Acquired CARD11 mutations promote BCR independence in Diffuse Large B-Cell Lymphoma.

Supplementary Methods

Patient Sample

The patient provided explicit written informed consent for the publication of photographic images. Biopsy samples were obtained with written informed consent for tissue banking and DNA sequencing. The study protocol was approved by the East of England Cambridge South Research Ethics Committee (approval reference number 07/MRE05/44). Samples were provided by the Cambridge Blood and Stem Cell Biobank, which is supported by the Cambridge NIHR Biomedical Research Centre, Wellcome Trust – MRC Stem Cell Institute and the Cambridge Experimental Cancer Medicine Centre, UK.

DNA extraction

600 µl Lysis Buffer and 15 µl Proteinase K solution were used to resuspend mononuclear cells from disaggregated bone marrow trephine and incubated at 65° for 15 min. Genomic DNA was precipitated using isopropanol and resuspended in TE Buffer.

Whole exome sequencing and analysis

Whole exome sequencing was performed on tumor biopsy samples from diagnosis, pre-treatment, relapse and germline control (Novogene). For whole exome sequencing, libraries were generated using the Agilent SureSelectXT Reagent Kit+Agilent SureSelect Human All ExonV6 and sequenced on the Illumina NovaSeq 6000 using 150 base paired-end sequencing (Q30≥80%).

All samples were aligned to the hg38 version of the reference human genome using bwa 0.7.15 in alt contig aware mode as described by the authors[1]. The generated SAM file was compressed into a BAM file and sorted by genomic position using samtools 1.3.1[2]. The sorted

29 BAM files were subject to Base Quality Score Recalibration and Indel Realignment as
30 specified in the Genome-Analysis Toolkit (GATK)[3] best practices[4,5]. For somatic variant calling
31 the following GATK's MuTect2[6] was used. A panel of normals (PON) was generated using the
32 germline samples with GATK's (version 4.0.3.0) *Mutect2* and *CreateSomaticPanelOfNormals*
33 algorithms. Variants were called in all tumours using the PON and the matched germline
34 sample with the GATK's *MuTect2* algorithm to generate a VCF file[6]. Finally, the VCF files were
35 filtered with GATK's *FilterMutectCalls* algorithm. The resulting VCF file was annotated and
36 prioritized using annovar[7]. CNVs were called using GATK's (version 4.1.4.0) best practices
37 for Somatic copy number variant discovery[3]. Copy ratio was adjusted for tumour purity. Gains
38 associated with 2 or more extra copies were defined as amplifications. Non-CNV Structural
39 variations were called using Delly according to their best practices[8]. The samples were
40 sequenced to an average of 96,176,220 reads per sample and an average coverage of 77.75x
41 (fold enrichment of 38.58).

42 **Sanger sequencing**

43 The CARD11 K215T region was analyzed in the diagnosis, pre-treatment and relapse sample
44 by Sanger sequencing. Primers were designed to generate an amplicon size of 99bp and
45 ordered from IDT. PCR reaction was performed using 200ng genomic DNA with Q5® High-
46 Fidelity 2X Master Mix (NewEngland). The following PCR conditions were used: 98°C for 30s,
47 35 cycles of 98°C for 10s, 60°C for 20s and 72°C for 20s, and the final extension, 72°C for 2
48 min. The resulting PCR product was purified using a Qiaquick PCR purification kit (Qiagen)
49 and sequenced at GATC (Germany).

50 **Plasmid construction and viral transduction**

51 To generate the CARD11 WT overexpression plasmid, the CDS sequence for CARD11 WT
52 (NM_032415, NCBI) was designed as g-block and purchased from IDT. This was cloned into
53 pRCMV-TOP vector using Gibson Assembly. The CARD11 K215T overexpression plasmid
54 was generated from the pRCMV-TOP-CARD11 WT plasmid. Mutation sequence (K215T,

c.644A>C) was obtained from IDT as g-blocks and cloned into the SbfI and SfiI-digested pRCMV-TOP-CARD11 WT using Gibson Assembly. Capillary sequencing was used for plasmid verification. 293T cells (Clontech) were transfected with pRCMV-TOP (Empty Vector), pRCMV-TOP-CARD11 WT/K215T, and packaging plasmids using TransIT-293 (Mirus). Virus supernatant was harvested and filtered (0.48micron) 48 hours post transfection and spinoculated into U2932 cells maintained in 10% FCS RPMI.

Ibrutinib sensitivity assay

The cell line U2932 was transduced with empty vector or CARD11 WT or K215T and selected using puromycin. Cells were then treated with the indicated concentration of ibrutinib and cell viability measured at day 5 using the CellTiter-Glo® Luminescent Cell Viability Assay per manufacture's instruction. Luminescence was read with SpetraMax M2 microplate reader.

Data Availability

Whole exome sequencing files have been uploaded to EGA.

82 **Supplementary References**

83 1. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-
84 Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

85 2. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools.
86 *Bioinformatics* **25**, 2078–2079 (2009).

87 3. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for
88 analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–
89 1303 (2010).

90 4. DePristo, M. a. *et al.* A framework for variation discovery and genotyping using
91 next- generation DNA sequencing data. *Nat Genet* **43**, 491–498 (2011).

92 5. Auwera, G. A. Van Der *et al. From FastQ data to high confidence varant calls:*
93 *the Genome Analysis Toolkit best practices pipeline*. *Curr Protoc*
94 *Bioinformatics* vol. 11 (2014).

95 6. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and
96 heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).

97 7. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of
98 genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**,
99 1–7 (2010).

100 8. Rausch, T. *et al.* DELLY: Structural variant discovery by integrated paired-end
101 and split-read analysis. *Bioinformatics* **28**, 333–339 (2012).

102

103

104

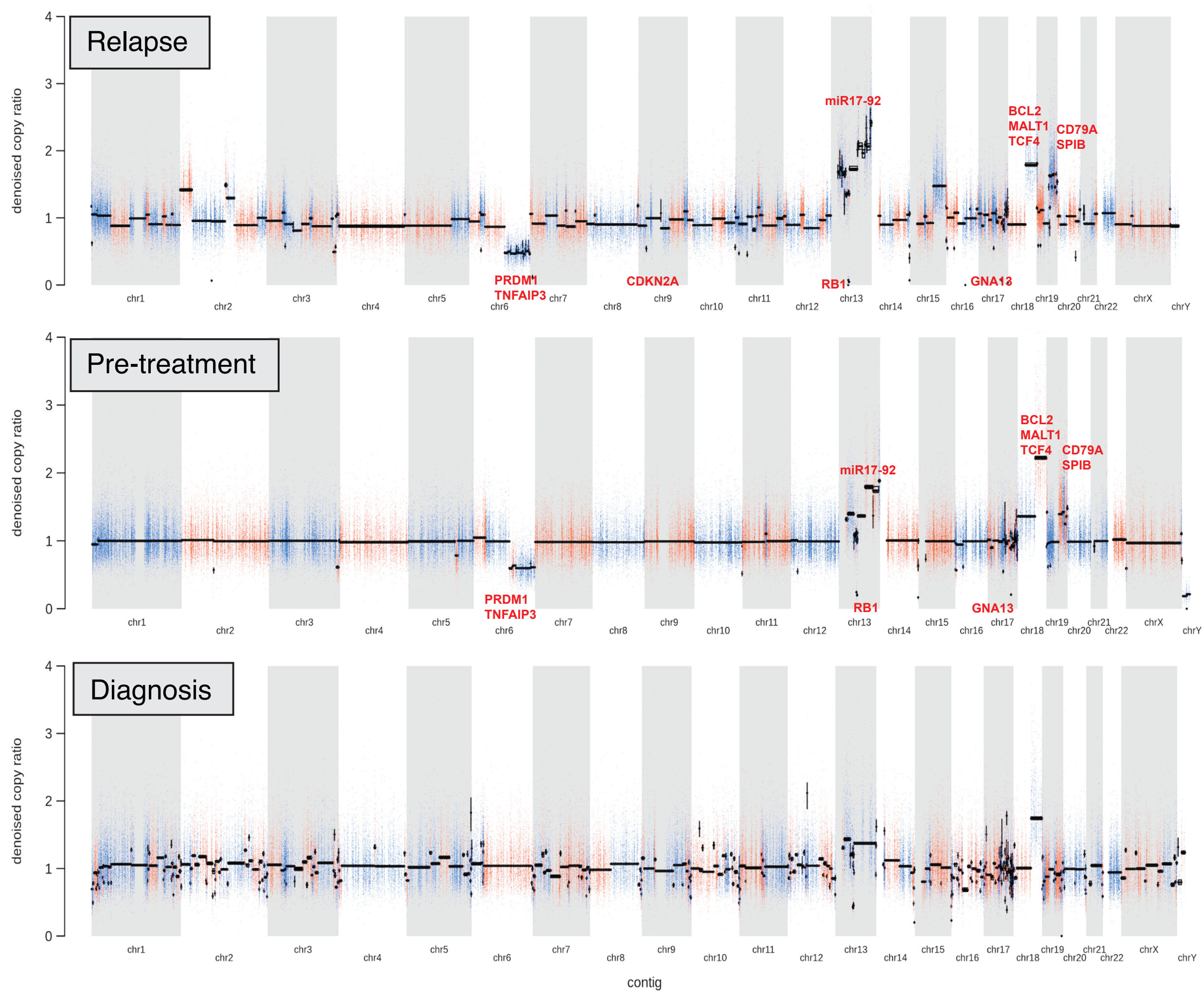
105

106

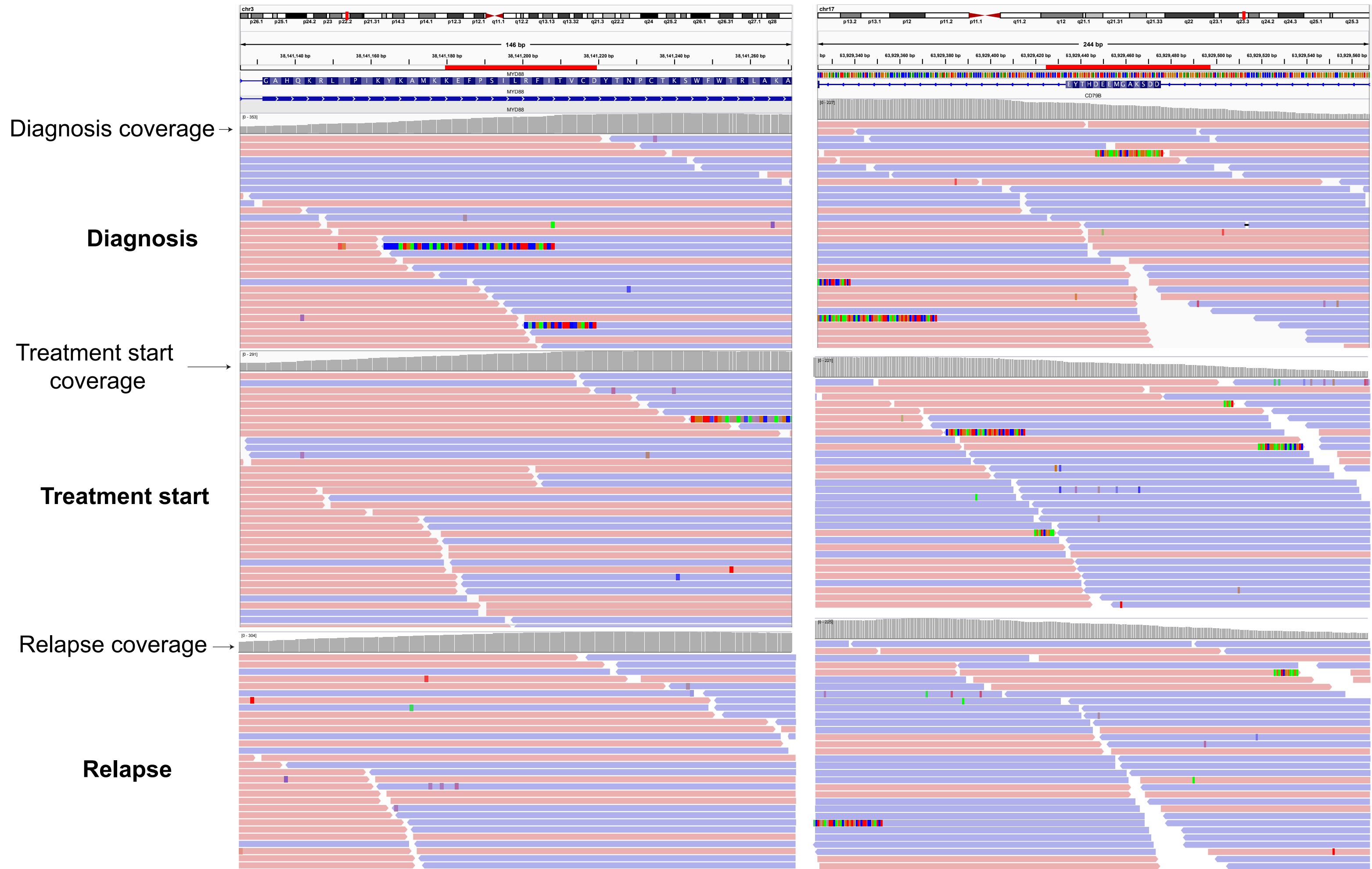
107 **Supplementary Data Files**

108 **Supplementary Data 1.** Summary of germline filtered variants with variant read count >5 and
109 allele fraction >0.1 identified from whole exome sequencing preformed on DNA from
110 diagnostic, pretreatment and relapse biopsies.

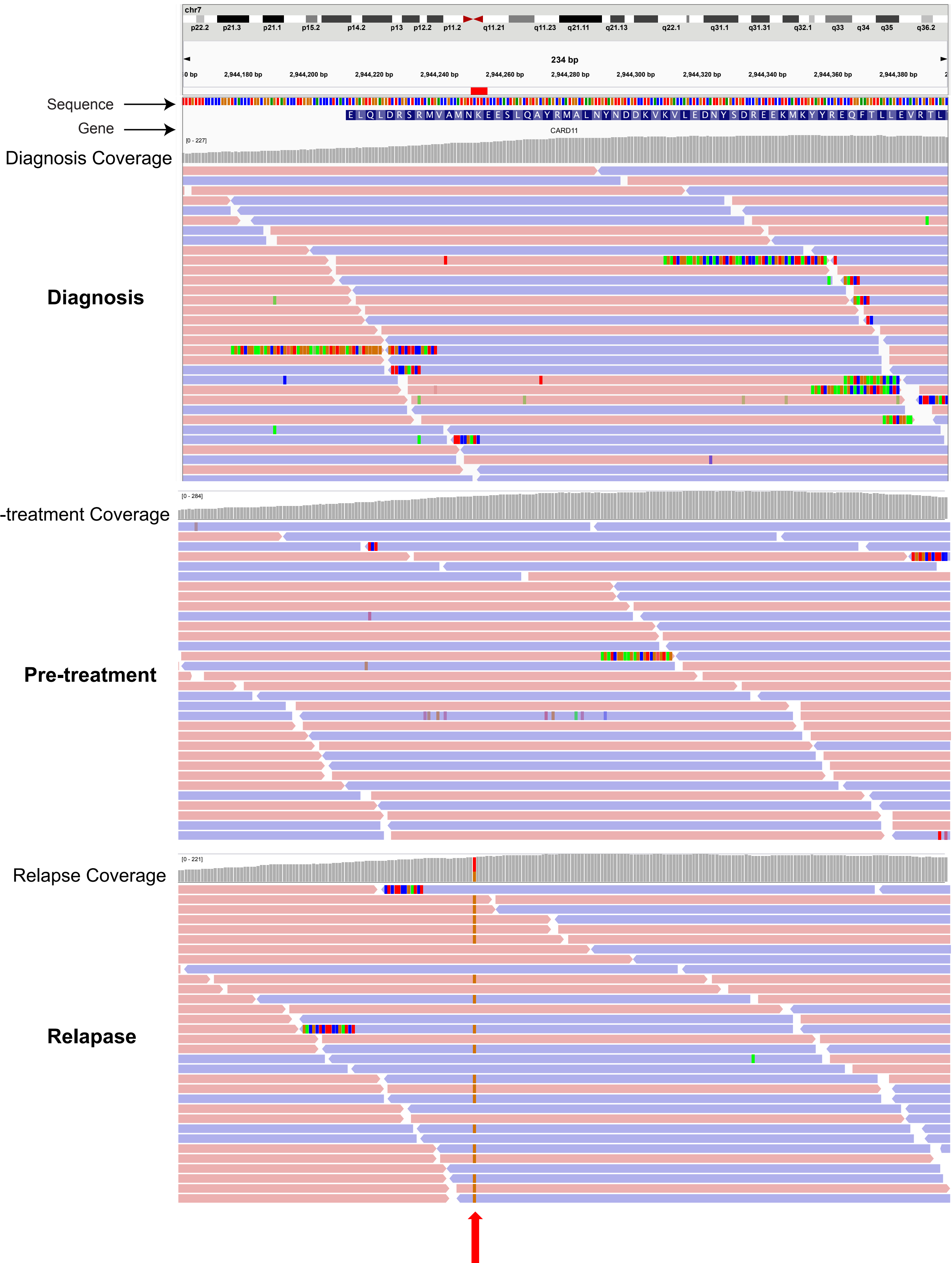
111 **Supplementary Data 2.** Summary of copy number changes detected from whole exome
112 sequencing at diagnosis, pretreatment and relapse.



Supplementary Figure 1
Copy number change detected at the indicated timepoints. The position of known lymphoma driver genes is indicated.



Supplementary Figure 2
Summary of exome sequencing reads mapping to mutation hotspot regions of *MYD88* and *CD79B* confirming that both genes are wild type at these genomic locations at each timepoint.



Supplementary Figure 3
Summary of exome sequencing reads mapping to the region of the *CARD11* locus containing detected variant at the indicated timepoints, showing acquisition of the *CARD11*_K215T mutation (red arrow) at relapse.