UNIVERSITY of York

This is a repository copy of *Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/183518/

Version: Published Version

Article:
Gong, Chun, Krupka, Joanna A, Gao, Jie et al. (34 more authors) (2021) Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis. Molecular Cell. pp. 4059-4075. ISSN: 1097-2765

https://doi.org/10.1016/j.molcel.2021.07.041

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

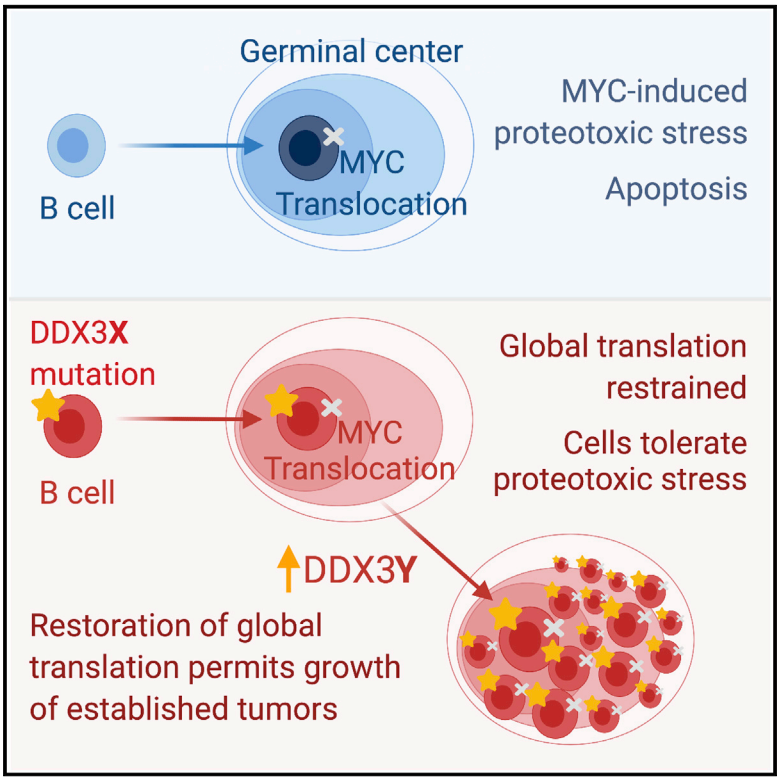
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Molecular Cell

Article

Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis

Graphical abstract



Authors

Chun Gong, Joanna A. Krupka, Jie Gao, ..., Ming-Qing Du, Shamith A. Samarajiwa, Daniel J. Hodson

Correspondence

djh1002@cam.ac.uk

In brief

Gong et al. show that during the early stages of lymphoma development, loss-of-function mutations in the RNA helicase DDX3X allow human B cells to tolerate the forced expression of MYC. In contrast, established tumors restore DDX3 helicase activity by ectopic expression of the Y-chromosome-encoded homolog DDX3Y.

Highlights

- Loss-of-function mutations of *DDX3X* are frequent in MYC-driven B cell lymphomas
- DDX3X promotes translation of mRNAs encoding the core protein synthesis machinery
- Loss of DDX3X buffers MYC-driven global protein synthesis and proteotoxic stress
- DDX3X loss is later rescued by ectopic expression of Y-chromosome-encoded DDX3Y

Gong et al., 2021, Molecular Cell *81*, 4059–4075
October 7, 2021 © 2021 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2021.07.041

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis

Chun Gong,[1,2,23] Joanna A. Krupka,[1,2,3,23] Jie Gao,[1,2] Nicholas F. Grigoropoulos,[4] George Giotopoulos,[1,2] Ryan Asby,[1,2] Michael Screen,[5] Zelvera Usheva,[1,2] Francesco Cucco,[6] Sharon Barrans,[7] Daniel Painter,[8] Nurmahirah Binte Mohammed Zaini,[4] Björn Haupl,[9,10,11] Susanne Bornelöv,[1] Igor Ruiz De Los Mozos,[12,13] Wei Meng,[14] Peixun Zhou,[15,17] Alex E. Blain,[15,16,17] Sorcha Forde,[6] Jamie Matthews,[6] Michelle Guet Khim Tan,[19]

(Author list continued on next page)

[1]Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Puddicombe Way, Cambridge CB2 0AW, UK
[2]Department of Haematology, University of Cambridge, Cambridge CB2 0AW, UK
[3]MRC Cancer Unit, University of Cambridge, Hutchison/MRC Research Centre, Cambridge CB2 0XZ, UK
[4]Department of Haematology, Singapore General Hospital, Singapore, Singapore
[5]Immunology Programme, The Babraham Institute, Cambridge CB22 3AT, UK
[6]Division of Cellular and Molecular Pathology, Department of Pathology, University of Cambridge, Cambridge CB20QQ, UK
[7]Haematological Malignancy Diagnostic Service, St. James's Institute of Oncology, Leeds LS9 7TF, UK
[8]Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York, York YO10 5DD, UK
[9]Department of Medicine II, Hematology/Oncology, Goethe University, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany
[10]German Cancer Research Center and German Cancer Consortium, Heidelberg, Germany

(Affiliations continued on next page)

SUMMARY

DDX3X is a ubiquitously expressed RNA helicase involved in multiple stages of RNA biogenesis. *DDX3X* is frequently mutated in Burkitt lymphoma, but the functional basis for this is unknown. Here, we show that loss-of-function *DDX3X* mutations are also enriched in *MYC*-translocated diffuse large B cell lymphoma and reveal functional cooperation between mutant DDX3X and MYC. DDX3X promotes the translation of mRNA encoding components of the core translational machinery, thereby driving global protein synthesis. Loss-of-function DDX3X mutations moderate MYC-driven global protein synthesis, thereby buffering MYC-induced proteotoxic stress during early lymphomagenesis. Established lymphoma cells restore full protein synthetic capacity by aberrant expression of DDX3Y, a Y chromosome homolog, the expression of which is normally restricted to the testis. These findings show that DDX3X loss of function can buffer MYC-driven proteotoxic stress and highlight the capacity of male B cell lymphomas to then compensate for this loss by ectopic DDX3Y expression.

INTRODUCTION

Burkitt lymphoma (BL) is an aggressive non-Hodgkin lymphoma, with a 3:1 male:female incidence ratio (Morton et al., 2006; Smith et al., 2015). BL arises from the germinal center (GC) stage of B cell development, where B cells undergo somatic hypermutation of the immunoglobulin genes associated with cycles of intense proliferation and selection (Basso and Dalla-Favera, 2015). The oncoprotein MYC is required for this process, but its expression is transient and limited to a minority of GC B cells undergoing positive selection (Calado et al., 2012; Dominguez-Sola et al., 2012). The GC is also associated with class switch recombination (CSR), a process that involves double-stranded breakage and rejoining of the immunoglobulin genes. An unwanted by-product of CSR is the risk of translocation of the *MYC* oncogene into the immunoglobulin loci leading to sustained and high-level expression of MYC. Translocation between *MYC* and one of the immunoglobulin genes is observed in >95% of BL (Swerdlow et al., 2016). *MYC* translocation is also seen in 10% of diffuse large B cell lymphoma (DLBCL), where it confers a poor prognosis (Savage et al., 2009). Many cases of MYC-translocated DLBCL fall into the molecular high-grade (MHG) transcriptional subtype (Painter et al., 2019; Sha et al., 2019). Those cases of MHG without detectable *MYC* rearrangement frequently possess cryptic alterations leading to the deregulation of MYC expression (Hilton et al., 2019). Mouse models reveal that high-level MYC expression alone is insufficient to drive lymphomagenesis, and further cooperating genetic events are required (Sander et al., 2012).

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

CellPress
OPEN ACCESS

Molecular Cell
Article

G.A. Amos Burke,[20] Siu Kwan Sze,[14] Philip Beer,[21] Cathy Burton,[7] Peter Campbell,[21] Vikki Rand,[15,17] Suzanne D. Turner,[6,18] Jernej Ule,[12,13] Eve Roman,[8] Reuben Tooze,[7,22] Thomas Oellerich,[9,10,11] Brian J. Huntly,[1,2] Martin Turner,[5] Ming-Qing Du,[6] Shamith A. Samarajiwa,[3,24] and Daniel J. Hodson[1,2,24,25,*]

[11]Frankfurt Cancer Institute, Goethe University Frankfurt, 60596 Frankfurt, Germany
[12]The Francis Crick Institute, London NW1 1AT, UK
[13]Department for Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London WC1N 3BG, UK
[14]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore, Singapore
[15]National Horizons Centre, Teesside University, 38 John Dixon Lane, Darlington DL1 1HG, UK
[16]Wolfson Childhood Cancer Research Centre, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, UK
[17]School of Health & Life Sciences, Teesside University, Middlesbrough TS1 3BA, UK
[18]CEITEC, Masaryk University, Brno, Czech Republic
[19]Department of Clinical Translational Research, Singapore General Hospital, Outram Road, Singapore 169856, Singapore
[20]Department of Paediatric Oncology, Addenbrooke's Hospital, Cambridge, UK
[21]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
[22]Section of Experimental Haematology, Leeds Institute of Molecular Medicine, University of Leeds, Leeds LS2 9JT, UK
[23]These authors contributed equally
[24]Senior author
[25]Lead contact
*Correspondence: djh1002@cam.ac.uk
https://doi.org/10.1016/j.molcel.2021.07.041

DDX3X is a ubiquitously expressed ATP-dependent RNA helicase. *DDX3X* is located on the X chromosome and escapes X chromosome inactivation (Berletch et al., 2011). The Y chromosome homolog DDX3Y shares 92% amino acid conservation with DDX3X. *DDX3Y* is widely transcribed; however, in normal adult tissues, protein expression is restricted to the testis (Ditton et al., 2004; Foresta et al., 2000; Rauschendorf et al., 2011). DDX3X is reported to function at multiple stages of RNA biology, including transcription, nuclear export, stress granule dynamics, and mRNA translation (Sharma and Jankowsky, 2014; Soto-Rifo and Ohlmann, 2013). It has been proposed to act as both a translational activator and a repressor (Shih et al., 2008). The mutation of *DDX3X* has been reported in BL (Bouska et al., 2017; Grande et al., 2019; López et al., 2019; Richter et al., 2012; Schmitz et al., 2012), chronic lymphocytic leukemia (CLL) (Ojha et al., 2015; Takahashi et al., 2018), natural killer (NK)-T cell lymphoma (Jiang et al., 2015), and medulloblastoma (Jones et al., 2012; Pugh et al., 2012; Robinson et al., 2012). However, its role in malignancy remains controversial, and it has been classified as both a tumor suppressor and an oncogene (He et al., 2018; Soto-Rifo et al., 2012). Here, we investigate the role of *DDX3X* mutations in BL and MYC-driven DLBCL. We reveal that somatic loss-of-function mutations of *DDX3X* facilitate the early stages of MYC-driven lymphomagenesis by buffering the effects of MYC on translation and global protein synthesis. In contrast, established lymphoma cells restore full translation by the ectopic expression of DDX3Y. Thus, the sequential loss of DDX3X, followed by the gain of DDX3Y, allows cells to adopt a level of protein synthesis matched to the stage-specific needs of the developing tumor.

RESULTS

DDX3X is preferentially mutated in MYC-driven lymphomas

We applied targeted sequencing to 39 cases of BL. The most frequently mutated genes were *MYC*, *ID3*, *TP53*, *CCND3*, *DDX3X*, *ARID1A*, *FOXO1*, and *SMARCA4* (Figure 1A; Table S1). While *MYC*, *ID3*, *TP53*, and *CCND3* have been intensively investigated, little is known about the role of *DDX3X* mutation in lymphoma. Mutation of *DDX3X* was found in 12 of 39 cases (30.8%) of BL (11 males and 1 female). In contrast, when the same targeted panel and variant calling strategy was applied to 928 cases of DLBCL, we detected *DDX3X* mutation in only 5.2% (Lacy et al., 2020), a figure similar to that of other recent DLBCL sequencing studies (Figure 1B) (Chapuy et al., 2018; Reddy et al., 2017; Schmitz et al., 2018). DDX3X mutations were predominantly clonal (Figure S1A).

DDX3X mutation was recently noted to be enriched among the MHG transcriptional subtype of DLBCL (Sha et al., 2019). To extend this observation, we analyzed mutation data from a previously characterized UK cohort of DLBCL enriched for *MYC*-rearranged cases (Cucco et al., 2020). *DDX3X* mutation was found in 12.8% of cases with *MYC* rearrangement, but only 4.2% of cases lacked *MYC* rearrangement ($p = 0.001$; chi-square test) (Figure 1C). Among 558 cases with gene expression, *DDX3X* mutation was enriched in MHG (16.7%) compared to GCB (4.3%) or ABC (2.2%) DLBCL ($p = 0.001$; chi-square test) (Figure 1D). We then reanalyzed publicly available RNA sequencing (RNA-seq) data from 553 DLBCL patients enrolled in the GOYA trial (McCord et al., 2019) (Figures 1E and S1B–S1D; Table S1). Once again, we found *DDX3X* mutation to be enriched in the MHG DLBCL subtype (19.0%), compared to GCB (8.1%) and ABC (3.0%) ($p < 10^{-5}$; chi-square test). Meta-analysis, including cases sequenced in this and previously published studies of BL and DLBCL with available sex data, confirmed that DDX3X mutation was more frequent in male patients (relative risk [RR] = 1.23, 95% confidence interval [CI] 1.15–1.33, $p = 0.0002$; random effects model) (Figures S1E and S1F).

Examining the distribution of mutations from this and previous studies (Figure 1F; Table S1) revealed a predominance of mutations within the C-terminal helicase domain. Some of these mutations, including R475 and R534, are shared with medulloblastoma and known to abolish helicase activity (Epling et al., 2015; Floor et al., 2016; Lennox et al., 2020). However, in contrast to

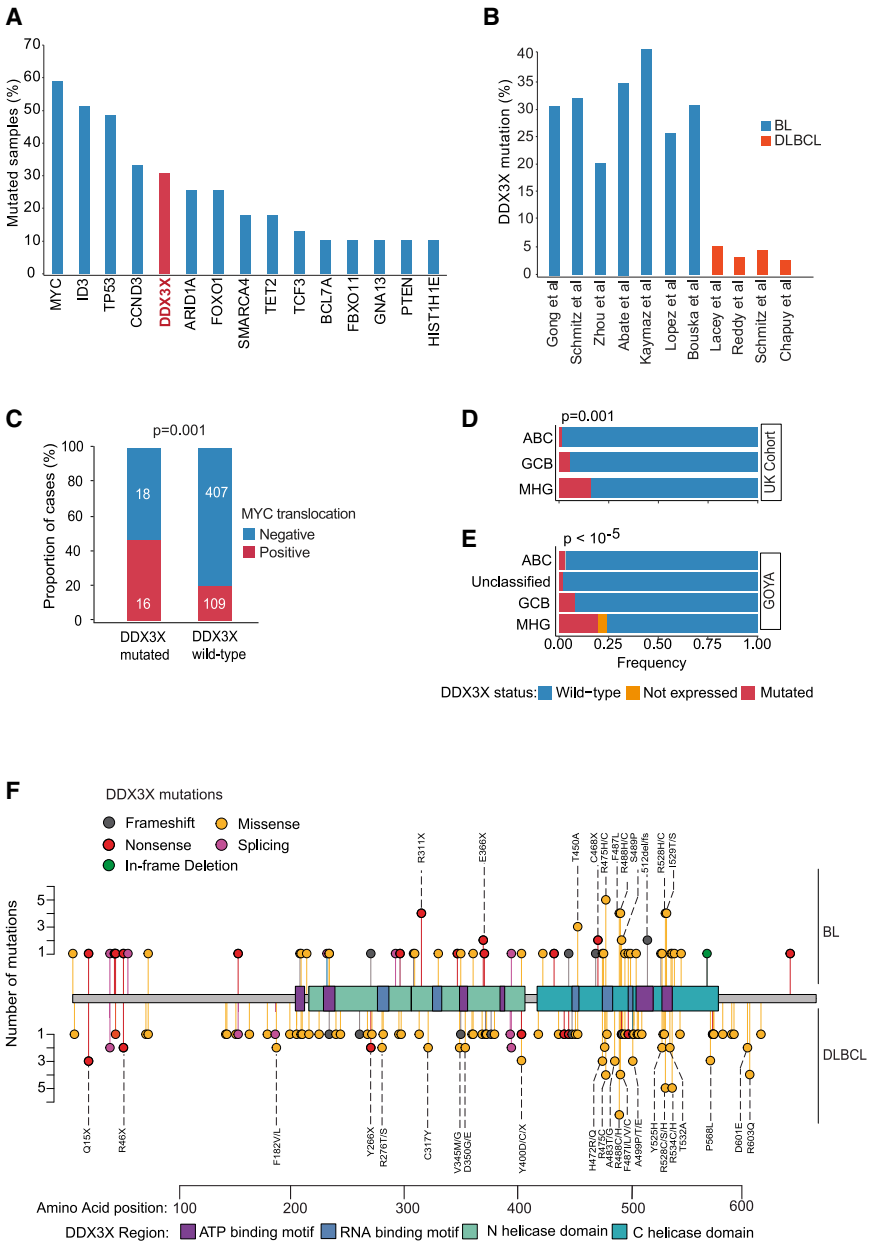


Figure 1. DDX3X is preferentially mutated in MYC-driven lymphomas

(A) Mutation frequency (%) for the indicated genes detected using a 293-gene panel applied to 39 cases of Burkitt lymphoma (BL).

(B) Frequency of DDX3X mutation across published sequencing studies of BL and DLBCL.

(C) Proportion of cases with MYC rearrangement detected by fluorescence *in situ* hybridization (FISH) from a cohort of 550 cases of DLBCL stratified by DDX3X mutation status.

(D) Frequency of DDX3X mutation in 558 cases of DLBCL stratified by transcriptional subtype.

(E) Frequency of DDX3X mutation determined by analysis of RNA-seq in 553 cases of DLBCL enrolled in the GOYA trial, stratified by transcriptional subtype.

(F) Distribution of DDX3X mutations identified in this and published studies of BL and DLBCL.

The p values were calculated by the chi-square test.

See also Figure S1.

medulloblastoma (Jones et al., 2012; Pugh et al., 2012; Robinson et al., 2012), nonsense and frameshift mutations were frequent in lymphoma, suggesting that the functional impact of *DDX3X* mutation may differ between cancer types (Figures 1F and S1G).

Analyses across multiple patient cohorts reveal that *DDX3X* is preferentially mutated in MYC-driven B cell lymphomas. This suggests that the loss of DDX3X helicase activity may cooperate with MYC during lymphomagenesis.

DDX3X mutation cooperates with MYC in *ex vivo*-cultured human GC B cells

Forced expression of wild-type (WT) or mutant DDX3X or DDX3Y in lymphoma cell lines induced cell death (Figures S2A–S2G). In contrast, HEK293 cells tolerated DDX3X overexpression without toxicity (Figure S2H), underscoring the potential for tissue and context-specific effects of DDX3X. Since BL arises from GC B cells, we used a strategy to culture and transduce primary human GC B cells in a coculture system designed to mimic the GC microenvironment (Caeser et al., 2019, 2021). Genetic manipulation of these cells allows us to model the early stages of human GC lymphomagenesis. GC B cells tolerated the forced expression of DDX3X (Figure S2I). Furthermore, a competitive advantage was seen in cells cotransduced with both MYC and the dominant-negative DDX3X mutant (K230E) (Yedavalli et al., 2004) or with the two most common lymphoma helicase mutations, R475C and R528C (Figures 2A and S2J–S2L). This competitive advantage in MYC-transduced cells was not seen when cells were cotransduced with WT DDX3X or DDX3Y (Figure 2A). Furthermore, no competitive advantage was seen in mutant DDX3X-transduced cells when the alternative lymphoma oncogene BCL6 was used in place of MYC (Figure 2A).

GC B cells cultured *ex vivo* remained viable in culture for only 7–10 days. Extended culture can be achieved by the cotransduction of BCL2 and MYC (Caeser et al., 2019). Therefore, we cotransduced cells with MYC, BCL2, and DDX3X (WT or mutant). A competitive advantage was again seen in cells expressing MYC and mutant DDX3X (Figure 2B), but was not observed with WT DDX3X, nor when BCL6 was used in place of MYC (Figure 2B). Overall, these experiments reveal a cooperative effect between MYC expression and loss-of-function, or dominant-negative, DDX3X helicase mutation.

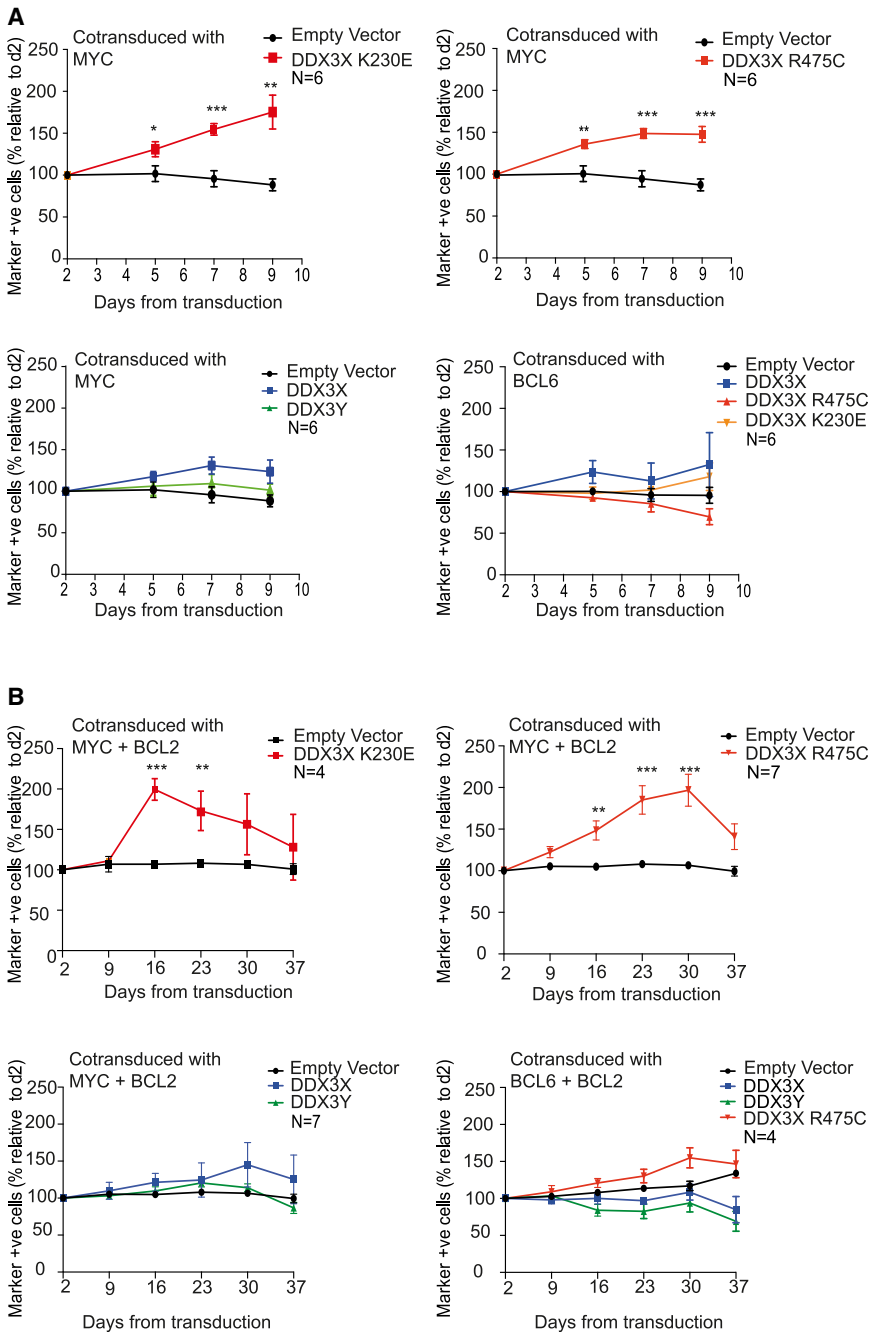


Figure 2. DDX3X mutation cooperates with MYC in *ex vivo* human GC B cells

(A) Human GC B cells were expanded *ex vivo* and cotransduced with MYC plus the indicated DDX3X construct at day 0. The frequency of DDX3X-transduced cells (marker positive) as a proportion of all MYC-transduced cells was monitored by flow cytometry and is shown over time relative to day 2. Where indicated, BCL6 was used in place of MYC.

(B) Human GC B cells were cotransduced with MYC-2A-BCL2 plus the indicated DDX3X construct at day 0. The frequency of DDX3X-transduced cells (marker positive) as a proportion of all MYC-transduced cells is shown over time relative to day 2. Where indicated, BCL6-2A-BCL2 was used in place of MYC-2A-BCL2.

Data show means ± SEMs for replicate cultures from 4–7 human donors, as indicated; *p < 0.01, **p < 0.01, ***p < 0.001, independent t test.

See also Figure S2.

Protein and mRNA interactomes implicate DDX3X in the regulation of mRNA translation in lymphocytes

Previous reports across multiple tissue types have implicated DDX3X in multiple stages of RNA biogenesis. To establish the role of DDX3X in human lymphoid cells, we sought to identify DDX3X-interacting proteins. We performed immunoprecipitation (IP) and mass spectrometry (MS) of endogenous DDX3X from two lymphoma cell lines, U2932 and Mutu, which express WT and R475S helicase mutant DDX3X, respectively (Figures 3A–3C and S3A–S3C; Table S2). We detected strong enrichment for the components of the translation initiation machinery, including eukaryotic initiation factor 4E (eIF4E), eIF4A, eIFG4, and almost all of the subunits of the eIF3 complex (Figures 3A–3C). These interactions were detected in both the WT and the helicase mutant cell lines. CoIP of DDX3X from control and R475C CRISPR-edited U2932 clones and immunoblotting for eIF3a and eIF3b showed no evidence that the helicase mutation altered these protein-protein interactions (Figures S3D and S3E). These findings are consistent with a previously described ability of DDX3X to promote mRNA translation (Lee et al., 2008; Soto-Rifo et al., 2012). Other interacting proteins (including DDX1, ATXN2L, NUFIP2, PDCD4, USP10, UPF1, and EWSR1) are known components of stress granules, cytoplasmic messenger ribonucleoprotein foci associated with stalled translation initiation (Buchan and Parker, 2009; Jain et al., 2016). These findings suggest that in lymphoid cells, DDX3X may influence mRNA translation, either through direct involvement in translation initiation or by remodeling of stress granules.

To establish whether DDX3X acts globally or on a specific subset of mRNAs, we performed individual nucleotide resolution crosslinking IP (iCLIP). Previous studies of DDX3X have examined DDX3X-bound transcripts in HEK cells (Calviello et al., 2021; Oh et al., 2016; Valentin-Vega et al., 2016). To establish the identity of transcripts bound by endogenous DDX3X in lymphoid cells, we performed iCLIP in two lymphoma cell lines U2932 (female, DDX3X-WT) and Mutu (male, DDX3X-R475S), as well as non-malignant human GC B cells from female tonsil tissue (Figures S3F–S3I). Crosslinking sites mapped

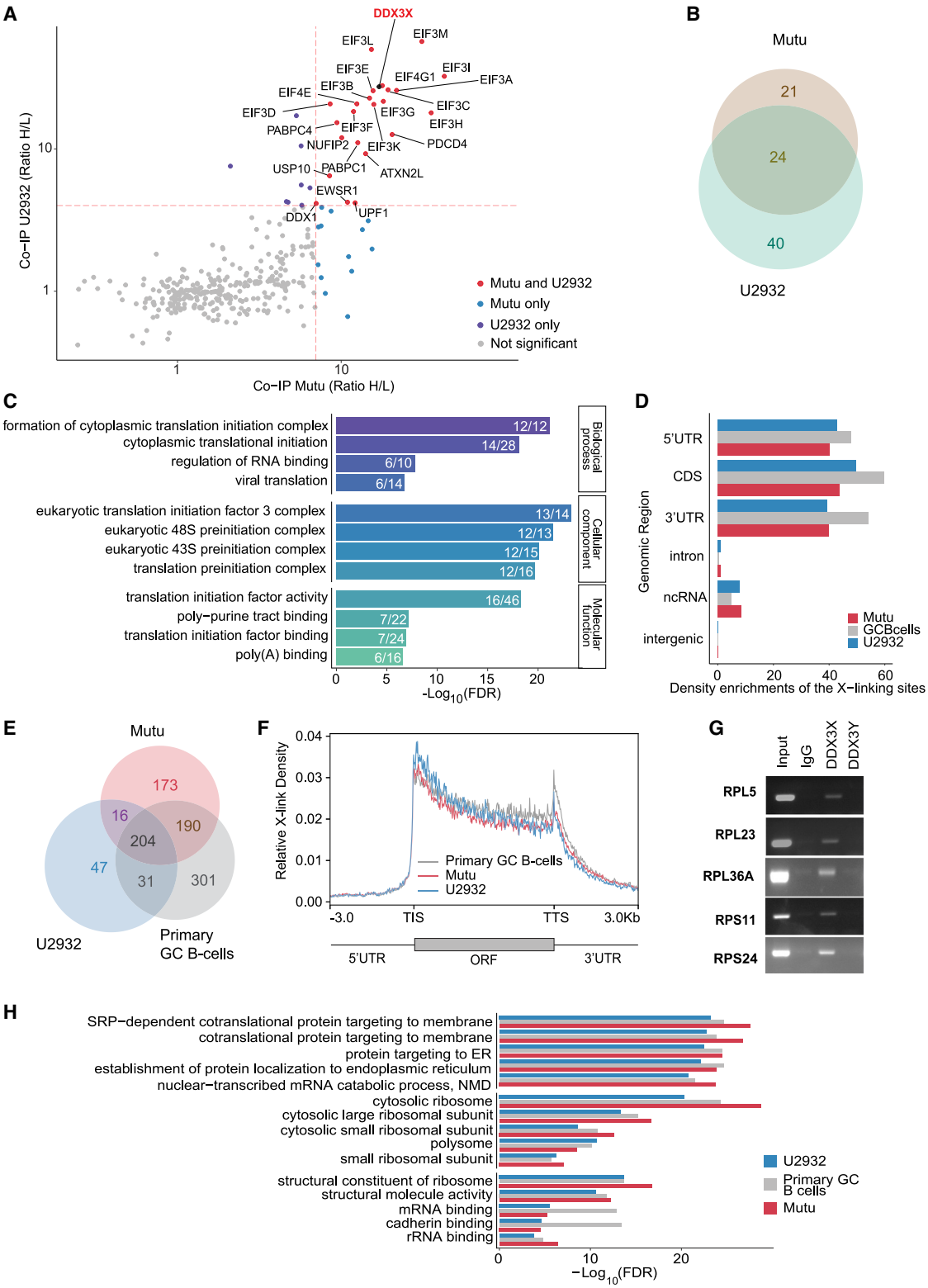


Figure 3. Protein and mRNA interactomes in lymphocytes implicate DDX3X in the regulation of translation

(A) DDX3X-interacting proteins were identified by stable isotope labeling by/with amino acids in cell culture-mass spectrometry (SILAC-MS) following immunoprecipitation of endogenous DDX3X in U2932 and Mutu. Scatterplot shows log2 SILAC ratios of interacting proteins. Proteins significantly enriched in both cell lines are labeled.

(legend continued on next page)

predominantly to mature protein-coding mRNA transcripts (Figure 3D). We found considerable overlap between transcripts bound by DDX3X in all three cell types (Figures 3E and S3H), suggesting that the R475 mutant did not abolish interaction with target mRNA transcripts. Metagene analysis showed the greatest enrichment of iCLIP crosslinks at translation initiation sites (TISs) (Figure 3F). DDX3X bound to a large number of transcripts across a broad range of mRNA expression (Figure S3J). However, strong enrichment was seen for mRNAs encoding components of the core protein synthesis machinery, in particular, the protein subunits of the ribosome (Figures 3G, 3H, and S3K; Table S3). While not excluding involvement in alternative aspects of RNA processing, our findings suggest that DDX3X may influence the translation of a cohort of mRNA transcripts encoding components of the core protein synthesis machinery.

DDX3X regulates ribosome biogenesis and global protein synthesis

To establish how DDX3X affects the translation of its mRNA targets, we performed transcriptome-wide translational profiling (ribo-seq) in cell lines following short hairpin RNA (shRNA) knockdown of DDX3X (Figure 4A). Using 2 separate DDX3X shRNAs, we profiled 8 replicate knockdowns in the WT cell line U2932 and 4 replicate knockdowns in the R475S mutant cell line Mutu. Ribo-seq data quality was confirmed by mRNA ribosome footprint length, evidence of triplet codon periodicity, high consistency between replicate experiments, and the abrupt dropoff of footprints at the stop codon (Figures S4A–S4F). Overall, 90 genes showed decreased translational efficiency (TE) (TE-down) and 70 genes showed increased TE (TE-up) after knockdown of DDX3X in the WT cell line U2932 (Figures 4B and 4C; Table S4). TE-up genes were not enriched among iCLIP targets and were not associated with any significantly enriched Gene Ontology (GO) term. However, TE-down genes were strongly enriched among iCLIP target genes (adjusted $p < 10^{-6}$; Fisher test) (Figures 4D, 4E, S4G, and S4H). Furthermore, TE-down genes were strongly enriched for transcripts encoding components of the core translational machinery (Figures 4C, 4F, and 4G; Table S4). GO analysis of TE-down transcripts revealed terms associated with translation, in particular, protein constituents of the ribosome (Figures 4F and 4G). TE was reduced for almost all of the transcripts encoding protein subunits of the ribosome (Figure 4H), an effect that appeared specific to components of the cytosolic ribosome. In contrast to the depletion of WT DDX3X, the depletion of R475S mutant DDX3X from the cell line Mutu had little effect on mRNA translation, consistent with the expected loss-of-function nature of this mutation.

To validate the conclusion of our ribo-seq experiments, we performed proteomic profiling. This confirmed the reduced abundance of almost all of the ribosome proteins following shRNA depletion of DDX3X in U2932 (Figures S4I and S4J; Table S5). GO analysis of changing proteins was enriched for terms related to translation, protein synthesis, and the ribosome (Figure S4I). In an orthogonal approach, we interrogated the ProteomeHD database, which uses data from 5,288 mass spectrometry runs across multiple tissue types and biological conditions to infer a co-regulation map of the human proteome (Kustatscher et al., 2019). This revealed 81 proteins co-regulated with DDX3X (Figure S4K; Table S5). Of these, 27 were components of the GO term cytosolic ribosome (adjusted p = 7.2E−42). Eight further proteins were components of the core translational machinery. The most enriched GO terms were ribonucleoprotein (RNP) complex, mRNA metabolism, translation initiation, and cytosolic ribosome. The analysis of iCLIP, translational profiling, proteomic, and ProteomeHD results converged onto the regulated translation of mRNAs encoding components of core translation and protein synthesis machinery.

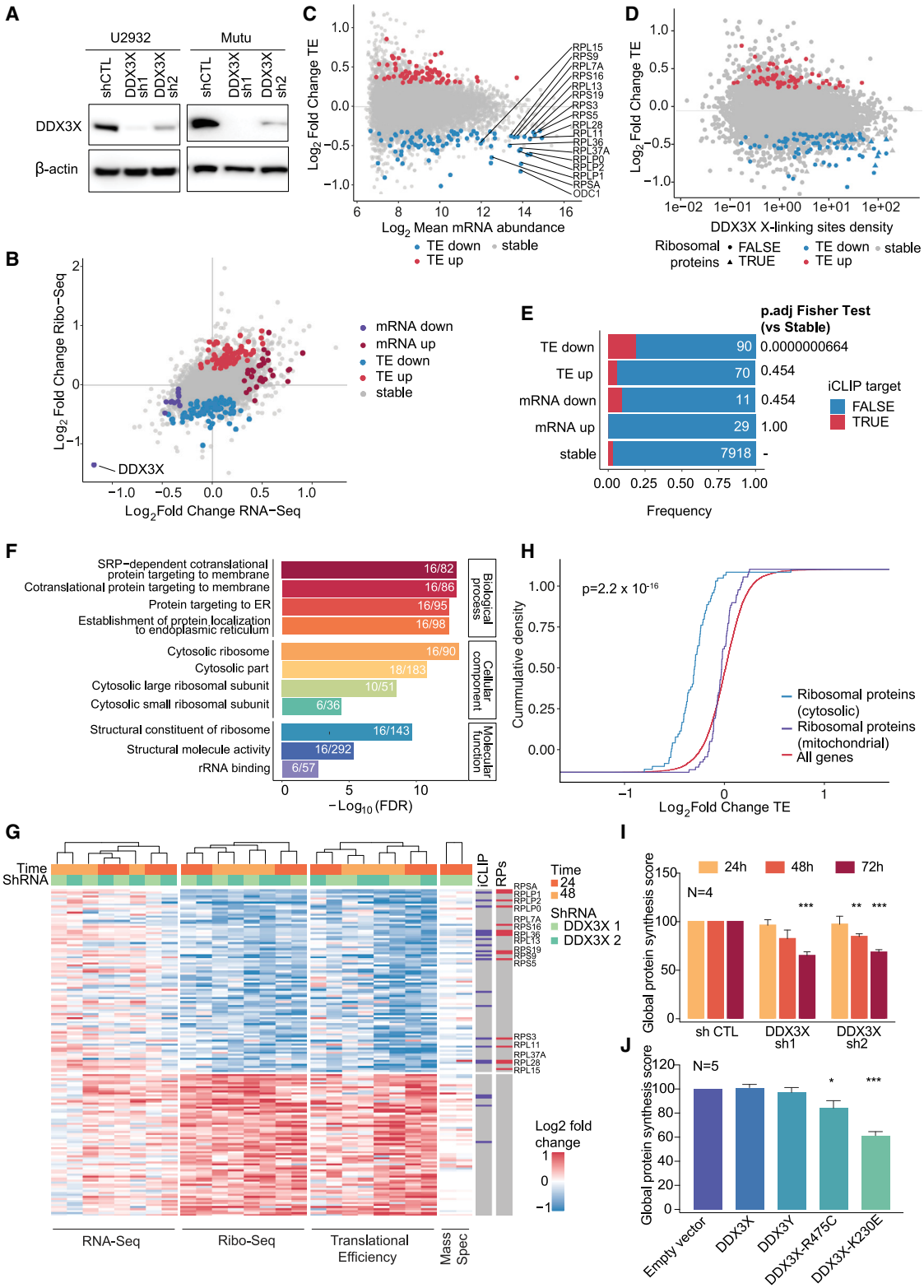
The reduced translation of components of protein synthesis machinery predicted a reduction in global protein synthesis following the depletion of DDX3X. Accordingly, we saw reduced global protein synthesis by quantifying O-propargyl-puromycin (OPP) incorporation in U2932 following the depletion of DDX3X (Figure 4I). Suppression of global protein synthesis was also seen in human GC B cells transduced with helicase mutant or dominant-negative DDX3X (Figure 4J). These data reveal that by promoting the translation of mRNAs encoding components of the translation machinery, DDX3X activity determines global protein synthesis capacity.

DDX3X buffers proteotoxic stress in MYC-transduced primary GC B cells

This impairment of protein synthesis capacity in *DDX3X* mutant cells sits in apparent contrast to the effect of MYC, which promotes ribosome biogenesis and global protein synthesis. We sought to confirm the effects of MYC in human GC B cells. Transduction of MYC alone into primary *ex vivo* GC B cells led to increased global protein synthesis but also increased apoptosis (Figure S5A). Neither effect was seen after the transduction of MYC into established lymphoma lines (Figure S5B), nor was it seen in primary GC B cells transduced with BCL6 (Figure S5C). To overcome this apoptosis and allow us to determine the transcriptional response to MYC, we cotransduced GC B cells with MYC and BCL2. As controls, cells were transduced with either BCL2 alone or BCL6 and BCL2. RNA-seq showed the upregulation of a signature of ribosome biogenesis in MYC-transduced

(B) Venn diagram showing overlap of DDX3X-interacting proteins in Mutu and U2932.
(C) Gene Ontology (GO) enrichment of DDX3X-interacting proteins identified in both cell lines.
(D) Density of iCLIP crosslink sites mapping to the indicated genetic features is shown for the indicated cell types.
(E) Overlap between DDX3X-bound transcripts detected in iCLIP experiments in lymphoma cell lines and primary human GC B cells.
(F) Metagene summary of crosslink density across DDX3X-bound mRNA transcripts, showing length-scaled coding region and 3 kb of the 5′- and 3′-untranslated regions. ORF, open reading frame; TIS, translation initiation site; TTS, translation termination site.
(G) Validation by RNA immunoprecipitation and RT-PCR of mRNA transcripts bound by DDX3X in U2932. DDX3Y is included as a negative control in this female cell line.
(H) GO enrichment of DDX3X-bound transcripts identified by iCLIP in the indicated cell types.
See also Figure S3.

CellPress
OPEN ACCESS



(legend on next page)

cells (Figure 5A). Gene set enrichment analysis (GSEA) revealed the induction of the unfolded protein response (UPR), suggestive of proteotoxic stress (Figure 5B). We confirmed the induction of endoplasmic reticulum (ER) stress in MYC-transduced human GC B cells by immunoblot for the ER stress marker phosphorylated eIF2a (p-eIF2α) (Figure 5C) and PCR and qRT-PCR for the spliced isoform of XBP1 (XBP1s), which undergoes alternative splicing in response to activation of the UPR (Figures 5D and 5E).

To determine whether MYC-induced proteotoxicity was causally related to the observed MYC-induced apoptosis in human B cells, we treated MYC-transduced human GC B cells with rapamycin, a consequence of which is to suppress ribosome biogenesis. While a higher concentration led to increased cell death, a lower concentration of rapamycin abolished MYC-induced increase in global protein synthesis and was associated with a modest reduction in MYC-induced apoptosis (Figures 5F, S5D, and S5E).

We hypothesized that DDX3X mutation may allow GC B cells to tolerate the high-level expression of MYC by moderating MYC-induced global protein synthesis and proteotoxic stress. We therefore cotransduced human GC B cells with MYC plus either WT or mutant DDX3X. Mutant DDX3X abrogated both the MYC-induced increase in global protein synthesis and MYC-induced apoptosis (Figures 5G, 5H, and S5F–S5H). To confirm that the effects of mutant DDX3X were due to loss of function, we used CRISPR-Cas9 to delete DDX3X from human GC B cells, followed by the forced expression of MYC and BCL2. DDX3X deletion reduced the impact of MYC on global protein synthesis, and ER stress judged by XBP1s and p-eIF2α expression (Figures 5I–5K). Our findings suggest that increased MYC expression in human GC B cells is associated with increased global protein synthesis and proteotoxic stress. These effects are buffered by the loss of DDX3X activity.

DDX3X mutation alters the ER stress response

We used CRISPR editing to create the R475C helicase mutant and control U2932 clones (Figure S6A). RNA-seq changes in CRISPR-edited homozygous mutant clones were compared with those from RNA-seq in shRNA-depleted cells. Samples with the strongest knockdown of DDX3X showed the greatest gene expression overlap with the R475C-edited clones (Figure S6B), reinforcing that R475C mutation should be considered loss of function. Strikingly, GSEA revealed how R475C mutant clones were associated with a transcriptional profile opposite to that seen in MYC-transduced GC B cells (Figures 5B and 6A; Table S6). Among the most downregulated gene sets in DDX3X-mutant clones were "MYC targets V1," "mTORC1 signaling," and "UPR" (Figure 6A). The UPR genes *ERN1* (encoding IRE1) and *XBP1* were strongly downregulated (Figure 6B). These data support the conclusion that loss-of-function *DDX3X* mutations act to oppose the effects of MYC, at least in part by constraining the ability of MYC to drive global protein synthesis and proteotoxic stress.

Proteomic profiling of R475C mutant clones showed a reduced abundance of proteins related to the ER and ER stress (Figures 6C and 6D; Table S6). Nearly one-third of downregulated proteins were related to the ER or ER stress (Figure 6C). We exposed WT and R475C mutant clones to the ER stress inducer thapsigargin. While equivalent at baseline, we saw reduced eIF2α phosphorylation following exposure to thapsigargin (Figure 6E), suggesting a reduced ER stress response in DDX3X mutant cells.

Analysis of gene expression data from two published BL studies (Grande et al., 2019; Schmitz et al., 2012) recapitulated the changes seen in CRISPR-edited clones. DDX3X mutant biopsies showed reduced expression of the gene sets "MYC targets," "mTORC1 signaling" and "UPR" (Figure 6F; Table S6) and reduced expression of the UPR genes *ERN1* and *XBP1* (Figures 6G and S6C).

It was previously reported that a reduced basal level of ER stress may render cells more sensitive to pharmacological inducers of ER stress (Huber et al., 2013). Therefore, we tested the sensitivity of R475C-edited clones and DDX3X mutant cell lines to the ER stress-inducing agents thapsigargin and tunicamycin. This revealed increased sensitivity of DDX3X mutant cells to pharmacological inducers of ER stress (Figures 6H and 6I).

Our data do not exclude the possibility that DDX3X regulates alternative tumor suppressor pathways; however, together, they support the conclusion that mutant DDX3X acts to buffer the increase in global protein synthesis and the resulting

Figure 4. DDX3X regulates ribosome biogenesis and global protein synthesis

(A) Knockdown of DDX3X using 2 independent shRNAs.

(B) Scatterplot comparing changes in mRNA abundance (RNA-seq) with changes in ribosome footprint density (ribo-seq) following shRNA depletion of DDX3X in U2932. Data are from 8 replicate knockdowns using 2 different shRNAs. Transcripts with altered translational efficiency (TE) or mRNA abundance are indicated by color.

(C) Scatterplot showing changes in TE plotted against mRNA abundance. Significantly changing genes are colored.

(D) Scatterplot showing changes in TE plotted against crosslinking density from iCLIP experiments.

(E) Bar chart showing the proportion of transcripts, with differential translation identified as direct targets of DDX3X in iCLIP experiments. The number of genes within each category is indicated. Adjusted p values (Fisher test) are shown and reflect the comparison of each category with stable genes.

(F) GO enrichment of genes with reduced TE following DDX3X depletion.

(G) Heatmap showing fold change in RNA-seq, ribo-seq, and TE across 8 replicate knockdowns for all differentially translated genes. Protein abundance changes are shown (Mass Spec). DDX3X targets identified from iCLIP and genes encoding ribosomal proteins (RPs) are indicated by purple or red bands, respectively.

(H) Cumulative distribution of TE change in U2932 following shRNA depletion of DDX3X is plotted for genes encoding cytosolic or mitochondrial ribosome proteins, or all other genes. The p value was calculated using the Kolmogorov-Smirnov test.

(I) Global protein synthesis quantified by OPP incorporation at the indicated time points following shRNA depletion of DDX3X in U2932 and normalized to control shRNA. Data show means ± SEMs; *$p < 0.05$, ***$p < 0.001$; ANOVA with multiple comparison testing; n = 4 replicate experiments.

(J) OPP incorporation 48 h after overexpression of WT or mutant DDX3 constructs in human GC B cells normalized to empty vector (EV) control. Data show means ± SEMs; *$p < 0.05$, ***$p < 0.001$; ANOVA with multiple comparison testing; n = 5 replicate human donors.

See also Figure S4.

CellPress
OPEN ACCESS

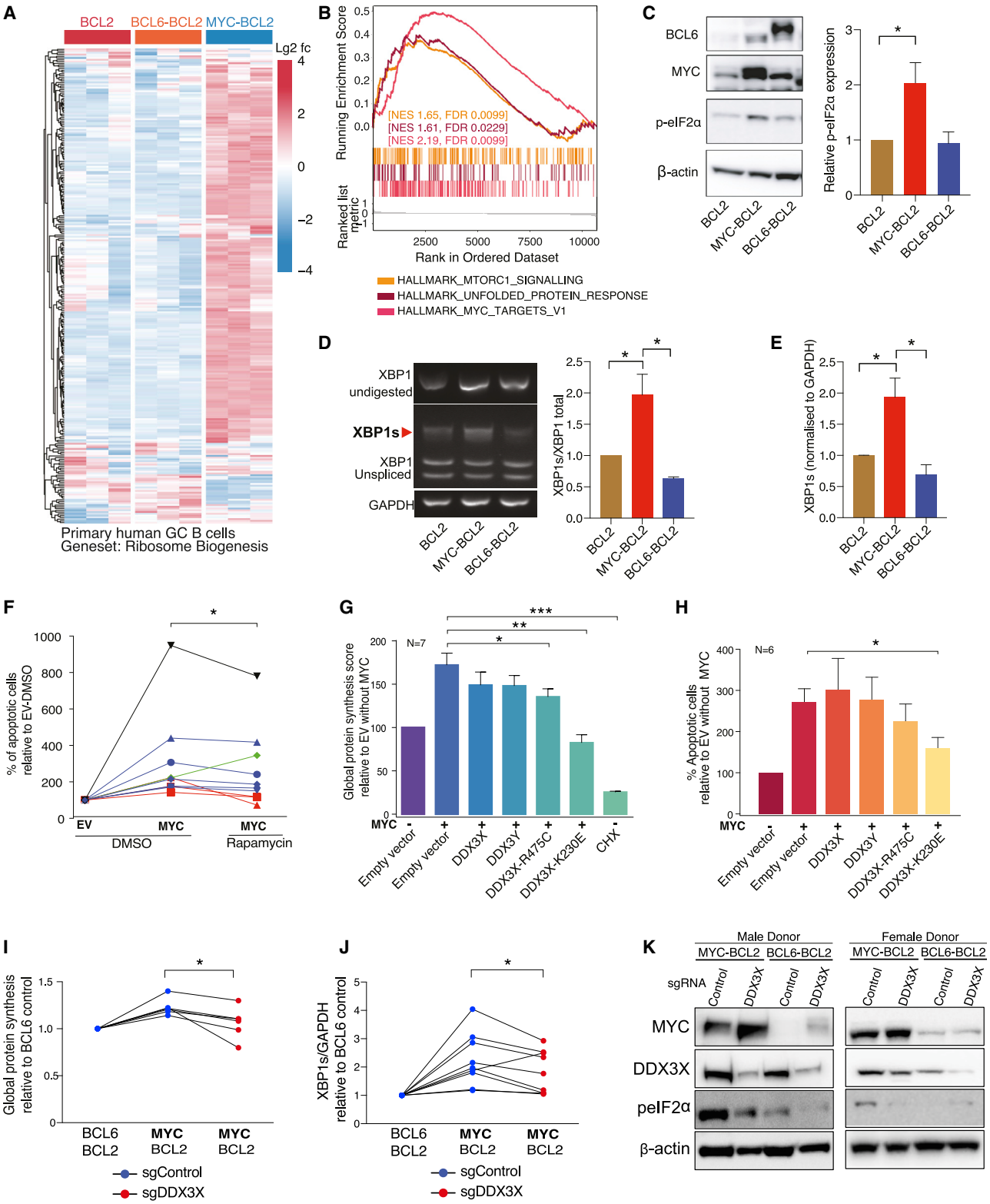


Figure 5. DDX3X buffers proteotoxic stress in MYC-transduced human GC B cells

(A) RNA-seq from human GC B cells transduced with BCL2, BCL6-2A-BCL2, or MYC-2A-BCL2 showing expression change of genes belonging to the gene set "Ribosome Biogenesis" (GO: 0042254).

(legend continued on next page)

proteotoxic stress that would otherwise occur as a consequence of MYC expression. Furthermore, these data reveal a potential therapeutic vulnerability of mutant cells to drugs that induce ER stress.

Upregulation of DDX3Y rescues loss of DDX3 helicase activity

The negative impact of DDX3X mutation on MYC-driven protein synthesis appeared inconsistent with previous data showing how increased translation is required for MYC-driven lymphoma (Barna et al., 2008). We hypothesized that this inconsistency may reflect stage-specific requirements during lymphomagenesis and that DDX3X mutant cells may later acquire mechanisms to compensate for their reduced translational capacity.

Given the enrichment of *DDX3X* mutation in male lymphoma (Figures S1E and S1F) and previous reports of *DDX3X* mutation in other male-skewed cancers (Alkallas et al., 2020; Dunford et al., 2017), we considered the possibility that the Y chromosome homolog *DDX3Y* may compensate in male cells (Figure S7A). Although widely transcribed, DDX3Y protein is not expressed in normal adult tissue outside of the testis (Ditton et al., 2004; Foresta et al., 2000; Rauschendorf et al., 2011). Using a DDX3Y-specific antibody, we saw no expression of DDX3Y protein in normal male GC B cells. However, DDX3Y was detected in all of the male lymphoma cell lines tested (Figure 7A). Furthermore, we observed the expression of DDX3Y protein in all four male BL patient-derived xenografts and all five male primary BL biopsies tested (Figures 7B and 7C). The specificity of our DDX3Y antibody was evidenced by reduced expression after shRNA knockdown in male cell lines (Figure S7B) and the absence from any female cell or biopsy (Figures 7A,–7C). Consistent with previous reports that DDX3Y is regulated at the level of translation (Jaroszynski et al., 2011), we observed no correlation with mRNA abundance (Figures S7C–S7E). We therefore cloned the 5′ UTR of human and mouse DDX3Y into reporter constructs and expressed these in primary GC B cells and in lymphoid cell lines. This revealed a striking suppression of reporter expression by the human but not the mouse DDX3Y 5′ UTR (Figure S7F). This reinforces the potential for translational regulation to determine DDX3Y protein expression and highlights potential interspecies differences in the regulation of this protein.

We found no evidence from our iCLIP or ribo-seq data that the DDX3Y transcript was a direct target of DDX3X. Furthermore, the deletion of *DDX3X* in male GC B cells did not induce the expression of DDX3Y protein *in vitro*, assessed 5 days after deletion (Figure 7D). To establish whether DDX3X depletion may lead to the induction of DDX3Y in a longer-term, *in vivo* tumorigenesis experiment, we deleted *DDX3X* from primary human GC B cells and then transduced them with MYC-2A-BCL2. Cells were then injected into immunodeficient mice, an approach we have previously used to create experimental tumors (Caeser et al., 2019). Tumors formed in all of the mice and were harvested at 7 weeks after injection. While modest depletion of DDX3X was seen at the time of injection, complete loss or truncation was seen in most tumors 7 weeks later, suggesting the selection of DDX3X-deleted clones during the process of MYC-driven lymphomagenesis (Figure 7D). In contrast, DDX3Y, while not detected in GC B cells at the time of injection, showed abundant expression in all tumors (Figure 7D). Furthermore, the expression of DDX3Y was greater in DDX3X-depleted cells. This suggests that deregulated translation of DDX3Y may be a feature of transformed B cells and that its expression is indirectly induced or selected in DDX3X-depleted cells.

RNA IP revealed that DDX3Y bound many of the same mRNA targets as DDX3X (Figure 7E), suggesting that it may function similarly to DDX3X. The knockdown of DDX3Y reduced global protein synthesis in male but not female cell lines (Figures 7F and S7G). Furthermore, the expression of a DDX3Y construct was able to rescue the defects in both global protein synthesis and cell growth when *DDX3X* was deleted by CRISPR in cultured female murine B cells (Figures 7G and 7H). Knockdown of DDX3Y was toxic to almost all male cell lines tested (Figure 7I), suggesting that male lymphoma cells are addicted to the aberrant expression of DDX3Y. These data suggest that the ectopic expression of DDX3Y protein in established male lymphomas compensates for DDX3X loss and restores protein synthetic capacity. DDX3Y is thus an induced tumor-essential gene.

DISCUSSION

DDX3X mutations have been reported in a variety of malignancies. Here, we have characterized the role of DDX3X and its mutation in MYC-driven B cell lymphoma. Initial studies of *DDX3X* mutation in medulloblastoma focused on the activation of WNT signaling by helicase-dead DDX3X. However, the pattern of mutation in lymphoma differs from that in medulloblastoma, suggesting that the effect of DDX3X mutation may differ between

(B) Gene set enrichment analysis (GSEA) of RNA-seq from human GC B cells transduced with MYC-2A-BCL2 compared to BCL2 alone, showing enrichment of gene sets related to MYC, UPR, and mammalian target of rapamycin complex 1 (mTORC1) signaling.

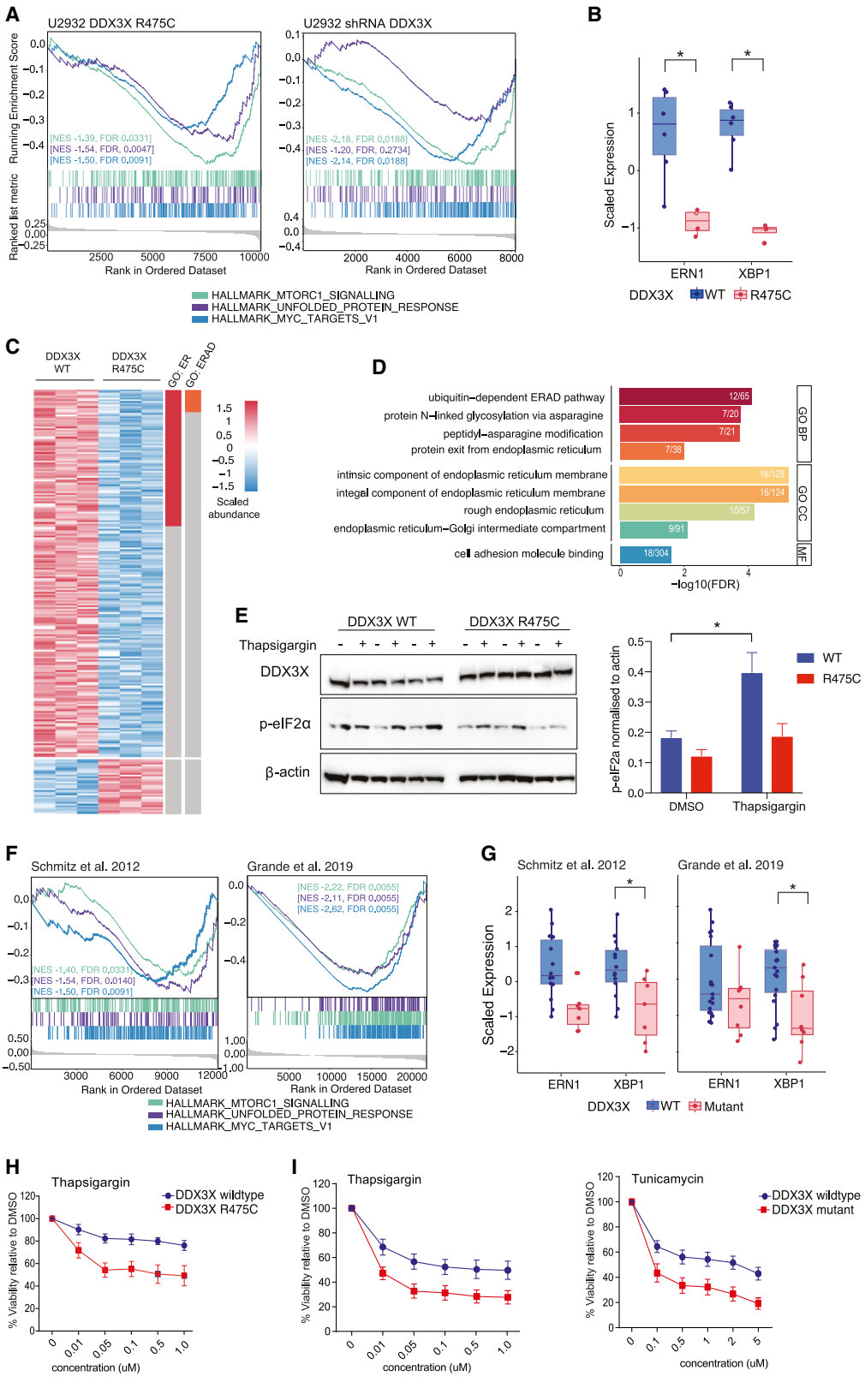
(C–E) Assessment of ER stress markers in human GC B cells 5 days after transduction with the indicated constructs. Data show immunoblot for p-eIF2α (C), spliced XBP1 by RT-PCR and Pst1 digestion (D), or spliced XBP1 by isoform-specific qRT-PCR normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (E). Bar charts show means ± SEMs relative to BCL2 control for 3 human donors; *$p < 0.05$, independent t test.

(F) Effect of rapamycin (50 nM) on MYC-induced apoptosis quantified by flow cytometric staining for cleaved poly(ADP-ribose) polymerase (PARP) and caspase-3 in human GC B cells transduced with MYC or EV control. *$p < 0.05$; Wilcoxon paired t test, 10 replicate donors.

(G and H) Global protein synthesis by OPP incorporation (G) or apoptosis (H) in cultured human GC B cells 48 h after transduction with MYC and either EV, WT, or mutant DDX3. Values are normalized to EV control without MYC. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ANOVA with multiple comparison testing, 6–7 replicate donors, as indicated.

(I–K) Global protein synthesis by OPP assay (I), spliced XBP1 by qRT-PCR (J), or p-eIF2α expression (K) in human GC B cells depleted of DDX3X by CRISPR deletion, then transduced with MYC-2A-BCL2 or BCL6-2A-BCL2. Significance analyzed by paired t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; paired t test. Immunoblots are representative of 5 donors.

See also Figure S5.



(legend on next page)

these two malignancies. Furthermore, our data suggest that the requirement for DDX3X activity may change throughout the development of the tumor and may be accompanied by significant cellular adaptation. These observations underscore the importance of the experimental system used. The use of primary human GC B cells grown *ex vivo* in a co-culture designed to mimic the GC microenvironment allowed us to model the initial stages of human lymphomagenesis in a manner that was not possible in established cell lines.

Previous reports have proposed individual target mRNAs, including Krüppel-like factor 4 (KLF4) (Cannizzaro et al., 2018) and microphthalmia-associated transcription factor (MITF) (Phung et al., 2019), whose translation is regulated by DDX3X. We did not detect KLF4 expression, nor did we observe changes in MITF translation in our experiments in lymphoid cells. In medulloblastoma, mutant DDX3X promotes WNT activity in a helicase-independent fashion through an association with CSNK1E protein (Cruciat et al., 2013; Pugh et al., 2012). We did not find evidence of this association in lymphoid cells, perhaps due to the tissue specificity of these interactions. Our data do not exclude the possibility that DDX3X may exert other functions beyond translation; however, multiple independent lines of investigation converged on the regulation of mRNAs' encoding components of the core translational machinery and a consequent effect on protein synthesis and ER stress. This leads us to conclude that during lymphomagenesis, loss or mutation of DDX3X acts to buffer proteotoxic stress associated with the translocation of *MYC* in a GC B cell.

The role of DDX3X in mRNA translation has been the subject of previous studies. DDX3X promotes the translation of target transcripts by binding to eIF3 and the 40S subunit to facilitate joining of the 60S subunit and formation of the 80S complex (Geissler et al., 2012; Lee et al., 2008). DDX3X also plays a non-redundant role with eIF4a in the unwinding of structured regions within the 5′ UTR of target transcripts to allow joining of the 43S pre-initiation complex (Soto-Rifo et al., 2012). Finally, studies of the yeast homolog Ded1 show how it is able to nucleate stress granule formation in a helicase-independent manner. In contrast, Ded1 helicase activity is required to release transcripts from stress granules, thereby promoting their translation (Hilliker et al., 2011). Recent studies using overexpressed mutant DDX3X reveal increased stress granule formation (Lennox et al., 2020; Valentin-Vega et al., 2016). Therefore, helicase mutant DDX3X may reduce the access of its target transcripts to the translating ribosome by sequestration in stress granules. While we do not define the precise mechanism, our data show how the loss of DDX3X is associated with the reduced translation of target transcripts encoding the core translational machinery, thereby leading to reduced global protein synthesis.

The requirement to limit global protein production has been reported in other malignancies. Suppressed translation by eEF2 kinase allows developing neuroblastoma cells to survive nutrient deprivation (Leprivier et al., 2013). This effect was specific to the MYC-N-driven tumor models (Delaidelli et al., 2017). The loss of Runx1 in myelodysplasia leads to reduced ribosome biogenesis and attenuation of the UPR, providing stress resistance and competitive advantage to *Runx1*-deleted hematopoietic stem cells (Cai et al., 2015). The enzyme protein kinase R-like ER kinase (PERK) is activated by mutation in human medulloblastoma; mouse models show that PERK is essential for the suppression of translation during tumor initiation, but, in a manner similar to our proposed DDX3X model, PERK becomes dispensable in an established tumor (Ho et al., 2016). Together with our findings, these reports exemplify how global protein translation and the cellular response to ER stress must be maintained within strict boundaries, the limits of which may change depending upon the stage of tumor development. This stage- and tissue-specific requirement may explain conflicting descriptions of DDX3X as a tumor suppressor in some contexts and as an oncogene in others.

Our conclusions may appear inconsistent with findings from mouse models of MYC-driven lymphoma. Haploinsufficiency for *Rpl24* reduces global protein synthesis and abrogates the development of MYC-induced lymphoma (Barna et al., 2008). However, genetic ablation of *Rpl24* leads to a permanent suppression of translation. In contrast, mutation of *DDX3X* can be rescued by ectopic DDX3Y, allowing a cell to titrate helicase activity to suit the stage of tumorigenesis. While we do not exclude subtle differences in the target repertoire, our data, combined with the almost complete amino acid conservation in the helicase RNA-interacting domain, support a model in which the ectopic expression of DDX3Y acts to rescue the translational defect associated with the loss of DDX3X. Recently published data support a redundancy between DDX3X and DDX3Y (Venkataramanan et al., 2020). This potential to reactivate DDX3Y may contribute to the sex bias seen in BL and other male-skewed cancers in which DDX3X is

Figure 6. DDX3X mutation alters endoplasmic reticulum (ER) stress response

(A) GSEA of RNA-seq comparing *DDX3X* R475C-edited or control clones (left), and DDX3X shRNA knockdown experiments (right).

(B) Relative expression of the UPR marker transcripts ERN1 (encoding IRE1) and XBP1 mRNA in RNA-seq from *DDX3X* R475C-edited clones. Statistical significance from differential expression analysis (DESeq2).

(C) Heatmap showing proteins with altered abundance in proteomic profiling of DDX3X R475C-edited clones. Proteins included in the GO terms ER (GO: 0005783) and ER-associated protein degradation pathway (ERAD) (GO: 0036503) are indicated by red and orange highlighting, respectively.

(D) GO terms enriched among proteins with decreased expression in DDX3X R475C-edited clones. BP, biological process; CC, cellular component; MF, molecular function.

(E) Immunoblot showing eIF2α phosphorylation in control and CRISPR-edited *DDX3X* R475C clones in the presence or absence of thapsigargin. Bar chart shows a summary of p-eIF2α quantification by densitometry across 3 *DDX3X* WT and mutant clones. *$p < 0.05$, paired t test.

(F and G) GSEA of RNA-seq data from the indicated studies reanalyzed to compare cases of sporadic BL with either WT or mutant *DDX3X*. Relevant gene sets downregulated in the presence of mutant *DDX3X* (F) or the relative abundance of the UPR transcripts ERN1 and XBP1 (G). Statistical significance is from DESeq2.

(H) Sensitivity of *DDX3X* R475C-edited U2932 or control clones to the ER stress-inducing agent thapsigargin.

(I) Sensitivity of cell lines carrying WT (n = 5) or mutant DDX3X (n = 5) to thapsigargin and tunicamycin.

See also Figure S6.

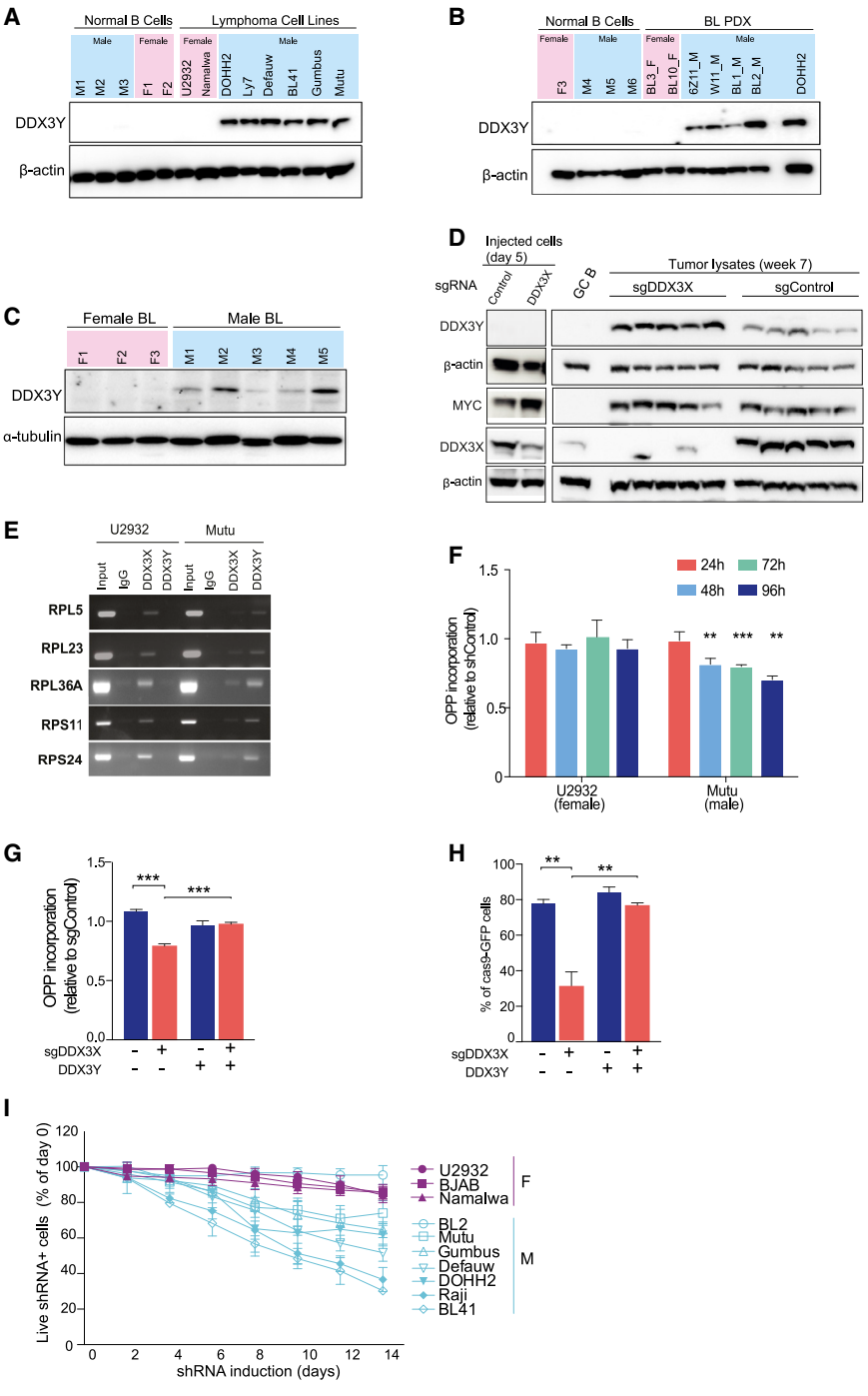


Figure 7. Upregulation of DDX3Y in established tumors rescues loss of DDX3 helicase activity

(A–C) Immunoblots showing DDX3Y protein abundance in lysates from normal GC B cells and lymphoma cell lines (A), normal GC B cells and BL patient-derived xenografts (BL PDX) (male cell line DOHH2 included as a positive control) (B), or BL patient biopsies (C).

(D) Immunoblot showing expression of DDX3X and DDX3Y in experimental tumors created by CRISPR deletion of *DDX3X* (or control) from normal human GC B cells from a male donor, followed by transduction with MYC-2A-BCL2 and engraftment into immunodeficient mice. The resulting tumors were harvested 7 weeks after injection. Expression is also shown for lysates harvested from cells immediately before injection (day 5 after transduction) and in unmanipulated GC B cells from the same donor (GC B).

(E) RNA-IP using antibodies specific for DDX3X, DDX3Y, or isotype control followed by RT-PCR for the indicated transcripts from the female line U2932 or the male line Mutu.

(F) Global protein synthesis by OPP assay in male and female cell lines at the indicated time points after shRNA depletion of DDX3Y. Graphs show means ± SEMs of 3 replicate experiments.

(G and H) Global protein synthesis measured by OPP assay (G) or competitive growth assay (H) in female Cas9-transgenic cultured murine B cells transduced with single-guide RNA (sgRNA) targeting *Ddx3x* or control, and cDNA encoding either human DDX3Y or EV. Graphs show means ± SEMs from 3 mice.

(I) Knockdown of DDX3Y in lymphoma cell lines. The percentage of shRNA-transduced cells was followed over time and presented as % shRNA$^+$ cells relative to day 0. Bars show means ± SEMs of 3–4 replicate experiments per cell line.

*p < 0.05, **p < 0.01, ***p < 0.001; ANOVA with multiple comparison testing.

See also Figure S7.

commonly mutated. In male lymphomas, our shRNA experiments demonstrate addiction to DDX3Y in almost every cell line tested. The absence of *DDX3Y* mutation in any cell line tested and the rarity of *DDX3Y* mutation across published data are consistent with a requirement for the ongoing expression of functional DDX3Y in male lymphoid tumors.

Our findings have relevance to the treatment of MYC-driven lymphoma and identify DDX3Y as an attractive therapeutic target. While essential for male lymphoma lines, DDX3Y protein is not expressed in normal adult human cells (Ditton et al., 2004; Foresta et al., 2000; Rauschendorf et al., 2011). Ectopic DDX3Y expression was observed and required in lymphoma cell lines, even in the absence of *DDX3X* mutation, suggesting it may be a therapeutic target independent of *DDX3X* status. Germline deletion of *DDX3Y* is observed in a proportion of infertile males and has no phenotype beyond azoospermia (Foresta et al., 2000; Rauschendorf et al., 2011), suggesting the toxicity of therapeutic DDX3Y inhibition would be low. Two inhibitors of DDX3 have been developed (Bol et al., 2015; Brai et al., 2016). However, both target the DDX3 helicase domains, which are conserved between DDX3X and DDX3Y. Specific inhibitors or degraders of DDX3Y have yet to be developed. A previous report described

the ability of CD8 T cells to specifically recognize aberrantly expressed DDX3Y protein in leukemic stem cells, suggesting the potential to harness ectopic DDX3Y expression as a target for immunotherapy (Rosinski et al., 2008).

In summary, our data reveal a complex stage-specific requirement for DDX3X helicase activity. We propose a two-stage model whereby mutant DDX3X moderates the ability of MYC to drive upregulation of the core translational machinery. As a result, MYC-induced global protein synthesis is buffered, allowing cells to tolerate the sustained, high-level expression of MYC. Following full transformation, tumor cells increase translational capacity by the ectopic expression of DDX3Y. These findings suggest that drugs that disrupt the delicate balance of translation and proteotoxic stress, or those that specifically target DDX3Y, represent attractive therapeutic strategies for MYC-driven lymphoma.

LIMITATIONS OF THE STUDY

The model we propose arises predominantly from experiments in *ex vivo* human GC B cells and lymphoma cell lines rather than *in vivo* models of MYC-driven murine lymphoma. Despite the recurrent finding of frequent *DDX3X* mutation in human BL, *ddx3x* mutation has not been observed in murine models of BL (Lefebure et al., 2017; Sander et al., 2012). Several factors may explain this observation. First, and in contrast to the human, is the apparent constitutive expression of DDX3Y protein in mouse hematopoietic and lymphoid cells. In keeping with this, deletion of *ddx3x* is not associated with a phenotype in male mice (Szappanos et al., 2018). Our own reporter experiments suggest potential species differences in the regulation of DDX3Y expression. Second, mice express a third DDX3 family gene, *d1pas1*, that is not present in the human genome. *D1pas1* expression was proposed to explain previously noted differences between *DDX3Y* loss in mouse and humans (Matsumura et al., 2019). These interspecies differences make it unlikely that our proposed two-stage "loss and rescue" model would be recapitulated in murine models of BL. However, we anticipate that future mouse models will elucidate further mechanistic understanding about the role of DDX3X in MYC-driven lymphoma.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
  - Primary human B cell culture
  - Human lymphoma cell line culture
  - Murine B cell culture
  - Xenograft models
  - Plasmids
  - Custom-made DDX3Y antibody
- METHOD DETAILS
  - Sequencing and variant ID
  - Variant annotation
  - DDX3X SNV calling and cell of origin assignment from RNA sequencing data
  - Individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP)
  - RNA-sequencing & ribosome profiling
  - Analysis of Ribo-seq and RNA-seq data
  - SILAC-based mass spectrometry for DDX3X Interactomes
  - Proteomic profiling
  - RNA-immunoprecipitation (RNA-IP)
  - Transfection and transduction
  - CRISPR knockout in primary B cells
  - Targeted CRISPR editing of U2932 cell lines
  - Protein extraction and immunoblotting
  - Total RNA extraction, cDNA synthesis and RT-qPCR
  - Apoptosis assay
  - Assessment of spliced XBP1
  - Click-IT OPP assay
  - Cell viability assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
- GRAPHICS DESIGN

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2021.07.041.

ACKNOWLEDGMENTS

D.J.H. was supported by fellowships from the MRC (MR/M008584/1) and Cancer Research UK (CRUK) (RCCFEL\100072). C.G. was supported by a fellowship from the Kay Kendall Leukaemia Fund (KKLF1398). J.A.K. was supported by a CRUK studentship (C9685/A25163). The Hodson laboratory is supported by Blood Cancer UK (grant ID 15022), the Kay Kendall Leukaemia Fund (KKL1144), the Evelyn Trust (ET19/27), and the Addenbrooke's Charitable Trust (900186). Research in the Hodson and Huntly laboratories was funded in part by the Wellcome Trust, which supported the Wellcome – MRC Cambridge Stem Cell Institute (203151/Z/16/Z) (for the purpose of Open Access, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission), the CRUK Cambridge Major Centre (C49940/A25117), and the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014) (the views expressed here are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care). B.J.H. was supported by CRUK (C18680/A25508) and the European Research Council (647685). Targeted lymphoma sequencing was funded by a grant from Blood Cancer UK to E.R. N.F.G. was funded by the Singapore National Medical Research Council (NMRC/TA/0051/2016). S.F., J.M., and S.D.T. were supported by the Alex Hume Foundation. J.U. was supported by the European Research Council under the European Union's Seventh Framework Programme (617837-Translate). The Francis Crick Institute receives its core funding from CRUK (FC001110), the UK Medical Research Council (FC001110), and the Wellcome Trust (FC001110). V.R. was supported by a Blood Cancer UK Senior Bennett Fellowship (12005) and the North East Promenaders Against Cancer (NEPAC), the JGW Patterson Foundation, and the Little Princess Trust. The Turner lab is supported by the BBSRC (BBS/E/B/000C0428) and a Wellcome Investigator award (200823/Z/16/Z). S.S. received core funding from the MRC (MC_UU_12022/10). We thank the patients, clinicians, and staff who participated in the collection of material and clinical data at the Children's Cancer and Leukaemia Group (CCLG) Tissue Bank. We are grateful to Joanna Baxter and Kay Elston from the Cambridge

Blood and Stem Cell Bank and Alice Mitchell and Jessica Bewick from the ENT Department, Cambridge University Hospitals, for the supply of tonsil tissue. We thank Claudia Ribeiro de Almeida, George Vassiliou, and Graham Packham for critical reading of the manuscript. Graphical abstract was generated with BioRender.com.

AUTHOR CONTRIBUTIONS

C.G. designed and performed the experiments, analyzed the data, and led the project. J.A.K. analyzed the RNA-seq, ribo-seq, iCLIP, and proteomic experiments, supervised by S.S. J.G., N.F.G., G.G., M.S., Z.U., R.A., N.B.M.Z., W.M., P.Z., S.K.S., B.J.H., and T.O. performed and analyzed the experiments. N.F.G., S.F., J.M., M.G.K.T., G.A.A.B., and S.DT. provided the reagents. F.C. and M.-Q.D. contributed DLBCL sequencing data. S. Barrans, D.P., C.B., R.T., and E.R. contributed the patient samples. S. Bornelöv, I.R.D.L.M., A.E.B., P.B., P.C., J.U., V.R., M.T., and B.J.H. analyzed and interpreted the data. D.J.H. conceived the project, obtained the funding, designed the experiments, analyzed the data, directed the research, and wrote the manuscript, with contributions from C.G. and J.A.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 29, 2021
Revised: May 17, 2021
Accepted: July 28, 2021
Published: August 25, 2021

REFERENCES

Alkallas, R., Lajoie, M., Moldoveanu, D., Hoang, K.V., Lefrançois, P., Lingrand, M., Ahanfeshar-Adams, M., Watters, K., Spatz, A., Zippin, J.H., et al. (2020). Multi-omic analysis reveals significantly mutated genes and DDX3X as a sex-specific tumor suppressor in cutaneous melanoma. Nat. Cancer *1*, 635–652.

Barna, M., Pusic, A., Zollo, O., Costa, M., Kondrashov, N., Rego, E., Rao, P.H., and Ruggero, D. (2008). Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. Nature *456*, 971–975.

Basso, K., and Dalla-Favera, R. (2015). Germinal centres and B cell lymphomagenesis. Nat. Rev. Immunol. *15*, 172–184.

Berletch, J.B., Yang, F., Xu, J., Carrel, L., and Disteche, C.M. (2011). Genes that escape from X inactivation. Hum. Genet. *130*, 237–245.

Bol, G.M., Vesuna, F., Xie, M., Zeng, J., Aziz, K., Gandhi, N., Levine, A., Irving, A., Korz, D., Tantravedi, S., et al. (2015). Targeting DDX3 with a small molecule inhibitor for lung cancer therapy. EMBO Mol. Med. *7*, 648–669.

Bouska, A., Bi, C., Lone, W., Zhang, W., Kedwaii, A., Heavican, T., Lachel, C.M., Yu, J., Ferro, R., Eldorghamy, N., et al. (2017). Adult high-grade B-cell lymphoma with Burkitt lymphoma signature: genomic features and potential therapeutic targets. Blood *130*, 1819–1831.

Brai, A., Fazi, R., Tintori, C., Zamperini, C., Bugli, F., Sanguinetti, M., Stigliano, E., Esté, J., Badia, R., Franco, S., et al. (2016). Human DDX3 protein is a valuable target to develop broad spectrum antiviral agents. Proc. Natl. Acad. Sci. USA *113*, 5388–5393.

Buchan, J.R., and Parker, R. (2009). Eukaryotic stress granules: the ins and outs of translation. Mol. Cell *36*, 932–941.

Caeser, R., Di Re, M., Krupka, J.A., Gao, J., Lara-Chica, M., Dias, J.M.L., Cooke, S.L., Fenner, R., Usheva, Z., Runge, H.F.P., et al. (2019). Genetic modification of primary human B cells to model high-grade lymphoma. Nat. Commun. *10*, 4543.

Caeser, R., Gao, J., Di Re, M., Gong, C., and Hodson, D.J. (2021). Genetic manipulation and immortalized culture of ex vivo primary human germinal center B cells. Nat. Protoc. *16*, 2499–2519.

Cai, X., Gao, L., Teng, L., Ge, J., Oo, Z.M., Kumar, A.R., Gilliland, D.G., Mason, P.J., Tan, K., and Speck, N.A. (2015). Runx1 Deficiency Decreases Ribosome Biogenesis and Confers Stress Resistance to Hematopoietic Stem and Progenitor Cells. Cell Stem Cell *17*, 165–177.

Calado, D.P., Sasaki, Y., Godinho, S.A., Pellerin, A., Köchert, K., Sleckman, B.P., de Alborán, I.M., Janz, M., Rodig, S., and Rajewsky, K. (2012). The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. Nat. Immunol. *13*, 1092–1100.

Calviello, L., Venkataramanan, S., Rogowski, K.J., Wyler, E., Tejura, M., Thai, B., Krol, J., Filipowicz, W., Landthaler, M., and Floor, S.N. (2021). DDX3 depletion selectively represses translation of structured mRNAs. Nucleic Acids Res. *49*, 5336–5350.

Cannizzaro, E., Bannister, A.J., Han, N., Alendar, A., and Kouzarides, T. (2018). DDX3X RNA helicase affects breast cancer cell cycle progression by regulating expression of KLF4. FEBS Lett. *592*, 2308–2322.

Chapuy, B., Stewart, C., Dunford, A.J., Kim, J., Kamburov, A., Redd, R.A., Lawrence, M.S., Roemer, M.G.M., Li, A.J., Ziepert, M., et al. (2018). Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat. Med. *24*, 679–690.

Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. *44*, e71.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Cruciat, C.M., Dolde, C., de Groot, R.E., Ohkawara, B., Reinhard, C., Korswagen, H.C., and Niehrs, C. (2013). RNA helicase DDX3 is a regulatory subunit of casein kinase 1 in Wnt-β-catenin signaling. Science *339*, 1436–1441.

Cucco, F., Barrans, S., Sha, C., Clipson, A., Crouch, S., Dobson, R., Chen, Z., Thompson, J.S., Care, M.A., Cummin, T., et al. (2020). Distinct genetic changes reveal evolutionary history and heterogeneous molecular grade of DLBCL with MYC/BCL2 double-hit. Leukemia *34*, 1329–1341.

Curk, T. (2019). iCount: protein-RNA interaction analysis. https://icount.readthedocs.io/en/latest/.

Delaidelli, A., Negri, G.L., Jan, A., Jansonius, B., El-Naggar, A., Lim, J.K.M., Khan, D., Oo, H.Z., Carnie, C.J., Remke, M., et al. (2017). MYCN amplified neuroblastoma requires the mRNA translation regulator eEF2 kinase to adapt to nutrient deprivation. Cell Death Differ. *24*, 1564–1576.

Ditton, H.J., Zimmer, J., Kamp, C., Rajpert-De Meyts, E., and Vogt, P.H. (2004). The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. Hum. Mol. Genet. *13*, 2333–2341.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dominguez-Sola, D., Victora, G.D., Ying, C.Y., Phan, R.T., Saito, M., Nussenzweig, M.C., and Dalla-Favera, R. (2012). The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat. Immunol. *13*, 1083–1091.

Dunford, A., Weinstock, D.M., Savova, V., Schumacher, S.E., Cleary, J.P., Yoda, A., Sullivan, T.J., Hess, J.M., Gimelbrant, A.A., Beroukhim, R., et al. (2017). Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. Nat. Genet. *49*, 10–16.

Epling, L.B., Grace, C.R., Lowe, B.R., Partridge, J.F., and Enemark, E.J. (2015). Cancer-associated mutants of RNA helicase DDX3X are defective in RNA-stimulated ATP hydrolysis. J. Mol. Biol. *427*, 1779–1796.

Fabregat, A., Korninger, F., Viteri, G., Sidiropoulos, K., Marin-Garcia, P., Ping, P., Wu, G., Stein, L., D'Eustachio, P., and Hermjakob, H. (2018). Reactome graph database: efficient access to complex pathway data. PLoS Comput. Biol. *14*, e1005968.

Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., and Doudna, J.A. (2016). Autoinhibitory Interdomain Interactions and Subfamily-specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. J. Biol. Chem. *291*, 2412–2421.

Foresta, C., Ferlin, A., and Moro, E. (2000). Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. Hum. Mol. Genet. *9*, 1161–1169.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. *47* (D1), D766–D773.

Geissler, R., Golbik, R.P., and Behrens, S.E. (2012). The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. Nucleic Acids Res. *40*, 4998–5011.

Grande, B.M., Gerhard, D.S., Jiang, A., Griner, N.B., Abramson, J.S., Alexander, T.B., Allen, H., Ayers, L.W., Bethony, J.M., Bhatia, K., et al. (2019). Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. Blood *133*, 1313–1324.

He, Y., Zhang, D., Yang, Y., Wang, X., Zhao, X., Zhang, P., Zhu, H., Xu, N., and Liang, S. (2018). A double-edged function of DDX3, as an oncogene or tumor suppressor, in cancer progression (review). Oncol. Rep. *39*, 883–892.

Hilliker, A., Gao, Z., Jankowsky, E., and Parker, R. (2011). The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. Mol. Cell *43*, 962–972.

Hilton, L.K., Tang, J., Ben-Neriah, S., Alcaide, M., Jiang, A., Grande, B.M., Rushton, C.K., Boyle, M., Meissner, B., Scott, D.W., and Morin, R.D. (2019). The double-hit signature identifies double-hit diffuse large B-cell lymphoma with genetic events cryptic to FISH. Blood *134*, 1528–1532.

Ho, Y., Li, X., Jamison, S., Harding, H.P., McKinnon, P.J., Ron, D., and Lin, W. (2016). PERK Activation Promotes Medulloblastoma Tumorigenesis by Attenuating Premalignant Granule Cell Precursor Apoptosis. Am. J. Pathol. *186*, 1939–1951.

Huber, A.L., Lebeau, J., Guillaumot, P., Pétrilli, V., Malek, M., Chilloux, J., Fauvet, F., Payen, L., Kfoury, A., Renno, T., et al. (2013). p58(IPK)-mediated attenuation of the proapoptotic PERK-CHOP pathway allows malignant progression upon low glucose. Mol. Cell *49*, 1049–1059.

Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. Methods *65*, 274–287.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat. Protoc. *7*, 1534–1550.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

Jaroszynski, L., Zimmer, J., Fietz, D., Bergmann, M., Kliesch, S., and Vogt, P.H. (2011). Translational control of the AZFa gene DDX3Y by 5′UTR exon-T extension. Int. J. Androl. *34*, 313–326.

Jiang, L., Gu, Z.H., Yan, Z.X., Zhao, X., Xie, Y.Y., Zhang, Z.G., Pan, C.M., Hu, Y., Cai, C.P., Dong, Y., et al. (2015). Exome sequencing identifies somatic mutations of DDX3X in natural killer/T-cell lymphoma. Nat. Genet. *47*, 1061–1066.

Jones, D.T., Jäger, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y.J., Pugh, T.J., Hovestadt, V., Stütz, A.M., et al. (2012). Dissecting the genomic complexity underlying medulloblastoma. Nature *488*, 100–105.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol. *17*, 909–915.

Kustatscher, G., Grabowski, P., Schrader, T.A., Passmore, J.B., Schrader, M., and Rappsilber, J. (2019). Co-regulation map of the human proteome enables identification of protein functions. Nat. Biotechnol. *37*, 1361–1371.

Lacy, S.E., Barrans, S.L., Beer, P.A., Painter, D., Smith, A.G., Roman, E., Cooke, S.L., Ruiz, C., Glover, P., Van Hoppe, S.J.L., et al. (2020). Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. Blood *135*, 1759–1771.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Lee, C.S., Dias, A.P., Jedrychowski, M., Patel, A.H., Hsu, J.L., and Reed, R. (2008). Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. Nucleic Acids Res. *36*, 4708–4718.

Lefebure, M., Tothill, R.W., Kruse, E., Hawkins, E.D., Shortt, J., Matthews, G.M., Gregory, G.P., Martin, B.P., Kelly, M.J., Todorovski, I., et al. (2017). Genomic characterisation of Eμ-Myc mouse lymphomas identifies Bcor as a Myc co-operative tumour-suppressor gene. Nat. Commun. *8*, 14581.

Lennox, A.L., Hoye, M.L., Jiang, R., Johnson-Kerner, B.L., Suit, L.A., Venkataramanan, S., Sheehan, C.J., Alsina, F.C., Fregeau, B., Aldinger, K.A., et al. (2020). Pathogenic DDX3X Mutations Impair RNA Metabolism and Neurogenesis during Fetal Cortical Development. Neuron *106*, 404–420.e8.

Leprivier, G., Remke, M., Rotblat, B., Dubuc, A., Mateo, A.R., Kool, M., Agnihotri, S., El-Naggar, A., Yu, B., Somasekharan, S.P., et al. (2013). The eEF2 kinase confers resistance to nutrient deprivation by blocking translation elongation. Cell *153*, 1064–1079.

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

Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. *47*, e47.

López, C., Kleinheinz, K., Aukema, S.M., Rohde, M., Bernhart, S.H., Hübschmann, D., Wagener, R., Toprak, U.H., Raimondi, F., Kreuz, M., et al.; ICGC MMML-Seq Consortium (2019). Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. Nat. Commun. *10*, 1459.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. https://doi.org/10.14806/ej.17.1.200.

Matsumura, T., Endo, T., Isotani, A., Ogawa, M., and Ikawa, M. (2019). An azoospermic factor gene, Ddx3y and its paralog, Ddx3x are dispensable in germ cells for male fertility. J. Reprod. Dev. *65*, 121–128.

McCord, R., Bolen, C.R., Koeppen, H., Kadel, E.E., 3rd, Oestergaard, M.Z., Nielsen, T., Sehn, L.H., and Venstrom, J.M. (2019). PD-L1 and tumor-associated macrophages in de novo DLBCL. Blood Adv. *3*, 531–540.

McGlincy, N.J., and Ingolia, N.T. (2017). Transcriptome-wide measurement of translation by ribosome profiling. Methods *126*, 112–129.

Morton, L.M., Wang, S.S., Devesa, S.S., Hartge, P., Weisenburger, D.D., and Linet, M.S. (2006). Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. Blood *107*, 265–276.

Nojima, T., Haniuda, K., Moutai, T., Matsudaira, M., Mizokawa, S., Shiratori, I., Azuma, T., and Kitamura, D. (2011). In-vitro derived germinal centre B cells differentially generate memory B or plasma cells in vivo. Nat. Commun. *2*, 465.

Oh, S., Flynn, R.A., Floor, S.N., Purzner, J., Martin, L., Do, B.T., Schubert, S., Vaka, D., Morrissy, S., Li, Y., et al. (2016). Medulloblastoma-associated DDX3 variant selectively alters the translational response to stress. Oncotarget *7*, 28169–28182.

Ojha, J., Secreto, C.R., Rabe, K.G., Van Dyke, D.L., Kortum, K.M., Slager, S.L., Shanafelt, T.D., Fonseca, R., Kay, N.E., and Braggio, E. (2015). Identification of recurrent truncated DDX3X mutations in chronic lymphocytic leukaemia. Br. J. Haematol. *169*, 445–448.

Painter, D., Barrans, S., Lacy, S., Smith, A., Crouch, S., Westhead, D., Sha, C., Patmore, R., Tooze, R., Burton, C., and Roman, E. (2019). Cell-of-origin in diffuse large B-cell lymphoma: findings from the UK's population-based Haematological Malignancy Research Network. Br. J. Haematol. *185*, 781–784.

Phelan, J.D., Young, R.M., Webster, D.E., Roulland, S., Wright, G.W., Kasbekar, M., Shaffer, A.L., 3rd, Ceribelli, M., Wang, J.Q., Schmitz, R., et al. (2018). A multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature *560*, 387–391.

Phung, B., Cieśla, M., Sanna, A., Guzzi, N., Beneventi, G., Cao Thi Ngoc, P., Lauss, M., Cabrita, R., Cordero, E., Bosch, A., et al. (2019). The X-Linked DDX3X RNA Helicase Dictates Translation Reprogramming and Metastasis in Melanoma. Cell Rep. *27*, 3573–3586.e7.

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., et al. (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature *488*, 106–110.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

Rauschendorf, M.A., Zimmer, J., Hanstein, R., Dickemann, C., and Vogt, P.H. (2011). Complex transcriptional control of the AZFa gene DDX3Y in human testis. Int. J. Androl. *34*, 84–96.

Reddy, A., Zhang, J., Davis, N.S., Moffitt, A.B., Love, C.L., Waldrop, A., Leppa, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M.L., et al. (2017). Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. Cell *171*, 481–494.e15.

Richter, J., Schlesner, M., Hoffmann, S., Kreuz, M., Leich, E., Burkhardt, B., Rosolowski, M., Ammerpohl, O., Wagener, R., Bernhart, S.H., et al.; ICGC MMML-Seq Project (2012). Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat. Genet. *44*, 1316–1320.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., et al. (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43–48.

Rosinski, K.V., Fujii, N., Mito, J.K., Koo, K.K., Xuereb, S.M., Sala-Torra, O., Gibbs, J.S., Radich, J.P., Akatsuka, Y., Van den Eynde, B.J., et al. (2008). DDX3Y encodes a class I MHC-restricted H-Y antigen that is expressed in leukemic stem cells. Blood *111*, 4817–4826.

Sander, S., Calado, D.P., Srinivasan, L., Köchert, K., Zhang, B., Rosolowski, M., Rodig, S.J., Holzmann, K., Stilgenbauer, S., Siebert, R., et al. (2012). Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis. Cancer Cell *22*, 167–179.

Savage, K.J., Johnson, N.A., Ben-Neriah, S., Connors, J.M., Sehn, L.H., Farinha, P., Horsman, D.E., and Gascoyne, R.D. (2009). MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood *114*, 3533–3537.

Schmitz, R., Young, R.M., Ceribelli, M., Jhavar, S., Xiao, W., Zhang, M., Wright, G., Shaffer, A.L., Hodson, D.J., Buras, E., et al. (2012). Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature *490*, 116–120.

Schmitz, R., Wright, G.W., Huang, D.W., Johnson, C.A., Phelan, J.D., Wang, J.Q., Roulland, S., Kasbekar, M., Young, R.M., Shaffer, A.L., et al. (2018). Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N. Engl. J. Med. *378*, 1396–1407.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Sendoel, A., Dunn, J.G., Rodriguez, E.H., Naik, S., Gomez, N.C., Hurwitz, B., Levorse, J., Dill, B.D., Schramek, D., Molina, H., et al. (2017). Translation from unconventional 5′ start sites drives tumour initiation. Nature *541*, 494–499.

Sha, C., Barrans, S., Care, M.A., Cunningham, D., Tooze, R.M., Jack, A., and Westhead, D.R. (2015). Transferring genomics to the clinic: distinguishing Burkitt and diffuse large B cell lymphomas. Genome Med. *7*, 64.

Sha, C., Barrans, S., Cucco, F., Bentley, M.A., Care, M.A., Cummin, T., Kennedy, H., Thompson, J.S., Uddin, R., Worrillow, L., et al. (2019). Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy. J. Clin. Oncol. *37*, 202–212.

Sharma, D., and Jankowsky, E. (2014). The Ded1/DDX3 subfamily of DEAD-box RNA helicases. Crit. Rev. Biochem. Mol. Biol. *49*, 343–360.

Shih, J.W., Tsai, T.Y., Chao, C.H., and Wu Lee, Y.H. (2008). Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. Oncogene *27*, 700–714.

Smith, A., Crouch, S., Lax, S., Li, J., Painter, D., Howell, D., Patmore, R., Jack, A., and Roman, E. (2015). Lymphoma incidence, survival and prevalence 2004-2014: sub-type analyses from the UK's Haematological Malignancy Research Network. Br. J. Cancer *112*, 1575–1584.

Soto-Rifo, R., and Ohlmann, T. (2013). The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. Wiley Interdiscip. Rev. RNA *4*, 369–385.

Soto-Rifo, R., Rubilar, P.S., Limousin, T., de Breyne, S., Décimo, D., and Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. EMBO J. *31*, 3745–3756.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G.A., Zelenetz, A.D., and Jaffe, E.S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood *127*, 2375–2390.

Szappanos, D., Tschismarov, R., Perlot, T., Westermayer, S., Fischer, K., Platanitis, E., Kallinger, F., Novatchkova, M., Lassnig, C., Müller, M., et al. (2018). The RNA helicase DDX3X is an essential mediator of innate antimicrobial immunity. PLoS Pathog. *14*, e1007397.

Takahashi, K., Hu, B., Wang, F., Yan, Y., Kim, E., Vitale, C., Patel, K.P., Strati, P., Gumbs, C., Little, L., et al. (2018). Clinical implications of cancer gene mutations in patients with chronic lymphocytic leukemia treated with lenalidomide. Blood *131*, 1820–1832.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. *11*, 2301–2319.

Valentin-Vega, Y.A., Wang, Y.D., Parker, M., Patmore, D.M., Kanagaraj, A., Moore, J., Rusch, M., Finkelstein, D., Ellison, D.W., Gilbertson, R.J., et al. (2016). Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. Sci. Rep. *6*, 25996.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinformatics *43*, 11.10.11–11.10.33.

van Schadewijk, A., van't Wout, E.F., Stolk, J., and Hiemstra, P.S. (2012). A quantitative method for detection of spliced X-box binding protein-1 (XBP1) mRNA as a measure of endoplasmic reticulum (ER) stress. Cell Stress Chaperones *17*, 275–279.

Venkataramanan, S., Calviello, L., Wilkins, K., and Floor, S.N. (2020). DDX3X and DDX3Y are redundant in protein synthesis. bioRxiv. https://doi.org/10.1101/2020.09.30.319376.

Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M.; Cancer Genome Atlas Research Network (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.

Yedavalli, V.S., Neuveut, C., Chi, Y.H., Kleiman, L., and Jeang, K.T. (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell *119*, 381–392.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Zhang, X., Smits, A.H., van Tilburg, G.B.A., Ovaa, H., Huber, W., and Vermeulen, M. (2018). Proteome-wide identification of ubiquitin interactions using UbIA-MS. Nat. Protoc. *13*, 530–550.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
DDX3X	Bethyl Laboratories	Cat #A300-474A,; RRID: AB_451009
DDX3X	Santa Cruz	Cat# sc-130736; RRID:AB_2092882
β-actin	Sigma	Cat # A5441; RRID: AB_476744
Rabbit IgG isotype control	Invitrogen	Cat # ab27478; RRID: AB_2616600
Cleaved PARP	BD Bioscience	Cat# 564129; RRID:AB_2738611
Active caspase 3 antibodies	BD Bioscience	Cat# 560626; RRID:AB_1727414
DDX3Y	This paper	N/A
peIF2a	Cell signaling	Cat# 3597S; RRID: AB_390740
eIF3a	Novus	Cat# NBP1-18891; RRID:AB_1625664
eIF3b	Cambio	Cat# A301-761A; RRID:AB_1210995
Phospho-eIF2a (Ser51)	Cell signaling	Cat# 3597S; RRID: AB_390740
MYC	Santa Cruz	Cat# SC-764; RRID: AB_631276
BCL6	Santa Cruz	Cat# SC-368; RRID: AB_2258974
Bacterial and virus strains		
NEB stable Competent E Coli	New England Biolabs	Cat # C3040
Biological samples		
Sporadic BL for targeted sequencing	UK Haematological Malignancies Research Network.	N/A
Human germinal center B cells	Cambridge Blood and Stem Cell Biobank, University of Cambridge	N/A
Burkitt lymphoma xenografts	This paper	N/A
Primary BL samples for western blot	Children's Cancer and Leukaemia Group Tissue Bank	N/A
Chemicals, peptides, and recombinant proteins		
Tunicamycin	Sigma	Cat# T7756-1mg, CAS 11089-65-9
Thapsigargin	Fisher scientific	Cat# 10798352, CAS 67526-95-8
Rapamycin	Cambridge Bioscience	Cat# SM83-5, CAS 53123-88-9
TransIT-293	Cambridge Bioscience	Cat# MIR2700
Alt-R® S.p. Cas9 Nuclease V3	IDT	Cat# 1081058
Alt-R® Cas9 Electroporation Enhancer	IDT	Cat# 1075916
Critical commercial assays		
NucleoSpin RNA extraction kit	Machery-Nagel	Cat# 740955.250
SuperScript III First-Strand Synthesis SuperMix	Invitrogen	Cat# 18080400
NEBNext Poly(A) mRNA magnetic isolation module	New England Biolabs	Cat# NEB E7490
RNA clean up and concentration kit	Norgen Biotek	Cat# 23600
Click-iT Plus OPP Alexa Fluor 647 Protein Synthesis Assay	Fisher Scientific	Cat# C10458
Cell Titer-Glo Luminescent Cell Viability Assay kit	Promega	Cat# G7572
P3 primary cell 4D nucleofector kit S	Lonza	Cat# V4XP-3032

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Targeted BL sequencing data	This paper	EGAS00001004649
Ribo-Seq and RNA-Seq sequencing data (DDX3X shRNA in U2932 and Mutu)	This paper	GSE143393
RNA-Seq sequencing data comparing DDX3X R475C knockout versus control edited cell line clones	This paper	GSE144983
Western blot data	This paper	doi: 10.17632/gsd2w927yy.1
Code used for analysis of RiboSeq data	This paper	GitHub https://doi.org/10.5281/zenodo.5082127
Experimental models: Cell lines		
U2932	Gift from Dr. Louis Staudt	RRID: CVCL_1896
Mutu	Gift from Dr. Louis Staudt	RRID: CVCL_ZY05
BL41	Gift from Dr. Louis Staudt	RRID: CVCL_1087
BJAB	Gift from Dr. Louis Staudt	RRID: CVCL_5711
BL2	Gift from Dr. Louis Staudt	RRID: CVCL_1966
Gumbus	Gift from Dr. Louis Staudt	RRID: CVCL_2051
Namalwa	Gift from Dr. Louis Staudt	RRID: CVCL_0067
Defauw	Gift from Dr. Louis Staudt	RRID: CVCL_0083
DOHH2	Gift from Dr. Louis Staudt	RRID: CVCL_1179
Raji	Gift from Dr. Louis Staudt	RRID: CVCL_0511
Lenti-X 293T	Takara	Cat# 632180, RRID:N/A
Oligonucleotides		
shDDX3X_1 sequence: 5′-GATCCCCGGA TCTCGTAGTGATTCAAGATTCAAGAGA TCTTGAATCACTACGAGATCCTTTTTA-3′	This paper	N/A
shDDX3X_2 sequence: 5′-GATCCCCG GTAGAATAGTCGAACAAGATTTCAAG AGAATCTTGTTCGACTATTCTACC TTTTTA-3′	This paper	N/A
shDDX3Y sequence: 5′-GATCCCCG CCAGCAGTATTCTTCAGTAATTCAA GAGATTACTGAAGAATACTGCTG GCTTTTTA-3′	Sigma	N/A
Mouse Ddx3x sgRNA sequence: GAAAGGGGGCAGATTCGCTGG	This paper	N/A
Human DDX3X sgRNA for CRISPR mutagenesis: CACCGTCCAT GGAGACCGTTCTCAG	This paper	N/A
Non-Pam strand template_WT for CRISPR mutagenesis sequence: A*G*C*GGAAC TGGTGAAGGGCCTCTTCTCTATCAC TCTGAGAACGGTCTCCATGGATACT AGTACATGCGTATCCTTCATGGTA TAAGAAATCCTCCAGAGAATC*T*G*C	This paper	N/A
Non-Pam strand template_mutant for CRISPR mutagenesis sequence: A*G*C*GGAACTGGTGAAGGGCCTC TTCTCTATCACTCTGAGAACAGTC TCCATGGATACTAGTACATGCGTA TCCTTCATGGTATAAGAAATCCTCC AGAGAAT*C*T*G	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PAM strand template with 3′ overhang_WT for CRISPR mutagenesis sequence: A*T*A*CCATGAAGGATACGCATGTAC TAGTATCCATGGAGACCGTTCTCA GAGTGATAGAGAAGAGGCCCTTC ACCAGTTCCGCTCAGGAAAAAG CCCAATTTTAGTGGCT*A*C*A	This paper	N/A
PAM strand template with 3′ overhang_ mutant for CRISPR mutagenesis sequence: A*T*A*CCATGAAGGATACGCATGTACTA GTATCCATGGAGACTGTTCTCAGAGTG ATAGAGAAGAGGCCCTTCACCAGTTC CGCTCAGGAAAAAGCCCAATTTTAGT GGCT*A*C*A	This paper	N/A
Forward primer for amplify genomic DDX3X for CRISPR mutagenesis screening: CAC TACAGCCCAGAACTCCTAGACTTAGAC	This paper	N/A
Reverse primer for amplify genomic DDX3X for CRISPR mutagenesis screening: ACA TGTTTCACATTTGAAATGTCCAGTCCTC TTGC	This paper	N/A
Forward sequencing primer for CRISPR mutagenesis screening: GTATGCCATGATTGCACCTG	This paper	N/A
Reverse sequencing primer for CRISPR mutagenesis screening: CCAGTCCTCTTGCTGCTACC	This paper	N/A
RPL5-Forward: attatgctcggaaacgcttg	This paper	N/A
RPL5-Reverse: acgggcataagcaatctgac	This paper	N/A
RPS24-Forward: ttttcctccttggctgtctg	This paper	N/A
RPS24-Reverse: atgacatccggtgtggtctt	This paper	N/A
RPL23-Forward:acacaggagccaaaaacctg	This paper	N/A
RPL23-Reverse: gctctggtttgcctttcttg	This paper	N/A
RPL28-Forward: ggaactgctccagtttcctg	This paper	N/A
RPL28-Reverse: ggatctccgcttaatgacca	This paper	N/A
RPS11-Forward: aagatggcggacattcagac	This paper	N/A
RPS11-Reverse: taccagtgaaggggcatttc	This paper	N/A
BTG1-Forward: gaggatggctccatctgtgt	This paper	N/A
BTG1-Reverse: tcgttctgcccaagagaagt	This paper	N/A
CSNK1E-Forward: ggctatccctccgaattctc	This paper	N/A
CSNK1E-Reverse: accgaatttcagcatgttcc	This paper	N/A
RPL13A-Forward:ctggaccgtctcaaggtgtt	This paper	N/A
RPL13A-Reverse: tggtacttccagccaacctc	This paper	N/A
RPL36A-Forward: ggctatggtgggcaaactaa	This paper	N/A
RPL36A-Reverse: ctcctcccagttcaaaatgc	This paper	N/A
RPS16-Forward: ctggagccagttctgcttct	This paper	N/A
RPS16-Reverse: tctccttcttggaagcctca	This paper	N/A
XBP1s-Forward: TGCTGAGTCCGCAGCAGGTG	van Schadewijk et al., 2012	N/A
XBP1s-Reverse: GCTGGCAGGCTCTGGGGAAG	van Schadewijk et al., 2012	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
XBP1 total-Forward: AAACAGAGTA GCAGCTCAGACTGC	van Schadewijk et al., 2012	N/A
XBP1 total-Reverse: TCCTTCTGGGTAGACCTCTGGGAG	van Schadewijk et al., 2012	N/A
GAPDH-Forward: GAAGGTGAAGGTCGGAGTC	This paper	N/A
GAPDH-Reverse: GAAGATGGTGATGGGATTTC	This paper	N/A
sgDDX3X: CGTGGACGGAGTGATTACGA	This paper	N/A
Non-targeting control sgRNA: TCAGCAAAGGACGAAACAAA	This paper	N/A
Software and algorithms		
GENCODE v.29	Frankish et al., 2019	https://www.gencodegenes.org
STAR 2.5.4a	Dobin et al., 2013	https://github.com/alexdobin/STAR
GATK 3.8 toolkit	Van der Auwera et al., 2013	https://gatk.broadinstitute.org/hc
Picard 2.20	N/A	http://broadinstitute.github.io/picard
ExAC database 0.3.1	N/A	http://exac.broadinstitute.org
DLBCL transcriptomic classification into ABC, GCB AND Unclassified	Reddy et al., 2017	N/A
DLBCL transcriptomic classification into MHG	Sha et al., 2019	N/A
rpart R package	Zhang et al., 2018	CRAN *rpart*
GTEx datasets (V8)	N/A	https://gtexportal.org/home/
TCGAbiolinks 2.18.0 R package	Colaprico et al., 2016	Bioconductor *TCGAbiolinks*
meta R package	N/A	CRAN *meta*
iMaps server	König et al., 2010	https://imaps.genialis.com/iclip/
iCount	Curk, 2019	https://github.com/tomazc/iCount
Cutadapt 1.16	Martin, 2011	https://cutadapt.readthedocs.io/en/stable
FastQC 0.11.5	N/A	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Bowtie2 2.3.4	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/
RiboWaltz R package	N/A	https://github.com/LabTranslationalArchitectomics/riboWaltz
Deeptools 3.3.0	Ramírez et al., 2014	https://deeptools.readthedocs.io/en/develop
GenomicFeatures R package	Lawrence et al., 2013	Bioconductor *GenomicFeatures*
DESeq2 R package	Love et al., 2014	Bioconductor *DESeq2*
Rsubread R package	Liao et al., 2019	Bioconductor *Rsubread*
clusterProfiler R package	Yu et al., 2012	Bioconductor *clusterProfiler*
Reactome	Fabregat et al., 2018	https://reactome.org
BWA	Li and Durbin, 2009	N/A
CaVEMan	Wellcome Sanger Institute	https://github.com/cancerit/CaVEMan
MSigDB v. 7.0	Subramanian et al., 2005	https://www.gsea-msigdb.org/gsea/msigdb/index.jsp
MaxQuant	Tyanova et al., 2016	
Proteome Discoverer 2.1	Thermo Fisher Scientific	N/A
DEP R package		Bioconductor *DEP*
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/
Graphpad Prism 7	Graphpad	N/A
Adobe Illustrator	Adobe	N/A
FlowJo	Treestar	N/A

RESOURCE AVAILABILITY

Lead contact

Requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr Daniel J Hodson (djh1002@cam.ac.uk).

Materials availability

Plasmids and stable cell lines uniquely generated in this study are available on request to the Lead Contact.

Data and code availability

Targeted sequencing data have been deposited to EGA: EGAS00001004649. RNA-seq and ribosome profiling data have been deposited to GEO: GSE144983, GSE143393. The data are publicly available as of the date of publication.

Original western blot images have been deposited at Mendeley and are publicly available as of the date of publication: https://doi.org/10.17632/gsd2w927yy.1. All bioinformatic code used to analyze ribosome profiling data has been uploaded to GitHub: https://doi.org/10.5281/zenodo.5082127, https://github.com/ashakru/lymphDDX3X. Any additional information required to reanalyze the data reported in the paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

DNA was collected from 39 cases of previously untreated sporadic BL (6 females and 33 males; aged between 3 and 86 years old). All cases were diagnosed at a central diagnostic laboratory HMDS, Leeds, and enrolled in the UK Haematological Malignancies Research Network (https://HMRN.org). Protein lysates from Burkitt Lymphoma xenografts used for immunoblotting were sourced from the Intergroup Trial for Children or Adolescents with B cell NHL or B-AL: Evaluation of Rituximab Efficacy and Safety in High-Risk Patients (NCT01516580), (REC reference: 13/EE/0202) with ethical approval for the generation of PDX models (REC number 07-Q0104-16). Diagnostic Burkitt lymphoma biopsy samples used for immunoblotting were sourced from the Children's Cancer and Leukaemia Group (CCLG) Tissue Bank. Project and ethics approval were granted via the CCLG Biological Studies Steering Group (REC 18/EM/0134; Biological Study 2012 BS 08). Patients were aged between 5 and 14 years old; sex is indicated in the relevant figure or legend.

Primary human B cell culture

Primary GC B cells were purified from freshly discarded tonsil tissue, from children (19 females and 25 males; aged between 2 and 16 years old) undergoing tonsillectomy at Addenbrooke's ENT Department, Cambridge with written informed consent of the patient/parent/guardian. Ethical approval for the use of human tissue was granted by the Health Research Authority Cambridgeshire Research Ethics Committee (REC no. 07/MRE05/44). GC B cells were purified from discarded tonsil tissue using the human B cell negative selection isolation Kit II (Miltenyi Biotec) modified to include negative selection antibodies to IgD and CD44, as described (Caeser et al., 2021).GC B cells were then cultured on irradiated CD40ligand-expressing feeder cells in Advanced RPMI medium (GIBCO) supplemented with 20% FBS, 1% penicillin/streptomycin and human IL21 (Peprotech). Primary cells used in these experiments were confirmed to be EBV negative by PCR (Caeser et al., 2021).

Human lymphoma cell line culture

Human Burkitt lymphoma and DLBCL cell lines were a gift from Dr Louis Staudt and were cultured in RPMI medium (GIBCO) supplemented with 10% testified tetracycline-free FBS (GIBCO). All cell lines used in this study were confirmed to be free from mycoplasma contamination and identity was verified using a 16-amplicon multiplexed copy number variant fingerprinting assay (Phelan et al., 2018). All cells were cultured at 37°C and 5% CO_2. The DDX3X mutation status, EBV status and sex of the cells used in this study is included as Table S7

Murine B cell culture

For experiments involving murine B cell cultures, resting splenic B cells were purified using mouse B cell negative selection isolation kit (Miltenyi Biotech) from R26-LSL-Cas9-GFP mb-1: cre/+ mice (aged 10-12 weeks; sex is indicated in the respective figures or legends) purchased from The Jackson Laboratory. Cells were culture cultured on 40LB cells (Nojima et al., 2011) in IL4 for 4 days followed by IL21 for 3 days.

Xenograft models

Primary B cells isolated from discarded tonsil tissue from a male donor (aged 10 years) were electroporated with Cas9-sgRNA ribonucleoprotein complexes (targeting DDX3X or non-targeting control) (see CRISPR knockout in primary B cells). Cells were transduced with MYC-2A-BCL2 three days after electroporation. The edited cells were harvested five days after transduction and 5 million cells per mouse were resuspended in 125ul of HBSS solution, mixed with 125ul of Matrigel and injected subcutaneously into the

flanks of irradiated female NSG mice (aged 11 weeks) (Jackson). Tumor size was monitored by a technician blinded to tumor genotype. This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB-PPL number P846C00DB).

Plasmids

The coding sequences of DDX3X and DDX3Y were amplified by PCR from cDNA of non-malignant human GC B cells and cloned into pBNM-LyT2 and MSCV vectors using Gibson assembly (NEB). DDX3X mutations were constructed using G-blocks (IDT) and inserted using Gibson assembly. ShRNA vectors were generated by cloning annealed oligonucleotides (sequences shown in Key resources table) into the retroviral vector backbone pRSMX-PG (a gift from Dr Louis Staudt). Sequences were designed to avoid cross-targeting of DDX3X and DDX3Y. CRISPR guide RNAs were cloned into Bbs1-linearized pSpCas9(BB)-2A-GFP (PX458, Addgene, RRID:Addgene_48138) for electroporation. Guide RNAs to murine *ddx3x* were cloned into was cloned into the hU6-sgRNA-PGK-Thy1.1 retroviral vector.

Custom-made DDX3Y antibody

Custom-made rabbit polyclonal antibody specifically against DDX3Y was generated by custom antibody production services of GenScript Corporation, Piscataway, NJ, USA. DDX3Y-specific rabbit polyclonal antibody was prepared against synthetic peptide SHVVVKNDPELDQ (NP_001116137, residues 2-13 of human DDX3Y).

METHOD DETAILS

Sequencing and variant ID

DNA from 39 diagnostic biopsies of sporadic BL were sequenced using a 293-gene hybrid capture hematological panel. The panel has been previously described (Lacy et al., 2020). Sequencing libraries were generated (SureSelect XT, Agilent Technologies) using 50-200ng DNA and sequenced on Illumina HiSeq 2500 instruments using 75nt paired end sequencing. Paired end reads were aligned to the reference genome (GRCh37) using Burrows-Wheeler Aligner (Li and Durbin, 2009). The variant calling pipeline of the Cancer Genome Project, Wellcome Trust Sanger Institute, was used to call substitutions (CaVEMan: Cancer Variants Through Expectation Maximization1) and indels (Pindel2). For substitutions, CaVEMan was run using a composite normal control (as only tumor samples were sequenced in this study). Unmapped reads, PCR duplicates and off-target variants were removed. Post-processing was then performed to remove likely artifact, which involved the removal of variants meeting the following criteria: 1) variant base position supported by < 10 total reads. 2) variant supported by < 3 reads reporting the variant. 3) variant with an allele fraction < 0.05. 4) variant with a repeat length > 4 in a region present in > 10% of normal individuals. 5) less than one third of variant alleles at minimum base quality of 25. 6) composite normal control harbors 3% or more reads reporting the variant allele at minimum base quality of 15. 7) variant alleles lacking bidirectional support (< 2 supporting reads in each direction). 8) variants lacking bidirectional support where mean variant base quality was less than 21 (< 4% supporting reads in each direction). 9) variants falling within the second half of a read containing a GGC[AT]G motif in sequenced orientation where the mean base quality after the motif was less than 20. 10) mean mapping quality of the variant allele < 21. 11) variant falls within a simple or centromeric repeat. 12) more than 10% of reads reporting the variant contain an indel. 13) more than 80% of reads contain the variant allele at the same read position. 14) variant falls within a blacklisted region (based on coordinates) known to generate artifactual results. 15) variant is reported by $\geq$ 3 reads in $\geq$ 1% of samples in composite normal control sample.

For indels, Pindel was run using a composite normal control (as only tumor samples were sequenced in this study). Unmapped reads, PCR duplicates and off target variants were removed. Post-processing was then performed to remove likely artifact using the filters built into Pindel136. Variants meeting the following criteria were also removed: 1) variant base position supported by < 10 total reads. 2) variant supported by < 3 reads reporting the variant. 3) variant with an allele fraction < 0.05. 4) variant with a repeat length > 4 in a region present in > 10% of normal individuals.

Variant annotation

Germline variants were filtered by referencing against the ExAC database using non-TCGA samples (alleles present 10 or more times were excluded). Each variant was then annotated according to likely biological effect, as either a driver event, a passenger event, or a reflection of somatic hypermutation. Any gene disrupting events targeting tumor suppressor genes were classed as driver variants (frameshift, nonsense, essential splice [-2, −1, +1, +2, +5], loss of start, in-frame indel $\geq$ 2 codons). Missense variants and in-frame indels in either tumor suppressor genes or oncogenes were classed as driver alleles based on codon-level recurrence if any of the following conditions were met: (i) $\geq$ 5 samples with a variant at the codon in a cohort of 1,529 lymphoma cases, (ii) $\geq$ 10 cases reported across published datasets: COSMIC, AACR-GENIE v1.0 and PMID 28985567, or (iii) known driver event with biological support from published literature. Splice variants in oncogenes were annotated as driver events using the same rules as missense variants.

DDX3X SNV calling and cell of origin assignment from RNA sequencing data

Single Nucleotide Variant (SNV) calling was performed using paired-end RNA-Seq data from 553 GOYA trial patient samples according to GATK guideline: (https://gatk.broadinstitute.org/hc/en-us/articles/360035531192?id=3891). Briefly, RNA-Seq paired-end reads were mapped to the reference genome using STAR 2.5.4a 2-pass mode (Dobin et al., 2013). Read groups were added with

AddOrReplaceReadGroups and duplicated reads were identified with MarkDuplicates script from Picard tools https://broadinstitute.github.io/picard/. Sequences overhanging intronic regions were hardclipped and STAR mapping qualities were reassigned to match GATK software (Van der Auwera et al., 2013). Variant calling was performed with HaplotypeCaller GATK script, with phredScore 20 as minimal variant calling confidence. Variants clusters (at least 3 valid variants in a window of 35 bases) were removed to diminish the effect of RNA-Seq mapping errors. Standard variants quality filtering was applied with VariantFiltration GATK script: Fisher Strand values > 30.0 and Qual By Depth < 2.0 with. Individual SNVs were then annotated with gene names and their predicted consequences on protein function using VariantAnnotation Bioconductor package and gene models from GENCODE comprehensive gene annotations set (v.28) (Frankish et al., 2019). In order to identify samples with potential DDX3X mutation, ENSG00000215301 (gene ID) and ENST00000644876 (transcript ID) DDX3X models were used as a reference. All nonsense, frameshift and non-synonymous SNVs with the ratio of variant coverage: reference coverage > 0.2, localized in DDX3X helicase domain and not previously reported in the ExAC database (ver. 0.3.1, http://exac.broadinstitute.org) of common population variants were considered as valid hits (Figures S1A and S1B).

Classification of GOYA trial cases into four transcriptomic subtypes: ABC, GCB, Unclassified and Molecular High Grade (MHG) DLBCL was performed as previously described (Reddy et al., 2017; Sha et al., 2019). Briefly, read counts were TMM normalized, log2 transformed and z-scores were computed across the genes. The subtype score was obtained per sample by subtracting mean z-scores of ABC-signature genes (ABC score) from mean z-scores of GCB-signature genes (GCB score). Each sample was assigned to a subtype according to the following criteria: ABC – subtype score > 0.25 and GCB score < 0.75; GCB - subtype score < −0.25 and ABC score < 0.75; Unclassified - not assigned to ABC or GCB subtype. MHG cases were identified among the GCB group using BDC R package (Sha et al., 2015). Chromosome Y expression identification in the GOYA dataset was performed using decision tree algorithm implemented in *rpart* R package: https://cran.r-project.org/web/packages/rpart/index.html. The model was trained using RNA-Seq data from Genotype-Tissue Expression (GTEx) Project, which includes gene expression samples from 54 tissue sites in non-diseased individuals (11688 samples in total)(2013). The GTEx datasets (V8) were obtained from https://gtexportal.org/home/. Raw RNA-Seq counts were filtered and TMM normalized. Per gene scaled gene expression values were used as an input. The GTEx data were randomly split into a training and test set comprising, respectively, 80% (9333 samples) and 20% (2355) of the data. The algorithm running on default parameters achieved high performance: F1 score = 0.9973, AUC = 0.9973, 8 males were misclassified. KDM5D, DDX3Y, USP9Y, RPS4Y1, TXLNGY, XIST were identified as classifying genes. In order to assess the ability of the algorithm to classify cancer samples, GTEx trained model was benchmarked against cancer dataset. TCGA gene expression data (RNA-Seq only) (Weinstein et al., 2013) were downloaded using TCGAbiolinks, the Bioconductor package for integrative analysis with GDC data (Colaprico et al., 2016). Similarly, the dataset was split into a training (80%, 9198 samples) and test set (20%, 2336 samples). The algorithm achieved lower performance than in GTEx data: F1 score = 0.9624, AUC = 0.9779, 46 males and 37 females were misclassified on default parameters. When the GTEx trained model was tested on the TCGA test dataset, the number of misclassified females and males was 2 and 295, respectively. The 295 males classified as females showed remarkably lower expression of genes localized on chromosome Y, which could reflect previously reported loss of chromosome Y during oncogenesis (Dunford et al., 2017). In order to obtain high-confidence set of male DLBCL patient samples with chromosome Y expression, the GTEx trained model was used to classify samples in the GOYA dataset.

Meta-analysis of DDX3X mutation associated sex skew

A total of 6 previous studies with accessible mutation and sex data were included along with our mutation data from 39 cases of BL and our DDX3X variant annotation from RNA-Seq data downloaded from GOYA trial data were also included. Mantel-Haenszel random-effects model was used to calculate the overall Relative Risk (RR) and 95% CI. The RR in all studies had a range of 1.06–1.45 and RR of 1.23. The heterogeneity across studies was assessed by Cochran's Q test and Tau-squared (p value 0.8338, $\tau^2 = 0.0033$). All computations were performed using *meta* R package:
https://cran.r-project.org/web/packages/meta/index.html.

Individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP)

iCLIP was performed as previously described (Huppertz et al., 2014) with minor modification. In brief, cells were cross-linked at 300mJ and then lysed in 1ml of IP lysis buffer (Thermo scientific, Cat No. 87787). The lysate was incubated with 15 μL of Dynabeads at 4°C for 4h for pre-clearing and then further incubated with antibody (4 μg per sample) overnight at 4°C. Beads were washed sequentially with high salt buffer, PNK buffer and water. After washing, beads were re-suspended in 20 μL of IP lysis buffer and incubated with 5 μL of TurboDNase and 0.5U RNase I at 37°C for 3min and then on ice for another 3min. The beads were then washed sequentially with high salt buffer, PNK buffer and water. After on-bead RNase digestion, 3′-RNA dephosphorylation, 3′-Linker ligation and 5′-labeling with ^{32}P reaction was carried out as previously described (Huppertz et al., 2014). Protein was eluted from the beads by incubation at 70°C for 10min in 20 μL of 2x LDS loading dye supplemented with 5% 2-mercaptoethanol. The whole lysate was loaded onto SDS-PAGE gel and electrophoresis and transferring of proteins onto PVDF membranes was carried out as described in Immunoblotting section. Radioactive signal was visualized by exposing the membrane to X-ray film and the desired bands were cut from the membrane. Protein on the membrane was digested by Proteinase K in PK buffer and extracted by phenol-chloroform. RNA was precipitated overnight at 4°C and the pellet was washed in 80% Ethanol. Reverse transcription was performed using the primer as previously described (Huppertz et al., 2014). The cDNA was precipitated and loaded onto TBE-urea gel for purification. The bands between the size of 75bp and 200bp were cut out and extracted using D-tube Dialyzer (VWR, Cat No. 71504-3) by electrophoresis.

CellPress
OPEN ACCESS

The extracted cDNA was precipitated and ligated with oligo onto 5′- end. The libraries were amplified using 14 to 18 cycles of PCR with Phusion master mix and P3/P5 primers. Libraries were purified by TBE gel electrophoresis and the quality was assessed by Bioanalyzer before sequencing on an Illumina Hiseq-4000.

iCLIP data analysis

Replicates of iCLIP data were analyzed using iMaps web server (https://imaps.genialis.com/iclip/), following procedures described elsewhere (König et al., 2010). Briefly, Unique Molecular Identifiers were used to distinguish and remove PCR duplicates before removing experimental barcodes and Solexa adapters. We mapped the trimmed reads to GENCODE GRCh38 v.27 using *iCount mapstar* with default parameters (Curk, 2019; Dobin et al., 2013) (https://github.com/tomazc/iCount). First nucleotide after the UMI was assigned as the crosslink site defined by the truncated cDNA. Crosslink significant sites were determined by the *iCount peaks* finding algorithm (False Discovery Rate < 0.05), by weighting the enrichment of crosslinks versus shuffled random positions. Neighboring cDNA start position less than 15 nt apart were join to form high confidence crosslink clusters with *iCount clusters* function (Curk, 2019). Genes with more than 4 cross-linking peaks in at least one experiment were considered as valid DDX3X targets.

RNA-sequencing & ribosome profiling

Experiments were performed in two cell lines (U2932 and Mutu) using scrambled control or two separate shRNAs against DDX3X. Following shRNA transduction cells were selected for 48hours with puromycin and then induced with doxycycline. Cells were harvested at 24 and 48 hours after shRNA induction. For each sample, we prepared parallel sequencing libraries for RNA-Seq and RiboSeq from the same lysate. For Mutu we analyzed four replicate comparisons. For U2932 the experiment was repeated on two separate occasions, providing eight replicate comparisons.

Ribosome profiling was conducted as previously described (Ingolia et al., 2012) with minor modifications. 5 million cells per sample were treated with 100μg/ml of cycloheximide and immediately centrifuged and lysed in 300μl of buffer containing 20mM Tris-HCl, pH7.4, 150mM NaCl, 5mM MgCl2, 1% NP40, 1mM DTT and 100μg/ml cycloheximide. 100μl of the lysate were reserved for paired RNA-Sequencing and the rest treated with DNase I and RNase I. Ribosome monomers were purified using Microspin S-400 columns. The ribosome protected RNA fragments (RPF) were extracted using RNA clean up and concentration kit (23600, Norgen Biotek). RPF were resolved in 15% Novex TBE-Urea gels, stained with SYBR gold and fragments with 26-34nt were excised from the gel. The RNA was extracted from the gel by electrophoresis using D-tube (MWCO 3.5 kDa,71506-3, Merck Chemical). Precipitated RNA was dephosphorylated by T4 polynucleotide kinase and ligated to universal miRNA Cloning Linker (NEB) using T4 RNA ligase 2 truncated. cDNA was prepared with SuperScript III and reverse transcription primer containing a degenerate 5-nucleotides molecular barcode sequence. The cDNA was resolved by polyacrylamide gel electrophoresis, excised and extracted by D-tube(71504-3, Merk). The extracted cDNA was then circularized and PCR amplified. The final library was separated from PCR primers by electrophoresis and extracted with D-tube(71504-3,Merk) before sequencing on an Illumina Hi-seq4000.

For RNA-seq, total RNA was extracted using NucleoSpin RNA extraction kit (Machery-Nagel, Cat No. 740955.250) according to manufacturer's protocol. 500ng of total RNA was used to prepare RNA-seq libraries using NEBNext Poly(A) mRNA magnetic isolation module (Cat No. NEB E7490) as per the manufacturer's instruction. Final libraries were amplified by PCR for 12 cycles, purified with AMPure XP beads and analyzed by Agilent Bioanalyser before sequencing on an Illumina Hi-seq4000.

Analysis of Ribo-seq and RNA-seq data

Sequencing reads pre-processing and alignment

Raw FASTQ files were stripped of adaptor sequence using Cutadapt 1.16 (Martin, 2011) reads shorter than 15 nucleotides were discarded. After quality check with FastQC 0.11.5 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, Ribo-Seq reads were additionally filtered by rRNA using Bowtie2 2.3.4 with seed length 23 (Langmead and Salzberg, 2012). The remaining reads were then mapped to the human genome (GRCh38) using STAR (Dobin et al., 2013) with parameters as follows:–outFilterType BySJout–alignSJoverhangMin 8–outFilterMultimapNmax 20–alignSJDBoverhangMin 1–outFilterMismatchNmax 999–outFilterMismatchNoverReadLmax 0.04–alignIntronMin 20–alignIntronMax 1000000–alignMatesGapMax 1000000. The reference human rRNA index was constructed from RefSeq database. STAR genome index was built with GENCODE v.29 comprehensive gene annotation set. Ribo-Seq specific quality check: sequencing reads length distribution, triplet nucleotide periodicity and open reading frame enrichment was performed using *RiboWaltz* R package
https://github.com/LabTranslationalArchitectomics/riboWaltz.

Metagene analysis

A metagene analysis for scaled density of Ribo-Seq reads or iCLIP hits relative to start and stop codon was performed using deeptools (Ramírez et al., 2014). The coverage of sequencing reads was normalized per sample by the total number of uniquely mapped reads (CPM) excluding sex chromosomes. Scaled coverage per each transcript was computed using *computeMatrix* function with parameters as follows: scale-regions bs 20 -m 5000 -b 3000 -a 3000 -p 40–metagene–exonID CDS–transcriptID transcript –skipZeros. Transcripts models were built using GENCODE v.29 basic gene annotation set (Frankish et al., 2019).

Read counting

Trimmed gene models were built using GENCODE v.29 comprehensive gene annotation set with *GenomicFeatures* R package (Lawrence et al., 2013). The first 30nt and last 30nt of each CDS region were removed to reflect translation elongation intensity, as described previously (McGlincy and Ingolia, 2017). Trimmed CDS models representing different transcript isoforms were merged

by gene. As is expected from Ribo-seq, 28-30 nt long fragments with evident triplet nucleotide periodicity relative to start and stop codon were the most abundant and were selected for further analysis. Localization of ribosomal P-site was determined by offset between 5′end of fragments spanning translation start site and annotated start codon, which was 12 nt for read lengths of 28-30 nt. Per sample per gene P-sites counts matrices were build using *Genomic Ranges* R package allowing for assignment of a uniquely mapped read to more than one overlapping features. At least 25 overlapped bases were required to assign a read to a gene. Corresponding RNA-Seq samples were counted using the same gene models. Genes with low counts were filtered out with a threshold of minimum 128 counts.

Differential translation analysis was performed as previously described (Sendoel et al., 2017). Briefly, genes differentially expressed were identified using DESeq2. RNA-Seq and Ribo-Seq derived read counts were analyzed separately. Median of ratios normalized expression values were obtained from DESeq2. Translation efficiency (TE) was computed by dividing normalized gene expression measured with Ribo-Seq by mRNA relative abundance. We applied the following rules to define differentially translated genes:

i. A significant change in RiboSeq (Adjusted p value < 0.1)
ii. TE log2 fold change threshold ± 0.3
iii. No change in RNA-Seq (log2fc < 0.5)

Differential expression and downstream analysis of RNA-seq

Uniquely mapped reads were assigned to genes using *featureCounts* function from Rsubread package (Liao et al., 2019) allowing for assignment of a read to more than one overlapping features. At least 25 overlapped bases were required to assign a read to a gene. Differential expression analysis was performed using standard workflow from *DESeq2* package (Love et al., 2014). Identified upregulated and downregulated genes were used to perform gene ontology analysis with *enrichGO* function from *clusterProfiler* package (Yu et al., 2012). Enrichment scores were computed using in-house scripts by taking the ratio between the number of differentially expressed genes overlapping with a gene ontology set and the number of background genes assigned to this gene set. Pathway analysis was performed using browser-based Reactome Pathway Database (Fabregat et al., 2018). A list of all expressed genes detected in RNA-Seq was used as a background set for over-representation testing. Gene Set Enrichment Analysis (GSEA) was performed using the *GSEA* function from *clusterProfiler* package. Gene expression measurements were normalized using variance stabilizing transformation, as implemented in DESeq2 package, and analyzed for enrichment in hallmark gene set from MSigDB v. 7.0 (Subramanian et al., 2005).

SILAC-based mass spectrometry for DDX3X Interactomes

One million of cells were cultured in SILAC medium devoid of lysine and arginine supplemented with 10% h.i. dialyzed FCS and heavy (13C615N4 L-arginine and 13C615N2 L-lysine) or regular (light) amino acids (12C614N4 L-arginine and 12C614N2 L-lysine) (Thermo Scientific) for at least 5 generations. The cells cultured in light medium were lysed and immunoprecipitated with IgG (Abcam) and the cells cultured in heavy medium were lysed and immunoprecipitated with DDX3X (Bethyl Laboratories, specific for DDX3X) or DDX3Y (Sigma, recognizes both DDX3X and DDX3Y (Figure S3)), antibodies using the same procedures as described in RNA-IP section. After the immunoprecipitation, Dynabeads Protein A beads were boiled with 25 μL of LDS loading buffer (Invitrogen) supplemented with β-mercaptoethanol and the light and heavy samples were mixed at 1:1 ratio. The total 50 μL lysate were run on SDS-PAGE gel and stained with Coomassie Brilliant Blue. Each gel lane was cut into 23 pieces, which were subjected to in-gel protein digestion with trypsin (16h, 37°C) after reduction with dithiothreitol and alkylation with iodoacetamide.

Peptide samples were analyzed by LC-MS/MS on a Q Exactive HF orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an Ultimate 3000 RSLCnano HPLC system (Dionex / Thermo Fisher Scientific). The peptides were first trapped on a precolumn (ReproSil-Pur 120 C18-AQ, 5 μm; Dr. Maisch GmbH; 100 μm x 5 cm) and separated on an analytical column (ReproSil-Pur 120 C18-AQ, 1.9 μm; Dr. Maisch GmbH; 280 × 0.075 mm) with a 80-min linear gradient of 2%–40% solvent B [80% (vol/vol) ACN, 0.1% FA] versus solvent A (0.1% FA in water) at a constant flow rate of 300 nL·min–1. Eluting peptides were analyzed by data-dependent acquisition using a top 30 MS/MS method with a survey scan resolution setting of 120,000 FWHM and an MS/MS resolution setting of 15,000 FWHM at 200 m/z. The 30 most abundant precursor peptide ions within the *m/z* 350-1600 range (charge states 2 to 5, intensity > 4x10^4) were selected for higher-energy collision-induced dissociation (HCD) with a normalized collision energy setting of 30% and an isolation width of 1.6 *m/z*. Target values for automatic gain control (AGC) and maximum ion injection times for MS and MS/MS were set to 1 × 10^6 in 50ms and 1 × 10^5 in 50ms, respectively. Selected precursor mass-to-charge ratio values were dynamically excluded from fragmentation for 20 s.

Processing of LC/MS data

Raw data files from LC-MS/MS measurements were analyzed with the MaxQuant software (MPI for Biochemistry) (Tyanova et al., 2016). Fragment ion spectra were searched using the Andromeda search engine (Cox et al., 2011) against the UniProtKB human reference protein database supplemented with frequently observed contaminants and setting trypsin as enzyme for protein digestion. After initial recalibration, precursor and fragment ion mass tolerances of 6 and 20 ppm were set, respectively. Oxidation of methionine and protein *N*-terminal acetylation were allowed as variable modifications and carbamidomethylation of cysteine was defined as fixed modification. Minimal peptide length was set to seven amino acids, with a maximum of two missed tryptic cleavages. On

both the peptide and protein level the maximum false discovery rate (FDR) was set to 1% using a forward-and-reverse concatenated decoy database search strategy. SILAC multiplicity was set to double labeling (Lys+0/Arg+0, Lys+8/Arg+10), requiring at least two ratio counts for peptide quantitation and enabling the "re-quantify" option.

Proteomic profiling

A total of 100 μg protein from each condition was subjected to in-solution digestion before labeling the resultant peptides using the TMT-6plex Isobaric Label Reagent Set (Thermo Scientific, Rockford, IL, USA) according to the manufacturer's protocol. The labeled samples were combined prior to fractionation on a Xbridge C18 column (4.6 × 250 mm, Waters, Milford, MA, USA) and subsequent analysis by LC-MS/MS. The fractionated peptides were separated and analyzed using a Dionex Ultimate 3000 RSLCnano system coupled to a Q Exactive instrument (Thermo Fisher Scientific, MA, USA). Separation was performed on a Dionex EASY-Spray 75 μm × 10 cm column packed with PepMap C18 3 μm, 100 Å (Thermo Fisher Scientific) using solvent A (0.1% formic acid) and solvent B (0.1% formic acid in 100% ACN) at flow rate of 300 nL/min with a 60 min gradient. Peptides were then analyzed on a Q Exactive apparatus with an EASY nanospray source (Thermo Fisher Scientific) at an electrospray potential of 1.5 kV. A full MS scan (350–1,600 m/z range) was acquired at a resolution of 70,000 and a maximum ion accumulation time of 100ms. Dynamic exclusion was set as 30 s. The resolution of the higher energy collisional dissociation (HCD) spectra was set to 350,00. The automatic gain control (AGC) settings of the full MS scan and the MS2 scan were 5E6 and 2E5, respectively. The 10 most intense ions above the 2,000 count threshold were selected for fragmentation in HCD, with a maximum ion accumulation time of 120ms. An isolation width of 2 m/z was used for MS2. Single and unassigned charged ions were excluded from MS/MS. For HCD, the normalized collision energy was set to 30%. The underfill ratio was defined as 0.3%. Raw data files from the three technical replicates were processed and searched using Proteome Discoverer 2.1 (Thermo Fisher Scientific). The raw LC-MS/MS data files were loaded into Spectrum Files (default parameters set in Spectrum Selector) and TMT 6-plex was selected for the Reporter Ion Quantifier. The SEQUEST algorithm was then used for data searching to identify proteins using the following parameters; missed cleavage of two; dynamic modifications were oxidation (+15.995 Da) (M) and deamidation (+0.984 Da) (N). The static modifications were TMT-6plex (+229.163 Da) (any N terminus and K) and Carbamidomethyl (+57 Da) (C). The false discovery rate for protein identification was < 1%. The Normalization mode was set based on total peptide amount. Differential peptides abundance analysis was performed using standard workflow from *DEP* R package with default settings (Zhang et al., 2018).

RNA-immunoprecipitation (RNA-IP)

Cells were harvested and lysed in IP-lysis buffer (Thermo scientific, Cat No. 87787). The lysate was pre-cleared by incubating with 15 μL of Dynabeads Protein A (Invitrogen, Cat No. 1002D) for 4h at 4°C. Afterward, 50 μL of lysate was saved as Input. The remaining lysate was divided equally into 3 portions for incubation overnight at 4°C with IgG (Abcam, Cat No. ab27478, RRID:AB_2616600), DDX3X (Bethyl Laboratories) and DDX3Y (custom antibody). The next day, the lysate was incubated with 15 μL of Dynabeads Protein A for 4h at 4°C. Beads were then washed three times with high salt buffer (as per iCLIP protocol). RNA was extracted from beads or Input with Trizol (Life Technologies, Cat No. 15596026) according to manufacturer's protocol. mRNA was converted to cDNA using SuperScript III First-Strand Synthesis SuperMix and PCR carried out using DreamTaq Green PCR Master Mix (Thermo scientific, Cat No. K1080) according to the manufacturer's protocol. The primer sequences were listed in the Key resources table.

Transfection and transduction

Transfection of HEK293T cells for virus production and transduction of lymphoma cell lines or primary GC B cells were performed as previously described (Caeser et al., 2019, 2021). Retroviral packaging plasmid pHIT60 and GaLV envelope plasmid were used as follows: 1 μg pHIT60 (gag-pol), 1 μg GaLV_WT (envelope) and 4 μg of retroviral construct were used to transfect each 10 cm^2 dish of HEK293T, after mixing with 1 mL of Opti-MEM media (Invitrogen) and 18 μL of TransIT-293 (Mirus). For lentivirus transfections, packaging plasmids pCMVDeltaR8.91 and GaLV_MTR envelope were used as follows: 8.3 μg pCMVDeltaR8.91 (gag-pol), 2.8 μg GaLV_MTR and 11 μg of a lentiviral construct per 10 cm^2 dish, incubated with 1 mL of Opti-MEM media (Invitrogen) and 33 μL of TransIT-293 (Mirus). For transduction of cell lines, pMD2.G (VSV-G envelope) was used instead of GaLV_MTR. pMD2.G was a gift from Didier Trono (Addgene plasmid # 12259; http://addgene.org/12259; RRID:Addgene_12259). The packaging line Lenti-X 293T Cell Line (Clontech Laboratories) was used for all transfections. Viral supernatant was harvested after 24 to 48hrs and filtered through a 0.45 μM filter. Target cells were centrifuged (1500 × g, 90 min at 32°C) in the presence of viral supernatant, 10 μg/ml Polybrene (INSIGHT biotechnology) and 25 μM HEPES (ThermoFisher). Viral supernatant was replaced with fresh media immediately after centrifugation for retroviral infection or after > 4 hours if transducing lentiviral constructs.

To generate viral supernatant for mouse B cell transduction, PlatE cells were transfected with 1ug retroviral plasmid and 3ul 293T Trans-it reagent in a 12 well plate. Media was replaced with B cell culture media after 24 hours and harvested after 48 hours. Transduction occurred on day 3 of culture. Media was replaced with viral media supplemented to 4ug/ml polybrene and IL4 and cells were infected by centrifugation 1000xg, 45 min at 32oC. Viral media was replaced with fresh media after 4 hours.

CRISPR knockout in primary B cells

The knockout of DDX3X in primary B cells was achieved by delivering ribonucleoprotein (RNP) complex into the cells by electroporation using Amaxa 4D electroporator as previously described. Primary B cells were stimulated in coculture with YK6-CD40Lg-IL21 feeder cells for two days before electroporation. 5ug of recombinant cas9 protein (IDT) was mixed with either DDX3X sgRNA (IDT,

sequence CGTGGACGGAGTGATTACGA) or non-targeting control sgRNA (IDT, sequence TCAGCAAAGGACGAAACAAA) at a molar ratio of 2:1 and electroporated into 10^6 cells resuspended in P3 Primary Cell Buffer (Lonza, V4XP-3032) using the Amaxa 4D nucleofector (program EO-117). Cells were transduced with MSCV-MYC-t2A-BCL2-hCD2 or MSCV-BCL6-t2A-BCL2-hCD2 three days after electroporation. Cells were harvested for immunoblot or RT-qPCR five days post-transduction.

Targeted CRISPR editing of U2932 cell lines

sgRNAs were cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene, RRID: Addgene_48138) and the plasmid was electroporated into the cells together with annealed double strand template using Amaxa nucleofector. Both control and R475C templates also contained silent mutations to disrupt the PAM site as well as introducing a SpeI restriction enzyme digestion site used for screening targeted cells. Two days after electroporation, GFP-positive cells were sorted by FACS. Genomic DNA was extracted and the modified segment of DDX3X was amplified followed by SpeI digestion to detect template incorporation. Cells were then single cell cloned into 96 well plates and expanded. Colonies were screened using SpeI digestion and correct targeting verified by Sanger sequencing. Sequences for templates and sgRNAs are listed in Key resources table.

Protein extraction and immunoblotting

Cell pellets or frozen tissues were homogenized, washed with PBS and lysed in 1X RIPA (Cell Signaling, Cat No. 9806S) buffer on ice. 10 μg of protein was loaded into each well of the gel. Electrophoresis was done with Bolt 4% to 12% Bis-Tris gel (Invitrogen). Membranes were blocked in 5% milk followed by incubation with primary antibodies including DDX3X (Bethyl Laboratories, A300-474A, RRID: AB_451009), β-actin (Sigma, Cat No. A5441, RRID: AB_476744), DDX3Y (Sigma, WH0008653M1, RRID: AB_1841225), DDX3Y (custom made) overnight at 4°C. Membranes were washed with 1x TBST and incubated with secondary antibodies for 90mins. ECL substrate (Bio-Rad, Cat No. 170-5061) was added onto the membranes and the images were captured by Azure Biosystem C300 and exported as JPEG.

Total RNA extraction, cDNA synthesis and RT-qPCR

Total RNA was extracted using NucleoSpin RNA extraction kit (Machery-Nagel, Cat No. 740955.250) according to manufacturer's protocol. Up to 2 μg of RNA was converted to cDNA using SuperScript III First-Strand Synthesis SuperMix (Invitrogen, Cat No. 18080400) according to manufacturer's protocol. DDX3X and DDX3Y were quantified by RT-qPCR using custom-made Taqman probes with Taqman master mix (Invitrogen, Cat No. 4461881). PGK1 was used as internal control for normalization (Cat No. 43337657).

Apoptosis assay

Cells were harvested 48h post-transduction, fixed in fixation buffer (BD Bioscience, Cat No. 564129), permeabilized in 1x PermWash Intracellular staining buffer (Biolegend, Cat No. 421002) and stained with Cleaved PARP (BD Bioscience, Cat No. 564129) and active caspase 3 antibodies (BD Bioscience, Cat No. 560626). The stained cells were then analyzed on flow cytometer and the data were analyzed by FlowJo.

Assessment of spliced XBP1

The XBP1s splicing assay was performed as previously described (van Schadewijk et al., 2012). XBP1s abundance was evaluated either by RT-qPCR quantification using the XBP1s-specific primers, or by PCR amplification of XBP1 cDNA followed by PstI1 restriction enzyme digestion, which cuts only the unspliced isoform. The products were visualized by agarose gel electrophoresis.

Click-IT OPP assay

Cells were harvested either the specified time points and global translation was measured using Click-iT Plus OPP Alexa Fluor 647 Protein Synthesis Assay Kit (Fisher Scientific, Cat No. C10458) according to manufacturer's protocol. The cells were then analyzed by flow cytometer and the data were analyzed by FlowJo.

Cell viability assay

Cell lines were treated with the drugs indicated for 48h using the concentrations indicated while control cells were treated with DMSO. Cell viability was measured using Cell Titer-Glo Luminescent Cell Viability Assay kit (Promega, Cat No. G7572) and analyzed on Spectramax M5. Results were calculated as percentage of luminesce reading relative to DMSO control.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of the band intensity was done using ImageJ (Schneider et al., 2012). All the t tests were performed in Graphpad Prism 7 and described in the figure legends.

GRAPHICS DESIGN

Visual Abstract was generated with BioRender.com

Molecular Cell, Volume *81*

Supplemental information

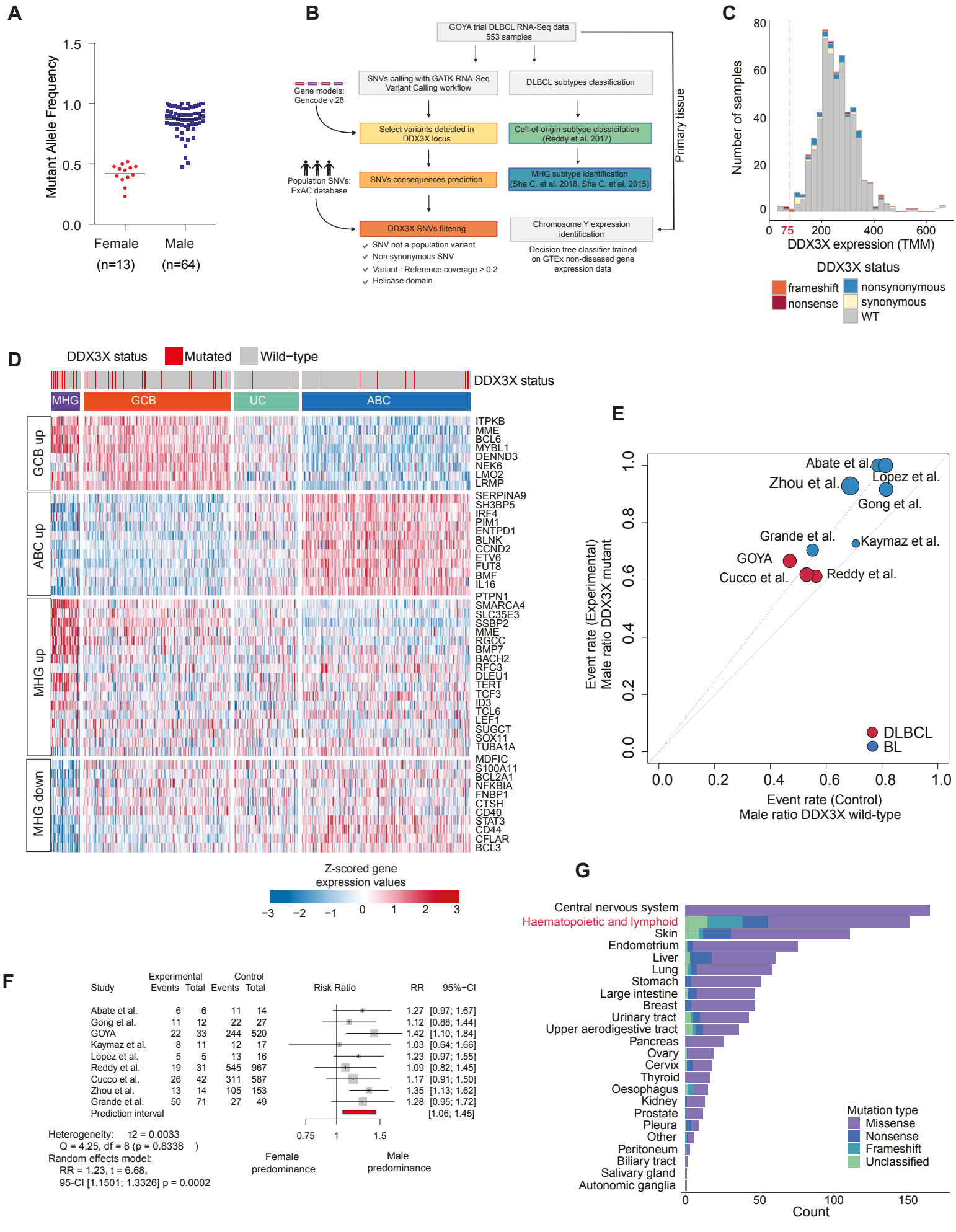
Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis

Chun Gong, Joanna A. Krupka, Jie Gao, Nicholas F. Grigoropoulos, George Giotopoulos, Ryan Asby, Michael Screen, Zelvera Usheva, Francesco Cucco, Sharon Barrans, Daniel Painter, Nurmahirah Binte Mohammed Zaini, Björn Haupl, Susanne Bornelöv, Igor Ruiz De Los Mozos, Wei Meng, Peixun Zhou, Alex E. Blain, Sorcha Forde, Jamie Matthews, Michelle Guet Khim Tan, G.A. Amos Burke, Siu Kwan Sze, Philip Beer, Cathy Burton, Peter Campbell, Vikki Rand, Suzanne D. Turner, Jernej Ule, Eve Roman, Reuben Tooze, Thomas Oellerich, Brian J. Huntly, Martin Turner, Ming-Qing Du, Shamith A. Samarajiwa, and Daniel J. Hodson

Supplemental Information

Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis.

Chun Gong, Joanna A Krupka, Jie Gao, Nicholas F Grigoropoulos, George Giotopoulos, Ryan Asby, Michael Screen, Zelvera Usheva, Francesco Cucco, Sharon Barrans, Daniel Painter, Nurmahirah Binte Mohammad Zaini[4], Björn Haupl, Susanne Bornelöv, Igor Ruiz De Los Mozos, Wei Meng, Peixun Zhou, Alex E Blain, Sorcha Forde, Jamie Matthews, Michelle Guet Khim Tan, G. A. Amos Burke, Siu Kwan Sze, Philip Beer, Cathy Burton, Peter Campbell[2], Vikki Rand[15,16], Suzanne D Turner, Jernej Ule, Eve Roman, Reuben Tooze, Thomas Oellerich, Brian J Huntly, Martin Turner, Ming-Qing Du, Shamith Samarajiwa and Daniel J Hodson



Supplementary Figure 1.

Figure S1. DDX3X is preferentially mutated in MYC-driven lymphomas. Related to Figure 1.

A) DDX3X mutant allele frequency by sex. Data are taken from this study and three other BL sequencing studies for which sex and MAF was available (Grande et al., Lopez et al. & Zhou et al.)

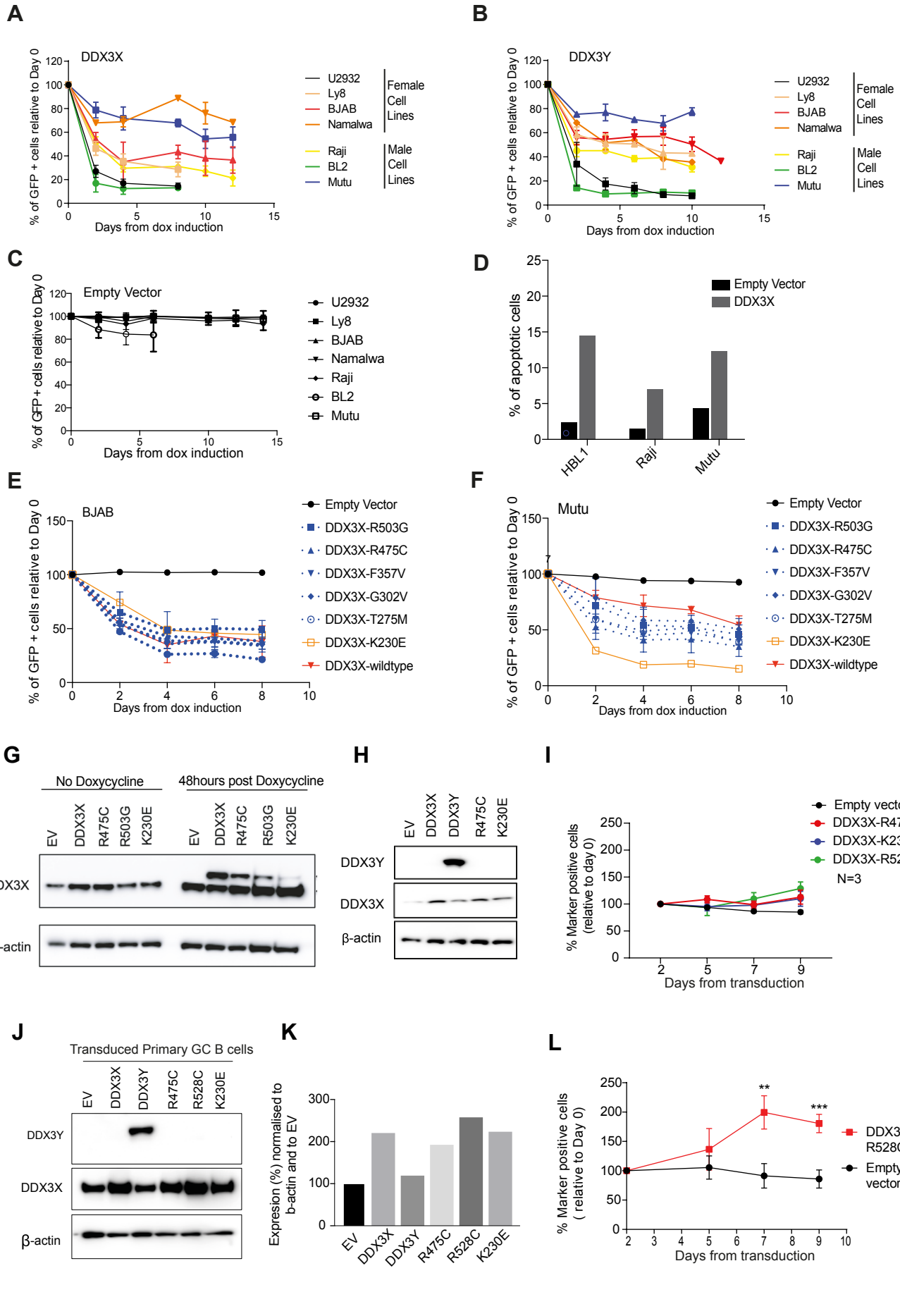
B) Flow chart showing the pipeline for analysis of publicly available RNA-Seq data downloaded from GOYA trial.

C) Histogram of DDX3X mRNA expression in GOYA RNA-Seq samples. DDX3X expression level below 75 (TMM) was considered as not expressed.

D) Heatmap showing DDX3X mutation status by transcriptional subtype in 553 DLBCL cases from GOYA trial. Top bar shows DDX3X mutation status. Rows represent genes included in gene expression signatures that were used to assign transcriptional subtypes.

E,F) Forrest plot and L'Abbé plot showing sex skew of DDX3X across this and other studies of BL and DLBCL.

G) Barplot showing distribution of DDX3X mutation types in different cancer types included in COSMIC database (v.89).



Supplementary Figure 2.

Figure S2. Transduction of DDX3X into lymphoma cell lines and primary human B cells. Related to Figure 2.

A-C) Retroviral constructs expressing doxycycline-inducible DDX3X WT **(A)**, DDX3Y **(B)** or empty vector **(C)** were transduced into the indicated lymphoma cell lines and the relative frequency of the transduced (GFP $^{+}$) cell population monitored over time by flow cytometry. Day zero corresponds to the first addition of doxycycline, used to induce expression of DDX3X/Y.

D) Apoptosis was quantified by flow cytometry for active caspase-3 and cleaved PARP in cell lines transduced with empty vector or DDX3X.

E,F) The indicated DDX3X mutants were transduced into BJAB **(E)** or Mutu **(F)** and relative frequency of transduced cells monitored over time by flow cytometry.

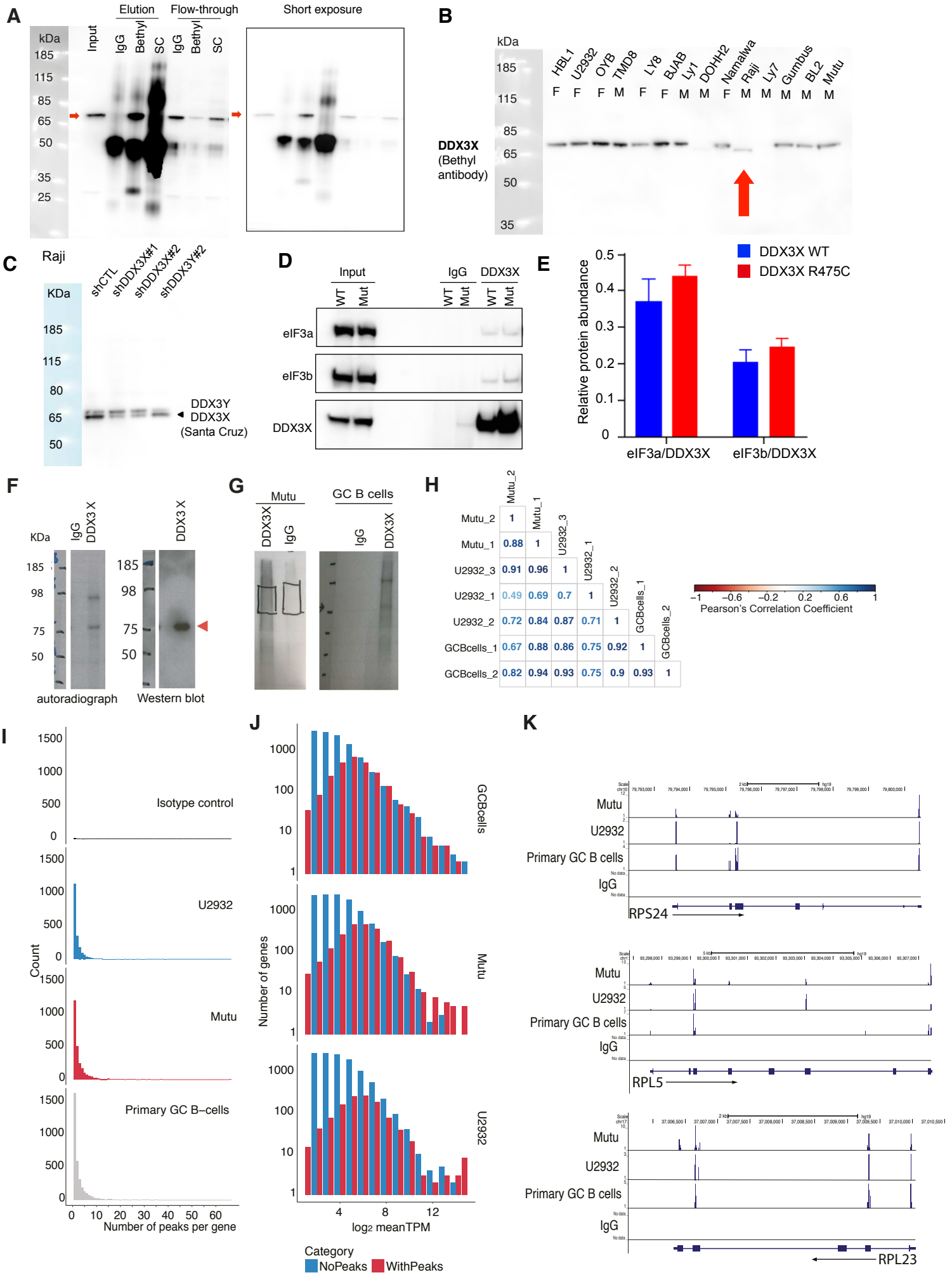
G) Immunoblot showing expression of inducible, tagged DDX3X constructs in U2932. Lower band is endogenous DDX3X. Upper band visible 48 hours after doxycycline induction is tagged DDX3X.

H) Immunoblot showing expression of DDX3X and DDX3Y proteins from retroviral constructs transduced into HEK293T cells

I) Primary GC B cells were transduced with the indicated DDX3X constructs and the frequency of transduced cells monitored over time by flow cytometry.

J,K) Immunoblot **(J)** and densitometry quantification **(K)** showing expression of the indicated DDX3X and DDX3Y constructs in female primary GC B cells.

L) Primary GC B cells were co-transduced MYC and either helicase-mutant R528C DDX3X or control vector. Results show mean+/-SEM for three replicate human donors.



Supplementary Figure 3.

Figure S3. Identification of protein and mRNA interactome of DDX3X in lymphocytes. Related to Figure 3.

A) Antibodies from Bethyl and Santa Cruz (SC) were tested with isotype control (IgG) for ability to immunoprecipitate endogenous DDX3X from lymphoma cell lines. Long and short exposures are shown.

B) Specificity of Bethyl DDX3X antibody for DDX3X is shown using a panel of cell lines that includes the cell line Raji that expresses a truncated DDX3X.

C) Evidence that Santa Cruz DDX3X antibody recognizes DDX3X and DDX3Y shown using Raji cell line (in which DDX3X and DDX3Y appear as differently sized bands) and shRNAs directed against DDX3X or DDX3Y.

D,E) Co-immunoprecipitation of DDX3X from U2932 clones in which the DDX3X locus was CRISPR-edited at R475 with either silent mutation (WT) or R475C mutation. Pulldown of eIF3a and eIF3b is seen in the immunoblot **(D)**, and densitometry quantification of three replicate experiments **(E)**.

F) Autoradiograph and immunoblot from iCLIP lysates precipitated with Bethyl anti-DDX3X antibody.

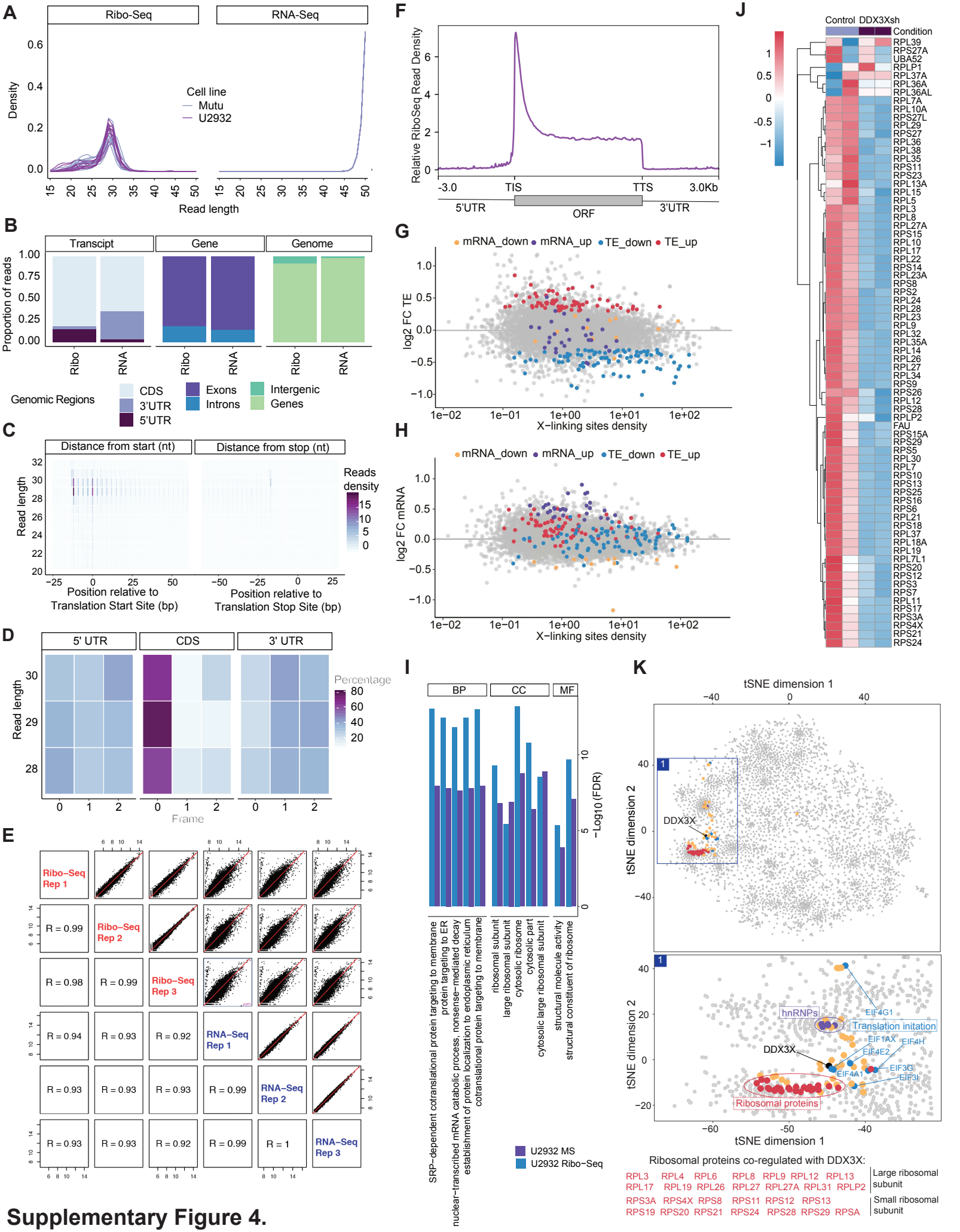
G) Autoradiograph showing radiolabeled DDX3X-bound transcripts and position of bands excised for iCLIP.

H) Summary of overlap between 7 replicate iCLIP experiments across the indicated cell types. Numbers and colors correspond to Pearson's Correlation Coefficient between technical replicates.

I) Distribution of number of iCLIP peaks mapped to each gene for the indicated cell types and isotype control (IgG).

J) Number of genes identified as DDX3X targets (red) in iCLIP experiments as a function of mRNA abundance in respective RNA-Seq experiments.

K) UCSC genome browser tracks showing iCLIP peaks in the indicated cell types and pooled IgG isotype control.



Supplementary Figure 4.

Figure S4. Ribosome profiling reveals that DDX3X promotes translation of mRNAs encoding components of the core protein synthesis machinery. Related to Figure 4.

A) Distribution of read length in Ribo-Seq and RNA-Seq experiments. Ribo-Seq shows a characteristic 28-31 nt footprint corresponding to the length of the ribosome protected mRNA fragment.

B) Distribution of mapped reads from Ribo-Seq and RNA-Seq experiments to gene features, showing the expected restriction of Ribo-Seq reads to the CDS and 5'UTR, and almost complete exclusion of reads mapping to the 3'UTR.

C) Heatmap of Ribo-Seq read frame usage by read length, showing read frame restriction in the 28-31nt ribosome protected fragments, characteristic of ribosome position on the mRNA transcript.

D) Heatmap of Ribo-Seq read frame usage by gene feature, showing how the characteristic read frame bias within the CDS but not the UTRs.

E) Example scatter plots showing consistency between three typical replicates of RiboSeq and RNA-Seq. Pearson correlation coefficient is shown.

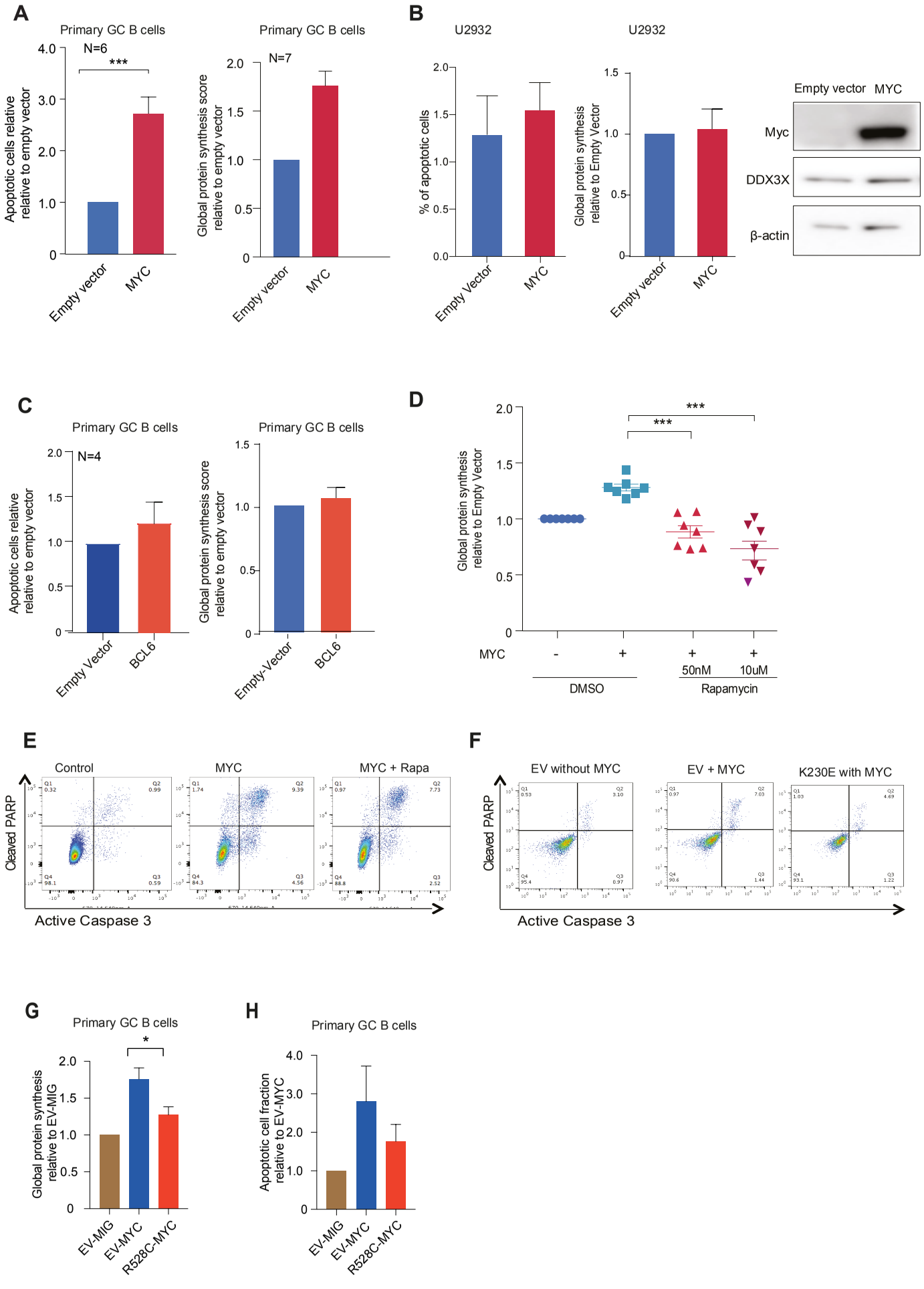
F) Metagene plot showing distribution of mapped Ribo-Seq reads to regions of the transcript, with the characteristic peak at the translation initiation site (TIS) and abrupt drop-off at translation termination site (TTS).

G,H) Scatter plots showing the relationship between transcripts altered at level of TE **(G)** or mRNA abundance **(H)** after DDX3X knockdown, and x-linking density from DDX3X iCLIP experiments.

I) GO terms enriched amongst proteins with reduced abundance in proteomic profiling (MS) following DDX3X depletion. Results from RiboSeq experiments are included for comparison.

J) Heatmap showing altered abundance of ribosomal proteins in mass spectrometry analysis performed following shRNA DDX3X depletion in U2932.

K) Map of the coregulated human proteome plotted using data downloaded from Proteome-HD(Kustatscher et al., 2019). The 81 proteins identified as being statistically coregulated with DDX3X are colored. Ribosome proteins identified as being coregulated with DDX3X are shown in red.



Supplementary Figure 5.

Figure S5. DDX3X buffers proteotoxic stress in MYC transduced primary GC B cells. Related to Figure 5.

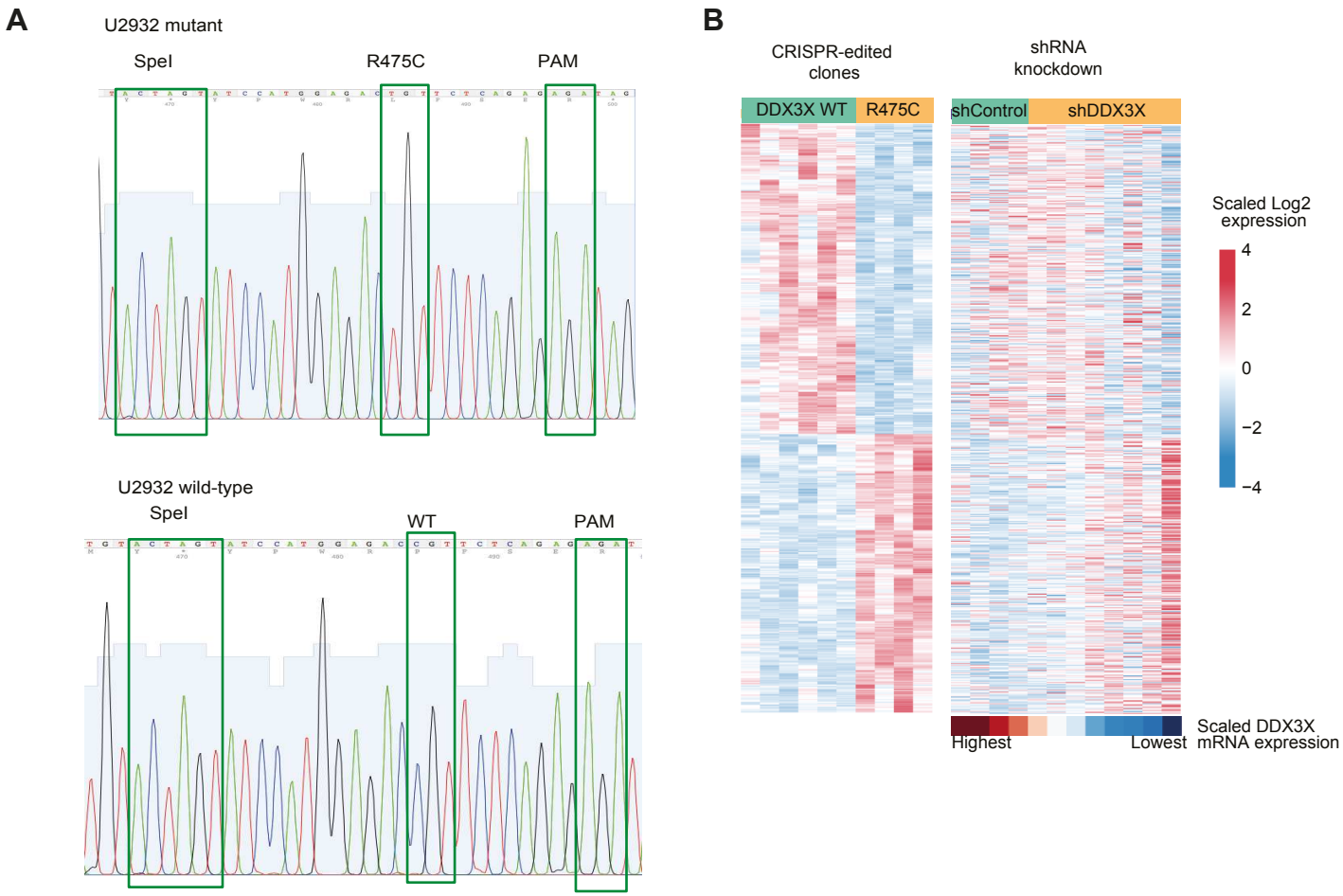
A,B) Quantification of apoptosis (flow cytometry for active caspase-3 and cleaved PARP) and global protein synthesis (OPP assay) in human germinal center B cells **(A)** or the cell line U2932 **(B)** 48 hours after transduction with MYC or empty vector control.

C) Quantification of apoptosis and global protein synthesis in human GC B cells transduced with either BCL6 or control vector.

D) Effect of rapamycin on MYC-induced global protein synthesis in human GC B cells.

E,F) Representative flow cytometry plots showing of MYC-induced apoptosis in human GC B cells treated with 50nM rapamycin **(E)**, or transduced with dominant negative DDX3X (K230E) **(F)**.

G,H) Bar chart showing the effect of the DDX3X R528C helicase-domain mutant on MYC-induced global protein synthesis **(G)** and MYC-induced apoptosis **(H)**. Mean+SEM, *$p<0.05$ independent t-test, n=4.



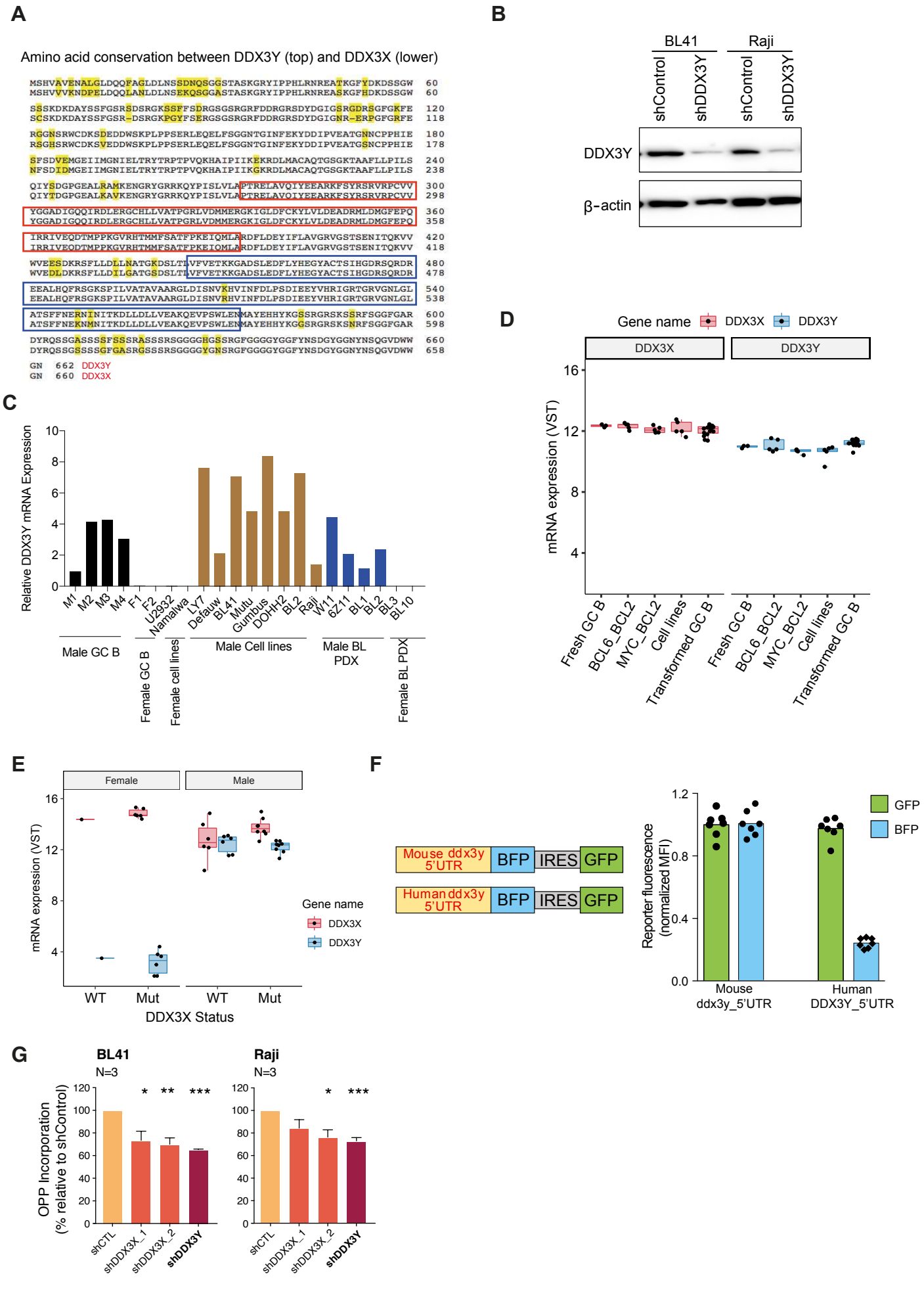
Supplementary Figure 6.

Figure S6. Generation and analysis of DDX3X R475C edited clones. Related to Figure 6.

A) Sanger sequencing of CRISPR-edited DDX3X R475C and control U2932 clones.

B) Heatmap showing genes that are differentially expressed between control and homozygous R475C edited clones (left). The same genes are shown for shRNA knockdown (right). For shRNA experiments samples are ordered left to right by DDX3X mRNA expression. The top bar indicates the expression of DDX3X mRNA showing how stronger knockdown recapitulates the signature seen in R475C-edited clones.

C) Heatmap showing mRNA expression in RNA-Seq of GSEA core enrichment genes in the gene set Hallmark Unfolded Protein Response in BL biopsies in two published Burkitt lymphoma RNA-Seq data sets and in U2932 DDX3X R475C edited clones.



Supplementary Figure 7.

Figure S7. Upregulation of DDX3Y in male lymphoma samples. Related to Figure 7.

A) Alignment of DDX3X protein and DDX3Y protein. Position of N and C terminal helicase domains are highlighted in red and blue boxes respectively.

B) Immunoblot showing knockdown of DDX3Y by shRNA in the male lymphoma cell lines BL41 and Raji.

C) qRT-PCR for DDX3Y by in human GC B cells, lymphoma cell lines and BL xenografts normalized to PGK1.

D) Analysis of published RNA-Seq data (Caeser et al. 2019) showing expression of DDX3X or DDX3Y in freshly purified human GC B cells, cultured GC B cells transduced with either BCL6-2A-BCL2 or MYC-2A-BCL2 (5 days), established lymphoma cell lines, or GCB cells transformed following oncogene transduction and engraftment into immunodeficient mice. All data is from male donors.

E) Analysis of published RNA-Seq data (Grande et al, 2019) showing expression of DDX3X and DDX3Y in Burkitt lymphoma patients stratified by *DDX3X* mutation status.

F) Design of a bicistronic fluorescent reporter to assess the contribution of the 5'UTR to the expression of human and murine DDX3X. Bar chart shows expression of GFP and BFP when reporter was transduced into primary GC B cells (4 donors) and three cell lines.

G) Global protein synthesis measured by OPP incorporation following shRNA knockdown of DDX3Y in male cell lines BL41 and Raji. Bars show mean+SEM. *p<0.05, **p<0.01, ***p<0.001, two tailed students' t-test, n=3.