medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Patterns of oncogene co-expression at single cell resolution in cancer influence survival

Authors: Michal Marek Hoppe, PhD[1]; Patrick Jaynes, PhD[1]; Fan Shuangyi, MD[2]; Yanfen Peng, MSc[1]; Phuong Mai Hoang, PhD[1]; Liu Xin, MA[3]; Sanjay De Mel, MBBS[3,4]; Limei Poon, MBBS[3,4]; Esther Chan, MBBS[3,4]; Joanne Lee, MBBS[3,4]; Ong Choon Kiat, PhD[5]; Tiffany Tang, MBBS[6]; Lim Soon Thye MBBS[6]; Chandramouli Nagarajan, MBBS[7]; Nicholas Francis Grigoropoulos, MBBS, PhD[7]; Soo-Yong Tan, MBBS DPhil[2,4]; Susan Swee-Shan Hue, MBBS[2,4]; Sheng-Tsung Chang, MD[8], Shih-Sung Chuang, MD[8]; Shaoying Li, MD[9]; Joseph D. Khoury, MD[9]; Hyungwon Choi, PhD[10]; Carl Harris III, PhD[11]; Alessia Bottos, PhD[11]; Laura J. Gay, PhD[12]; Hendrik F. P. Runge, BSc[12]; Ilias Moutsopoulos, BSc[12]; Irina Mohorianu, PhD[12]; Daniel J. Hodson, MBBChir PhD[12]; Pedro Farinha, MD[13]; Anja Mottok, MD[13]; David W. Scott, MBChB PhD[13]; Gayatri Kumar, PhD[14]; Kasthuri Kannan, PhD[14]; Wee-Joo Chng, MBChB PhD[1,4,10]; Yen Lin Chee, MBChB PhD[3,4]; Siok-Bian Ng, MBBS[1,2,4]; Claudio Tripodo, MD PhD[15*]; Anand D. Jeyasekharan, MBBS PhD [1,3,4,10*]

Affiliations:
[1]Cancer Science Institute of Singapore, National University of Singapore, Singapore
[2]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[3]Department of Haematology-Oncology, National University Health System, Singapore
[4] NUS Centre for Cancer Research (N2CR), Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore
[5]Division of Cellular and Molecular Research, National Cancer Centre Singapore, Singapore
[6]Division of Medical Oncology, National Cancer Centre Singapore, Singapore
[7]Department of Haematology, Singapore General Hospital, Singapore
[8]Department of Pathology, Chi-Mei Medical Center, Tainan City, Taiwan
[9]Department of Hematopathology, Division of Pathology and Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, USA
[10]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[11]F. Hoffmann-La Roche Ltd, Basel, Switzerland.
[12]Wellcome MRC Cambridge Stem Cell Institute, Cambridge, UK
[13]BC Cancer Research Centre, Vancouver, Canada
[14]Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, USA
[15]Tumor Immunology Unit, University of Palermo, Palermo, Italy

* To whom correspondence is to be addressed
Dr Anand D. Jeyasekharan; Cancer Science Institute of Singapore, Centre for Translational Medicine (MD6) #13-01G, 14 Medical Drive, Singapore 117599; csiadj@nus.edu.sg;

Prof. Claudio Tripodo; University of Palermo School of Medicine, Istituto di Patologia Generale Corso Tukory 211, 90134, Palermo, Italy; claudio.tripodo@unipa.it

Running title: Cellular co-expression of oncogenes in DLBCL
Keywords: Intratumor Heterogeneity, DLBCL, Single-cell analysis, oncogene

This manuscript includes:
Main Text: 2604 (not including Funding, Acknowledgements and COI)
Figures 1 to 5
Supplementary tables 1 to 10
Supplementary figures 1 to 9

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background:

Cancers often overexpress multiple clinically relevant oncogenes. However, it is not known if multiple oncogenes within a cancer combine uniquely in specific cellular sub-populations to influence clinical outcome. We studied this phenomenon using the prognostically relevant oncogenes MYC, BCL2 and BCL6 in Diffuse Large B-Cell Lymphoma (DLBCL).

Methods:

Quantitative multispectral imaging simultaneously measured oncogene co-expression at single-cell resolution in reactive lymphoid tissue (*n*=12) and four independent cohorts (*n*=409) of DLBCL. Mathematically derived co-expression phenotypes were evaluated in DLBCLs with immunohistochemistry (*n*=316) and eight DLBCL cohorts with gene expression data (n=4186). Bulk and single-cell RNA sequencing was performed on patient-derived B-cells with induced co-expression of MYC, BCL2 and BCL6.

Results:

Unlike in non-malignant lymphoid tissue where the co-expression of MYC, BCL2 and BCL6 in a B-cell is limited, DLBCLs show multiple permutations of oncogenic co-expression in malignant B-cells. The percentage of cells with a unique combination MYC+BCL2+BCL6- (M+2+6-) consistently predicts survival in contrast to that of other combinations (including M+2+6+). An estimated percentage of M+2+6- cells can be derived from any quantitative measurement of the component individual oncogenes, and correlates with survival in immunohistochemistry and gene expression datasets. Comparative transcriptomic analysis of DLBCLs and transformed patient-derived B-cells identifies cyclin D2 (CCND2) as a potential BCL6-repressed regulator of proliferation in the M+2+6- population.

Conclusions:

Unique patterns of oncogene co-expression at single-cell resolution affect clinical outcomes in DLBCL. Similar analyses evaluating oncogenic combinations at the cellular level may impact diagnostics and target discovery in other cancers.

244/ 250

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

Oncogene overexpression is common in cancer. The concomitant increase in oncogenic proteins (oncoproteins) influences both prognosis and treatment[1]. Notable examples routinely assessed in clinical practice include HER2 in breast cancer and ALK in lung cancer. However, cancers often overexpress more than one oncogene. Whether multiple oncogenes interact at the single-cell level to influence clinical outcome remains an important unresolved question. This is particularly relevant since cancers are a heterogenous mosaic of tumor cell sub-populations[2], and oncogenes show clinically significant intratumor heterogeneity (ITH) in expression[1]. While only a subset of cells within a tumor co-express all oncogenes measured at the whole-tumor scale, clinical techniques for oncogene estimation in cancer (such as immunohistochemistry, IHC) study them in isolation. It is therefore still not known if subsets of cells within a given cancer expressing specific combinations of oncogene drive eventual clinical phenotypes.

We aimed to address this question using multiplexed fluorescent immunohistochemistry (mfIHC), a technique that can simultaneously and quantitatively evaluate a set of proteins with single-cell resolution. This allows measurement of single-cell oncogene co-expression from sufficient samples for robust multivariate correlations with clinical outcome. We chose Diffuse Large B-cell Lymphoma (DLBCL) as a model to evaluate the clinical impact of ITH in oncogene co-expression. DLBCL is the most common aggressive lymphoma worldwide[3], and overexpression of the oncogenes MYC, BCL2 and BCL6 influence pathogenesis and prognosis[4-6]. However, there is significant variability among studies regarding the prognostic significance of these oncogenes with debate on appropriate cut-off thresholds to define “positivity”. These considerations offer an ideal scenario to evaluate whether these oncogenes show differential co-expression at the single cell level in DLBCL, and to investigate how they collaborate or influence each other at the cellular level to affect survival.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Methods

Samples and datasets

Non-malignant samples included chronic tonsillitis (*n*=10) and reactive lymph nodes (*n*=2) collected at the National University Hospital (NUH) Singapore (DSRB protocol 2015/00176). DLBCL samples in tissue microarray (TMA) format were from NUH (*n*=152), Chi-Mei Medical Center, Taiwan (CMMC; *n*=150), Singapore General Hospital, Singapore (SGH; *n*=67) and the MD Anderson Cancer Center, USA (MDA; *n*=40), and the British Columbia Cancer Agency (BCA; *n*=274) - used for validation[7]. Datasets GSE117556 (*n*=928)[8], GSE125966 (*n*=553), GSE31312 (*n*=498)[9], GSE10846 (*n*=420)[10], GSE32918 (*n*=140)[11] and GSE87371 (*n*=221)[12] were obtained from the Gene Expression Omnibus (GEO), while those from Reddy *et al* (*n*=773)[13] and Schmitz *et al* (*n*=480)[14] were from the database of Genotypes and Phenotypes (dbGaP). The GOYA dataset (GSE125966) was analysed independently by F. Hoffmann-La Roche Ltd.

Quantitative multiplexed fluorescent immunohistochemistry (mfIHC) and scoring

Quantitative multiplexed fluorescent immunohistochemistry (mfIHC) was performed using sequential OPAL-TSA staining as described previously[15] (Supplementary table1 and Supplementary table2). Images were acquired using the Vectra imager and analysed using inForm2.4.8 (Akoya Biosciences, USA). DAPI nuclear staining and CD20 membrane staining were used to segment cells. A pathologist-evaluated fluorescent intensity threshold defined “positive” and “negative” cells for CD20, MYC, BCL2 and BCL6 per tissue microarray (TMA) core. For every cell in a given case, information on CD20, MYC, BCL2 and BCL6 allowed categorization into one of the following permutations: M+2+6+, M+2+6-, M+2-6+, M+2-6-, M-2+6+, M-2+6-, M-2-6+ and M-2-6- (where M is MYC, 2 is BCL2, and 6 is BCL6 and + or – indicates the presence or absence of the oncogene) within the CD20+ B-cell compartment.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Survival analysis

Where used, all variables satisfied proportional hazards assumptions for Cox proportional hazard (Cox PH) models. Hazard ratios are displayed per unit (in 5% increments) of percentages of subpopulations as continuous variables. For Kaplan-Meier analyses, the cohorts were dichotomized at an optimal cut-off. Statistical tests were two-sided and $p \leq 0.05$ was considered as statistically significant. Statistical tests were performed using SPSS23, GraphPad Prism9 and R.

Supplementary methods/ table

The supplementary methods contains details on patient characteristics, probabilistic subpopulation percentage prediction, transformation of RNA data, semi-quantitative IHC, mutational and gene expression analysis, and manipulation of tonsil-derived GC B cells.

Data Sharing Statement

RNAseq data is deposited to GEO[16] (https://www.ncbi.nlm.nih.gov/geo/) GSE203446. Requests for primary image data/protocols can be made to Dr Anand Jeyasekharan via email: csiadj@nus.edu.sg.

Results

Physiological patterns of MYC, BCL2 and BCL6 co-expression are disrupted in DLBCL

We first compared the co-expression of the oncogenes MYC, BCL2 and BCL6 between reactive lymphoid tissue and DLBCL (Supplementary figure1A). Consistent with known physiology, BCL6 and BCL2 expression was restricted to reactive lymphoid germinal centres (GC) and extra-GC regions respectively, while MYC showed sparse positivity in the GC CD20+ cells[17] (Figure1A, Supplementary figure1B). A binary +/- map for each oncogene facilitates the quantitation of “sub-populations” of cells based on MYC BCL2 and BCL6 co-expression (Supplementary figure1C). The repertoire of sub-populations defined by MYC/BCL2/BCL6 in reactive lymphoid B-cells was limited, with M-2-6+ being dominant within

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the GC, and M-2+6- being dominant outside the GC (Figure1B). Hardly any cells displayed co-expression of all three oncogenes MYC, BCL2 and BCL6 in either compartment, consistent across several reactive lymphoid tissues analysed (Figure1C, Supplementary figure1D).

In contrast, DLBCL cells frequently co-express these three oncogenes (Figure1D, Supplementary figure1E, and Supplementary figure2). Even within cases characterized by high overall levels of MYC, BCL2 and BCL6 expression, these three oncogenes were not always found in the same cells, underscoring intratumoral heterogeneity (ITH) in DLBCL. The percentage of each sub-population was variable between patients but was remarkably similar in overall distribution and clustering among four cohorts (Figure1E, Supplementary figure2). The percentages of sub-populations were stable across distinct TMA cores from the same patient, confirming them to show patient-specific profiles (Figure1F, Supplementary figure1F). These sub-populations were not associated with clinico-pathological features such as age, gender, International Prognostic Index (IPI) Risk Group, cell-of-origin (COO, IHC Hans algorithm[18]), nor were they associated with MYC/BCL2/BCL6 translocation status (Supplementary figure3) - confirming previous observations that translocations do not account for the majority of MYC, BCL2 and BCL6 overexpression in DLBCL[19]. Pair correlation function (PCF) analyses on single-cell resolved image data with spatial coordinates[20] show that subpopulations cluster in space and do not display a random (Poisson) distribution (Supplementary figure4 and Supplementary Results), in keeping with local ecosystems for proliferation and evolution within a growing tumor.

Cells co-expressing MYC and BCL2 without BCL6 confer poor survival in DLBCL

We next evaluated the relationship between MYC/BCL2/BCL6 sub-populations and prognosis, using pretreatment biopsies of R-CHOP treated DLBCL cases from 3 cohorts (NUH, *n*=98, SGH, *n*=41; MDA, *n*=36). To avoid arbitrary cut-offs, we evaluated the percentage of cells with oncogenic combinations as a continuous variable in a univariate cox proportional hazards (Cox PH) analysis for overall survival (OS). Though there was expected variability between cohorts in the prognostic impact of MYC, BCL2 and BCL6 as individual oncogenes[21], the

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

percentage of M+2+6- cells stood out for consistent poor prognostic significance (Figure2). This association of high M+2+6- percentage and poor outcome remained statistically significant in a multivariate Cox PH model adjusted for clinically relevant DLBCL clinicopathological parameters of IPI Risk Group, COO and MYC FISH status (Supplementary table3; Supplementary table4). These results suggest that the prognostic impact of these oncogenes in DLBCL is driven by a unique sub-population of cells expressing MYC and BCL2 without BCL6.

A probabilistic model accurately predicts MYC, BCL2 and BCL6 co-expression in DLBCL

As the percentage of MYC+BCL2+BCL6- (M+2+6-) cells correlates with poor survival, we wanted to check if this percentage could be inferred mathematically. We first evaluated the relationship between these oncogenes in DLBCL at the single cell-level (Figure3A). In contrast to mutually exclusive expression within specific topological compartments in reactive lymphoid tissues, MYC, BCL2 and BCL6 do not show any correlation with each other in DLBCL (Figure3B). This suggests that independent gene regulatory mechanisms drive expression of these oncogenes in DLBCL, and that cellular co-expression therefore is a stochastic outcome (i.e. based on probability alone). An implication of this phenomenon is that the percentage of any sub-population are predictable using a probabilistic model, if the percentages of each component oncogene are known. In this model, the percentage of any given subpopulation is derived by multiplying proportions for presence or absence of each individual marker comprising the sub-population (Supplementary Methods). We validated this hypothesis using our single-cell resolved mfIHC data, observing a highly concordant correlation between observed and predicted percentages for each sub-population (Figure3C, Supplementary figure5).

An extension of these findings is that any quantitative data on MYC, BCL2 and BCL6 allows estimation of the percentage of each sub-population characterized by their co-expression. Visual scoring of MYC, BCL2 and BCL6 percentage on chromogenic IHC remains the

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

reference method for assessment of these oncogenes in clinical practice. We checked if our algorithm could estimate prognostic MYC, BCL2, BLC6 sub-populations from clinical-grade pathologist scores for MYC, BCL2 and BCL6 chromogenic IHC in a well-characterized cohort of DLBCL[22]. High M+2+6- percentage, whether directly measured by mfIHC or inferred from visual chromogenic IHC scores for single oncogenes, remained prognostic in this cohort (Figure3D), confirming the applicability of this algorithm to clinical IHC scoring in DLBCL.

Estimation of MYC, BCL2 and BCL6 co-expressing sub-populations can be extended to gene expression data

We hypothesized that if the percentage of MYC/BCL2/BCL6 sub-populations could be inferred from individual oncogene components on IHC, they can be inferred from any quantitative data measuring MYC/BCL2/BCL6 expression- including gene expression profiling (GEP)[23]. mRNA levels of MYC, BCL2 and BCL6 do not correlate with each other (Figure4A-B), confirming their independent regulation at the transcriptional level as well. Remarkably, across eight distinct cohorts of DLBCL patients with GEP data, mRNA-inferred M+2+6- percentage proved again to be the only subpopulation consistently associated with poor survival (Figure4C; Supplementary Methods, Supplementary figure6 and Supplementary table5). This was further corroborated in a multivariate Cox PH analysis correcting for IPI Risk Group and COO (where the data was available; Supplementary table6). This GEP-focused analysis of M+2+6- was further independently replicated and validated in the GOYA clinical trial cohort[24] (Supplementary figure7). Using these large GEP datasets, we establish an optimal dichotomization cut-off for M+2+6- at ≥15% to reliably stratify M+2+6- high patients for poor prognosis (Figure4D, Supplementary table7).

***CCND2* is enriched in the M+2+6- subpopulation and enhances its proliferative potential**

Inference of oncogene co-expression from GEP datasets allows a comparative analysis with other molecular characteristics in DLBCL. Inferred M+2+6- percentage extent was consistently associated with the unfavourable MCD and A53 genetic subtypes of DLBCL (Figure5A)[25]. We

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

also compared the inferred M+2+6- sub-population percentage with the expression of 15314 consensus genes across 3873 patients (Figure5B, Supplementary table8). Sixty-seven genes consistently correlated either positively or negatively with M+2+6- percentage extent (Figure5C), the biological significance of which will require detailed study in the future. To narrow down in the first instance, we leveraged on the observation that M+2+6- percentages robustly predict for survival while M+2+6+ percentages do not. We compared gene expression of DLBCL with gene expression of primary human tonsil-derived GC B-cells immortalised by either the overexpression of MYC and BCL2 (M+2+6-) or MYC, BCL2 and BCL6 (M+2+6+)[26]. Since BCL6 is a transcriptional repressor[27], we hypothesised that the absence of BCL6 could influence the transcriptional profile of the M+2+6- subpopulation (Figure5D, Supplementary table9). Cross-comparing genes enriched in the M+2+6- GC B cells with genes correlated with the M+2+6- population from clinical GEP datasets, we found that *CCND2* was highly enriched in both groups (Figure5E). Single-cell RNA sequencing of a tonsil-derived GC B-cell sample clearly demonstrated an inverse correlation between *CCND2* and *BCL6* expression (Figure5F), confirming prior observations that CCND2 is transcriptionally repressed by BCL6[28,29]. We then transduced *CCND2* in M+2+6+ immortalized B cells, which were characterized by low background levels of CCND2 expression (Figure5G). The transduced M+2+6+/CCND2^{High} population started as a relatively small fraction, but rapidly expanded over time eventually outgrowing the M+2+6+/CCND2^{Low} population (Figure5H), confirming that increased CCND2 can confer a fitness advantage to cells with MYC and BCL2. Cyclin D2 expression has been reported as a marker for an adverse outcome in DLBCL[18,30]. These results may explain, in part, why DLBCL patients with a high percentage extent of the M+2+6- have a poorer prognosis, although other contributing genes and cell-extrinsic features will need evaluation in-vivo. Of general significance however, these results illustrate how estimation of oncogene co-expression phenotypes through gene expression data may uncover novel biological insight.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Discussion

In this paper, we show for the first time that sub-populations of tumor cells expressing combinations of oncogenes influence patient prognosis. We also show that (under conditions of independently regulated expression) these sub-populations can be inferred from quantitative single oncogene data, with a remarkable concordance to actual observed single-cell co-expression on multiplex IHC. We show two applications of predicting oncogenic sub-populations in the setting of DLBCL. First, the M+2+6- sub-population can be predicted from clinical chromogenic IHC scores, offering a refined method for utilizing MYC, BCL2 and BCL6 expression for prognostic use in DLBCL. Secondly, by estimating subpopulation percentages from GEP datasets, we demonstrate feasibility of identifying molecular features associated with a poor prognostic oncogene combination from the multitude of gene expression studies available for a given disease. Such features could identify therapeutic targets, or offer biological insight- as with our demonstration that cell cycle regulator CyclinD2 (*CCND2)* may play a role in the aggressive phenotype of M+2+6- cells. CyclinD2 promotes the G1/S transition of haematopoietic cells[31], enhances cytokine induced-proliferation[32], and is stabilized by EBV infection[33], highlighting the rationale for further studies of *CCND2* in DLBCL pathogenesis and evolution.

Our single-cell resolved quantitative imaging confirms that ITH in co-expression of MYC, BCL2 and BCL6 exists in almost every case of DLBCL. This co-expression shows distinct spatial organization with non-random clustered patterns, suggesting that forces beyond genetic heterogeneity shape DLBCL evolution. These findings also suggest that quantitative assessment of the M+2+6- sub-population refines the MYC-BCL2 “double expressor lymphoma” (DEL)[4,34,35] classification. As prior DEL classifications do not take DLBCL ITH[36,37] into account, it was not known if DELs represent two distinct and co-existing clonal phenotypes within a lymphoma- one expressing MYC and the other BCL2. Nor was it understood why the poor outcome of DEL is exacerbated when BCL6 expression is absent[6,38]. These issues are addressed by the description of the M+2+6- subpopulation, evaluating the phenomenon at

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

single-cell resolution. DEL remains relevant in the era of genetic DLBCL classification[39] and novel targeted therapies. For example- patients with the DEL phenotype show improved survival on the Polatuzumab arm in comparison to the R-CHOP arm of the POLARIX trial[40]. Additional studies are required to clarify the relevance of the M+2+6- population in this setting.

This is a proof-of-concept study with several limitations. The prognostic impact of M+2+6- percentages derive from retrospective analyses and need prospective validation, along with the development of a standardized cut-off for the M+2+6- percentage. Our initial analyses from large GEP datasets suggest that $\geq$15% M+2+6- percentage extent may be a suitable starting point that is discriminatory for survival. The model we describe for prediction of oncogenic co-expression only holds true when the expression of the oncogenes are independently regulated, and will need validation in the setting of other oncogenes/cancers. Nonetheless, our demonstration that the M+2+6- phenotype confers poor survival in four empirically evaluated cohorts (mfIHC) and eight inferred cohorts (GEP) of DLBCL underscore the clinical importance of evaluating the ITH generated by the co-expression of oncogenes, and suggest that similar studies in other cancer types will be informative. Oncogene ITH occurs at multiple molecular levels in cancer: genetic[41,42], epigenetic, transcriptomic and proteomic[43] , and affects clinical phenotypes[1,44]. Single-cell approaches to evaluate genetic[45], transcriptomic[46,47] and proteomic ITH have provided valuable insight into microevolutionary processes operating in cancer[48,49]. However, due to high experimental costs, the number of patients represented in scRNAseq and mass cytometry datasets are invariably small, thus precluding clinically meaningful multivariate analyses. Here we demonstrate that multiplexed microscopy through mfIHC, though limited in multiplexing potential compared to scRNAseq, is well suited to measure the clinical impact of ITH in large patient cohorts.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgements

The authors acknowledge a Yong Siew Yoon Research Grant to ADJ from the National University Cancer Institute of Singapore towards the purchase of the Vectra 2 multispectral imaging system microscope.

Funding

ADJ was supported by the Singapore Ministry of Health's National Medical Research Council Transition Award (NMRC/TA/0052/2016). Work in ADJ's laboratory is funded by a core grant from the Cancer Science Institute of Singapore, National University of Singapore through the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiative, as well as by Singapore Ministry of Health's National Medical Research Council 'Singapore lYMPHoma translatiONal studY (SYMPHONY)' Open Fund Large Collaborative Grant (MOH-000205-03). CT was supported by the Italian Foundation for Cancer Research (AIRC) Investigator Grant IG ID.22145; 5x1000 Grant ID.22759. The work in DJH's lab was supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. D.J.H. was supported by a fellowship from Cancer Research UK (CRUK) (RCCFEL\100072) and received core funding from Wellcome (203151/Z/16/Z) to the Wellcome-MRC Cambridge Stem Cell Institute and from the CRUK Cambridge Centre (A25117). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Competing interests

ADJ has received consultancy fees from Turbine Ltd, AstraZeneca, Antengene, Janssen MSD and IQVIA; travel funding from Perkin Elmer; and research funding from Janssen and AstraZeneca The other co-authors have no relevant conflicts of interest to declare.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Authorship Contributions

Conceptualization- MMH, ADJ

Methodology (Experimental and analytical design)- MMH, FS, HC, WJC, SBN, CT, ADJ

Project administration (Coordination of sample provision/ethics approvals) - LX, SDM, LP, AB

Investigation (Quantitative mfIHC staining/ imaging) FS, YP, PMH

Data curation (Collection and curation of clinical data): LX, SDM, LP, EC, JL, YLC, DWS, OCK, TT, LST, CN, NFG, JDK

Formal analysis (Pathological review of cases): SYT, SSSH, STC, SSC, SBN, PF, AM, SL

Formal analysis (Multiplex analysis) MMH, YP, FS

Formal analysis (GC B cell manipulation/FACS analysis): LJG, HFPR, DJH

Formal analysis (Bioinformatics/statistical analysis) MMH, HC, CH, GK, KK, IrM, IIM, ADJ

Interpretation of findings- MMH, SBN, DJH, CT, ADJ

Writing – original draft: PJ, MMH, ADJ

Writing – review & editing: PJ, MMH, WJC, SBN, CT, ADJ

Underlying data verification: MMH, ADJ

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell 2020;37:471-84.
2. Maley CC, Aktipis A, Graham TA, et al. Classifying the evolutionary and ecological features of neoplasms. Nature Reviews Cancer 2017;17:605-19.
3. Miao Y, Medeiros LJ, Li Y, Li J, Young KH. Genetic alterations and their clinical implications in DLBCL. Nat Rev Clin Oncol 2019;16:634-52.
4. Johnson NA, Slack GW, Savage KJ, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol 2012;30:3452-9.
5. Hu S, Xu-Monette ZY, Tzankov A, et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. Blood 2013;121:4021-31.
6. Horn H, Ziepert M, Becher C, et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. Blood 2013;121:2253-63.
7. Ennishi D, Jiang A, Boyle M, et al. Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma. J Clin Oncol 2019;37:190-201.
8. Sha C, Barrans S, Cucco F, et al. Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy. J Clin Oncol 2019;37:202-12.
9. Visco C, Li Y, Xu-Monette ZY, et al. Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

International DLBCL Rituximab-CHOP Consortium Program Study. Leukemia 2012;26:2103-13.

10. Lenz G, Wright G, Dave SS, et al. Stromal Gene Signatures in Large-B-Cell Lymphomas. N Engl J Med 2008;359:2313-23.

11. Barrans SL, Crouch S, Care MA, et al. Whole genome expression profiling based on paraffin embedded tissue can be used to classify diffuse large B-cell lymphoma and predict clinical outcome. Br J Haematol 2012;159:441-53.

12. Dubois S, Viailly PJ, Bohers E, et al. Biological and Clinical Relevance of Associated Genomic Alterations in MYD88 L265P and non-L265P-Mutated Diffuse Large B-Cell Lymphoma: Analysis of 361 Cases. Clin Cancer Res 2017;23:2232-44.

13. Reddy A, Zhang J, Davis NS, et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. Cell 2017;171:481-94.e15.

14. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med 2018;378:1396-407.

15. Hoppe MM, Jaynes P, Wardyn JD, et al. Quantitative imaging of RAD51 expression as a marker of platinum resistance in ovarian cancer. EMBO Mol Med 2021;13:e13366.

16. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 2002;30:207-10.

17. Dominguez-Sola D, Victora GD, Ying CY, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol 2012;13:1083-91.

18. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 2004;103:275-82.

19. Meriranta L, Pasanen A, Alkodsi A, Haukka J, Karjalainen-Lindsberg M-L, Leppä S. Molecular background delineates outcome of double protein expressor diffuse large B-cell lymphoma. Blood Adv 2020;4:3742-53.

20. Gavagnin E, Owen JP, Yates CA. Pair correlation functions for identifying spatial correlation in discrete domains. Phys Rev E 2018;97:062104.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

21. Li L, Li Y, Que X, et al. Prognostic significances of overexpression MYC and/or BCL2 in R-CHOP-treated diffuse large B-cell lymphoma: A Systematic review and meta-analysis. Sci Rep 2018;8:6267.

22. Ennishi D, Mottok A, Ben-Neriah S, et al. Genetic profiling of MYC and BCL2 in diffuse large B-cell lymphoma determines cell-of-origin–specific clinical impact. Blood 2017;129:2760-70.

23. Zhang L, Xia B, Guo S, et al. Coexpression of MYC and BCL-2 Predicts Inferior Survival in Pgi-DLBCL Treated with CHOP-like Regimen/Rituximab. Blood 2014;124:5404-.

24. Vitolo U, Trněný M, Belada D, et al. Obinutuzumab or Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Previously Untreated Diffuse Large B-Cell Lymphoma. J Clin Oncol 2017;35:3529-37.

25. Wright GW, Huang DW, Phelan JD, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. Cancer Cell 2020;37:551-68.e14.

26. Caeser R, Di Re M, Krupka JA, et al. Genetic modification of primary human B cells to model high-grade lymphoma. Nature Communications 2019;10:4543.

27. Cattoretti G, Chang CC, Cechova K, et al. BCL-6 protein is expressed in germinal-center B cells. Blood 1995;86:45-53.

28. Hirata Y, Ogasawara N, Sasaki M, et al. BCL6 degradation caused by the interaction with the C-terminus of pro-HB-EGF induces cyclin D2 expression in gastric cancers. British Journal of Cancer 2009;100:1320-9.

29. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP, Staudt LM. BCL-6 Represses Genes that Function in Lymphocyte Differentiation, Inflammation, and Cell Cycle Control. Immunity 2000;13:199-212.

30. Hans CP, Weisenburger DD, Greiner TC, et al. Expression of PKC-beta or cyclin D2 predicts for inferior survival in diffuse large B-cell lymphoma. Modern Pathology 2005;18:1377-84.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

31. Ando K, Ajchenbaum-Cymbalista F, Griffin JD. Regulation of G1/S transition by cyclins D2 and D3 in hematopoietic cells. Proceedings of the National Academy of Sciences 1993;90:9571-5.

32. Sasaki Y, Jensen CT, Karlsson S, Jacobsen SEW. Enforced expression of cyclin D2 enhances the proliferative potential of myeloid progenitors, accelerates in vivo myeloid reconstitution, and promotes rescue of mice from lethal myeloablation. Blood 2004;104:986-92.

33. Pei Y, Singh RK, Shukla SK, Lang F, Zhang S, Robertson ES. Epstein-Barr Virus Nuclear Antigen 3C Facilitates Cell Proliferation by Regulating Cyclin D2. J Virol 2018;92:e00663-18.

34. Herrera AF, Mei M, Low L, et al. Relapsed or Refractory Double-Expressor and Double-Hit Lymphomas Have Inferior Progression-Free Survival After Autologous Stem-Cell Transplantation. J Clin Oncol 2017;35:24-31.

35. Green TM, Young KH, Visco C, et al. Immunohistochemical Double-Hit Score Is a Strong Predictor of Outcome in Patients With Diffuse Large B-Cell Lymphoma Treated With Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone. J Clin Oncol 2012;30:3460-7.

36. Wang Y, Feng W, Liu P. Genomic pattern of intratumor heterogeneity predicts the risk of progression in early stage diffuse large B-cell lymphoma. Carcinogenesis 2019;40:1427-34.

37. Ye X, Wang L, Nie M, et al. A single-cell atlas of diffuse large B cell lymphoma. Cell Reports 2022;39.

38. Iqbal J, Greiner TC, Patel K, et al. Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. Leukemia 2007;21:2332-43.

39. Xu-Monette ZY, Wei L, Fang X, et al. Genetic Subtyping and Phenotypic Characterization of the Immune Microenvironment and MYC/BCL2 Double Expression

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Reveal Heterogeneity in Diffuse Large B-cell Lymphoma. Clinical Cancer Research 2022;28:972-83.

40. Tilly H, Morschhauser F, Sehn LH, et al. Polatuzumab Vedotin in Previously Untreated Diffuse Large B-Cell Lymphoma. New England Journal of Medicine 2021;386:351-63.

41. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the Evolution of Non–Small-Cell Lung Cancer. New England Journal of Medicine 2017;376:2109-21.

42. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. New England Journal of Medicine 2012;366:883-92.

43. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 2017;168:613-28.

44. Aparicio S, Caldas C. The Implications of Clonal Genome Evolution for Cancer Medicine. New England Journal of Medicine 2013;368:842-51.

45. Xu X, Hou Y, Yin X, et al. Single-Cell Exome Sequencing Reveals Single-Nucleotide Mutation Characteristics of a Kidney Tumor. Cell 2012;148:886-95.

46. Lambrechts D, Wauters E, Boeckx B, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. Nature Medicine 2018;24:1277-89.

47. Azizi E, Carr AJ, Plitas G, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell 2018;174:1293-308.e36.

48. Dentro SC, Leshchiner I, Haase K, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. Cell 2021;184:2239-54.e39.

49. Iacobuzio-Donahue CA, Litchfield K, Swanton C. Intratumor heterogeneity reflects clinical disease course. Nature Cancer 2020;1:3-6.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figures

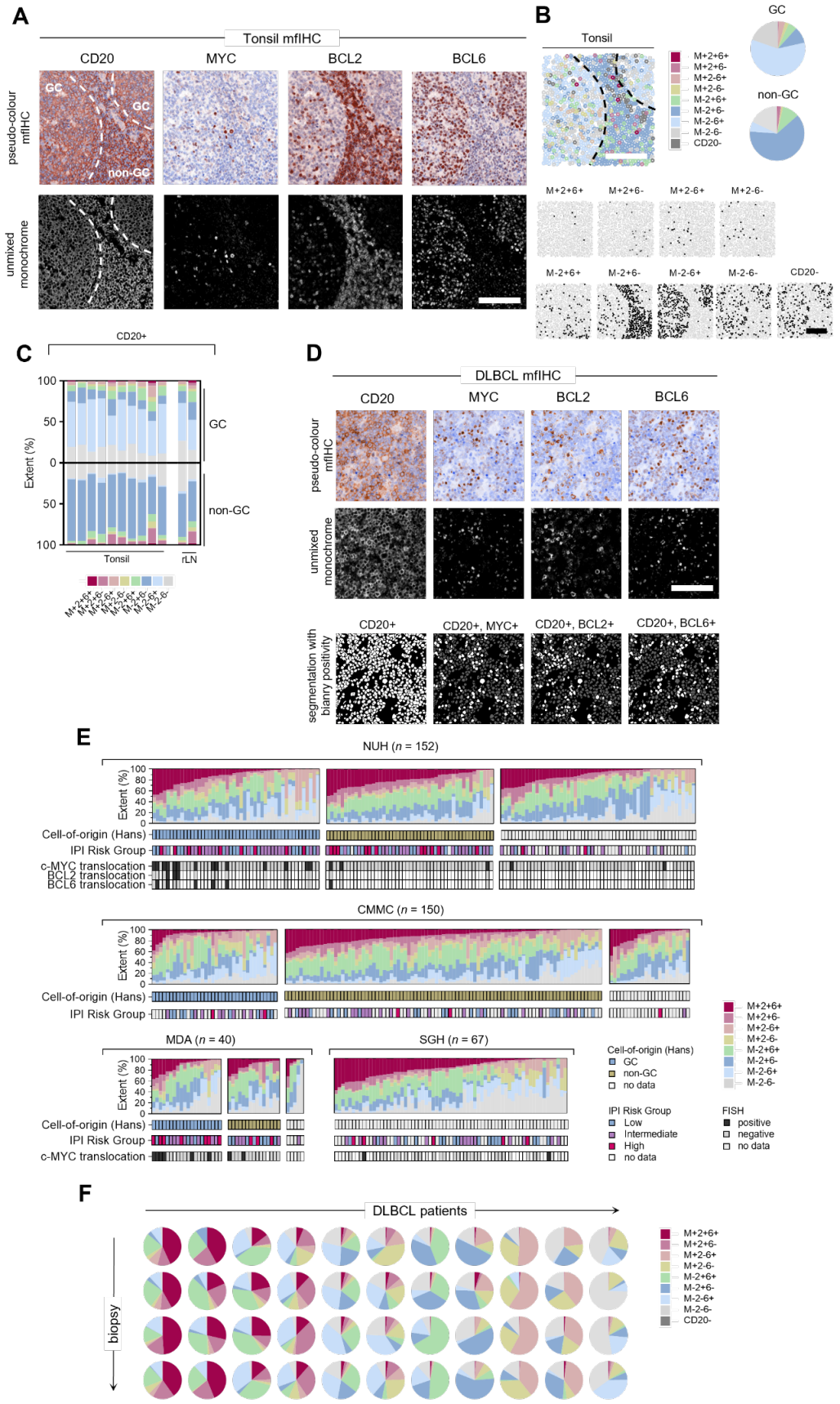


Figure 1

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure1. Quantitative single-cell analysis of MYC, BCL2 and BCL6 protein expression in B-cells in normal tissues and diffuse large B-cell lymphoma. A, Spectrally unmixed multiplexed fluorescent images for CD20, MYC, BCL2 and BCL6 and nuclear counterstaining in tonsil tissue. The germinal center (GC) and non-germinal center (non-GC) zones are indicated. Scale bar is 100µm. Marker-positivity is indicated. B, Spatial map of MYC/BCL2/BCL6 sub-populations, i.e. possible permutations of MYC/BCL2/BCL6-positivity and -negativity within the CD20-positive cell population in a tonsil image. Scale bars are 100µm. C, Quantitation of sub-population extent within CD20-positive cells in normal samples resolved spatially between the GC and non-GC zones. D, Example of MYC/BCL2/BCL6/CD20 multiplexed fluorescent immunohistochemistry (mfIHC) staining in diffuse large B-cell lymphoma (DLBCL). Spectrally unmixed false-coloured and monochrome images for CD20, MYC, BCL2 and BCL6 are shown (top and middle panels). Cell-segmentation and single-marker positivity masks are shown within the CD20-positive cell population (bottom panel). Scale bar is 100µm. E, Summary of percentage extent of sub-populations across patients from National University Hospital (NUH), Chi-Mei Medical Center (CMMC), MD Anderson (MDA) and Singapore General Hospital (SGH). Relevant clinicopathological features are indicated, see also Supplementary figure2. Patients are grouped according to Cell-of-origin based on the Hans algorithm and sorted arbitrarily according to extent of the triple-positive M+2+6+ sub-population extent. IPI Risk Group - International Prognostic Index Risk Group, GC - germinal center B-cell like lymphoma. FISH - fluorescence in situ hybridization. F, Intra-patient spatial stability of sub-populations - proportion of sub-populations measured across four spatially distinct biopsies from the same patient (rows). Biopsy comparison overview is shown across eleven example DLBCL patients (columns).

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

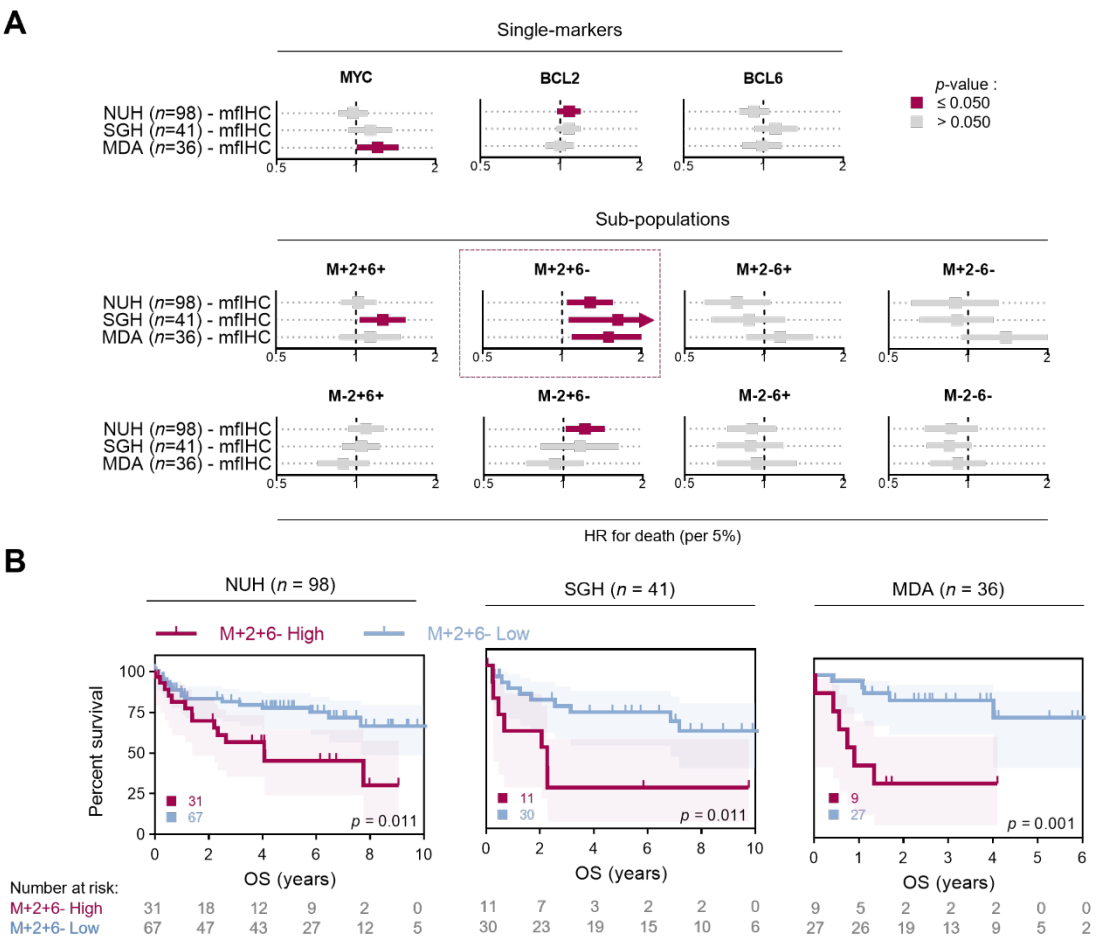
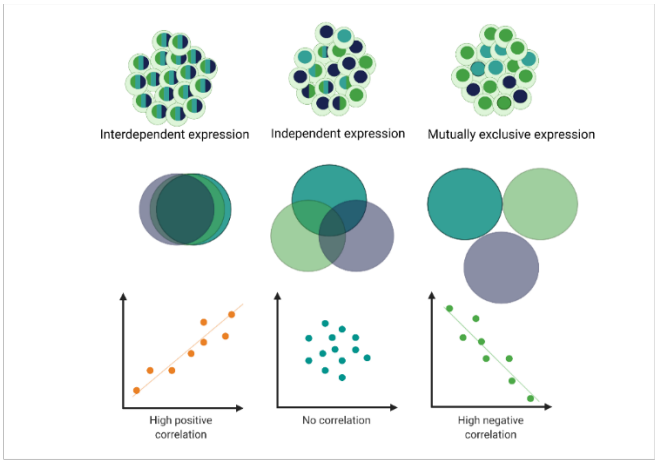


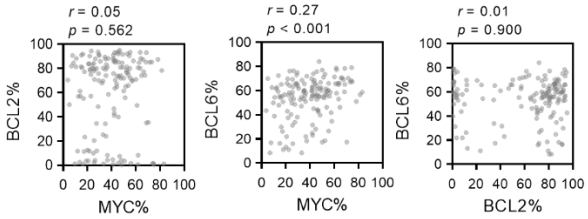
Figure2. Prognostic significance of sub-populations after R-CHOP therapy. A, Univariate Cox proportional hazards model analysis for MYC, BCL2 and BCL6 single-marker and sub-populations percentage extent in the NUH, SGH and MDA cohorts. Percentage-extent was used as a continuous variable in the model at 5% increments (see Methods). OS – overall survival. Hazard ratio (HR) with 95% confidence interval per 5%-positivity increment is shown. B, Kaplan-Meier overall survival analysis of the NUH, SGH and MDA cohorts stratified into M+2+6- High and Low groups. Log- rank test, shading denotes 95% confidence interval. An optimal dichotomization cut-off was used for stratification; total patient numbers in each group are shown.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

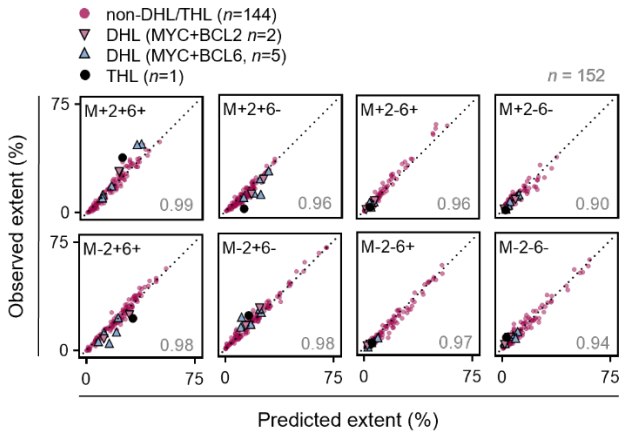
A



B



C



D

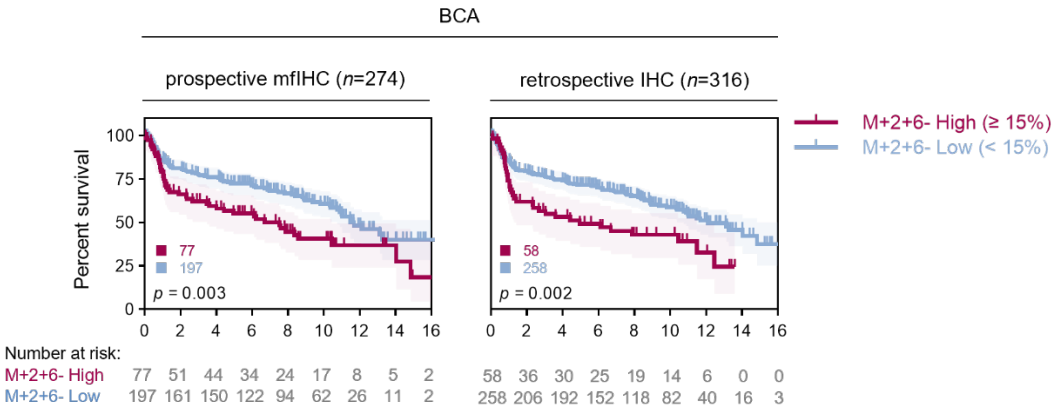


Figure 3

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure3. MYC, BCL2 and BCL6 protein expression is independently distributed in B-cells and therefore can be inferred form quantitative data. A, Schematic of possible relationships between expression of three oncogenes in a population. The distribution of these oncogenes can either reflect interdependent expression, independent/stochastic expression, or mutually exclusive expression. These relationships result in the percentage extent of oncogenes in the tumor being either strongly positive correlated, not correlated or strongly negatively correlated, respectively. B, Correlation of MYC, BCL2 and BCL6 percentage extent in the NUH cohort. Spearman correlation. C, Good correlation between predicted sub-population percentage extent based on single-marker positivity and observed percentage extent in the NUH cohort. Cases of double-hit lymphoma (DHL, MYC+BCL2+ translocations or MYC+BCL6+ translocations) or triple-hit lymphoma (THL) are highlighted. Spearman rho for each correlation is shown. Similar graphs for the CMMC, SGH and MDA cohorts can be found in Supplementary figure5. D, Kaplan-Meier overall survival analysis of the BCA cohort stratified into M+2+6- High and Low groups across an absolute value of 15 percentage extent positivity in prospective mfIHC validation (*left*) and retrospective IHC images (*right*). Log-rank test, shading denotes 95% confidence interval.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

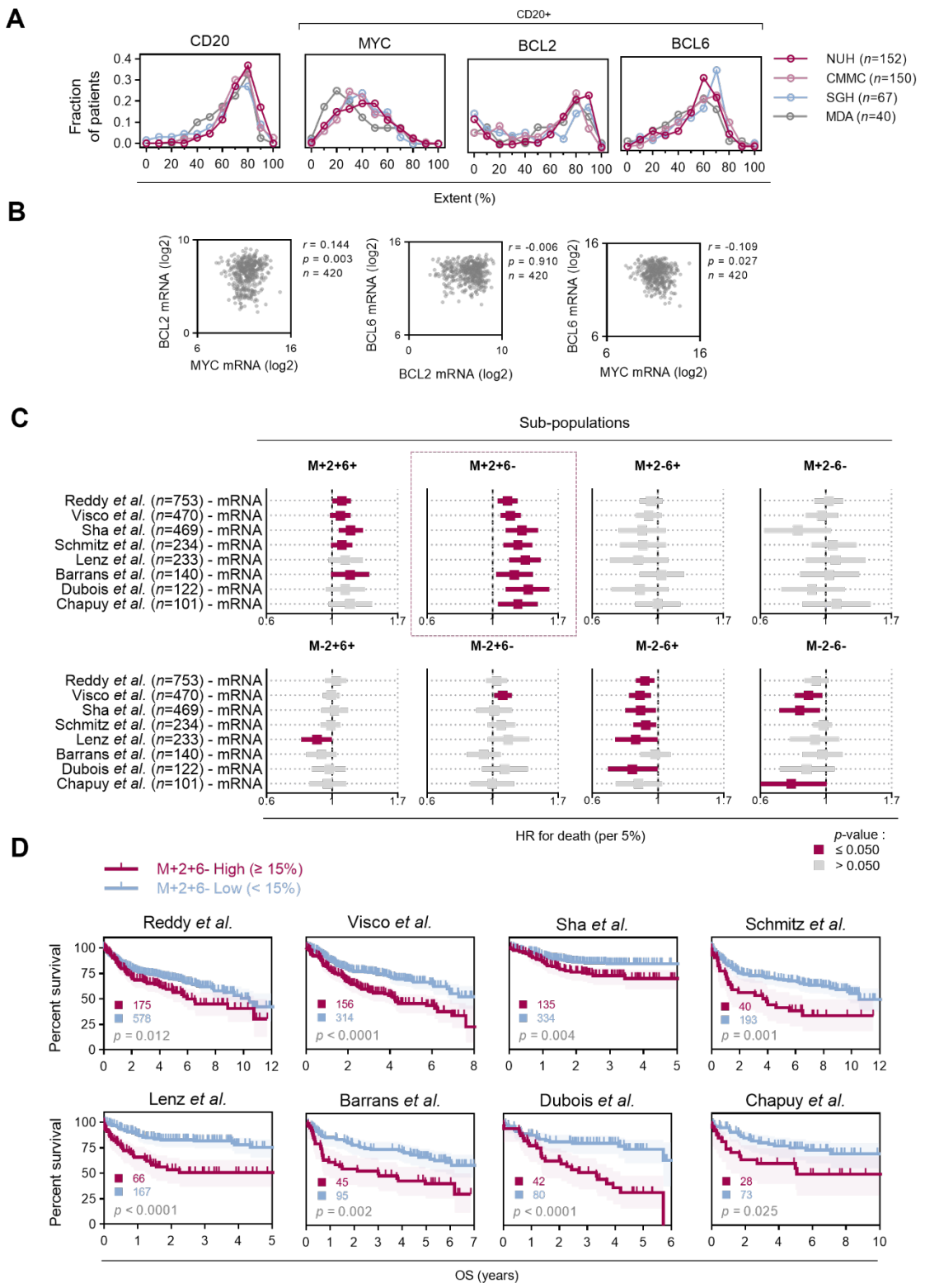


Figure 4

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure4. Validation of prognostic significance of inferred sub-populations extent after R-CHOP therapy in GEP datasets. A, Distribution of single-marker positivity in four DLBCL cohorts as assessed by mfIHC (See also Supplementary figure8A). B, Correlation of MYC, BCL2 and BCL6 mRNA expression in the Lenz *et al.* cohort. Spearman correlation. C, Impact of inferred MYC, BCL2 and BCL6 sub-populations in gene-expression profiling (GEP) datasets. Univariate Cox proportional hazards model analysis; percentage-extent was used as a continuous variable in the model at 5% increments. D, Kaplan-Meier overall survival analysis of GEP cohorts stratified uniformly across absolute 15% M+2+6- subpopulation extent into -High and -Low groups. Log- rank test, shading denotes 95% confidence interval. Total patient numbers in each group are shown.

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

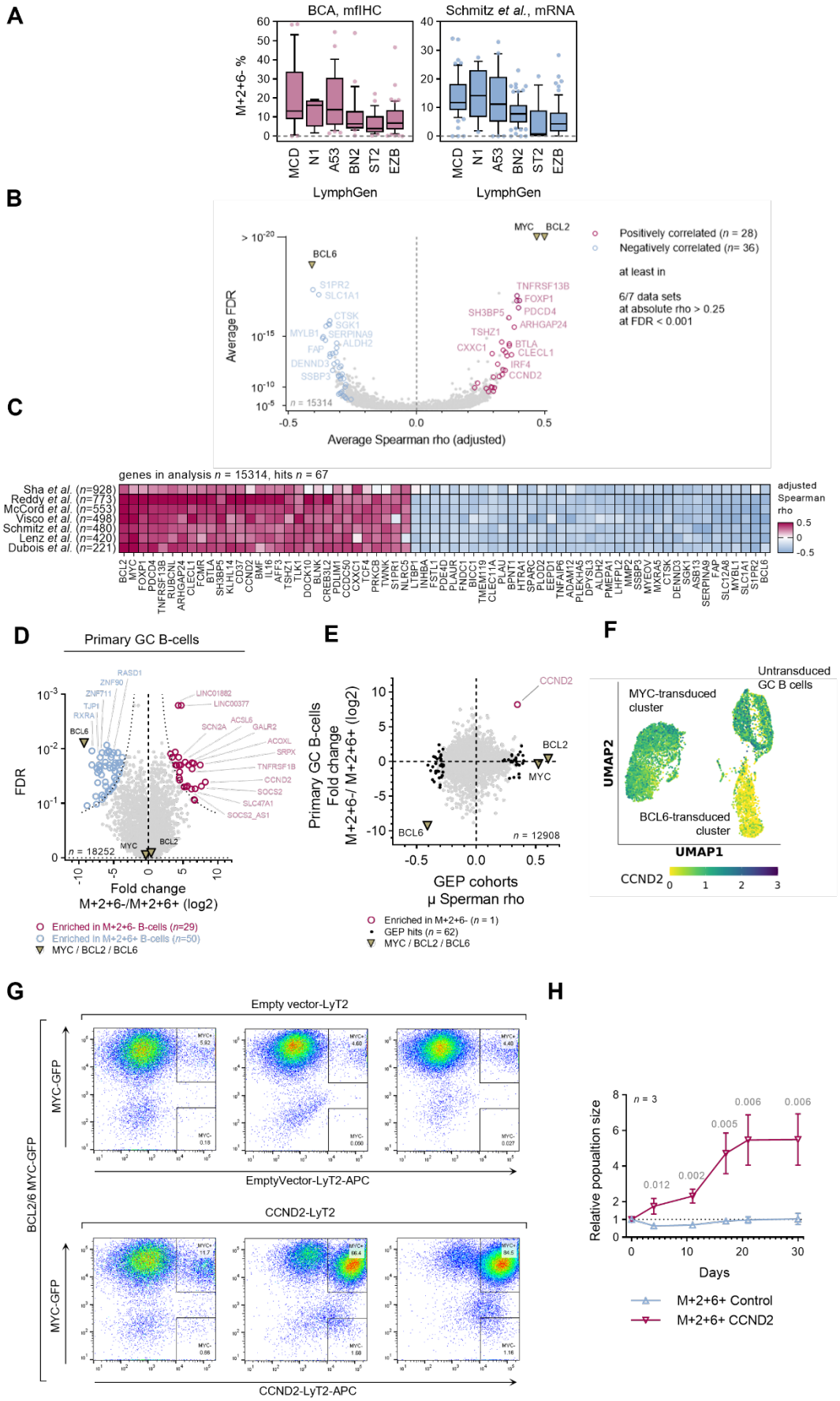


Figure 5

medRxiv preprint doi: https://doi.org/10.1101/2020.10.20.20216101; this version posted July 25, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure5. Molecular features of M+2+6- High cases and derivation and validation of cyclin D2 oncogene A, Correlation of the M+2+6- percentage extent evaluated by mfIHC and mRNA inference with genetic subtypes (LymphGen classification). B, Volcano plot of direct correlation of gene mRNA expression and inferred M+2+6- percentage extent across seven GEP cohorts. Genes highly correlated with M+2+6- percentage extent in at least 6 out of 7 GEP cohorts at absolute Spearman rho ≥0.25 and FDR ≤0.001 are show. C, Sixty-seven hits from panel (B) criteria are shown, full analysis is available in Supplementary table8. D, Differential gene-expression between primary germinal center (GC) B-cells overexpressing M+2+ and M+2+6+. Full analysis available in Supplementary table9. E, Genes highly enriched in M+2+6- cells - correlation of results from panel (B) and (D). F, Single-cell RNA-seq of GC primary B-cells transduced either with MYC or BCL6. Expression of CCND2 is indicated in colour. G. Representative FACS plots documenting to the expansion over time of the CCND2 positive GC B-cell population in CCND2 overexpressing GC B cells (CCND2-LyT2) and non-CCND2 overexpressing GC B-cells (Empty vector LyT2). All GC B cells co-overexpress BCL2, BCL6, and MYC-GFP. H, Proliferation analysis of primary GC B-cells overexpressing M+2+6+ and CCND2.