによる多検体 DNA ライブラリー調整法

トランスポゼースにより PCR 増幅産物とアダプターオリゴの間で組換えが起こり、断片化と共にアダプター配列が挿入される。この配列に相補的な配列および Index 配列を含むプライマーを用いた PCR をユニークな Index の組み合わせとしておこなうことで簡単に複数のライブラリーが調整可能である。

表 1 NGS による HLA タイピング法と対象遺伝子

Year	Author	Locus	PCR amplified region	Sequencer
2009	Gabriel C <i>et al.</i>	HLA-A, -B	exons 1, 2, 3, and 4	GS FLX (Roche)
2009	Bentley G <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQA1, -DQB1, DPB1	exons 2, 3, and 4 of A, B, and C exon 2 of DRB1, DPB1, DQA1	GS FLX (Roche)
2010	Lind C <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQB1	exons 2 and 3 of DQB1 entire gene of A, B, and C exons 2 - 3 of DRB1 and DQB1	GS FLX (Roche)
2010	Lank SM <i>et al.</i>	HLA-A, -C, -B	exons 2, 3, and 4 of A, B, and C	GS FLX (Roche)
2011	Erlach <i>et al.</i>	HLA-A, -C, -B	exons 2, 3, and 4 of A, B, and C	GS FLX (Roche)
2011	Holcomb CL <i>et al.</i>	HLA-A, -C, -B, -DPB1, -DQA1, -DQB1, -DRB1, -DRB3/4/5	exons 2, 3, and 4 of A, B, and C exon 2 of DRB1, DRB3/4/5, DPB1, DQA1	GS FLX (Roche)
2012	Wang C <i>et al.</i>	HLA-A, -C, -B, -DRB1	exons 2 and 3 of DQB1 exons 1 - 7 of A, B, and C exons 2 - 5 of DRB1	GAIIX (Illumina) HiSeq2000 (Illumina) MiSeq (Illumina)
2012	Shiina T <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQA1, -DQB1, -DPB1	entire gene (2 amplicons for DRB1 and DPB1)	GS Junior (Roche) ionPGM (Thermo)
2012	Lank SM <i>et al.</i>	HLA class I, all HLA class II loci	exons 1 - 7 of HLA class I (two amplicons) exons 1 - 4 of HLA class II	GS Junior (Roche) ionPGM (Thermo)
2013	Moonsamy PV <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DRB3/4/5	exon 2 of DRB1, DRB3/4/5, DQB1 exon2	GS Junior (Roche)
2013	Ringquist S <i>et al.</i>	HLA-DRB1	exons 2, 3, and 4 of A	GS FLX (Roche)
2013	Danzer M <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DRB3/4/5, -DQB1, -DPB1	exons 1, 2, 3, and 4 of B exons 1, 2, 3, 4, and 7 of C exon 2 and 3 of DRB1, DRB3/4/5, DQB1	GS FLX (Roche)
2013	Hosomichi K <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQB, -DPB1	exon 2 of DPB1	MiSeq (Illumina)
2013	Trachtenberg EA <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DRB3/4/5, -DQA1, -DPB1	entire gene exons 2, 3, and 4 of A, B, and C	GS FLX (Roche)
2014	Ozaki Y <i>et al.</i>	HLA-DRB1, -DRB3/4/5	exons 2 and 3 of DQB1	GS Junior (Roche)
2014	Hajeer AH <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQB1	exons 2 - 6 of DRB1, DRB3/4/5 exons 2 and 3 of A, B, and C exon 2 of DRB1	GS FLX (Roche) GS Junior (Roche)
2014	Hosomichi K <i>et al.</i>	HLA-B	exons 2 and 3 of DQB1	MiSeq (Illumina)
2014	Smith AG <i>et al.</i>	HLA-DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPB, -DPB1	entire gene exons 2 and 3 of DQA1, DQB1, DRB1, DRB3/4/5 exon 2 of DPA1 and DPB1	MiSeq (Illumina)
2014	Ehrenberg PK <i>et al.</i>	HLA-A, -C, -B, -DRB1	entire gene	MiSeq (Illumina)
2014	Zhou M <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQB1	exons 1 - 7 of A, B, and C (4 amplicons) exon 2 of DRB1	HiSeq2000 (Illumina)
2015	Lan JH <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DQB1	exons 2 and 3 of DQB1	MiSeq (Illumina)
2015	Ozaki Y <i>et al.</i>	HLA-A, -C, -B, -DRB1, -DRB3/4/5, -DQA1, -DQB1, -DPB1	entire gene (2 amplicons for DRB1) entire genes of A, B, and C exons 2 - 4 of DRB1, DRB3/4/5, DQA1, DQB1 exons 2 - 6 of DPB1	ionPGM (Thermo)

最大 96 種類 (Index1:8 種類 x Index2:12 種類) のサンプルを同時に解析することができるという特徴を持つ。また、他の次世代シーケンサーの DNA ライブラリー調整と異なり、開始 DNA 量必要量がわずか 20ng というのも特徴の一つである。さらに、後のデータ解析で詳述するが調整した DNA ライブラリーのインサートサイズが 200bp ~ 1kb 以上と非常に幅が広いことも、本手法での特徴と長所となっている。

Phase-defined sequencing による HLA 遺伝子配列決定

ロングアンプリコンシーケンスのリード情報を得た後、データ解析により HLA タイピングを行うが、我々は Phase-defined sequencing という物理的な相決定に基づく完全な遺伝子配列決定法を確立している (図 3)。まず得られたシーケンスリードを各 PCR 増幅領域のリファレンス配列に BWA および Samtools によってマッピングし、ディプロタイプとしての解析結果を得た後、Picard および Genome Analysis TK によりリファレンス配列と異なる塩基を検出する。次に我々が独自に開発したアルゴリズムにより、検出された一塩基置換 (SNV) および挿入欠失 (Indel) のうち、SNV のみをシーケンスリードのペア情報を元に相を分類し、この相を分けられたハプロタイプの SNV を元に 2 つの HLA 遺伝子配列を作成する。このとき、1 つのペアエンドのシーケンスリードによって 2 つ以上の SNV 情報を物理的な情報として結合し、相の特定が可能となる。2 つ以上の SNV 間の距離は様々であるが、HLA 遺伝子が非常に多様性に富んでいるという特徴と Nextera DNA Sample Preparation Kit で調整したライブラリーのインサートサイズが 200bp ~ 1kb 以上と非常に幅が広いことが様々な距離の 2 つ以上の SNV の相の決定を可能としている。相の決定の後、2 つのハプロタイプに分けたアライメントを bam ファイルとして出力、ハプロタイプの配列、すなわち遺伝子の上流および下流領域も含めた 2 つの HLA 遺伝子配列を得る (図 3)。最も多様なアレルを有する *HLA-B* の例を挙げると約 4kb に渡って相の特定に成功している (図 4)。ここで得られた 2 つのコンセンサス配列が 2 つの HLA 遺伝子のシーケンスであり、これを IMGT/HLA データベースに検索することで HLA アレルが新規か既知か、既知であれば HLA アレル番号を得ることができる。現状では IMGT/HLA データベースに登録されている HLA アレルの塩基配列は完全長であるものが少なく、我々が決定した数 k から数十

kb の塩基配列に対して短いものでは数百 bp と極めて短い。今後、IMGT/HLA データベースにおける HLA 遺伝子配列を完全長遺伝子配列の登録として拡充させる必要がある。

DDBJ の p-galaxy による解析パイプライン

上述の Phase-defined sequencing は bwa や samtools といった定番のプログラムおよび独自に開発したスクリプトによって実行するが、その全ての行程がコマンドラインによる操作で行われる。作業の工程が多く複雑なため、より手軽に誰でも解析が行えるように、この解析パイプラインを DDBJ の p-galaxy (<https://p-galaxy.ddbj.nig.ac.jp>) にて使用可能にしている。ウェブベースのグラフィカルインターフェイスを持つ p-galaxy 上で解析を実行することで、必要項目を画面上で選択し、実行するのみでバイオインフォマティクスの専門知識を必要とせず誰でも解析結果を得ることが可能となっている。

全ゲノムまたは全エクソンシーケンスデータからの HLA タイピング

ここまで *HLA-B* について多検体の解析を行うプロトコルを紹介した。一方で、現在の NGS によるヒトゲノム解析は全ゲノムおよび全エクソンシーケンスによる手法が広く使われており、全ゲノムおよび全エクソンシーケンスを行った検体の HLA タイピング法も重要なツールの一つである。全ゲノムおよび全エクソンシーケンスデータからの HLA タイピングツールとしては、HLAminer や Omixon Target HLA Typing など多くのプログラムが開発されている (表 2)。手法としてはデータベース上の既知の HLA 遺伝子配列に対して検索を行い、独自のスコアにより評価することで最も確からしい HLA アレルを推定する手法である。

おわりに

本稿ではヒトでの知見のみを述べたが、MHC 遺伝子は家畜においても経済有用形質や疾病への抵抗性との関連が強く、遺伝育種上重要な遺伝子の一つである。また、NGS を用いたウシ MHC 遺伝子のタイピング法も開発されている (Lee ら 2011)。すでに動物においても広く NGS による解析が用いられてきているが、本稿で紹介した MHC 遺伝子配列決定法は MHC アレル配列データベースを用いずに MHC 遺伝子の完全長配列を決定する手法であることから、様々な種における応用が可能な手法であり、NGS の高い解析能力が

図 3 完全長 HLA 遺伝子配列決定の解析パイプライン
解析パイプラインの工程 (A) とパイプラインによって操作されるシーケンスリードの概要 (B) を示した。HLA 遺伝子のアレルの組み合わせにより、検出される多型・変異の数は異なるため、様々な場合を想定したチェックポイントを経てハプロタイプの完全長遺伝子配列へと相を分けるプログラムとしている。

図 4 Phase-defined シーケンスの解析実例
リファレンス配列とシーケンスリード間で異なる塩基が検出されており、縦に 1 色のみはホモ接合、2 色はヘテロ接合を示す。相を分ける前にヘテロ接合として検出された多型・変異は相を分けた後は全てホモ接合となり、ハプロタイプの完全長遺伝子配列が決定される。

表2 NGSによるHLAタイピングプログラム

Year	Author	HLA typing software	URL	Read type
2012	Warren, R. L. <i>et al.</i>	HLAminer	http://www.begsc.ca/platform/bioinfo/software/hlaminer	WGS/WES/RNA-seq/amplicon
2012	Boegel, S. <i>et al.</i>	seq2HLA	http://tron-mainz.de/tron-facilities/computational-medicine/seq2hla/	RNA-seq
2013	Liu, C. <i>et al.</i>	ATHLATES	https://www.broadinstitute.org/scientific-community/science/projects/viral-genomics/athlates	WGS/WES/amplicon
2013	Kim, H. J. <i>et al.</i>	HLAforest	http://code.google.com/p/hlaforest	RNA-seq
2013	Bai, Y. <i>et al.</i>	PHLAT	https://sites.google.com/site/projectphlat/	WGS/WES/RNA-seq
2013	Hosomichi, K. <i>et al.</i>	Phase-defined HLA sequencing	https://p-galaxy.ddbj.nig.ac.jp	Amplicon
2014	Szolek, A. <i>et al.</i>	OptiType	http://omictools.com/optitype-s6206.html	WGS/WES/RNA-seq
2015	Huang, Y. <i>et al.</i>	HLAreporter	http://paed.hku.hk/genome/software.html	WGS/WES
2015	Wittig, M. <i>et al.</i>	HLAssign	http://www.ikmb.uni-kiel.de/resources/download-tools/software/hlassign	Sequence capture
2015	Nariai, N. <i>et al.</i>	HLA-VBSeq	http://nagasakilab.csmi.org/hla/	WGS
-	-	Assign ATF (Conexio Genomics)	http://www.conexio-genomics.com	Amplicon
-	Major, E. <i>et al.</i>	Omixon Target HLA (Omixon)	http://www.omixon.com/hla/	WGS/WES/amplicon
-	-	NGSengine (Gen Dx)	http://www.gendx.com/products/ngsengine	Amplicon
-	-	GeMS (Scisco Genetics)	http://scisogenetics.com/services/integrated-genotyping-system/	Amplicon

らヒト以外の動物 MHC においても新規 MHC アレルの完全長配列も含めた、高精度の MHC 遺伝子のゲノム配列が可能であろう。

文献

- Bai Y, Ni M, Cooper B, Wei Y, Fury W. 2014. Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. *BMC Genomics*, 15: 1–16.
- Bentley G, Higuchi R, Hoglund B, Goodridge D, Sayer D, Trachtenberg EA, Erlich HA. 2009. High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens*, 74: 393–403.
- Boegel S, Löwer M, Schäfer M, Bukur T, de Graaf J, Boissguérin V, Türeci O, Diken M, Castle JC, Sahin U. 2012. HLA typing from RNA-Seq sequence reads. *Genome Medicine*, 4: 102.
- Danzer M, Niklas N, Stabentheiner S, Hofer K, Pröll J, Stückler C, Raml E, Polin H, Gabriel C. 2013. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. *BMC Genomics*, 14, 221.
- Ehrenberg PK, Geretz A, Baldwin KM, Apps R, Polonis VR, Robb ML, Kim JH, Michael NL, Thomas R. 2014. High-throughput multiplex HLA genotyping by next-generation sequencing using multi-locus individual tagging. *BMC Genomics*, 15: 864.
- Erlich RL1, Jia X, Anderson S, Banks E, Gao X, Carrington M, Gupta N, DePristo MA, Henn MR, Lennon NJ, de Bakker PI. 2011. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics*, 12: 42.
- Gabriel C, Danzer M, Hackl C, Kopal G, Hufnagl P, Hofer K, Polin H, Stabentheiner S, Pröll J. 2009. Rapid high-throughput human leukocyte antigen typing by massively parallel pyrosequencing for high-resolution allele identification. *Human Immunology*, 70: 960–964.
- Hajeer AH, Al Balwi MA, Aytül Uyar F, Alhaidan Y, Alabdulrahman A, Al Abdulkareem I, Al Jumah M. 2013. HLA-A, -B, -C, -DRB1 and -DQB1 allele and haplotype frequencies in Saudis using next generation sequencing technique. *Tissue Antigens*, 82: 252–258.
- Holcomb CL, Höglund B, Anderson MW, Blake LA, Böhme I, Egholm M, Ferriola D, Gabriel C, Gelber SE, Goodridge D, Hawbecker S, Klein R, Ladner M, Lind C, Monos D, Pando MJ, Pröll J, Sayer DC, Schmitz-Agheguian G, Simen BB, Thiele B, Trachtenberg EA,

- Tyan DB, Wassmuth R, White S, Erlich HA. 2011. A multi-site study using high-resolution HLA genotyping by next generation sequencing. *Tissue Antigens*, 77: 206–217.
- Hosomichi K, Jinam T A, Mitsunaga S, Nakaoka H, Inoue I. 2013. Phase-defined complete sequencing of the HLA genes by next-generation sequencing. *BMC Genomics*, 14: 355.
- Hosomichi K, Mitsunaga S, Nagasaki H, Inoue I. 2014. A Bead-based Normalization for Uniform Sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. *BMC Genomics*, 15: 645.
- Huang Y, Yang J, Ying D, Zhang Y, Shotelersuk V, Hirankarn N, Sham PC, Lau YL, Yang W. 2015. HLAreporter: a tool for HLA typing from next generation sequencing data. *Genome Medicine*, 7: 25.
- Kim HJ & Pourmand N. 2013. HLA haplotyping from RNA-seq data using hierarchical read weighting. *PLoS ONE*, 8: e67885.
- Lan JH, Yin Y, Reed EF, Moua K, Thomas K, Zhang Q. 2015. Impact of three Illumina library construction methods on GC bias and HLA genotype calling. *Human Immunology*, 76: 166–175.
- Lank SM, Golbach BA, Creager HM, Wiseman RW, Keskin DB, Reinherz EL, Brusic V, O'Connor DH. 2012. Ultra-high resolution HLA genotyping and allele discovery by highly multiplexed cDNA amplicon pyrosequencing. *BMC Genomics*, 13: 378.
- Lank SM, Wiseman RW, Dudley DM, O'Connor DH. 2010. A novel single cDNA amplicon pyrosequencing method for high-throughput, cost-effective sequence-based HLA class I genotyping. *Human Immunology*, 71: 1011–1017.
- Lee BY, Hur TY, Jung YH, Kim H. 2011. Identification of BoLA-DRB3.2 alleles in Korean native cattle (Hanwoo) and Holstein populations using a next generation sequencer. 2011. *Animal Genetics*, 43:438-41.
- 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. 2010. Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. *Human Immunology*, 71: 1033–1042.
- Liu C, Yang X, Duffy B, Mohanakumar T, Mitra RD, Zody MC, Pfeifer JD. 2013. ATHLATES: accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Research*, 41: e142–e142.
- Moonsamy PV, Williams T, Bonella P, Holcomb CL, Höglund BN, Hillman G, Goodridge D, Turenchalk GS, Blake LA, Daigle DA, Simen BB, Hamilton A, May AP, Erlich HA. 2013. High throughput HLA genotyping using 454 sequencing and the Fluidigm Access Array™ system for simplified amplicon library preparation. *Tissue Antigens*, 81: 141–149.
- Nariai N, Kojima K, Saito S, Mimori T, Sato Y, Kawai Y, Yamaguchi-Kabata Y, Yasuda J, Nagasaki M. 2015. HLA-VBSeq: accurate HLA typing at full resolution from whole-genome sequencing data. *BMC Genomics*, 16(Suppl 2):S7
- Ozaki Y, Suzuki S, Kashiwase K, Shigenari A, Okudaira Y, Ito S, Masuya A, Azuma F, Yabe T, Morishima S, Mitsunaga S, Satake M, Ota M, Morishima Y, Kulski JK, Saito K, Inoko H, Shiina T. 2015. Cost-efficient multiplex PCR for routine genotyping of up to nine classical HLA loci in a single analytical run of multiple samples by next generation sequencing. *BMC Genomics*, 16: 318.
- Ozaki Y, Suzuki S, Shigenari A, Okudaira Y, Kikkawa E, Oka A, Ota M, Mitsunaga S, Kulski JK, Inoko H, Shiina T. 2014. HLA-DRB1, -DRB3, -DRB4 and -DRB5 genotyping at a super-high resolution level by long range PCR and high-throughput sequencing. *Tissue Antigens*, 83:10-6.
- Ringquist S, Bellone G, Lu Y, Roeder K, Trucco M. 2013. Clustering and alignment of polymorphic sequences for HLA-DRB1 genotyping. *PLoS ONE*, 8: e59835–7.
- Shiina T, Hosomichi K, Inoko H, Kulski, JK. 2009. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics*, 54: 15–39.
- Shiina T, Suzuki S, Ozaki Y, Taira H, Kikkawa E, Shigenari A, Oka A, Umemura T, Joshita S, Takahashi O, Hayashi Y, Paumen M, Katsuyama Y, Mitsunaga S, Ota M, Kulski JK, Inoko H. 2012. Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. *Tissue Antigens*, 80: 305–316.
- Smith AG, Pyo CW, Nelson W, Gow E, Wang R, Shen S, Sprague M, Pereira SE, Geraghty DE, Hansen JA. 2014. Next generation sequencing to determine HLA class II

- genotypes in a cohort of hematopoietic cell transplant patients and donors. *Human Immunology*, 75: 1040–1046.
- Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. 2014. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*, 30: 3310–3316.
- Trachtenberg EA. & Holcomb CL. 2013. Next-generation HLA sequencing using the 454 GS FLX system. *Methods in Molecular Biology*, 1034: 197–219.
- Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, Levinson D, Fernandez-Viña MA, Davis RW, Davis MM, Mindrinos M. 2012. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 109: 8676–8681.
- Warren RL, Choe G, Freeman DJ, Castellarin M, Munro S, Moore R, Holt RA. 2012. Derivation of HLA types from shotgun sequence datasets. *Genome Medicine*. 4, 95.
- Wittig M, Anmarkrud JA, Kässens JC, Koch S, Forster M, Ellinghaus E, Hov JR, Sauer S, Schimmeler M, Ziemann M, Görg S, Jacob F, Karlsen TH, Franke A. 2015. Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Research*, pii: gkv184.
- Zhou M, Gao D, Chai X, Liu J, Lan Z, Liu Q, Yang F, Guo Y, Fang J, Yang L, Du D, Chen L, Yang X, Zhang M, Zeng H, Lu J, Chen H, Zhang X, Wu S, Han Y, Tan J, Cheng Z, Huang C, Wang W. 2014. Application of high-throughput, high-resolution and cost-effective next generation sequencing-based large-scale HLA typing in donor registry. *Tissue Antigens*, 85: 20–28.