SA-S29-CLL-020

The Genomic Landscape of Chronic Lymphocytic Leukaemia: Clinical Implications

Kate Ridout, PhD,[*] Pauline Robbe, PhD,[*] Dimitris Vavoulis, BSc, MSc, MSc, PhD,[*] Jenny Klintman, MD, PhD,[*] Basile Stamatopoulos, PhD,[*] David Bruce, Adam Burns, Doriane Cavalieri, Helene Dreau, MSc and Anna Schuh, MD, PhD, MRCP, FRCPath,[**]

University of Oxford, Department of Oncology, Oxford OX37LJ, United Kingdom

[*] *These authors contributed equally*

[**] **Corresponding author:** anna.schuh@oncology.ox.ac.uk

Keywords

IgHV mutation status, TP53, complex karyotype, whole genome sequencing, subclonal evolution

Introduction

Chronic lymphocytic Leukaemia (CLL) is characterized by significant clinical and biological heterogeneity. Despite recent advances in therapy, CLL remains largely incurable.

While there has been a rapid increase in our knowledge of the complex genomic landscape of CLL and its spatial and temporal evolution from large scale sequencing efforts[1–5], this knowledge is only slowly translating into clinically relevant biomarkers. Moreover, samples from only 234 patients with CLL across 6 studies have undergone whole genome sequencing (WGS) so far and none of these were linked to clinical trials of uniformly treated patients. Potential advantages of WGS over whole exome sequencing (WES) include its ability to detect mutations in non-coding regions and to call all types of mutations including copy number changes and translocations in a single assay.

Specific acquired genetic abnormalities are associated with clinical outcome of CLL.

The immunoglobulin heavy variable (IgHV) mutation status defined by multiplex PCR and Sanger sequencing predicts time to first treatment and overall survival[6] and has recently also been linked to functional cure in patients treated with chemoimmunotherapy[7,8]. However, on its own, its predictive value is not strong enough to be used in clinical decision making for individual patients.

Both clonal and subclonal disruption of TP53 (i.e. deletions of chromosome 17p[9] and mutations in TP53[10,11]) are known to predict chemotherapy resistance. This is particularly relevant as effective targeted therapies are available for this patient group[12–15] and international guidelines recommend TP53 testing of all patients before each new line of treatment[16–19]. However, access to clinically accredited diagnostics that offer TP53 deletion and sensitive mutation testing remains an obstacle to widespread and equitable adoption of this important biomarker.

Finally, recent evidence points towards a role of complex karyotype in defining response to targeted agents, in particular venetoclax[20–22]. However, current methods for karyotyping of CLL cells are labour intensive and not standardized. Besides, this biomarker is strongly associated with TP53 disruption and showing its independent prognostic impact convincingly will require large number of patients.

In order to begin to address these issues, we initially evaluated deep targeted DNA sequencing, WGS and total RNAseq as alternatives to conventional diagnostics in a series of test and validation cohorts of patients with CLL and then subsequently initiated a whole genome sequencing programme for CLL in the UK's National Health Service in collaboration with Genomics England and Illumina.

Our aims were to overcome the limitations of current low- sensitivity multimodality diagnostics and to (1) increase the predictive value and precision of existing biomarkers by using WGS; (2) develop WGS as a comprehensive routine clinical diagnostics tool for use in the National Health Service; (3) discover novel predictors of clinical outcome for CLL including also drivers of high-grade transformation (Richter's Syndrome; RS) in the coding and non-coding space.

Methods

Patient Characteristics: Data on three cohorts will be presented. Ethical approval was obtained in accordance with the declaration of Helsinki (REC 09/H1306/54).

Cohort 1: 32 patients with CLL in need of therapy from a discovery series who had undergone total RNAseq and WGS. 780 patients from 3 independent validation cohorts undergoing analysis for IgLV3-21 re-arrangement.

Cohort 2: A total of 902 patients with CLL recruited into national clinical trials consented for WGS with Genomics England between 2014 and 2017. Data on the first 392 patients will be presented.

Cohort 3: 65 patients with Richter's Syndrome (RS) from the UK NCRN CHOP-OR study[23] and an independent cohort of Swedish patients with RS. For 17 of the 43 patients recruited into CHOP-OR, paired samples of CLL and RS were available and subjected to WGS. Samples of all patients underwent targeted RNA sequencing of the 800 most common cancer genes using Nanostring.

DNA extraction and Library Preparation for WGS

Performed as previously described[5].

WGS Bio-informatics

Somatic variant detection was performed using Strelka 2.4.7, Manta 0.28.0 and Canvas 1.3.1 (Illumina). For Variant filtering, SNVs and indels were filtered on DP>10 and VAF>0.05; CNAs and structural variants were filtered on variant quality. All relevant variant in genes/regions of interest were validated using targeted sequencing panel +/- Sanger sequencing or FISH +/- whole genome SNP array.

RNAseq analysis

Libraries were prepared using the TruSeq Stranded Total RNA Sample Preparation Kit (Illumina). The median number of paired-end reads was 60.5 million (range, 49.7-79.7 million). Data processing and analyses were conducted using bioinformatics pipelines utilizing the Computational Genomics Analysis and Training (CGAT) code collection (https://github.com/CGATOxford/CGATPipelines).

Clonal evolution

The clonal structure of each pair of blood and lymph node samples was studied using standard statistical methods [24,25]. Assuming that each tumor is a mixture of normal and cancer cells, we expressed, for each mutation, the expected VAF as a function of copy-number variation, tumor purity, and cancer cell fraction (CCF), i.e. the fraction of cancer cells harboring the mutation. Given this expression and the observed VAF, we applied a variational, non-parametric Bayesian clustering methodology, in order to infer, for each mutation, the expected CCF given the data, as well as the expected number of clones supported by the data in each sample. All mutations that had a CCF>85% with high probability (>0.95) were considered clonal. By tracking how the CCF for each mutation changes in time, we identified patterns of tumor evolution with potential clinical significance, including clonal expansion and contraction, linear and branched evolution, as well as the temporal ordering of any pair of mutations.

NanoString Sequencing

Total RNA was extracted from Formalin-Fixed Paraffin Embedded tumour biopsies using the AllPrep® DNA/RNA FFPE Kit from Qiagen according to the manufacturers instructions (Qiagen, Venlo, Netherlands). Total concentration and fragment analysis of the RNA was done using Agilent RNA 6000 Nano Reagents on the Bioanalyzer (Agilent, Santa Clara, CA, USA). A fragment-adjusted RNA concentration of 140ng was hybridized with the PanCancer Pathway Panel Code Set including 770 cancer genes and 30 CLL/RS-related genes for a minimum of 16 hours at 65 degrees C. Hybridized samples were processed on the nCounter *SPRINT* Profiler (NanoString, Seattle, WA, USA). Quality controls and normalisation using 40 housekeeping genes was done using the nSolver software (NanoString).

Pathway analysis

We examined 45 curated gene sets, including CLL drivers and DDR genes, all NanoString PanCancer pathways (including housekeeping genes), cell-cycle and apoptosis genes and all KEGG signaling pathways. For each pathway, we performed burden analysis and clonal analysis of component mutations, as well as analysis of differentially expressed component genes. For establishing differences between paired blood and lymphnode samples, we applied a Bayesian version of McNemar's exact test. Subsequently, pathways were ranked based on the calculated posterior probabilities and Bayes factors in support of the null hypotheses of no difference between blood and lymphnode samples.

Results and Discussion

IgHV mutation status

Using targeted deep sequencing we reveal multiple hypermutated or unmutated IgHV subclones or a mixture of both in 25% of patients[26]. We demonstrate that the predictive value of conventional IgHV mutation analysis can be improved according to this classification and failure rates are lower compared to conventional diagnostics. In addition, Total RNAseq data of 32 patients with an aggressive CLL identifies a biologically distinct group expressing IgLV3-21 light chain (Stamatopoulos et al, under review). This group has poor outcome with respect to treatment-free and overall survival independent of other markers including hypermutated IgHV in multivariate analysis of two additional independent cohorts of almost 800 patients. In conclusion, we establish that number and mutation status of IgHV subclones and types of IgHV and IgLV rearrangements refine conventional prognostication of CLL.

TP53 disruption

We show that WGS -similar to targeted deep NGS- improves sensitivity of TP53 mutation detection compared to Sanger sequencing[27]. Analysis of the TP53 locus using WGS also reveals cryptic structural changes, further increasing the negative predictive value of TP53 analysis. The Oxford MDC offer a free IsoStandard accredited NGS test with a sensitivity of 5% for all patients in the UK in need of TP53 mutation analysis. Phase IV audit data of this service shows that sensitive NGS-based analysis combined with rigorous pathogenicity assessment of TP53 variants abolishes the need for FISH testing.

We have developed a bio-informatics approach for digital karyotyping to overcome the weaknesses of conventional karyotyping of high failure rates, sample requirements for life cells and the challenges of standardization associated with conventional metaphase analysis of stimulated cells. The concordance for calling copy number aberrations between WGS and FISH and WGS and array was 82% and 93%, respectively[27].