INTRODUCING SINGLE MOLECULE REAL-TIME (SMRT®) SEQUENCING IN A CLINICAL ROUTINE SETTING.

saving the lives of people with blood cancer

Dario Merlo[1], Katy Latham[1], Neema P Mayor[1,2], James Robinson[1,2], Franco Tavarozzi[1], Shem Wallis-Jones[1], Reetinder Grewal[1], Francisco Boix-Giner[1], Davide Lepore[1], Nicola Brosnan[1], Will P Bultitude[1,2], Thomas R Turner[1] and Steven GE Marsh[1,2]

1 Anthony Nolan Research Institute, Royal Free Hospital, London UK; **2** UCL Cancer Institute, Royal Free Campus, London UK

INTRODUCTION

HLA typing plays a major role in the selection of haematopoietic stem cell donors and there is published evidence that higher resolution HLA matching is associated with increased overall survival. Presently, many high-throughput molecular HLA typing approaches like PCR-SSOP or SBT lead to ambiguous results as they fail in determining the phase of observed DNA sequence polymorphisms and in covering the entire length of the gene. They are also unable to achieve allele-level resolution data at primary typing resulting in unavoidable multi-step typing approach.

Anthony Nolan has implemented Pacific Biosciences' Single Molecule Real-Time (SMRT®) sequencing technology, a Third Generation Sequencing (TGS) method which ensures single allele resolution at primary typing, allowing a strategy for full gene sequencing and enabling complete phasing of DNA polymorphisms by sequencing single amplicons in isolation (Fig.1). The sequencing reaction occurs within the RSII sequencing system in SMRT® Cells, nanofabricated consumables patterned with 150,000 wells called zero-mode waveguides (ZMWs), whose unique structure enables real-time sequencing observation.

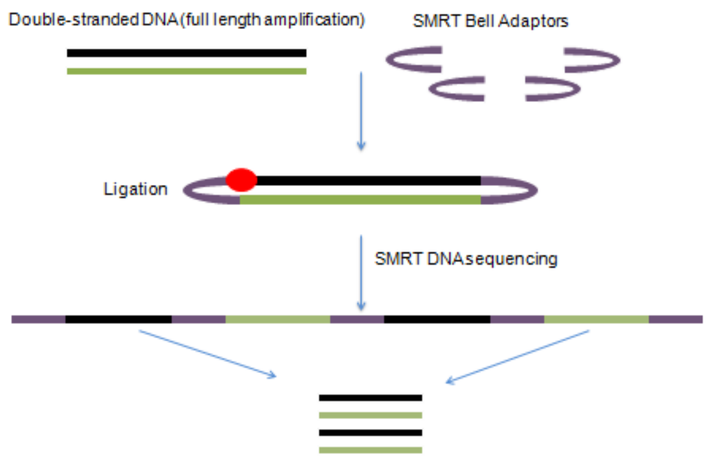

Figure 1: Single Molecule Real-Time (SMRT®) DNA sequencing method. SMRT Bell adaptors are ligated onto blunt-ended PCR full length amplicons to allow continuous sequencing of both sense and anti-sense strands of the PCR amplicon in a single read known as the Continuous Long Read (CLR). Post-sequencing Bioinformatics processes are used to break down the CLR into shorter sub-reads, which can be used to generate a consensus sequence. The higher the number of the reads, the higher the typing accuracy.

PURPOSE

Here we describe the outcome of the work to assess the viability of our semi-automated SMRT® sequencing pipeline to produce results in a time frame that supports the clinical application of HLA typing. This requires 75% of the reports to be sent within seven days to comply with our transplant centres' service agreement. All testing was performed after the validation of the SMRT® sequencing pipeline to Clinical Pathology Accreditation (CPA) standards.

MATERIALS AND METHODS

Between August and October 2015, 522 DNA samples extracted from blood specimens using Bio Robot MDx® (Qiagen) and Tecan Freedom EVO 200 MCA® robotic platform (Tecan) utilising a ReliaPrep® chemistry (Promega) were typed for HLA-A, -B, and -C loci, a total of 1,566 tests. Previous typing for the three loci was available for 55.9% (292/522) of the samples, a total of 876 tests. PCR amplification targeted whole gene amplification using gene-specific primers tagged with unique DNA barcodes (Fig.2), allowing multiplex processing. Eleven 48-plex libraries were prepared according to optimised Pacific Biosciences' recommended protocols and loaded on RSII sequencing system across six SMRT® runs. PCR and equinanogram pooling were automated using Micro Lab Star Automated Liquid Handler (Hamilton) whilst purification, enzymatic steps and quality control (QC) were performed manually. For the QC steps, size and concentration of all the amplicons in each plate were accurately and simultaneously measured using 96-capillary Fragment Analyzers (Advanced Analytical Technologies Inc.), which proved to be the most precise and quickest QC tool for this high-throughput technique. In-house software developed by the HLA Informatics Group within Anthony Nolan Research Institute, {AT}toolset, was used for post SMRT® analysis (see poster 90 for more details), to assign the HLA typing referring to WHO nomenclature version 3.22.0.

REFERENCES

Mayor *et al*. HLA Typing for the Next Generation (2015). PLoS ONE 10(5): e0127153

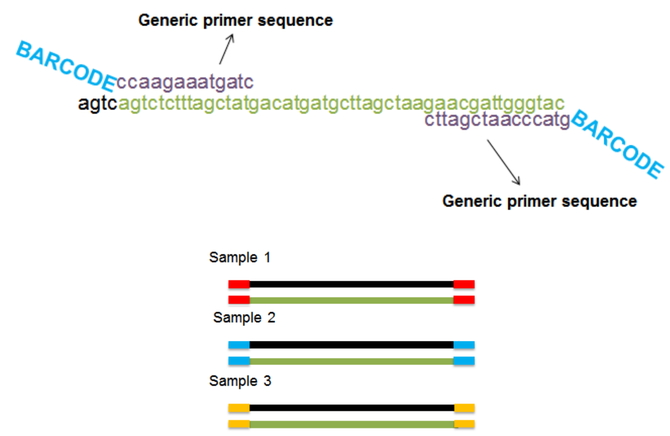

Figure 2: DNA barcode technology. Top: Generic primers (purple) allow amplification of a specific target region (green) i.e. a particular HLA locus. A specific, unique DNA sequence (the barcode in blue) has been previously added to the 5' end of the generic primer and it will be sequenced together with it. Bottom: All samples are amplified with primers tagged with different barcodes and sequenced in multiplex processing. Post-sequencing Bioinformatics analysis will allow the assignment of the typing to the correct sample .

RESULTS

All HLA typing results were at definitive allele resolution and reported within a timeframe of six working days (Fig.3). Figure 4 shows the test success rate and the level of concordance with previous typing. We observed an initial overall test failure rate of 13%, (204/1,566), of which 8.8% (138/1,566) were at PCR stage and a post-sequencing failure rate of 4.6% (66/1,428). Discrepancies with previous typing have been detected in 1.8% of the tests (16/876), of which 1.1% (10/876) were due to errors in former techniques because of the method limitation or to discovery of novel alleles. Four potential new cDNA sequences have been found during this work, currently under investigation and naming process. The remaining 0.7% (6/876) of the discrepancies were due to known allele drop-outs, however subsequent improvements in primers design have eliminated the issue.

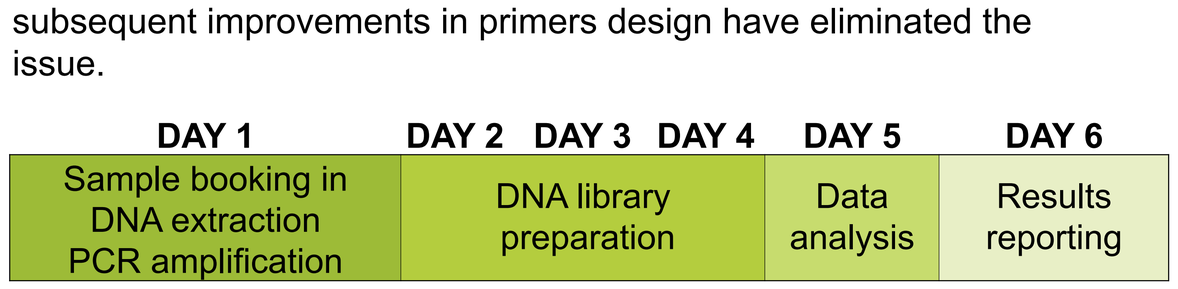

Figure 3: Semi-automated SMRT® sequencing clinical pipeline.

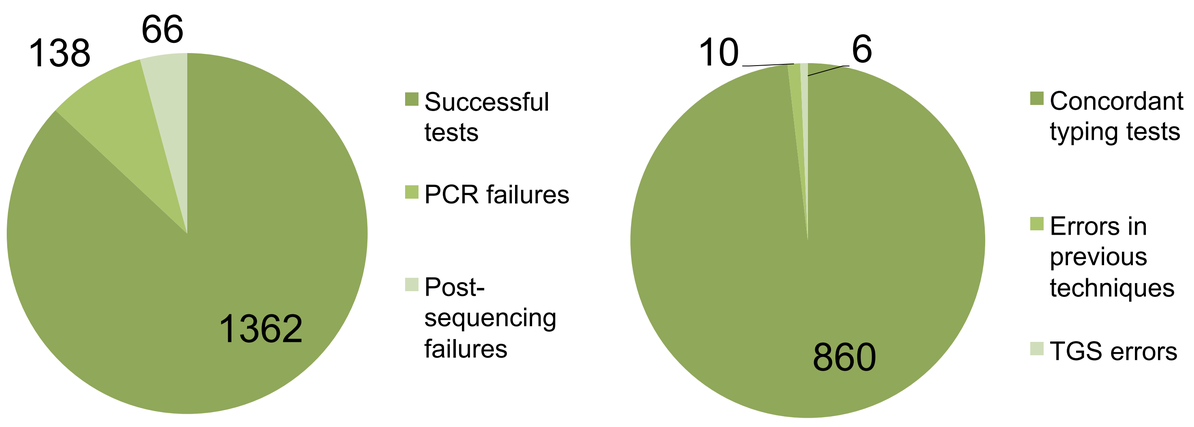

Figure 4: Test success rate and level of concordance with previous typing.

CONCLUSIONS

With an assay success rate of 87% (1,362/1,566) and six day turnaround times up to report preparation, SMRT® sequencing was confirmed as a viable method for HLA clinical typing. We have generated allele level data for multiplexed samples superior to current techniques in terms of resolution, cost effectiveness and time. An area for improvement is primer design and quality, in order to reduce PCR failures and avoid allelic imbalance; quantification and normalisation of genomic DNA prior to amplification will also ensure equal sample representation downstream. Furthermore, pooling algorithms taking into consideration the different size of different HLA amplicons have been designed for controlling allelic and loci imbalance. Future changes to the pipeline will include: full automation throughout the whole pipeline; integration with Laboratory Information Management System (LIMS) for sample and reagents tracking and including HLA class II typing, still within seven days.

Reg charity no 803716/SC038827 978PA/0416

None of the authors have any conflict of interest to declare