Short- and long-read next-generation sequencing approaches to characterize *HLA-A*30* null alleles differing in the length of homopolymer C region of exon 4

G. Montero Martin, I. Scott, B. Ostland, S. Rodriguez Hernandez, K. Gaines, N. Marcus, B.J. Franz
Vitalant HLA Laboratory, Phoenix, AZ, USA
Contact information: gmonteromartin@vitalant.org

Poster# 508

INTRODUCTION

The 5' end of exon 4 of *HLA-A*30:01:01:01* contains a homopolymer stretch of seven cytosines (7Cs), a motif that is common to *HLA-A* expressed alleles. In contrast, a group of HLA Class I null allele variants (such as *HLA-A*30:132N*) contains a C insertion extending this stretch to 8Cs, which leads to a frameshift mutation causing a premature stop codon downstream. A different *HLA-A*30* null allele variant, *HLA-A*30:130N*, displays an even longer 9Cs motif. NGS platforms have limitations [i) DNA polymerase slippage in the Illumina system; ii) signal versus noise detection technology in the ONT platform] to accurately characterize these homopolymer regions.

METHODS

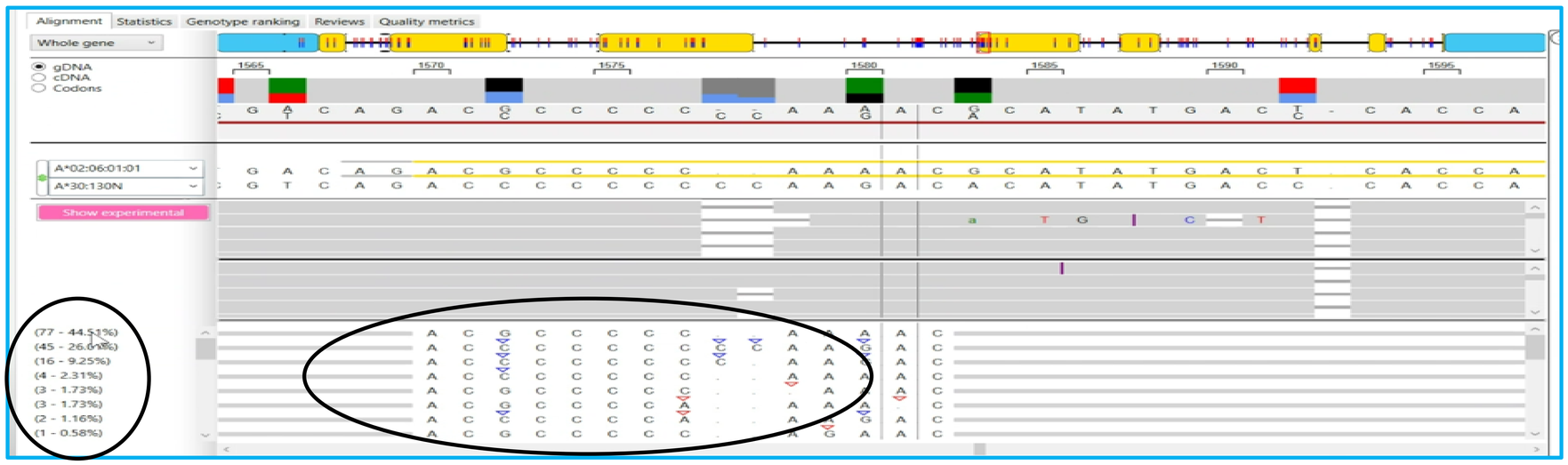
We evaluated three different NGS HLA genotyping strategies: i) AllType NGS (Thermo Fisher One Lambda) and ii) AlloSeq Tx17 (CareDx) respectively with Illumina system; and then iii) NGS-Turbo (GenDx) with ONT Nanopore system. We sequenced an *HLA-A*30* null allele variant (*HLA-A*30:130N* versus *HLA-A*30:132N*) in the presence of a second expressed allele (*HLA-A*02:06:01:01*) from a deceased donor whole blood specimen. We evaluated how the different assay chemistries (PCR-based versus Hybrid capture-based), sequencing platforms (Illumina MiSeq versus ONT MinION Mk1C), and HLA genotyping analysis software tools contribute to the accuracy of this heterozygous HLA genotype assignment.

RESULTS

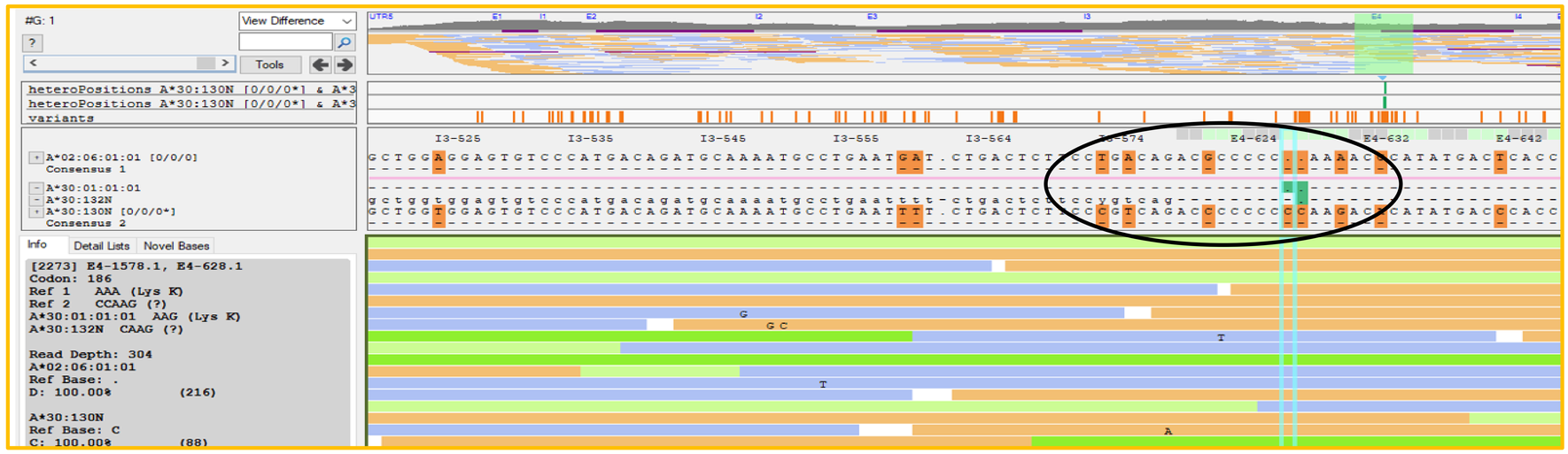
IPD-IMGT v.3.57.0 HLA Sequence Alignment: *HLA-A* 5' End of Exon 4

AA Codon	(C Homopolymer)	185							190
A*30:01:01:01	(7Cs)	AC	CCC	CCC	..AAG	ACA	CAT	ATG	ACC
A*02:06:01:01	(7Cs)	--	G--	---	..--A	--G	---	---	--T
A*30:130N	(9Cs)	--	---	---	CC---	---	---	---	---
A*30:132N	(8Cs)	--	---	---	C.---	---	---	---	---

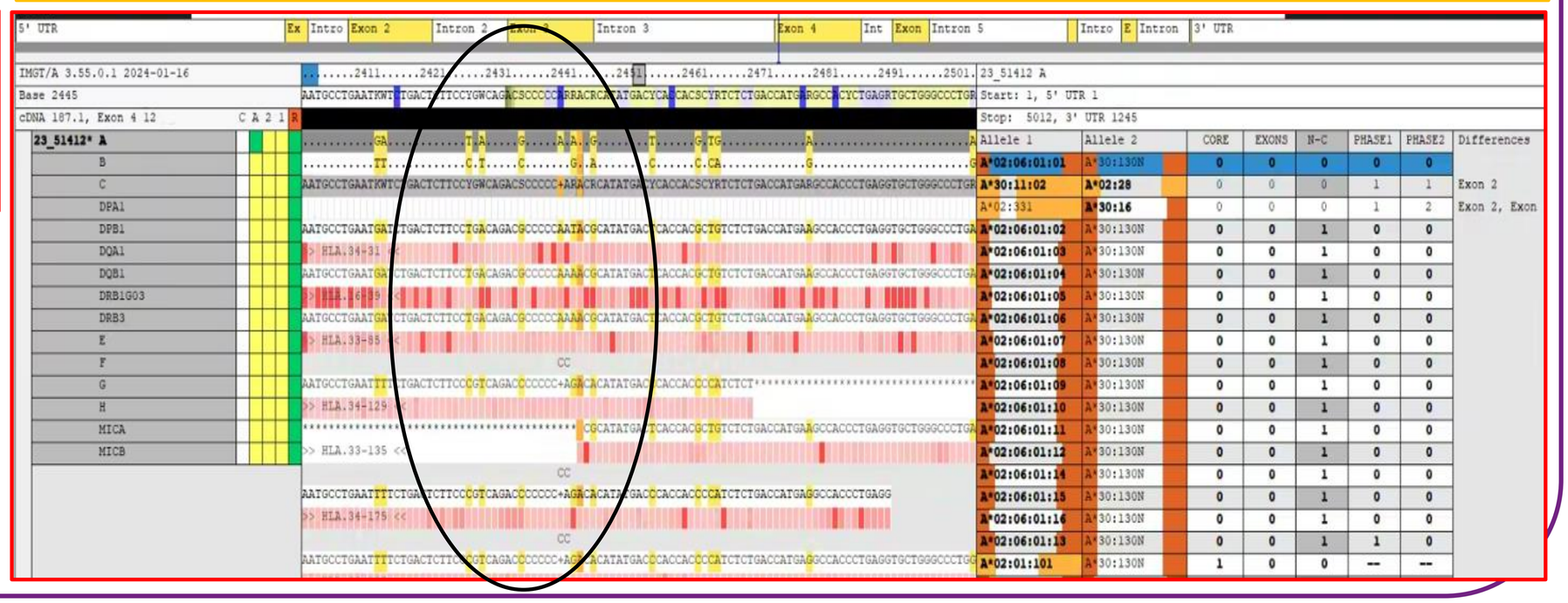
NGS-Turbo GenDx + ONT MinION Mk1C (SUP basecalling and NGSengine-Turbo 1.0 RUO)



AllType NGS One Lambda (kit and software) + Illumina MiSeq



AlloSeq Tx17 NGS CareDx (kit and software) + Illumina MiSeq



CONCLUSIONS

All three NGS HLA genotyping approaches provided concordant results for the *HLA-A*30:130N* allele (bearing the 9 Cs homopolymer motif) + *HLA-A*02:06:01:01*, where the corresponding bioinformatics component had a critical contribution. In particular, the ONT sequencing approach was as accurate as the other two Illumina-dependent protocols thanks substantially to: i) the most developed ONT basecalling tool, using Super-accuracy (SUP) instead of High-accuracy (HAC) system; ii) and by using a newer ONT-data adapted GenDx NGSengine-Turbo software version, NGSengine-Turbo 1.0 RUO instead of the former Prototype version.