Detection of clinically relevant molecular alterations in Chronic Lymphocytic Leukaemia by Nanopore sequencing

Adam Burns[1,2], David Bruce[1,2], Pauline Robbe[3], Adele Timbs[2], Basile Stamatopoulos[4], Ruth Clifford[5], Maria Lopopolo[6], Duncan Parkes[6], Kate Ridout[1,2], Anna Schuh[1,2]

[1] Department of Oncology, University of Oxford, Oxford, UK, [2] Oxford Molecular Diagnostic Centre, John Radcliffe Hospital, Oxford, UK [3] NDCLS, University of Oxford, Oxford, UK, [4] Université libre de Bruxelles, Brussels, Belgium, [5] Department of Haematology, University Hospital Limerick, Limerick, Ireland, [6] WTCHG, University of Oxford, UK

Background

Chronic lymphocytic leukaemia (CLL) is the most prevalent form of leukaemia in the Western world, and displays considerable biological and clinical heterogeneity. CLL can progress into either an aggressive, chemo-resistant, form with poor prognosis, or an indolent form with a similar life-expectancy to that of the general population. A number of prognostic markers, in particular IgHV mutation status, mutations in the *TP53* gene and deletions of the p-arm of chromosome 17, can be used to predict a patient's response to individual chemo-therapeutics and give an indication as to their long-term prognosis. Clinical guidelines recommend screening patients prior to the initial, and any subsequent, rounds of treatment. Current screening methods involve three separate assays, each of which is time-consuming and requires significant investment in equipment. Nanopore sequencing offers a rapid, low-cost alternative, with the potential to generate a full prognostic dataset in a single assay.

Aims

1. Demonstrate the feasibility of Nanopore sequencing to detect prognostic markers in chronic lymphocytic leukaemia.
2. Identify prognostic markers in a cohort of clinical CLL patients, and compare the performance of nanopore sequencing to existing diagnostic methods.

Methods

Samples

We selected genomic DNA from 11 CLL patients with known *TP53* mutation, del(17p) encompassing the *TP53* locus, and IgHV mutation status, including three with multiple IgHV rearrangements (Table 1).

Table 1. Sample details

Sample ID	IgHV			*TP53* Mutations	del(17p)
	VH Family	% Homology	Status		
CLL345	VH3-49	100	Unmutated	K132Q	-
CLL347	VH3-9	100	Unmutated	WT	-
CLL145	VH3-23	95.95	Mutated	R273H, A159P	-
CLL331	VH3-72	96.03	Mutated	R158SfsTer6	20Mb
CLL351	VH3-72	92.38	Mutated	WT	-
	VH3-13	91.81	Mutated	WT	-
CLL371	VH3-23	93.92	Mutated	WT	-
	VH3-30	90.88	Mutated	WT	-
ARC245	VH3-30	100	Unmutated	S109P	-
ARC603	VH3-7	100	Unmutated	R81X	19Mb
ADM293	VH1-69	100	Unmutated	G193V	-
ADM477	VH3-30	98.56	Unmutated	E457A	19Mb (cnLOH)
	VH1-3	99.65	Unmutated		
ADM455	VH3-21	96.7	Mutated	WT	-

Library Preparation

Individually barcoded libraries were prepared for IgHV and *TP53* per sample using previously described primers[1], modified to include Oxford Nanopore barcoding sequences. Following individual barcoding PCR reactions, the amplicons were prepared for 1D sequencing using the amplicon-by-ligation protocol (Oxford Nanopore Technologies). Whole genome sequencing libraries were prepared for each sample from 400ng gDNA using the Rapid Sequencing protocol (Oxford Nanopore Technologies).

Nanopore Sequencing

Pooled IgHV libraries were loaded onto a single flowcell, with *TP53* libraries sequenced on a second flowcell. Whole genome sequencing libraries were each loaded onto individual flowcells. All libraries were sequenced on a MinION instrument for 48 hours. A bespoke pipeline was used to demultiplex the samples, determine IgHV status and call and annotate *TP53* variants. The QDNASeq package was used to call copy-number alterations.

References

[1] Campbell. M.J. *et al*, Mol. Immunol. **29**, 193-203, 1992

Results - IgHV

After generating a consensus IgHV sequence for each sample, we were able to correctly identify the IgHV rearrangement in each case, including the presence of sub-clones in two cases. Furthermore, the level of germline homology was concordant with short-read sequencing data.

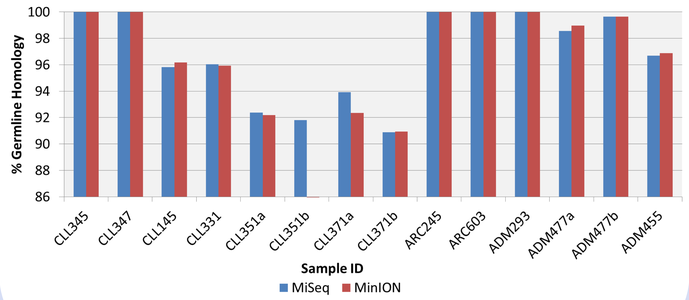

Figure 1. IgHV germline homology obtained by nanopore sequencing.

Results – *TP53*

In order to exclude as many false positive calls as possible, we set the minimum supporting read depth to >40, with minimum coverage requirement of 400x depth and a minimum quality score of Q15. A total of 56 SNVs were identified in our data, of which 8 were both exonic and non-synonymous (Table 2).

Table 2. Details of mutations detected in *TP53*.

Sample ID	AA Change	MiSeq				MinION			
		Read Depth	Ref Count	Alt Count	VAF	Read Depth	Ref Count	Alt Count	VAF
ADM293	G325V	960	480	480	0.50	4,781	2,619	2,162	0.45
ADM477	E285K	933	336	597	0.64	18,277	8,552	9,725	0.53
ARC245	S241P	4,015	2,288	1,727	0.43	4,563	3,337	1,226	0.27
ARC603	R213X	1,839	331	1,508	0.82	8,147	3,729	4,418	0.54
CLL145	R273H	3,485	1,828	1,657	0.48	16,666	8,790	7,876	0.47
CLL145	A159P	1,336	1,212	118	0.09	99,877	92,212	7,665	0.07
CLL345	K132Q	2,560	1,374	1,186	0.46	33,289	20,050	13,239	0.40
CLL331	V173A	-	-	-	-	12,779	8,748	4,031	0.31

Seven of the eight variants (87.5%) were confirmed by a TSCA panel run on a MiSeq. The eighth (V173A in CLL331), was confirmed as a false positive, with 97% of the supporting reads located on one strand, suggesting some motif specificity (Figure 2).

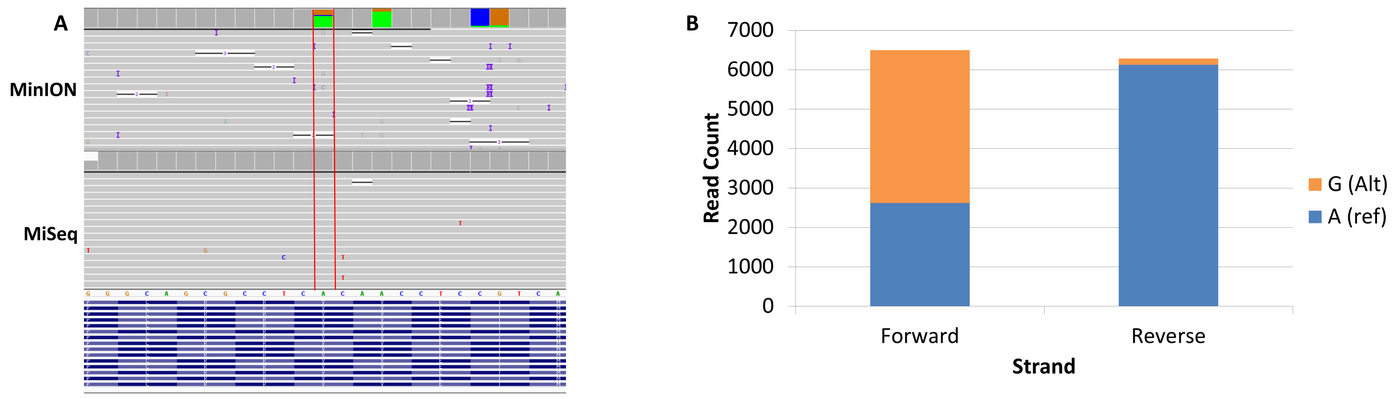

Figure 2. False positive V173A *TP53* mutation in CLL331. A) Variant can be clearly seen in nanopore data (top panel), but is missing from the MiSeq data (bottom panel). **B**) Distribution of variant supporting reads between strands.

Results - del(17p)

Following alignment to hg19, we achieved whole genome sequencing depth ranging from 0.3-3.4x. Using this low-coverage WGS data as input for the QDNAseq package in R, we were able to identify both the 20Mb deletion in CLL331, and the 19Mb deletion in ARC603 (Figure 3). This correlates well with data obtained in an earlier study from Illumina WGS data (Table 3).

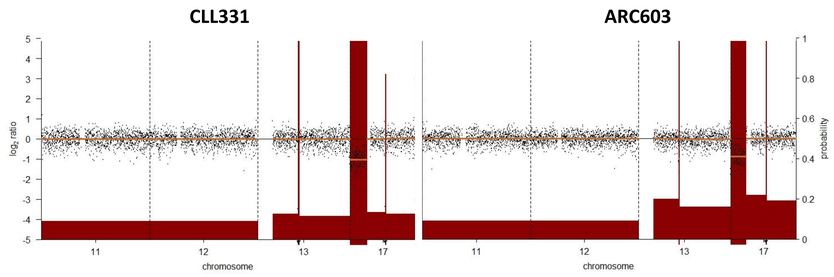

Figure 3. Detection of copy-number alterations by whole genome sequencing. Log2 ratio plots for both CLL331 and ARC603 across chromosomes 11, 12 13 and 17.

Table 3. Comparison of copy-number alterations detected using nanopore and short-read sequencing.

Sample ID	MinION			HiSeq		
	Start	End	Type	Start	End	Type
CLL331	1	22,262,112	Het Loss	1	20,638,596	Het Loss
ARC603	1	20,853,668	Het Loss	1	18,928,388	Het Loss

Conclusions

Here we demonstrate that characterisation of the IgHV locus in CLL cases is possible using the MinION platform, provided sufficient downstream analysis, including error correction, is applied. Furthermore, somatic SNVs in *TP53* can be identified, although small insertions and deletions are more problematic. Work is ongoing to try alternative variant callers to improve the analysis pipeline. The detection of copy-number alterations, specifically del(17p), is possible from low-coverage WGS on the MinION. This shows that Nanopore sequencing is a viable, low-cost alternative to established screening methods.