High-resolution HLA typing with NGSgo-MX6-1 by HiFi NGS data of a PacBio Sequel II platform

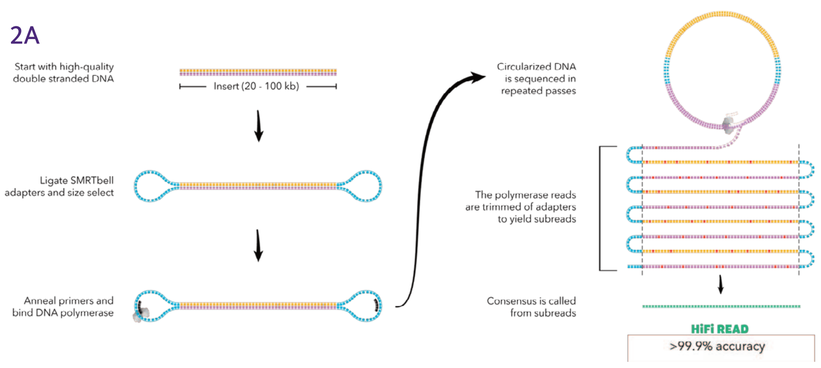
Reliable NGS-based HLA typing requires high-quality reads of sufficient length to resolve ambiguities. Short read sequencing has the advantage of being of high quality, however, it can sometimes hamper unambiguous HLA typing due to limited phasing. This study investigates whether HiFi sequencing on a PacBio Sequel II System allows for generating long, high-quality reads with full phasing, supporting reliable and unambiguous HLA typing.

1A

1B

Figure 1 (A) NGSgo-MX6-1 amplification strategy for HLA-A, B, C, DRB1, DQB1 and DPB1 and (B) genetic coverage.

2A



2B

PACBIO®

For Research Use Only

Figure 2 (A) PacBio SMRTbell library was created from these amplicons using SMRTbell Express Template Prep Kit 2.0 and barcoded overhang adapters. (B) Libraries were sequenced on a PacBio Sequel II System.

Materials & Methods

A 58 gDNA sample panel was amplified using the NGSgo-MX6-1 amplification strategy for HLA-A, B, C, DRB1, DQB1 and DPB1 from GenDx (Figure 1), which was processed in the PacBio library preparation workflow (Figure 2). Libraries were sequenced on a Sequel II System (Pacific Biosystems) and data was analyzed in NGSengine.

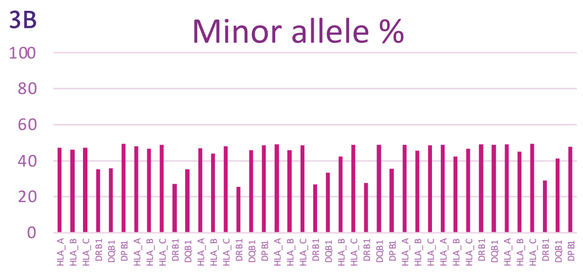
Results

High-quality data was generated for all samples. Figure 3A shows sequencing metrics of a representative sample from the 58 sample panel. The typing results of all loci in all samples were 100% concordant to the pre-type information. Each allele was recognized with a fixed allele ratio across the full amplicon region. Therefore, minor alleles were readily recognized by NGSengine (Figure 3B). The average noise of all core regions in all loci was 3.6% with 95% confidence interval limits of 3.3 and 3.9% (Figure 3C). Depth of coverage was constant across the complete amplicon. As promised by long-read sequencing all loci could be completed phased, even for samples with sparsely distributed heterozygous positions as identified in DPB1 (Figure 4), which resolves genotype ambiguities reported by short read sequencing.

3A

Locus	Map	Region	LRD	#HP	Noise	ΔSN	ESA	MM	
HLA-A	100%	Core+:	694	34	3.0%	45%	49%	0	C C A*23:01:01:01, A*30:01:01:01
HLA-B	100%	Core+:	352	16	4.3%	44%	49%	0	C C B*13:02:01:01, B*49:01:01:01 Bw4 , Bw4
HLA-C	100%	Core+:	446	11	1.6%	46%	47%	0	C C C*06:02:01:01, C*07:01:01:01
DRB1	94%	Core+:	2257	34	1.5%	33%	35%	0	C C DRB1*07:01:01:01, DRB1*11:04:01:01
DQB1	100%	Core+:	2842	25	2.4%	44%	46%	0	C - DQB1*02:02:01:01, DQB1*03:01:01:02
DPB1	99%	Core+:	1086	15	2.1%	41%	44%	0	C C DPB1*04:01:01:01, DPB1*17:01:01:01

3B



3C

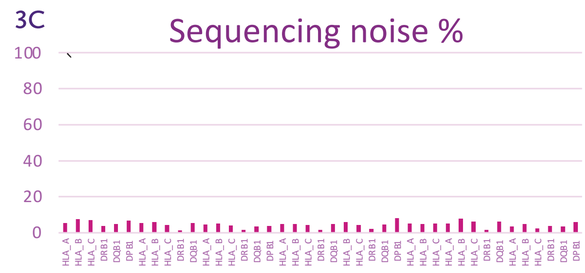


Figure 3. High quality data was generated for all samples. (A) NGSengine sample overview showing sequencing quality metrics (B) Estimated second allele (minor allele) % in 8 samples (C) Sequencing noise levels in the core regions for 8 samples.

4

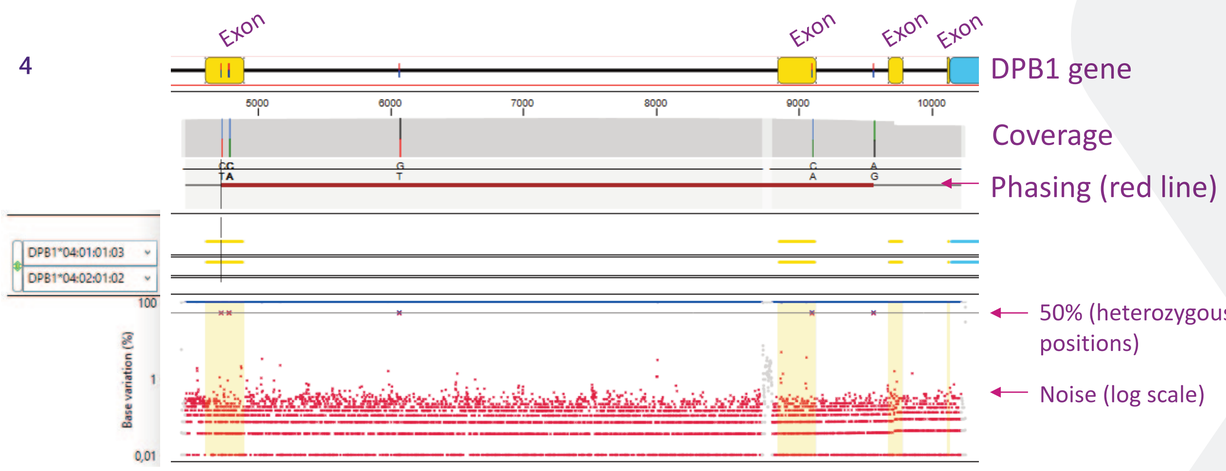


Figure 4. Complete phasing of heterozygous positions across the whole amplicon. NGSengine data plots for DPB1.

Conclusion

PacBio HiFi sequencing of amplicons generated with NGSgo-MX6-1 from GenDx results in high-quality long read sequences of HLA. The long reads contribute to a constant depth of coverage combined with full phasing and low noise levels, allowing for reliable HLA typing with limited ambiguities. This new long read sequencing strategy is an attractive alternative to current short read sequence technologies with limited phasing capacity.

GenDx.com

Ian McLaughlin[1], John Harting[1], Sake van Wageningen[2], Hanneke Merkens[2], Loes A.L. van de Pasch[2], Maarten T. Penning[2], Erik H. Rozemuller[2]

1 Pacific Biosciences, Menlo Park, California, USA
2 GenDx, Utrecht, the Netherlands

experts in transplant diagnostics