**RRM2B** is frequently amplified across multiple tumor types: non-oncogenic addiction and therapeutic opportunities

*Running Title: RRM2B amplifications in cancer*

Waleed Iqbal\(^\ddagger\), Elena V. Demidova\(^b\ddagger\), Samantha Serra\(^c\), Taha ValizadehAslani\(^d\), Gail Rosen\(^e\) and Sanjeevani Arora\(^f\)*

\(^a\)Cancer Prevention and Control Program, Fox Chase Cancer Center, Philadelphia, PA-19111; School of Biomedical Engineering and Health Systems, Drexel University, Philadelphia, PA-19104; mwi468@gmail.com

\(^b\)Cancer Prevention and Control Program, Fox Chase Cancer Center, Philadelphia, PA-19111; Kazan Federal University, Kazan-420008, Russia; Elena.Demidova@fccc.edu

\(^c\)Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA-19104; samanthaserrao1528@gmail.com

\(^d\)Department of Electrical and Computer Engineering, College of Engineering, Drexel University, Philadelphia, PA-19104; tv349@drexel.edu

\(^e\)Department of Electrical and Computer Engineering, College of Engineering, Drexel University, Philadelphia, PA-19104; glr26@drexel.edu

\(^\ddagger\)Authors contributed equally

\(^f\)*Correspondence:

**Sanjeevani Arora, PhD**

Assistant Professor

Cancer Prevention and Control Program

Fox Chase Cancer Center

333 Cottman avenue

P2149 Young Pavilion Building

Philadelphia, PA 19111

Phone (215) 214-3956

Email: Sanjeevani.Arora@fccc.edu
Abstract

*RRM2B* plays a crucial role in DNA replication, repair and oxidative stress. While germline *RRM2B* mutations have been implicated in mitochondrial disorders, its relevance to cancer has not been established. Here, using TCGA data, we investigated *RRM2B* alterations in cancer. We found that *RRM2B* is highly amplified in multiple tumor types, particularly in *MYC*-amplified tumors, and is associated with increased *RRM2B* mRNA expression. We also observed that the chromosomal region 8q22.3–8q24, is amplified in multiple tumors, and includes *RRM2B*, *MYC* along with several other cancer-associated genes. An analysis of genes within this 8q-amplicon showed that cases that have both *RRM2B*-amplified along with *MYC* have a distinct pattern of amplification compared to unaltered cases or cases that have amplifications in *RRM2B* or *MYC* only. These other 8q-proteins were shown to interact functionally within the RRM2B network of DNA repair, hypoxia and apoptosis regulating proteins. Notably, *RRM2B*-amplified tumors are characterized by mutation signatures of defective DNA repair and oxidative stress, and in some cancers also associated with poor clinical outcome. These findings suggest that some cancers may require RRM2B for cellular survival, providing novel therapeutic opportunities in these cancers.

Keywords

Chromosome 8; 8q-amplicon; *RRM2B*; cancer; *MYC*
**Introduction**

RRM2B plays an important role in regulating replication stress, DNA damage, and genomic stability [1,2]. *RRM2B* encodes a small subunit of p53-inducible ribonucleotide reductase (RNR). The RNR is a heterotetrameric enzyme responsible for the *de novo* conversion of ribonucleotide diphosphates into the corresponding deoxyribonucleotide diphosphates for DNA synthesis, thus playing an important role in maintaining deoxyribonucleotide pools [3]. The large subunit of the RNR complex consists of a dimer of the RRM1 protein, while the small subunit dimer is either RRM2 or RRM2B (varies depending on cellular conditions). P53-dependent induction of RRM2B expression by hypoxia leads to the exchange of the small RNR subunit from RRM2 to RRM2B, forming a new RNR complex that drives basal DNA replication, reduces replication stress, and maintains genomic stability [4,5]. These known functions of RRM2B suggest that *RRM2B* alterations may play a role in tumorigenesis [1,2].

RRM2B is located on chromosome 8q (8q23.1 [6]; in 2018, annotation changed to 8q22.3 [https://useast.ensembl.org/]). Germline missense and loss of function mutations in *RRM2B* have been associated with mitochondrial depletion syndrome (MDS), with distinct but variable clinical phenotypes [7,8]. At present, there are no known *RRM2B* germline alterations associated with cancer risk. However, somatic changes in *RRM2B*, including most typically amplifications, have been observed in multiple cancers including breast, liver, lung and skin cancer [9]. In a survey of the COSMIC database, *RRM2B* emerged as the most highly amplified DNA repair gene [9]. Additionally, TCGA studies of ovarian, breast, liver, and prostate cancer, have found that cases with *RRM2B* copy number variations (CNV) (amplifications and deletions) have decreased overall survival (OS) [9]. Similarly, increased metastasis and poor prognosis were correlated with *RRM2B* overexpression in head and neck cancer [10], esophageal cancer...
and lung sarcomatoid carcinoma [12]. Another study noted that elevated expression of \textit{RRM2B} is associated with better survival in advanced colorectal cancer [13].

\textit{RRM2B} amplification may be driven by selection of \textit{RRM2B} function, or \textit{RRM2B} may be amplified as a passenger, concurrent with selection for a nearby gene with a driver activity in cancer. In breast cancer, multiple genes localized in the 8q12.1 - 8q24.23 interval were found to be amplified, including \textit{RRM2B} [14]. Most \textit{RRM2B} amplifications are accompanied by \textit{MYC} amplifications, and these two genes are located in close proximity [15]. However, \textit{RRM2B} amplifications also occur independent of \textit{MYC} amplifications, albeit at a lower frequency [16]. While multiple studies have observed the amplification of the 8q region, currently the frequency and specificity of these amplifications is not known, and more specifically, the consequence of \textit{RRM2B} amplification with \textit{MYC} or as independent from \textit{MYC} amplification is not clearly understood [14,17-26].

Here, using TCGA data, we found that \textit{RRM2B}-amplified tumors not only exhibit increased \textit{RRM2B} expression in multiple cancers (such as breast, ovarian, and head and neck cancer), but also exhibit distinct mutation signatures. Comparison of tumors bearing single gene amplifications of \textit{RRM2B}, \textit{MYC} versus co-amplifications (\textit{RRM2B} and \textit{MYC}) indicated that \textit{RRM2B} amplifications may independently impact clinical outcomes in some cancers. An analysis of tumors with \textit{RRM2B} and \textit{MYC} co-amplifications or \textit{RRM2B} alone showed that several genes in the 8q22 – 8q24 region are highly amplified. Additionally, RRM2B and these 8q-proteins interact within the same cell regulatory mechanisms of DNA damage response and repair, hypoxia and apoptosis. Based on these results, we hypothesize that while \textit{MYC} may be the cancer driver, co-amplification of \textit{RRM2B} and other 8q-genes may be relevant for cancer cell survival. These finding present opportunities for novel therapeutic targeting strategies (such as those targeting DNA damage response and repair) for tumors carrying \textit{RRM2B} alterations.
Materials and Methods

Analysis of alteration frequencies using TCGA studies

TCGA tumor data was accessed using the cBioPortal website (http://www.cbioportal.org) [60]. Tumors that have been profiled for mutations as well as copy number variants from 30 TCGA studies were downloaded on 6/21/2018. For cancer types with the highest frequency of RRM2B amplifications, cases were segregated based on co-occurrence of TP53 mutations or MYC amplifications. The TP53 alterations were all mutations (missense, and truncating). The truncating mutations in TCGA are frameshift deletions, frameshift insertions, nonsense, splice site. All mutations were analyzed for co-occurrence, including those that were homozygous or heterozygous.

RNA expression analysis

RNA seq. V2 RSEM data was downloaded from OV, BRCA and HNSC studies on 11/7/2018, and 3/19/2020. The RNA seq. data was only analyzed for tumors that were also profiled for copy number variants. Cases were separated based on copy number variants type: deep deletion, shallow deletion, diploid, gain and amplification. For RRM2B there was only a single case of deep deletion out of all (n= 2181 cases from all studies), observed in BRCA, and was removed. For CCNE2, EI3FE, MTDH, MYC, RAD21, TP53INP1 and YWHAZ - there were no cases of deep deletion for any of the genes in the HNSC study, but there was one case each in EI3FE, RAD21 and YWHAZ with deep deletions in the BRCA study, and also one case each in EI3FE and MTDH as well as two cases in RAD21 for the OV study. In the supplementary figure 2, the box plot represents log2 mRNA expression values from RNA seq. V2 RSEM from cases with different copy number variants categories and statistical significance was tested in Graphpad Prism 8.0 using Mann-Whitney Non-Parametric T-test [62,63]. Additionally, plots of
log2 mRNA expression values based on relative linear copy number values were used to test for correlation between increasing copy number change and mRNA expression, then Pearson coefficient [64] and log rank [65] p-values for each cancer type were calculated. Graphpad Prism 8.0 was used for all statistical analysis. For **CCNE2, E13FE, MTDH, MYC, RAD21, TP53INP1 and YWHAZ**: the data were analyzed as above, but are presented as supplemental tables with expression data and statistics.

**Mutation signature analysis**

Tumor whole-exome sequence data for TCGA Pan-Cancer Atlas studies (for OV, BRCA, HNSC and LUAD), which includes spectra of individual tumors, were downloaded from the Synapse platform (synapse.org) on 11/5/2018. Next, IDs of subjects with different cancer types were downloaded from the cBioPortal website (cBioPortal.org), and subjects with amplification in **RRM2B** and **MYC** genes were selected. Presence or absence of these amplifications classified the subjects into four categories: 1) cases with amplification in both **RRM2B** and **MYC**, 2) cases with amplification in **RRM2B**, but not in **MYC**, 3) cases with amplification in **MYC**, but not in **RRM2B**, 4) and cases without amplification in both genes. For each specific cancer type, mutational spectra of each category were calculated by finding the average number of amplifications in each of the 96 mutation classes. To compare differences in these 96 mutation types, one-way ANOVA [66], two-way ANOVA [67,68], and Wilcoxon rank-sum test [67] were used to test for significant differences. In order to eliminate the possibility of false discoveries caused by multiple comparison, in each test, Benjamini–Hochberg correction [69] was applied to each group of 96 p-values (corresponding to 96 mutations). For each cancer type the patients were divided into the following mutation signature groups: 1) cases with amplification in both **RRM2B** and **MYC**, 2) cases with amplification in **RRM2B**, but not in **MYC**, 3) cases with amplification in **MYC**, but not in **RRM2B**, 4) and cases without amplification in both genes, and anova1, anova2 and Wilcoxon rank-sum functions in MATLAB R2018a were used to test for
significance. Each signature type was determined to be significant only if its P-value was below 0.05 in the one-way ANOVA test. The two-way ANOVA was used to distinguish the impact of factors such as co-amplifications with MYC. Additionally, the Wilcoxon rank-sum test was also performed on the data as another analysis of significance.

Survival analysis

Clinical data was downloaded for the OV, HNSC, LUAD and BRCA TCGA studies. Patients were divided into different groups based on having amplifications in RRM2B or MYC, in both or neither. The OV cases were all Stage 3 or 2 for RRM2B and MYC. The clinical data from each group was used to generate overall survival and disease/progression-free survival Kaplan-Meier curves [70] using GraphPad Prism 8.0, and each curve was compared with the other respective curves using the log-rank test [65] and significance was shown based on * for P-value <0.05, ** for P-value <0.01 and *** for P-value <0.001. The significance level was set at p<0.05. Additionally, Wilcoxon tests were also used [71], and significance was shown based on * for P-value <0.05, ** for P-value <0.01 and *** for P-value <0.001. The significance level was set at p<0.05.

Analysis of amplifications in 8q-genes

Data previously used for TCGA alteration analysis (cbioportal.org) were downloaded on 6/22/18 for our analysis of amplifications in 8q-genes. TCGA studies were used for BRCA, OV and HNSC cancers. The genes used for analysis are located on chromosome 8 from the region 8q11.2 – 8q24.3 and have been associated with cancer according to the Cancer Index (www.cancerindex.org). The gene names, their chromosomal location, and nucleotide start positions are in Supplementary Table 1. The amplification pattern of these other 8q-genes was compared in cases with co-amplifications of MYC and RRM2B, cases with single amplifications or unaltered cases. A univariate analysis was conducted to calculate the positive or negative
linear trend for the frequency of gene amplification, assessed at 95% confidence interval. The Pearson correlation coefficient and p-values were computed by using PROC CORR in SAS 9.4. The significance level was set at p<0.05.

Interaction and pathway enrichment analysis

Data for the RRM2B interaction network was downloaded from the BioGRID database [72] (10/25/2018) and analyzed by Cytoscape v. 3.6.1 [73]. Several interactions were added manually based on literature findings (RRM2B-FOXO3 [34], RRM2B-P21 [32], RRM2B-TP73 [74], RRM2B-E2F1 [33], RRM2B-MEK2 [75]).

WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) was used for gene set enrichment analysis to extract biological insights from the genes of interest [40]. The online WebGestalt tool was used, and an over representation analysis was performed. The enrichment results were prioritized by significant p-values, and FDR thresholds at 0.01.

Results

RRM2B is highly amplified in multiple cancers, with an amplification frequency similar to MYC, while alterations in other RNR genes are infrequent

Using cancer cases from TCGA, we observed that, RRM2B is frequently amplified, with the highest percentage observed in ovarian, breast, bladder, and liver cancers (21.54% - 14.5%), and a lower rate of amplifications in multiple other cancers (14% to 0.6%) (Figure 1A). Deletions of RRM2B were only observed in Non-Hodgkin’s lymphoma (~2%). In addition to amplifications, a low frequency of mutations (<2%) were observed in lung, endometrial, esophagogastric, cervical and pancreatic cancers (Figure 1A). Mapping of somatic RRM2B mutations to the RRM2B protein structure indicated the mutations are present in multiple
regions of the protein and do not always cluster at defined functional domains/regions [27-29], unlike the mutations observed in mitochondrial disorders, which typically result in reduced or eliminated function of the RRM2B protein [7,8] (Figure 1B).

Since RRM2B is regulated by TP53 and is typically co-amplified with MYC (Figure 1), we next compared the frequency of RRM2B alterations with alterations in TP53, and MYC. We also compared the frequency of RRM2B alterations along with the members of the RNR complex (RRM1, RRM2 [30]) (Figure 1C). As expected, TP53 was the most altered gene in the studied cancers, however, RRM2B amplifications did not always significantly co-occur with TP53. Interestingly, most cases with TP53 mutations (which were missense, and truncating mutations) did not have RRM2B amplifications (Supplementary Figure 1A). In comparison, a greater number of cases with MYC amplifications were observed to also have RRM2B co-amplifications. For ovarian (OV) and head and neck squamous cell carcinoma (HNSC), this was half of all cases with MYC amplifications, while in breast cancer (BRCA) this was ~90% cases (Supplementary Figure 1B). Finally, RRM1 and RRM2, the functional partners of RRM2B, were infrequently altered and did not co-occur with RRM2B amplifications (Figure 1C).

Tumors with RRM2B amplifications exhibit increased RRM2B expression

Previous studies have not shown whether tumors carrying amplified RRM2B have increased RRM2B expression, which might directly impact its role in cancer. Multiple studies have observed that gene amplifications do not always lead to increased expression [31]. Thus, to confirm increased expression, we used mRNA expression data from three TCGA studies (OV, BRCA, HNSC), and found that tumors with gains and amplifications had significantly increased RRM2B mRNA expression in all three cancer types studied (Supplementary Figure 2). RRM2B amplifications also significantly correlated with an increase in RRM2B mRNA expression in ovarian and breast cancer, and with an increased trend in HNSC (OV: Pearson correlation=0.64, Log rank p-value=0.05; BRCA: Pearson correlation=0.53, Log rank p-
**RRM2B amplification by itself or co-amplification with MYC is in an 8q-amplicon that is present in multiple cancer types**

In multiple cancers, we identified that *RRM2B* and/or *MYC* were amplified as part of an amplicon with multiple other 8q-genes (Supplementary table 1, list of genes and gene ontology classification). We queried the amplifications in OV (Figure 2), BRCA (Supplementary Figure 3), and HNSC (Supplementary Figure 4) cancers in TCGA.

We observed a strong Pearson correlation for amplifications in the 8q11-8q24 region in cases with *RRM2B* and *MYC* co-amplifications. The strongest correlation was observed for ovarian cancer (R²=0.7637, p-value <0.00001, Figure 2A). The cases with either *MYC* or *RRM2B* amplification alone (R²=0.4309, p-value= 0.000273, Figure 2B, and R²=0.0.0503, p-value= 0.905, Figure 2C) or those without either (R²=0.7637, p-value= 0.518, Figure 2D) showed weaker Pearson correlation. These amplifications observed (Figure 2A-C) are in the chromosome segment between *RRM2B* (8q22.3) and *MYC* (8q24.21). This amplicon contains 11 cancer relevant genes: *BACL, ANGPT1, EIF3E, EBAG9, TRSP1, RAD21, EXT1, TNFRSF11B, NOV, HAS2* and *RNF139* (Supplementary table 1).

Next, we examined the overall amplification frequency of 8q genes using TCGA data, without segregating cases based on *RRM2B* and/or *MYC* amplifications. We found, that the 8q11.3-8q24 amplicon is present in multiple tumor types (Figure 3). A positive correlation was observed for increased amplifications in the 8q11-8q24 region, which includes *RRM2B* and *MYC* (Pearson correlation range: R² =0.8976-0.0568). The strongest correlations were observed for skin (SKCM, R²=0.8976, p-value <0.00001), pancreatic adenocarcinoma (PAAD, R²=0.8714, p-value <0.00001), ovarian cancer (OV, R²=0.8582, p-value <0.00001), liver...
hepatocellular carcinoma (LIHC, $R^2=0.7814$, p-value $<0.00001$), and esophageal cancer (ESCA, $R^2=0.7376$, p-value $<0.00001$) (Figure 3). All other correlations were significant at p-value $<0.00001$, except for BLCA (p=0.525). Finally, we also observed that mRNA expression of multiple genes in the amplicon was also significantly increased (Supplementary Table 2).

**RRM2B protein interaction network includes co-amplified 8q-amplicon genes**

Since the expression of *RRM2B*, *MYC* and several other 8q-amplicon genes was increased, we next tested if the products of these genes within the 8q-amplicon interact with RRM2B. RRM2B is regulated by p53 and p21 [32] and transcriptional factors, such as, E2F1[33]. E2F1 regulates RRM2B in the absence of p53, and FOXO3 [34]. It binds the RRM2B promoter and induces its expression. Amplification in RRM2B may further impact the interacting proteins, which include proteins such as ATM, ATR, CHEK1, CHEK2, MDM2. These proteins are important for cell regulatory mechanisms such as DNA damage and response (DDR) pathway, cell cycle and apoptosis. Hypoxia pathway participants, VHL and HIF1α, interact with the above-mentioned DDR proteins, and with p53 and Myc.

We next found that several 8q22-24 gene products (Figure 4, in light purple) also interact or intersect with the proteins in the RRM2B network. For example, YWHAZ or 14-3-3 protein zeta/delta, is an adapter protein, which binds with phosphorylated FOXO3, which retains phosho-FOXO3 in the cytoplasm and prevents its activity as a transcription factor, and potentially prevents apoptosis [35]. Under certain conditions, such as irradiation, ATM-dependent activation of p53 involves dephosphorylation of p53 Ser376 residue, leading to formation of a binding site for YWNAZ, thus, increasing the affinity of p53 to bind with regulatory parts of genes such as *CDKN1A*, *GADD45*, *MDM2*, providing transcriptional regulation [36]. Another 8q-gene product found in the RRM2B network is EIF3E, which is required for several steps in the protein synthesis initiation. IF3E binds with the 40s ribosomal subunit to facilitate recruitment of proteins to initiate protein, and these complexes participate in ribosomal
disassembly [37]. EIF3E interactions with ATM and BRCA1 are essential for execution of DDR, and altered EIF3E has been observed in breast cancer [38]. EIF3E has been suggested to play a role in tumor promotion (due to its interactions with MYC, an oncogene) or tumor suppression (due to interactions with NDRG1, a tumor suppressor protein and a target of MYC) [39].

Finally, we performed a pathway enrichment analysis for all the interacting genes presented in Figure 4. Using WebGestalt tool [40], we found that the most significantly enriched pathways were signal transduction mediated by p53, response to DNA damage and other environmental stimuli, cell cycle checkpoints, DNA replication, response to oxygen levels and apoptotic signaling. The results are presented in Supplementary Table 3. Overall, these data suggest that amplifications in the 8q-region, observed in multiple cancers, may impact cancer cell survival due to their involvement and intersection in important cellular pathways.

RRM2B amplifications correlate with clinical outcome in ER+PR+HER2+ breast cancer

Previous studies have suggested that RRM2B alterations may impact patient prognosis in ovarian, breast, HNSC, and lung cancers [9-12]. However, these studies did not segregate cases based on relevant alterations, such as MYC. Here we compared survival differences in cases segregated based on amplifications in RRM2B alone, RRM2B with MYC or MYC alone in BRCA, OV, HNSC, and lung cancer (Figure 5, and Supplementary Figure 5).

In breast cancer, no significant differences in OS were observed (Figure 5A), which is a different result from a previous study in breast cancer [9]. However, in comparison to OS, significant differences were observed for DFS (Figure 5B). MYC amplification alone had the best DFS, compared to MYC and RRM2B co-amplified cases, and unaltered cases (p<0.0001). Additionally, RRM2B amplification cases tend to have worse DFS than MYC only cases (p=0.0519). To further explore the findings in breast cancer, we analyzed RRM2B amplifications separately in each of the three major subtypes of breast cancer (ER+PR+HER2+ (n=86),...
ER+PR+HER2- (n=334), and triple negative ER-PR-HER2- (n=100)). We observed that patients with ER+PR+HER2+ breast cancer and RRM2B amplifications had a significant decrease in OS (p=0.0178, Figure 5C and 5E), with no difference in DFS (Figure 5D, 5F). Additionally, when comparing OS in all three breast cancer subtypes with RRM2B amplifications, ER+ PR+ HER2+ patients (n=18) had significantly worse OS than ER- PR- HER2- patients (n=26) (p=0.0426). The previous studies that observed worse OS in breast cancer, based on RRM2B amplification, did not consider subtype differences, and may have only included the major breast cancer subtype (ER+PR+HER2+).

In ovarian cancer, individuals with MYC amplification alone were observed to have a significantly better OS compared to those with MYC and RRM2B co-amplification (p=0.0019), or those that were unaltered for both genes (p=0.014) (Supplementary Figure 5A, see legend for case numbers). There were not enough RRM2B cases alone to evaluate statistically, but the data trended towards worse OS. No significant differences were observed in OV cancer for disease-free survival (DFS) (Supplementary Figure 5B). In lung cancer (non-small cell lung cancer or LUAD), significant difference was only observed in DFS for RRM2B only cases compared to unaltered cases. RRM2B amplifications alone had significantly worse DFS compared to unaltered cases (p=0.0057). Overall, there were not enough cases for statistical evaluation. For OS, there were only 7 cases with amplifications in both RRM2B and MYC, 15 cases with amplifications in MYC only, 6 cases with amplifications in RRM2B only, and 195 unaltered cases with no amplifications in either gene (Supplementary Figure 5C). For DFS, there were only 7 cases with amplifications in both RRM2B and MYC, 13 cases with amplifications in MYC only, 6 cases with amplifications in RRM2B only and 162 unaltered cases with no amplifications in either gene (Supplementary Figure 5D). Finally, no significant differences were observed based on RRM2B amplification status with or without MYC amplification, in OS and DFS in HNSC (Supplementary Figure 5E and 5F). Additionally, p-
values for all comparisons (log-rank test) in all the above cancer types is available in Supplementary Table 4.

RRM2B amplifications exhibit a distinct tumor mutation signature but do not impact tumor mutation burden

As RRM2B plays an important role in DNA synthesis and maintaining genomic stability, we also tested whether RRM2B alterations impact tumor mutation burden (TMB). TMB is a biomarker of defective cellular pathways and mutagenic processes in the cell, and is emerging as a potential biomarker of cancer therapies, such as immune checkpoint blockade [41]. We first compared TMB (in OV, BRCA, and HNSC) in cases with and without RRM2B amplification, and did not find any differences in mutation rates, with similar rates observed for both (Supplementary Table 5).

Recently, distinct mutation signatures have been associated with defects in certain genes or pathways, and with certain endogenous and exogenous exposures [42]. Here, we tested whether RRM2B amplifications are associated with specific tumor mutational signatures. For this analysis, we used the recent PanCancer Atlas data, reporting mutational spectra for 9493 individual tumors, including 926 BRCA, 384 OV, and 461 HNSC and 524 LUAD [42]. Interestingly, breast cancer cases with RRM2B amplification alone were significantly associated with T>C and T>A mutations (P<0.05, Figure 6 and Supplementary Table 6).

A similar signature was observed in HNSC, but not in OV (Supplementary Figure 6). The HNSC and ovarian cancer signature could not be tested for significance as only few cases were present in the PanCancer Atlas data. For LUAD, a distinct signature of C>A mutations (significant by one-way ANOVA and two-way ANOVA tests as above, p<0.05; Supplementary Figure 7, and Supplementary Table 6) was observed for cases with only RRM2B amplifications.
The observed mutation signatures for *RRM2B* amplifications were most similar to those that have been recently described for defects in distinct DNA replication and repair pathways [42]. The mutation signatures observed have been associated with defective DNA base excision repair, and DNA mismatch repair. We also observed mutation signatures for presence of reactive oxygen species that can lead to extensive DNA base damage [43].

**Discussion**

Alterations in genes involved in DNA replication and/or repair pathways are known to increase genetic instability, and lead to cancer [9]. *RRM2B* is essential for DNA replication (nuclear and mitochondrial), repair, protection from oxidative damage, and overall maintenance of genetic stability. Despite these known functions, our current understanding of *RRM2B* is limited to its role in MDS. *RRM2B* alterations implicated in MDS are typically mutations that result in the reduction of mitochondrial dNTP pools [44]. The current understanding of *RRM2B* in cancer is very limited, and in this study, we provide a broad analysis of *RRM2B* alterations in cancer and their potential significance.

*RRM2B* is a major downstream target of p53, however, in contrast to p53 loss of function mutations or deletions, we observed that *RRM2B* is highly amplified across multiple tumor types. Interestingly, *RRM2B*, which is present on chromosome 8 is usually co-amplified with *MYC* oncogene (which is also on chromosome 8). Currently, from these observations it is challenging to establish *RRM2B* as a tumor promoter or suppressor, but these data may reflect potential addiction of cancer cells to *RRM2B*, and other genes in the 8q-amplicon observed in multiple cancers in this study. Our data warrant further investigation in cellular models to investigate role of *RRM2B* in cancer. In previous studies, *RRM2B* expression has been identified as a predictive marker of resistance to paclitaxel and gemcitabine in breast cancer, by using machine learning approaches [45]. Similarly, *RRM2B* has been identified as an essential
metabolic gene for renal tumors, and suggested as a potential therapeutic target to disrupt cancer cell metabolism [46].

We also found, in multiple cancers, that tumors with co-amplified RRM2B and MYC exhibit significant amplification of 8q22.3 – 8q24 region of chromosome 8. Interestingly, a previous study in breast cancer found frequent MYC co-amplifications with multiple genes in the 8q chromosomal region. This study suggested that these co-amplifications may explain the aggressive phenotypes of these tumors [14]. In high-grade ovarian serous carcinomas, it was found that MYC was one of the most amplified genes, and patients with such copy number changes had better overall survival [47]. In this study it was not possible to address the problem of inter- and intra- tumor heterogeneity which has implications for therapeutic resistance [48,49]. However, supportively, a recent study in the highly heterogenous ovarian serous grade carcinomas consistently found amplifications (starting from 8q22.3 region) in ovarian tumor biopsies from different regions of the ovary, and also in the relapsed tumor tissue [50]. Future studies addressing inter- and intra- tumor heterogeneity in other cancer types will further highlight the potential clinical significance of amplifications in the 8q-region.

Amplification of RRM2B, along with other genes in the 8q-amplicon could reflect tumor non-oncogenic addiction in multiple tumors [51,52]. While MYC is the main driver for these tumors, the amplification RRM2B and the other 8q-genes may be relevant for cancer cell survival, with clinical and therapeutic implications. A previous study has observed prognostic significance of RRM2B alterations in nasopharyngeal carcinoma [12]. Our analysis of clinical outcomes revealed that based on the cancer type or subtype, co-alteration of RRM2B with MYC, or alone may significantly impact patient outcome in breast cancer. The clinical implications of these data will be further illuminated by large sample size, and analysis of other independent prognostic factors of overall survival. For example, absence of residual tumor post-surgery is a major prognostic factor in advanced ovarian cancer [53].
Finally, we observed that several genes within the 8q-amplicon also interact with the RRM2B-functional network. This suggests that amplifications in this region could have a profound effect on DNA replication, repair, cell cycle and hypoxia. Pathway enrichment analysis suggests the importance of 8q-region genes in response to oxygen levels, DNA replication, and DNA damage response. Future studies in cellular models are needed to validate the findings from the network analysis. Finally, at least in some cancers, RRM2B amplifications may either drive or participate in molecular events that lead to differential tumor mutation signatures. Here, the observed mutation signatures were relevant to the role of RRM2B in DNA replication, and repair pathways. A previous study in mice found that overexpression of RRM genes together with defective mismatch repair pathway can lead to increased mutagenesis, and carcinogenesis [54]. While a regulated expression of RRM2B reduces oxidative stress (in a p53-dependent manner) [55], the impact of RRM2B overexpression on the levels of reactive oxygen species is not known. The observation of mutation signatures associated with increased reactive oxygen species suggest that RRM2B overexpression may impede the oxidative stress pathway. The observed mutation signatures in RRM2B-amplified cases may assist in identifying cancer types or subtypes for precision-targeting of tumors. An example, is combining PARP inhibitors, with existing chemotherapies and irradiation [56-59].

Overall, this study provides an analysis of RRM2B alterations in multiple tumors, majorly reflected as RRM2B amplifications in conjunction with MYC, and other genes on chromosome 8. Future genome-wide studies of other cancer datasets are warranted to confirm the results of the present study. Finally, studies in cellular models are required to delineate the role of RRM2B, and other 8q-chromosome genes in cancer cell maintenance, therapeutic targeting and clinical outcomes.

**Conclusions**
Overall, this study provides an in-depth analysis of *RRM2B* alterations in multiple tumor types, majorly reflected as *RRM2B* amplifications in conjunction with *MYC*, and other genes on chromosome 8. These cases exhibit a distinct 8q-amplification pattern as well as survival outcome differences and mutation signature differences, depending on cancer type and subtypes. Future genome-wide studies of other cancer datasets are warranted to confirm the results of the present study. Finally, studies in cellular models are required to delineate the role of *RRM2B*, and other 8q-chromosome genes in cancer cell maintenance, therapeutic targeting and clinical outcomes.
References


[31] Jia Y, Chen L, Jia Q, Dou X, Xu N, Liao DJ. The well-accepted notion that gene amplification contributes to increased expression still remains, after all these years, a reasonable but unproven assumption. J Carcinog 2016;15:3.


[36] Chernov MV, GR S. The p53 activation and apoptosis induced by DNA damage are reversibly inhibited by salicylate. Oncogene 1997;14:2503-10.


genomic regions of focal and recurrent copy number alteration in 3q26.2 and 8q24.3. Int J Cancer 2019;145:2670-81.


Declaration of Interest

The authors declare no conflict of interest.

Credit authorship contribution statement

**Waleed Iqbal:** Methodology, Formal Analysis, Investigation, Roles/Writing - original draft, Data Curation, Visualization, Validation. **Elena V. Demidova:** Methodology, Formal Analysis, Investigation, Roles/Writing - original draft, Writing - review & editing, Data curation, Visualization, Validation. **Samantha Serrao:** Formal Analysis, Investigation, Roles/Writing - original draft, Data curation. **Taha ValizadehAslani:** Software, Formal Analysis, Data curation, Visualization. **Gail Rosen:** Writing - review & editing, Methodology, Supervision, Funding Acquisition. **Sanjeevani Arora:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Resources, Funding Acquisition.

Acknowledgment

We are express gratitude to Dr. Erica Golemis at the Fox Chase Cancer Center for many helpful discussions and advice with this work.

Ethics Statement

Not applicable to this study.

Funding
W.I. was supported by Fox Chase Cancer Center Risk Assessment Program Funds and NSF #1650531 (to G.R.), S.A. and E.V.D. were supported by DOD W81XWH-18-1-0148 (to S.A.). T.V.D. and G.R. were supported by NSF #1650531 (to G.R).

Abbreviations

BLCA - Bladder urothelial carcinoma
BRCA - Breast invasive carcinoma
CESC - Cervical squamous cell carcinoma and endocervical adenocarcinoma
CNV - Copy number variations
COAD - Colon adenocarcinoma
DDR - DNA damage and response
DFS – Disease-free survival
ESCA - Esophageal carcinoma
HNSC - Head and neck squamous cell carcinoma
KIRC - Kidney renal clear cell carcinoma
LIHC - Liver hepatocellular carcinoma
LUAD - Lung adenocarcinoma
MDS - Mitochondrial depletion syndrome
MESO - Mesothelioma
NHL - Non-Hodgkin lymphoma
OV - Ovarian serous cystadenocarcinoma
OS - Overall survival
PAAD - Pancreatic adenocarcinoma
PCPG - Pheochromocytoma and paraganglioma
PRAD - Prostate adenocarcinoma
RNR - Ribonucleotide reductase
SARC - Sarcoma
SKCM - Skin cutaneous melanoma
STAD - Stomach adenocarcinoma
TMB - Tumor mutation burden
USEC - Uterine corpus endometrial carcinoma
Figure legends

Figure 1. Somatic alterations in *RRM2B* from cbioportal.org. A: Somatic alterations in *RRM2B* across TCGA studies: BLCA (Bladder Urothelial Carcinoma), BRCA (Breast invasive carcinoma), CESC (Cervical squamous cell carcinoma and endocervical adenocarcinoma), COAD (Colon adenocarcinoma), SKCM (Skin Cutaneous Melanoma), USEC (Uterine Corpus Endometrial Carcinoma), ESCA (Esophageal carcinoma), TGCT (Testicular Germ Cell Tumors), HNSC (Head and Neck squamous cell carcinoma), LIHC (Liver hepatocellular carcinoma), MESO (Mesothelioma), NHL (Non-Hodgkin lymphoma), LUAD (Lung adenocarcinoma), OV (Ovarian serous cystadenocarcinoma), PAAD (Pancreatic adenocarcinoma), PCPG (Pheochromocytoma and Paraganglioma), PRAD (Prostate adenocarcinoma), KIRC (Kidney renal clear cell carcinoma), SARC (Sarcoma). Amplifications (grey), mutations (green) and deletions (pink) are represented as percent frequency. B: *RRM2B* mutations in TCGA. 2D RRM2B protein stick figure showing the important domains of RRM2B (N-terminal swivel region, required for dimer stability, grey; ribonucleotide small subunit signature [conserved region in catalytic site between RRM2 and RRM2B], blue, and iron-binding residues required for catalytic activity, red) and the number of somatic mutations in RRM2B. C: Frequency of *TP53, MYC, RRM1, RRM2* and *RRM2B* alterations in ovarian (OV), breast (BRCA) and head and neck (HNSC) cases.

Figure 2. Amplification frequency of 8q-genes in ovarian cancer (OV). A: cases with co-amplification of *RRM2B* and *MYC*; B: cases *MYC* only amplification; C: cases *RRM2B* only amplifications; D: cases with neither (unaltered) were plotted as percent of frequency for amplifications in various 8q-region genes relevant for cancer (see Supplementary Table 1). *RRM2B* (red circle), *MYC* (yellow circle), and other genes (blue circle). The Pearson correlation (R2 value) for the data points is represented by a black trend line.
Figure 3. Amplification frequency of cancer relevant 8q-genes in Skin Cutaneous Melanoma (SKCM), Pancreatic adenocarcinoma (PAAD), Ovarian serous cystadenocarcinoma (OV), Head and Neck squamous cell carcinoma (HNSC), Liver hepatocellular carcinoma (LIHC), Cervical and endocervical cancers (CESC), Esophageal carcinoma (ESCA), Breast invasive carcinoma (BRCA), Stomach adenocarcinoma (STAD), Lung adenocarcinoma (LUAD), Colon adenocarcinoma (COAD), Bladder urothelial carcinoma (BLCA) amplifications were plotted as percent of frequency for amplifications. RRM2B (red circle), MYC (yellow circle), and other genes (blue circle). The Pearson correlation (R^2 value) for the data points is represented by a black trend line.

Figure 4. An interactive network of RRM2B and related genes. The network represents proteins that functionally interact or intersect with RRM2B (yellow) or other 8q-gene products (purple). Red lines represent direct interaction, and green lines – physical association (interactions classified according to the BioGRID database annotation).

Figure 5. Kaplan-Meier curves for overall survival (OS) and disease/progression-free survival (DFS) in breast cancer (BRCA) (A, B) and its subtypes (C-G) studies with RRM2B amplifications and/or MYC amplifications. A: OS in BRCA cases. The cases that were unaltered for both genes (black, n=737), cases with RRM2B amplifications (red, n=14), cases with MYC amplifications (blue, n=62) and co-amplifications (pink, n=210) were plotted. B: DFS in BRCA cases. The cases that were unaltered for both genes (black, n=678), cases with RRM2B amplifications (red, n=12), cases with MYC amplifications (blue, n=55) and co-amplifications (pink, n=184) were plotted. C: OS in ER+ PR+ HER2+ BRCA cases. The cases that were unaltered for RRM2B (black, n=68), cases with RRM2B amplifications (salmon, n=18). D: DFS in ER+ PR+ HER2+ cases. The cases that were unaltered for RRM2B (black, n=63), cases with RRM2B amplifications (red, n=13). E: OS in BRCA subtypes with RRM2B amplifications. ER+ PR+ HER2+ cases with amplifications in RRM2B (purple, n=18), ER- PR-
HER2- cases with amplifications in $RRM2B$ (blue, n=26), and ER+ PR+ HER2- cases with amplifications in $RRM2B$ (peach, n=36). **F:** DFS in BRCA subtypes with $RRM2B$ amplifications. ER+ PR+ HER2+ cases with amplifications in $RRM2B$ (purple, n=13), ER- PR-HER2- cases with amplifications in $RRM2B$ (blue, n=25), and ER+ PR+ HER2- cases with amplifications in $RRM2B$ (peach, n=34). The plots were compared using Log-rank test and significance is shown as follows: P-value *<0.05, ** for P-value <0.01 and *** for P-value <0.001. **G:** P-values for the graphs E and F.

**Figure 6. Mutation signature in breast cancer (BRCA) cases segregated by amplification type.** One-way ANOVA ($RRM2B$ amplifications only versus other groups) and two-way ANOVA (included group with $RRM2B$ and $MYC$ co-amplifications) analysis showed that the T>C and T>G mutations are statistically significant. **Top panels:** tumor whole-exome sequence data from the PanCancer Atlas studies was used to calculate the average frequency of the 96 trinucleotide context mutations in each group: unaltered cases, cases with only $RRM2B$ amplifications, cases with only $MYC$ amplifications, and cases with amplifications in both. **Bottom panels:** The statistical significance of each comparisons is represented by ANOVA tests as $-\log_{10}$(P-value) for each of the 96 trinucleotide context mutations. The $-\log_{10}$(P) visualizations are provided for: one-way ANOVA comparing $RRM2B$ only group to the all other groups, a two-way ANOVA comparing all groups with $RRM2B$ or $MYC$ amplifications. $-\log_{10}$(P) - the taller the bars are, the lower is the P-value and when the bars exceed the red line, P-values are less than 5%. P-values of each subfigure have been corrected using Benjamini–Hochberg procedure.
**A**

**RRM2B alterations**

Key (A, C):
- Deletions
- Amplifications
- Mutations

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>BLCA</th>
<th>BRCA</th>
<th>CESC</th>
<th>COAD</th>
<th>SKCM</th>
<th>USEC</th>
<th>ESCA</th>
<th>TGCT</th>
<th>Glioma</th>
<th>HNSC</th>
<th>LHC</th>
<th>Melanoma</th>
<th>CESC</th>
<th>GC</th>
<th>GESCATGCTGlioma</th>
<th>HNSC</th>
<th>N=504</th>
<th>OV</th>
<th>N=311</th>
<th>BRCA</th>
<th>N=962</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Frequency</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td></td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**B**

**RRM2B**

- N-terminal swivel region
- Iron binding residues
- Ribonucleotide small subunit signature

Key amino acids:
- G11A
- S25G
- S39Y
- I47T
- R121H
- M137L
- K150N
- N157S
- K166N
- K180I
- A189V
- A192S
- A204S
- F234I
- K248E
- N244S
- S316Y
- A349V

**C**

**OV (N=311)**

- TP53
- MYC
- RRM1
- RRM2
- RRM2B

**BRCA (N=962)**

- TP53
- MYC
- RRM1
- RRM2
- RRM2B

**HNSC (N=504)**

- TP53
- MYC
- RRM1
- RRM2
- RRM2B

*Iqbal, Demidova et al., Figure 1*
Ovarian Cancer (OV)

A  
RRM2B and MYC (N=61)  
$R^2=0.7637$

B  
MYC (N=67)  
$R^2=0.4309$

C  
RRM2B (N=8)  
$R^2=0.0503$

D  
Unaltered RRM2B and MYC (N=175)  
$R^2=0.0491$

---

**RRM2B**  
**MYC**  
**Other 8q-genes studied**

Iqbal, Demidova et al., Figure 2

---
Iqbal, Demidova et al., Figure 3
- 8q-located genes (other than RRM2B)
- direct interaction
- physical association

Iqbal, Demidova et al., Figure 4
Breast cancer (BRCA)

Overall survival

- Unaltered
- MYC + RRM2B
- MYC
- RRM2B

Disease-free survival

- Unaltered
- MYC + RRM2B
- MYC
- RRM2B

BRCA subtypes

C. ER+ PR+ HER2+

D. ER+ PR+ HER2+

E. RRM2B amplification

F. RRM2B amplification

G. Breast cancer subtypes

<table>
<thead>
<tr>
<th></th>
<th>OS</th>
<th>DFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ER- PR- HER2-</td>
<td>ER+ PR+ HER2-</td>
</tr>
<tr>
<td>ER+ PR+ HER2+</td>
<td>0.0426</td>
<td>0.0998</td>
</tr>
<tr>
<td>ER- PR- HER2-</td>
<td>0.1911</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ER- PR- HER2-</td>
<td>ER+ PR+ HER2-</td>
</tr>
<tr>
<td>ER+ PR+ HER2+</td>
<td>0.2888</td>
<td>0.3237</td>
</tr>
<tr>
<td>ER- PR- HER2-</td>
<td>0.9212</td>
<td></td>
</tr>
</tbody>
</table>

Iqbal, Demidova et al., Figure 5
Breast cancer (BRCA)

Iqbal, Demidova et al, Figure 6