Supplementary material for:

MicroRNA-155 Controls Vincristine Sensitivity and Predicts Superior Clinical Outcome in Diffuse Large B-cell Lymphoma

Hanne Due, Anna Amanda Schönherz, Laura Ryø, Maria Nascimento Primo, Ditte Starberg Jespersen, Emil Aagaard Thomsen, Anne Stidsholdt Roug, Min Xiao, Xiaohong Tan, Yuyang Pang, Ken H Young, Martin Bøgsted, Jacob Giehm Mikkelsen, and Karen Dybkær.

Supplementary Tables

Supplementary Table 1. Cell line specifications.

Cell line	AUC	Vincristine class	ABC/GCB	Endogenous miR-155
OCI-Ly19	53.99	Sensitive	UC	11.79
FARAGE	56.11	Sensitive	GCB	13.01
SU-DHL-5	57.86	Sensitive	GCB	11.20
MC-116	61.97	Intermediate	GCB	12.48
NU-DHL-1	70.72	Intermediate	ABC	10.40
OCI-Ly3	74.83	Intermediate	ABC	14.18
HBL-1	84.53	Intermediate	ABC	13.65
U2932	85.34	Intermediate	GCB	12.70
NU-DUL-1	90.25	Intermediate	UC	11.02
RIVA	108.99	Intermediate	ABC	12.32
OCI-Ly7	114.49	Resistant	GCB	6.66
SU-DHL-8	126.06	Resistant	GCB	8.14
DB	130.75	Resistant	GCB	8.10

DLBCL cell lines DB, NU-DHL-1, NU-DUL-1, MC-116, and SU-DHL-5 were purchased from DSMZ (German Collection of Microorganisms and Cell Cultures), while FARAGE, HBL-1, OCI-Ly3, OCI-Ly7, OCI-Ly19, RIVA, SU-DHL-8, and U2932 were kindly provided by Dr. Jose A. Martinez-Climent (Molecular Oncology Laboratory, University of Navarra, Pamplona, Spain). The cell lines are ranked according to vincristine sensitivity based on area under dose-response curve (AUC).[1] Division into tertiles defines 3 sensitive, 7 intermediate, and 3 resistant cell lines. Based on GEP, DLBCL cell lines were classified into ABC/GCB subclasses by Wright classification using published algorithms at hemaClass.org.[2] Endogenous miR-155 expression levels are measured by GeneChip miRNA 1.0.2 arrays. Data is RMA normalized, thus the specified values are in log2

scale. For technical validation of miRNA array data, the expression levels of miR-155 were determined by RT-qPCR (data not shown).

Supplementary Table 2. Patient characteristics.

Characteristic	In-house cohort	Meta-cohort
No. of patients	73	701
Sex		
Female	30 (41%)	297 (42%)
Male	43 (59%)	404 (58%)
Age at diagnosis		
Median	64	62
Range	20-87	17-82
IPI score		
0-1	21 (29%)	71 (10%)
2-3	36 (49%)	314 (45%)
4-5	12 (16%)	201 (29%)
NA	4 (6%)	115 (16%)
ABC/GCB		
ABC	32 (44%)	302 (43%)
GCB	32 (44%)	281 (40%)
UC	9 (12%)	118 (17%)

IPI, International prognostic index; NA, not available; UC, unclassified.

Supplementary Table 3. Vincristine response specific miRNAs.

miRNA	P-value	Log2(Fold change)	Fold change	Fold change description
hsa-miR-155	0.0008	4.37	20.67	Resistant down vs. sensitive
hsa-miR-148a	0.03	3.14	8.82	Resistant down vs. sensitive
hsa-miR-21	0.01	2.79	6.90	Resistant down vs. sensitive
hsa-let7b	0.02	2.75	6.75	Resistant down vs. sensitive
hsa-miR-21-star	0.006	2.29	4.90	Resistant down vs. sensitive
hsa-miR-23a	0.03	2.07	4.21	Resistant down vs. sensitive
hsa-miR-501-3p	0.04	1.99	3.98	Resistant down vs. sensitive
hsa-miR-24	0.02	1.89	3.71	Resistant down vs. sensitive
hsa-let7c	0.01	1.45	2.73	Resistant down vs. sensitive
hsa-miR-550	0.02	1.32	2.50	Resistant down vs. sensitive
hsa-miR-378-star	0.006	1.20	2.29	Resistant down vs. sensitive
hsa-miR-658	0.04	1.19	2.28	Resistant down vs. sensitive
hsa-miR-675	0.01	1.09	2.13	Resistant down vs. sensitive
hsa-miR-484	0.05	1.14	2.20	Resistant up vs. sensitive
hsa-miR-223	0.009	1.27	2.41	Resistant up vs. sensitive

Differentially expressed miRNAs detected comparing global miRNA expression profiles of vincristine sensitive and resistant DLBCL cell lines (p-values ≤ 0.05 and FC >|2|).

Supplementary Table 4. Enriched gene sets identified through GSEA analysis conducted for miR-155 vs. control

Gene sets enriched in miR-155	Total size	ES	NES	NOM p-val	FDR q-val	Enrichment Rank (ES)
E2F_TARGETS	188	0.8	2.83	0	0	2223
MYC_TARGETS_V1	171	0.78	2.74	0	0	2596
G2M_CHECKPOINT	189	0.74	2.63	0	0	2223
MYC_TARGETS_V2	55	0.84	2.5	0	0	1574
MTORC1_SIGNALING	185	0.53	1.91	0	0	2326
HYPOXIA	189	0.52	1.83	0	0	1880
GLYCOLYSIS	191	0.51	1.82	0	0	2538
SPERMATOGENESIS	128	0.53	1.81	0	0	2125
DNA_REPAIR	133	0.52	1.79	0	0	2328
MITOTIC_SPINDLE	195	0.47	1.7	0	0.001	1714
OXIDATIVE_PHOSPHORYLATION	175	0.46	1.62	0	0.003	3042
ESTROGEN_RESPONSE_LATE	194	0.38	1.38	0.004	0.039	3332
UV_RESPONSE_UP	153	0.39	1.36	0.012	0.044	2845
Gene sets enriched in control	**Total size**	**ES**	**NES**	**NOM p-val**	**FDR q-val**	**Enrichment Rank (ES)**
MYOGENESIS	193	-0.35	-1.23	0.052	0.117	3423
IL2_STAT5_SIGNALING	189	-0.35	-1.26	0.032	0.154	1693
APICAL_JUNCTION	194	-0.36	-1.26	0.041	0.166	2391
KRAS_SIGNALING_DN	186	-0.36	-1.29	0.029	0.171	3077
APOPTOSIS	146	-0.39	-1.35	0.017	0.108	1833
P53_PATHWAY	192	-0.38	-1.36	0.013	0.108	2756
PI3K_AKT_MTOR_SIGNALING	99	-0.43	-1.41	0.023	0.082	1345
IL6_JAK_STAT3_SIGNALING	83	-0.52	-1.68	0.002	0.004	2056
ALLOGRAFT_REJECTION	183	-0.52	-1.85	0	0	2002
INTERFERON_GAMMA_RESPONSE	180	-0.55	-1.93	0	0	1983
INTERREFON_ALPHA_RESPONSE	86	-0.67	-2.16	0	0	1982

Gene set enrichment analysis (GSEA) was conducted for transcriptional profiles of SU-DHL-5 cells transduced with LV/miR-155 and LV/MCS, respectively. GSEA was restricted to genes sets included in the *Hallmark* collection (50 gene sets) from the Molecular Signature Database. Gene sets with normalized p-value≤0.05 and FDR q-value≤0.25 were considered significantly enriched. Only significant gene sets were shown in the table. Abbreviations: Total size, number of genes included in the gene set; ES, enrichment score; NES, normalized enrichment score; NOM p-val, multiple test corrected p-value for gene set size normalized ES; FDR q-val, false discovery rate of normalized ES.

Supplementary Table 5. Enriched gene sets identified through GSEA analysis conducted for TuD-155 vs. control

Gene sets enriched in TuD	Total size	ES	NES	NOM p-val	FDR q-Val	Enrichment Rank (ES)
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Gene sets enriched in control	**Total size**	**ES**	**NES**	**NOM.p.val**	**FDR.q.Val**	**Enrichment Rank**
XENOBIOTIC_METABOLISM	191	-0.33	-1.28	0.043	0.139	2335
MITOTIC_SPINDLE	195	-0.33	-1.29	0.033	0.134	3184
UV_RESPONSE_UP	153	-0.36	-1.37	0.015	0.068	3270
P53_PATHWAY	192	-0.37	-1.42	0.007	0.045	3051
ADIPOGENESIS	187	-0.39	-1.53	0.001	0.014	4682
MYC_TARGETS_V2	55	-0.53	-1.73	0	0.001	3681
FATTY_ACID_METABOLISM	144	-0.46	-1.74	0	0.001	3457
DNA_REPAIR	133	-0.47	-1.74	0	0.001	4388
MTORC1_SIGNALING	185	-0.45	-1.76	0	0.001	3914
OXIDATIVE_PHOSPHORYLATION	175	-0.48	-1.84	0	0	5592
HYPOXIA	189	-0.49	-1.9	0	0	2180
GLYCOLYSIS	191	-0.49	-1.92	0	0	2787
G2M_CHECKPOINT	189	-0.54	-2.1	0	0	2787
MYC_TARGETS_V1	171	-0.59	-2.29	0	0	3600
E2F_TARGETS	188	-0.65	-2.51	0	0	3463

Gene set enrichment analysis (GSEA) was conducted for transcriptional profiles of SU-DHL-5 cells transduced with LV/TuD-155 and LV/MCS, respectively. GSEA was restricted to genes sets included in the *Hallmark* collection (50 gene sets) from the Molecular Signature Database. Gene sets with normalized p-value≤0.05 and FDR q-value≤0.25 were considered significantly enriched. Only significant gene sets were shown in the table. Abbreviations: Total size, number of genes included in the gene set; ES, enrichment score; NES, normalized enrichment score; NOM p-val, multiple test corrected p-value for gene set size normalized ES; FDR q-val, false discovery rate of normalized ES; n.d., not detected.

Supplementary Table 6. Negatively correlated genes associated with cell cycle processes.

		Prediction algorithms				
Gene Symbol	Gene Ontology Biological Process	TagetScan	miRDB	mircroRNA .org	MicroT-DCD	TarBase
WEE1	G2/M transition	+	+	+	+	IP, RA
CAB39 *RPS6KB1*	Cell cycle arrest G1/S transition	+	+	+	+	-
GSK3B	Re-entry into mitosis	+	-	+	+	PR
PARD3B *TFDP2*	Mitotic cell cycle	+	-	-	+	-
C7orf25 *PAK2* *RBBP4*	G1/S transition Mitotic cell cycle G2/M transition	-	-	-	+	-
38961, APPL1 CABLEES1, CCPG1, HAPECAM2, MAPRE2, PVRIG, RABGAP1, RASSF2, SGSM3,	Cell cycle	-	-	-	-	-
CSNK2A1, FAM89B, GARASP1, HISTH14, MAU2, MCPH1, NEK6, NUP188, NUP214, POLD4	Mitotic cell cycle	-	-	-	-	-
CLIP1, KIF2B	Mitotic cell cycle Microtubule	-	-	-	-	-
CAMK2D, CRLF3, CUL3, EIF4E, ITGB1, MARK4, PIM2, PPP3CA, PPP6C, PSMB9, PSMD9, SPDYA, UBA52, VIL1	G1/S transition	-	-	-	-	-
ANAPC10, CDK5RAP, CSNK1D, DCTN2, DYNC1|2, ENSA, EP300, FBXL15, PCM1, PPM1D, PPP1R12A, PPP2R2A, SDCCAG8, STK16, TUBG2, TUBGCP6	G2/M transition	-	-	-	-	-
CCNG2	Cell cycle checkpoint	-	-	-	-	-
CAB39L, SESN2	Cell cycle arrest	-	-	-	-	-

GEP of SU-DHL-5 cells transduced with LV/miR-155, LV/TuD-155, and the comparable negative control LV/MCS were investigated for identifying negatively correlated gene expressions. The 64 genes related to cell cycle processes were investigated as potential miR-155 targets using four well-documented miRNA-mRNA prediction algorithms: TargetScan v7.1, miRDB, microRNA.org, and MicroT-CDC.[3–6] Published experimental

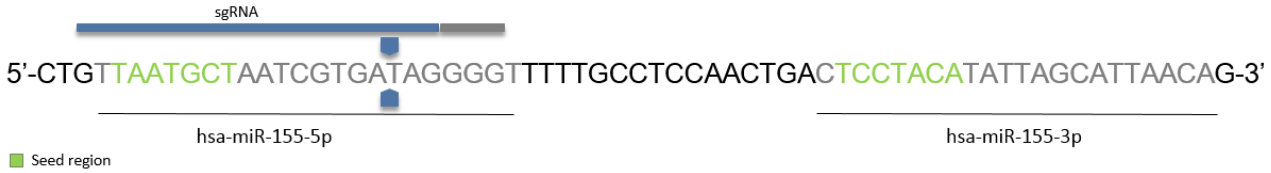
validated miRNA-mRNA interactions were identified from TarBase v7.0.[7] IP, immunoprecipitation; PR, proteomics; RA, reporter assay; - not identified; + identified. If more than one gene is noted in the Gene Symbol Column, +/- symbol is related to all the genes.

Supplementary Table 7. Simple and multiple Cox regression analyses conducted for: (A) all DLBCL patients, (B) ABC classified patients, and (C) GCB classified patients included in the R-CHOP restricted meta-cohort.

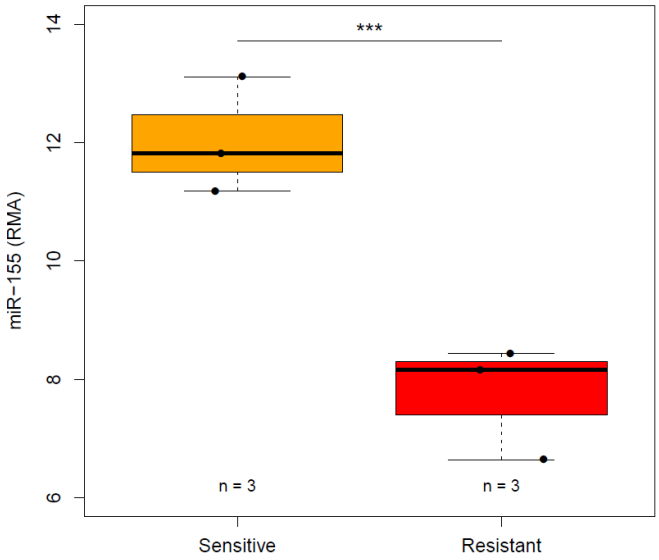
				Simple			Multiple		
		n	no.	HR	95% CI	*P*	HR	95% CI	*P*
A.									
All DLBCL	IPI								
	0-1	71	6	1			1		
	2-3	314	84	3.76	1.64-8.61	0.0017	3.29	1.43-7.56	0.0050
	4-5	201	106	9.61	4.22-21.88	7.19e-08	8.09	3.53-18.54	7.78e-07
	Subclass								
	ABC	242	105	1			1		
	GCB	248	58	0.45	0.33-0.62	1.09e-06	0.41	0.28-0.59	1.37e-06
	UC	96	34	0.73	0.50-1.08	0.11	0.52	0.34-0.81	0.0032
	miR-155								
	Continuous	586	197	0.95	0.86-1.04	0.28	-	-	-
B.									
ABC-DLBCL	IPI								
	0-1	15	2	1					
	2-3	125	38	3.57	0.86-14.89	0.080	-	-	-
	4-5	102	65	10.37	2.51-42.87	0.0012	-	-	-
	miR-155								
	Continuous	242	105	0.85	0.71-1.00	0.052	-	-	-
C.									
GCB-DLBCL	IPI								
	0-1	48	2	1			1		
	2-3	140	32	6.22	1.49-25.97	0.012	5.95	1.43-24.86	0.014
	4-5	60	24	12.53	2.96-53.06	0.0006	11.37	2.68-48.26	0.00098
	miR-155								
	Continuous	248	58	0.72	0.60-0.86	0.00044	0.74	0.62-0.89	0.0016

Array-based miR-155 expression (continuous) and outcome were analyzed by simple and multiple Cox proportional hazards regression analyses for overall survival. IPI score information was not available for all patients, thus cohort sizes are reduced in this setting (115 samples were removed). The multiple Cox proportional regression analysis was performed using an additive model with IPI (trichotomized; IPI 0-1, IPI 2-3, IPI 4-5), ABC/GCB (ABC, GCB, UC), and miR-155 expression (continuous) as independent confounders. Abbreviations: CI, 95% lower and upper confidence intervals; HR, hazard ratio; n, number of samples; no., number of events; - value not available since variables were only included in multiple Cox proportional hazards regression analysis if significant results were obtained in simple Cox proportional hazards regression analysis.

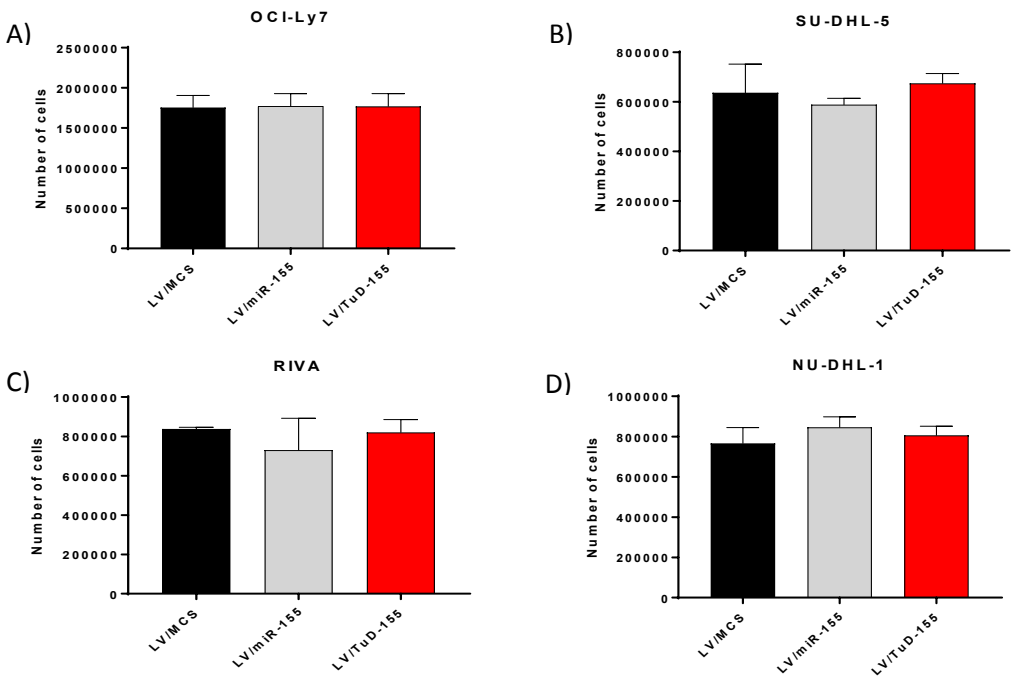
Supplementary Figures



Supplementary Figure 1. Schematic presentation of sgRNA for miR-155 knock-out by the CRISPR-Cas9 technology. The single guide RNA (sgRNA) targets the functional part of miR-155-5p of the miR-155 encoding gene *MIR155HG.*[8] The expected Cas9 cut site is marked by ▼. The sequences encoding miR-155-5p and miR-155-3p (miR-155-star) are marked in grey and seed regions are marked in green.

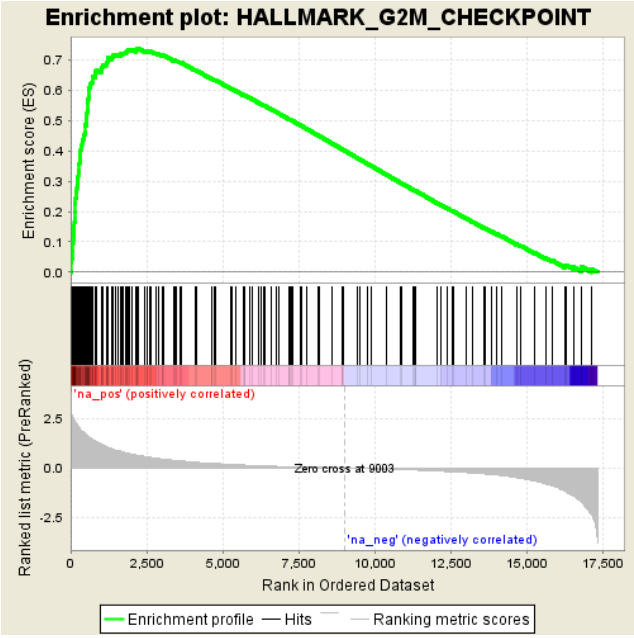


Supplementary Figure 2. miR-155 expression in DLBCL cell lines. Expression levels of miR-155 in vincristine sensitive (OCI-Ly19, FARAGE, SU-DHL-5) and resistant DLBCL cell lines (OCI-Ly7, SU-DHL-8, DB) determined by GeneChip miRNA 1.0.2 microarrays. *** $p<0.001$

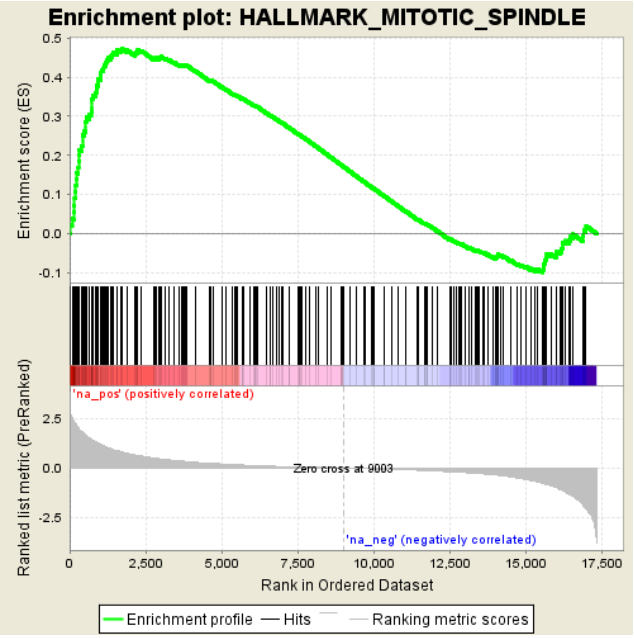


Supplementary Figure 3. Cell growth post-transduction. Cell proliferation was determined by the trypan blue exclusion method after 48 hours of growth in (A) OCI-Ly7, (B) SU-DHL-5, (C) RIVA, and (D) NU-DHL-1 cells transduced with LV/miR-155, LV/TuD-155, and LV/MCS. These cell lines were chosen for *in vitro* studies based on ABC/GCB classification, resistance/sensitivity to vincristine, and lentiviral transducibility.[9] The latter is of great importance, since B-cells generally are difficult to transduce.

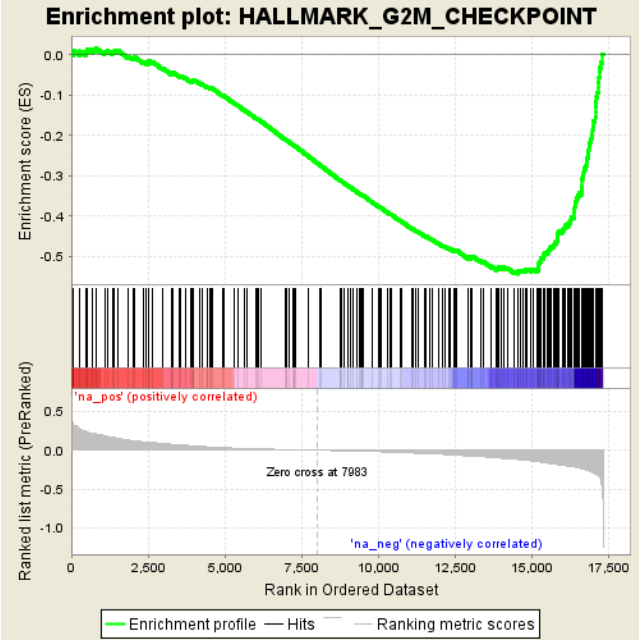
A)



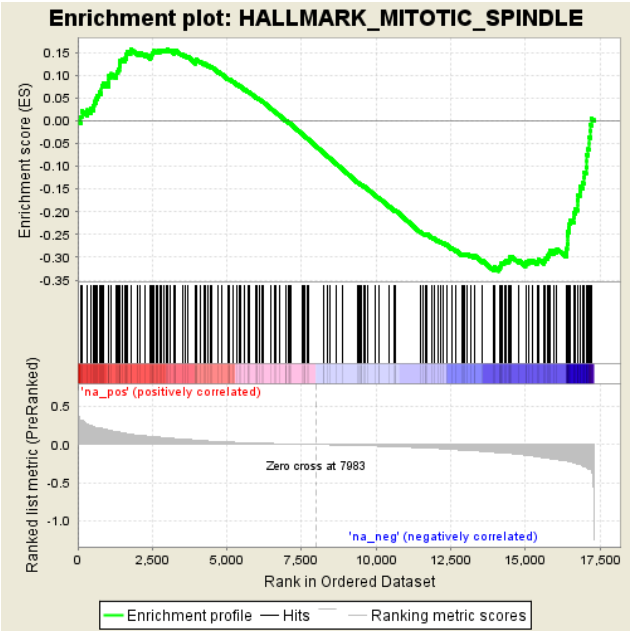
B)



C)

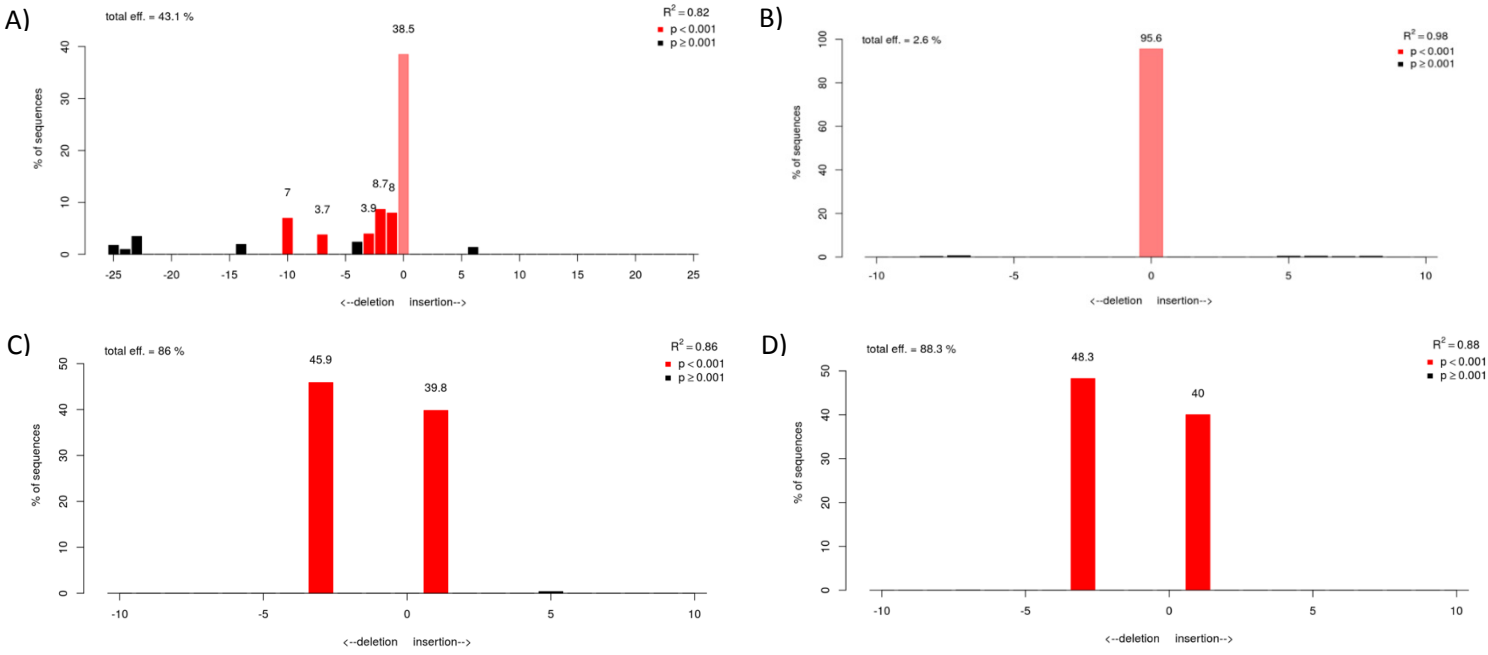


D)

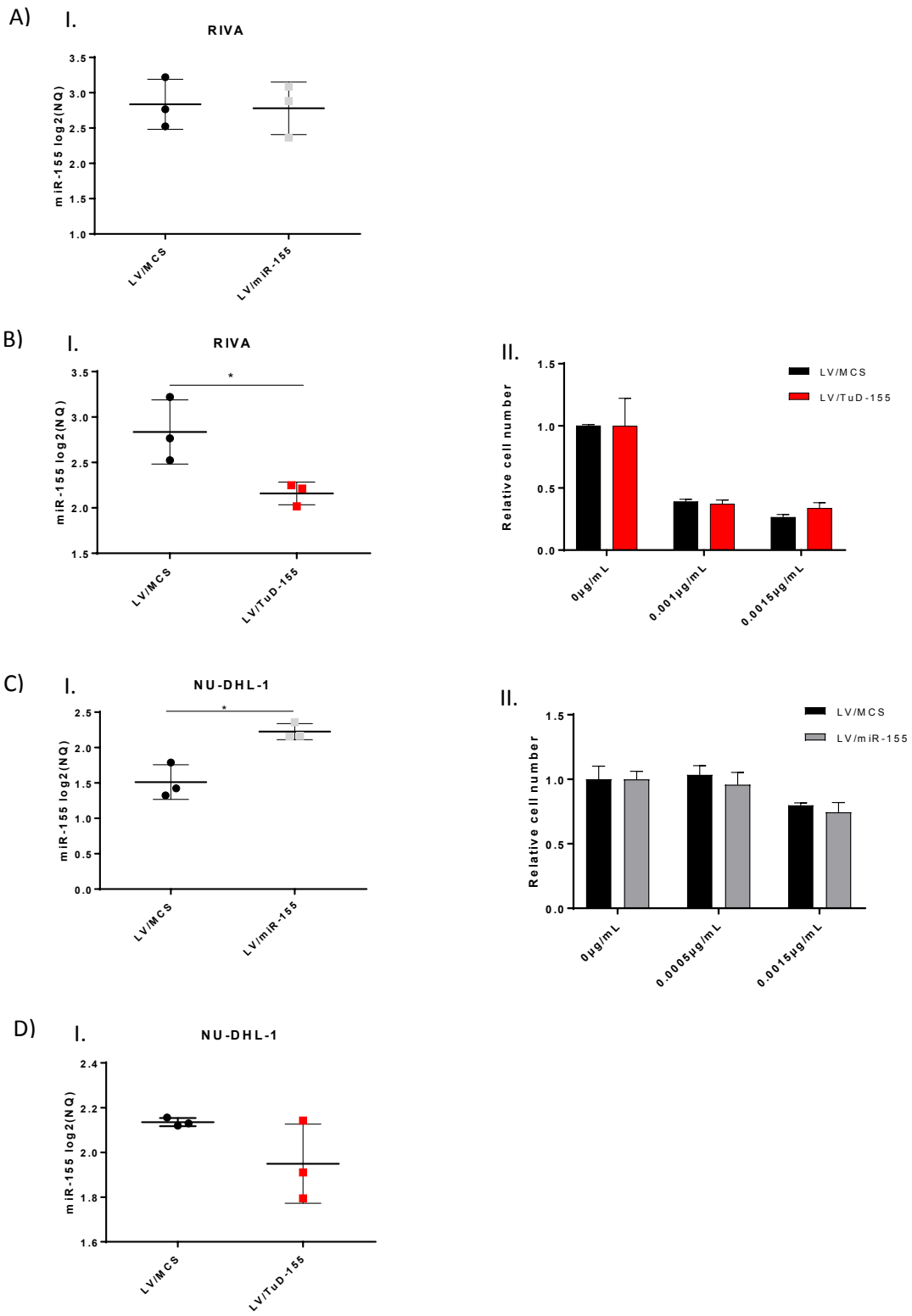


Supplementary Figure 4. Enrichment plots generated by GSEA analysis for (A+B) miR-155 vs. control and (C+D) TuD-155 vs. control. GSEA was conducted for transcriptional profiles of SU-DHL-5 cells transduced with LV/miR-155, LV/TuD-155 and LV/MCS (control), respectively. GSEA was restricted to genes sets included in the *Hallmark* collection (50 gene sets) from the Molecular Signature Database. Gene sets with normalized p-value≤0.05 and FDR q-value≤0.25 were considered significantly enriched. Significance of each enrichment score was calculated by 2000 permutation tests. Since vincristine functions as an antimitotic drug, only

significantly enriched gene sets associated with G2/M checkpoints and mitotic spindle assembly were depicted. The green line represents the running-sum statistic used to calculate the enrichment score (ES) of the gene set. The ES is represented as maximum deviation from zero encountered in the running-sum statistic. The vertical black bars beneath the enrichment score curve indicate the positions of gene set members and their expression profile (red, upregulated; blue, downregulated).

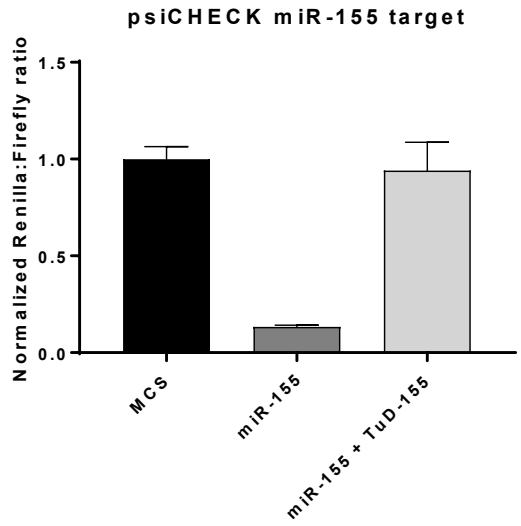


Supplementary Figure 5. Indel frequencies. Indel frequencies determined by TIDE analysis in: (A) OCI-Ly7 cells transduced with LV/CRISPR-sgRNA-miR-155. (B) OCI-Ly7 cells transduced with LV/CRISPR-sgRNA-control. (C+D) miR-155 knock-out clone 1 and 2, respectively, which are characterized by a single nucleotide insertion on one allele and a 3 nucleotide deletion on the other.

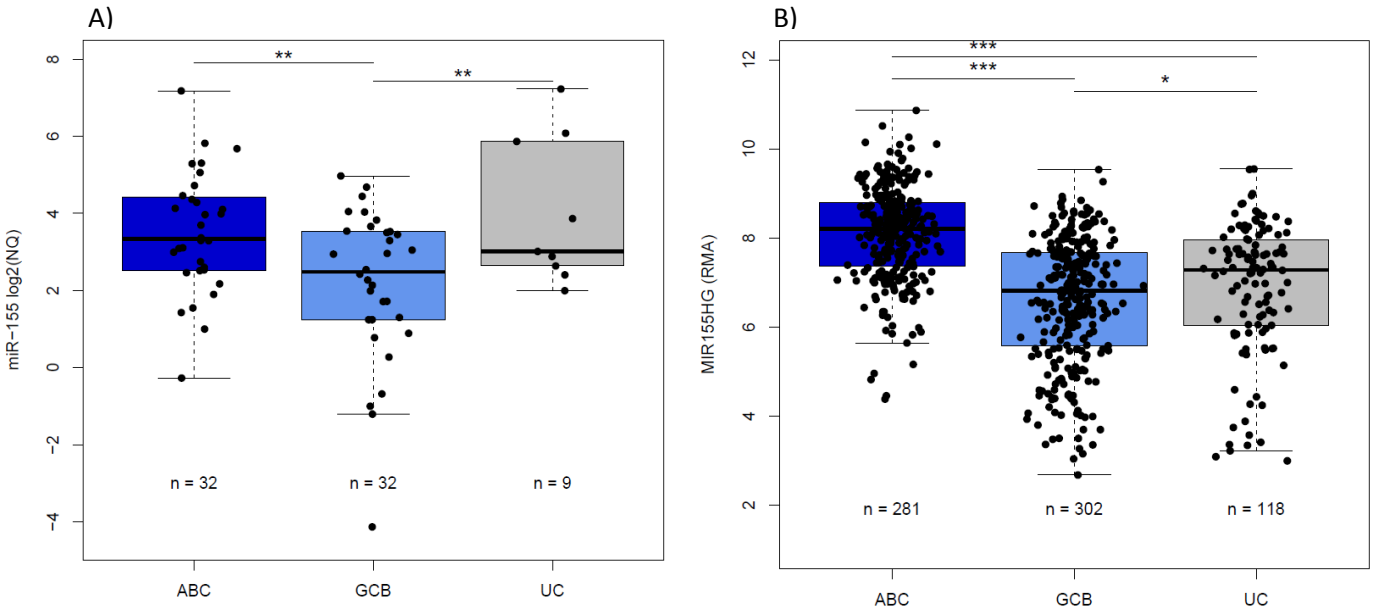


Supplementary Figure 6. Manipulation of miR-155 expression in ABC-DLBCL cell lines do not affect vincristine response. (AI+BI) Expression levels of miR-155 were determined upon lentiviral transductions of

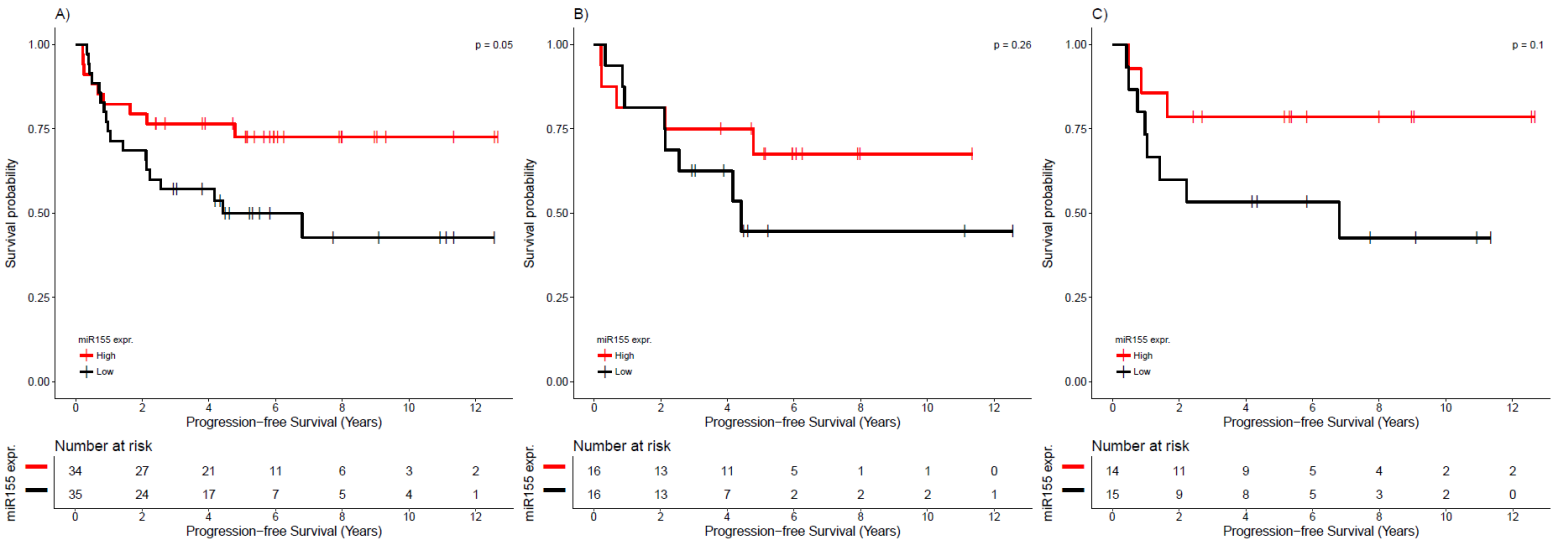
RIVA cells. (BII) Vincristine dose-response analysis was performed for TuD-155 transduced cells, since miR-155 was significantly down-regulated. (CI+DI) Similarly, NU-DHL-1 cells were transduced and miR-155 expression was measured and (CII) vincristine response was investigated in miR-155 over-expressing cells (LV/miR-155). Vincristine response is shown as number of cells relative to the no-drug condition. NQ, normalized quantity.



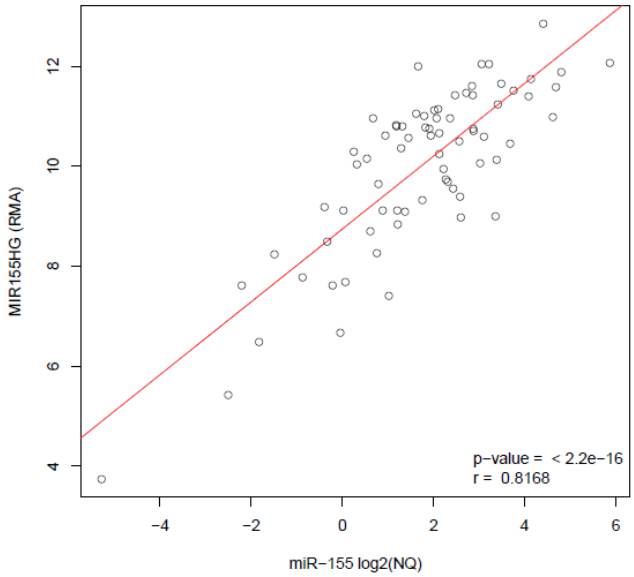
Supplementary Figure 7. Suppression of miR-155 by TuD-155. Dual luciferase reporter assays were performed in HEK293 cells co-transfected with psiCHECK-miRtarget, pCCL/U1-miRNA.PGK-eGFP and pCCL/PGK-eGFP-TuD.



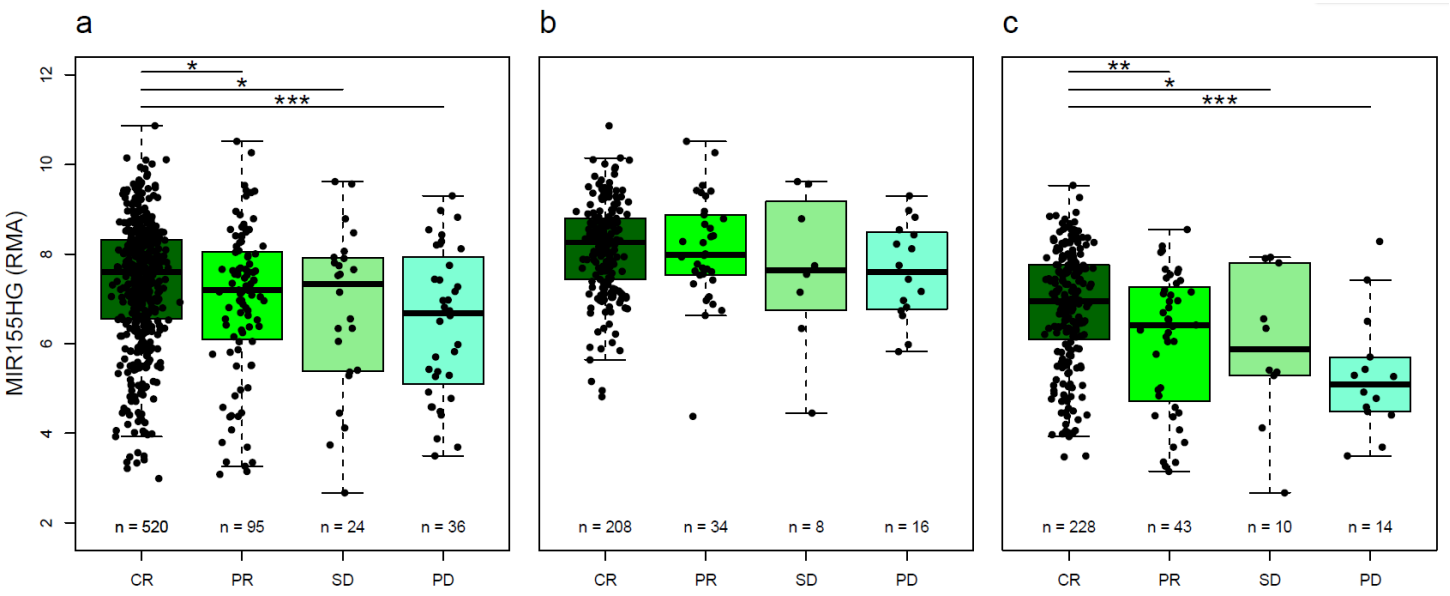
Supplementary Figure 8. miR-155 expression in ABC/GCB subclasses of DLBCL. (A) Expression levels of miR-155 in the ABC/GCB subclasses of the in-house cohort measured by RT-qPCR. (B) Expression levels of *MIR155HG* determined by microarray (Human Genome U133 Plus 2.0) in the meta-cohort. Significance levels: * p<0.05, ** p<0.01, ***p<0.001. NQ, normalized quantity; RMA, Robust Multichip Average normalized.



Supplementary Figure 9. Analysis of association between progression-free survival and miR-155 expression. Kaplan-Meier plots depicting progression-free survival of R-CHOP treated DLBCL patients in the in-house cohort. The analysis was conducted for (A) all DLBCL patients, (B) ABC classified patients, and (C) GCB classified patients. For each cohort, patients were dichotomized by median split of miR-155 expression.



Supplementary Figure 10. Correlation analysis between *MIR155HG* and miR-155 expression. In the in-house cohort of 73 samples, the mature miR-155 expression measured by RT-qPCR was well correlated to its precursor *MIR155HG* measured by microarray (HG-U133; 229437_at) (r=0.8, p<2.2e-16). NQ, normalized quantity; RMA, Robust Multichip Analysis normalized.



Supplementary Figure 11. MIR155HG expression levels in Cheson response evaluation classes. Response evaluations of patients in the validation cohort were extracted and investigated for association to MIR155HG expression. The analysis was performed for (A) all DLBCL patients, (B) ABC classified patients, and (C) for GCB classified patients. Significance levels: * p<0.05, ** p<0.01, ***p<0.001. CR, complete remission; PD, progressive disease; PR, partial remission; RMA, Robust Multichip Average normalized; SD, stable disease.

Supplementary References

1. Falgreen S, Dybkær K, Young KH, et al. Predicting response to multidrug regimens in cancer patients using cell line experiments and regularised regression models. *BMC Cancer*. 2015;15(1):235.

2. Falgreen S, Ellern Bilgrau A, Brøndum RF, et al. hemaClass.org: Online One-By-One Microarray Normalization and Classification of Hematological Cancers for Precision Medicine. *PLoS One*. 2016;11(10):e0163711.

3. TargetScanHuman 7.1.

4. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 2015;43(D1):D146–D152.

5. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. *Nucleic Acids Res.* 2007;36(Database):D149–D153.

6. Paraskevopoulou MD, Georgakilas G, Kostoulas N, et al. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* 2013;41(W1):W169–W173.

7. Vlachos IS, Paraskevopoulou MD, Karagkouni D, et al. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res.* 2015;43(D1):D153–D159.

8. sgRNA Designer: CRISPRko.

9. Ranjbar B, Krogh LB, Laursen MB, et al. Anti-Apoptotic Effects of Lentiviral Vector Transduction Promote Increased Rituximab Tolerance in Cancerous B-Cells. *PLoS One*. 2016;11(4):e0153069.