



Identification of a Splenic Marginal Zone Lymphoma Signature: Preliminary Findings With Diagnostic Potential

Jacob E. Robinson¹, Timothy C. Greiner², Alyssa C. Bouska², Javeed Iqbal² and Christine E. Cutucache^{1*}

¹ Department of Biology, University of Nebraska Omaha, Omaha, NE, United States, ² Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, United States

OPEN ACCESS

Edited by:

Basem M. William,
The Ohio State University,
United States

Reviewed by:

Miguel A. Piris,
University Hospital Fundación Jiménez
Díaz, Spain
Francesco Bertoni,
Institute of Oncology Research
(IOR), Switzerland

*Correspondence:

Christine E. Cutucache
ccutucache@unomaha.edu

Specialty section:

This article was submitted to
Hematologic Malignancies,
a section of the journal
Frontiers in Oncology

Received: 03 February 2020

Accepted: 06 April 2020

Published: 08 May 2020

Citation:

Robinson JE, Greiner TC, Bouska AC,
Iqbal J and Cutucache CE (2020)
Identification of a Splenic Marginal
Zone Lymphoma Signature:
Preliminary Findings With Diagnostic
Potential. *Front. Oncol.* 10:640.
doi: 10.3389/fonc.2020.00640

Splenic marginal zone lymphoma (SMZL) is a rare, indolent non-Hodgkin's lymphoma that affects 0.13 per 100,000 persons annually. Overall survival of SMZL is estimated to reach 8–11 years in most cases, but up to 30% of SMZL cases develop aggressive presentations resulting in greatly diminished time of survival. SMZL presents with a very heterogeneous molecular profile, making diagnosis problematic, and accurate prognosis even less likely. The study herein has identified a potential diagnostic gene expression signature with highly specific predictive utility, coined the SMZL-specific Gene Expression Signature (SSGES). Additionally, five of the most impactful markers identified within the SSGES were selected for a five-protein panel, for further evaluation among control and SMZL patient samples. These markers included EME2, ERCC5, SETBP1, USP24, and ZBTB32. When compared with control spleen and other B-cell lymphoma subtypes, significantly higher expression was noticed in SMZL samples when stained for EME2 and USP24. Additionally, ERCC5, SETBP1, USP24, and ZBTB32 staining displayed indications of prognostic value for SMZL patients. Delineation of the SSGES offers a unique SMZL signature that could provide diagnostic utility for a malignancy that has historically been difficult to identify, and the five-marker protein panel provides additional support for such findings. These results should be further investigated and validated in subsequent molecular investigations of SMZL so it may be potentially incorporated into standard oncology practice for improving the understanding and outlook for SMZL patients.

Keywords: SMZL, marginal zone lymphoma, signature, lymphoma, diagnostic

INTRODUCTION

Splenic marginal zone lymphoma (SMZL) comprises <2% of all lymphoid malignancies (1–5), yet remains incurable and challenging to diagnose accurately. SMZL typically progresses indolently, with an overall survival of 8–11 years among most patients (6–8), but up to 30% of cases will advance more aggressively, with 5–10% of cases progressing to diffuse large B-cell lymphoma (DLBCL), both of which result in decreased overall survival (9–13). SMZL presents extremely heterogeneously and, as a result, a myriad of differential diagnostic assessments are utilized in

an attempt to accurately identify the malignancy (13, 14). Due to inconsistencies among most evaluations, the definitive diagnosis of SMZL is typically only recognized upon review of spleen histology following a splenectomy (13); an operation that should be exercised as an absolute last resort due to the impending associated health risks (15, 16). What's more, even among cases properly diagnosed with SMZL, accurate genomic prognoses are not currently possible.

In addition to the rarity and symptomatic heterogeneity of SMZL, few molecular identifiers are consistently presented among cases. Chromosomal losses at 7q account for 25–40% of cases, and gains at 3q are seen in 10–20% of patients, and the most common genetic mutations are among NF- κ B regulators (30–40%), mutations of NOTCH pathway components (20–40%), and in KLF2 (30–40%) (17–21). While these mutations provide mechanistic indications for understanding pathogenesis, they lack clear diagnostic significance due to their variation among cohorts.

These aforementioned restraints have resulted in limited availability of SMZL clinical data, and consequently, large-scale genomic analyses of SMZL has been difficult. Thus, there exists a need to properly evaluate the potential for a diagnostic panel across multiple sample cohorts. Gene expression profiling and the associated analysis of its data have provided great potential for deciphering diagnostic panels among malignancies (22–24). Many panels have been identified among malignancies, and it has been shown that these panels can potentially provide clinical utility—from diagnosis, to prognosis, to treatment design. Glas et al. (25) demonstrated the validity of a 78 probe microarray for prognostic evaluation among a subset of breast cancer patients and utilized a platform that provided test results to patients in as little as 5 days. Gene expression panels have also been utilized in an effort to differentially diagnose non-Hodgkin lymphomas in preserved samples (26). These efforts demonstrate the demand and promise for the development of gene expression panels for diseases and disease states that are difficult to assess clinically, like SMZL.

The identification of novel genetic targets also allows for the evaluation of their utility as potential treatment targets. Increasing the number of targeted treatment options would enable the reduction of splenectomy usage as the primary treatment option in SMZL patients, thus improving patient experience. Splenectomy can result in a variety of complications in patients involving susceptibility to infection, cardiovascular complications, thrombosis, and increased oncogenesis potential (27–29). As these complications have become more apparent, the decision to conduct splenectomy has begun to be reconsidered but without reliable diagnoses little can be changed (22). Hence, a gene expression panel that can be evaluated from a splenic biopsy rather than full splenectomy may point to other treatment alternatives that have shown success in SMZL patients, without the adverse splenectomy consequences (30).

Furthermore, genetic markers that can be utilized for diagnostic and prognostic use and can provide clear differentiation of SMZL cases against related malignancies and control samples, offers significant utility in an era of precision medicine (22–24). This study identified genetic markers and

pathways that uniquely identify SMZL. Consequently, we report what we call the “SMZL Splenic Gene Expression Signature (SSGES)” as a gene expression panel with potential for use in SMZL diagnosis, and recommend EME2, ERCC5, ZBTB32, USP24, and SETBP1, for further investigation into their potential role in lymphomagenesis.

METHODS

Gene Expression Data Collection and Collation

Gene expression datasets and profiles were extracted from the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) database as well as from a collaborating laboratory. A description and record of all datasets obtained for this study can be found in (**Supplementary Table 1**). All datasets were downloaded as raw data (.CEL) and were uploaded concurrently into *BRB ArrayTools*. Additionally, samples were only selected for the gene expression classification if they were samples taken from a secondary lymphoid tissue site such as LN or Spleen, due to influence of the tumor microenvironment; consequently, any peripheral blood samples were removed. In total, 437 gene expression profiles from 13 different datasets were used for analysis in the present study, with 42 of those samples designated as SMZL.

Data Preparation

Data were downloaded from GEO as individual datasets. In order to minimize batch effects, data were only selected if the *Affymetrix GeneChip Human Genome U133 Plus 2.0 Array* was used for the microarray platform. The CEL files underwent MAS5.0 normalization, log₂ transformation, and quantile normalization. No microarray spot filters were applied to the dataset, but the gene filter: exclude if <10% of expression data values have at least a 1.5-fold change in either direction from the gene's median value or if percent missing exceeds 60% at any gene, was applied. Replicate spots on the array were averaged and multiple probes/probe sets were reduced to one per gene symbol. This was selected by the maximally expressed probe as measured by average intensity across array, as previously described (31). Once collated, the data was labeled based on sample ID, tissue, and sample type (CSP-control spleen, DLBCL-Diffuse Large B-cell Lymphoma, EMZL-Extranodal Marginal Zone Lymphoma, FL-Follicular Lymphoma, MCL-Mantle Cell Lymphoma, NMZL-Nodal Marginal Zone Lymphoma, or SMZL). GEPs were then assigned CSP, SMZL, or tissue B-cell lymphoma (TBCL) for signature comparisons.

Calculation and Selection of the SMZL-Specific Gene Expression Signature

Unsupervised clustering was conducted on all 437 samples using all available filtered genes. Sample type designations were assigned in order to differentiate each malignancy among this gene expression cohort. Genes were centered and scaled, and the samples were clustered by one minus correlation, based on average value, to determine linkage.

SMZL GEPs were then compared against CSP GEPs via univariate class comparison. No additional gene filters were applied, and the analysis was conducted at a significance level of 0.001. This univariate comparison was replicated between SMZL and TBCL samples.

All probes highly significantly, differentially expressed ($p < 1e-7$) in each respective class comparison were then selected for and cross compared to identify differentially expressed genes present in both datasets. Significance analysis of microarray (SAM) was then conducted on the most differentially expressed gene probes, as was recommended (32, 33). SAM was run using a target proportion of false discoveries at 0.01 and by completing 100 permutations. SMZL was first compared to CSP using the 164 genes highly differentially expressed, and the process was then repeated between SMZL and TBCL. These lists were cross-compared one final time, and 135 genes were identified to be uniquely, differentially expressed in SMZL against both, CSP and TBCL comparisons and identified as the SMZL Splenic Gene Expression Signature (SSGES) (Supplementary Table 2).

Multiple Methods prediction analysis was utilized through the *BRB Array Tools* platform. The SSGES was the designated gene list for predictive comparison. The alpha value for the multiple methods analysis was set to 0.01, and all predictive models were utilized and reported in the analysis. Functional pathway analysis was then conducted on both respective class comparisons. The functional and network analyses were generated using *Ingenuity Pathway Analysis* (34).

Gene Expression Replication Cohort

In order to verify the predictive value of the SSGES, a replication gene expression cohort was utilized. All datasets were acquired from GEO, and all normalization, filtering, and design methodologies followed procedures of the initial test dataset. All sample information is provided in (Supplementary Table 3). Predictive analysis was also conducted as described above for the replication cohort.

Protein Signature Extraction

For the development of a protein signature, the SSGES was condensed to the most impactful markers for evaluation in patient samples. This was completed using the total percent impact score (TPIS) in conjunction with *The Human Protein Atlas* in order to identify which markers would be optimal candidates for IHC (35).

This study proposes the calculation and implementation of a TPIS for ranking the differentially expressed markers. The TPIS for each differentially expressed gene was calculated by ranking each probe by its SAM score (achieved during the SAM analyses) through assigning the percent rank to each marker based on their score. For example, *IL7* presented the highest SAM score [$d(i) = 7.852$] when SMZL was compared with CSP, and so it was assigned an impact score of 1.0. *ZNF763* had the second highest SAM score [$d(i) = 7.196$] in the same comparison. In order to weigh each gene probe equally on its impact on total score, TPIS must be assigned based on SAM score instead of numerical order. Hence, *ZNF763* had a TPIS of 0.91645, as that is equivalent to the percent rank of SAM score. This was done for all up-regulated

gene probes as well as all down-regulated gene probes for SMZL vs. CSP and SMZL vs. TBCLs. The sum of these scores were then taken to identify the TPIS for both up- and down-regulated genes in the signature (Supplementary Table 2). In order to further evaluate the markers potential as protein assay candidates, the genes identified on the SSGES were then researched on *The Human Protein Atlas* for protein, spleen-specific expression in control tissues (35). Additionally, the chromosomal location of the genes was investigated and considered to evaluate if chromosomal aberrations could be responsible for the alteration in gene expression. Targets for immunohistochemistry (IHC) were identified by taking into account each of the aforementioned TPIS, chromosomal characteristics, and protein expression, as reported by *The Human Protein Atlas*. This resulted in identification of five key markers (*ZBTB32*, *EME2*, *SETBP1*, *ERCC5*, and *USP24*) that could be feasibly evaluated by IHC, based on resource and time availability.

Immunohistochemistry Patient Samples

Tissue samples were provided by the Lymphoma Study Group at the University of Nebraska Medical Center (UNMC). Tissues were provided as either unstained slides or as paraffin embedded tissue blocks. There were 17 total samples provided, 6 SMZL, 6 control spleen, and 5 other B-cell lymphoma splenic samples [1 DLCL-C, 1 DLCL-NC, 1 FL, 2 follicular medium cell (FM)]. Full listing with available prognostic information can be found in (Supplementary Table 4). Each case was previously diagnosed by a physician then confirmed by pathologist upon removal via histology. Cases were also validated via IHC of immunophenotypic markers characteristic of SMZL (Supplementary Figure 1). This study had full approval for its completion through the institutional review board (IRB) at UNMC, and of the Fred and Pamela Buffett Cancer Center, and the Nebraska Lymphoma Study Group Tissue Bank #441-18-EP.

Immunohistochemistry

IHC was conducted using the *Roche Diagnostics ULTRA Discovery* IHC platform housed at the UNMC Tissue Science Facility. Each slide contained 5 microns thick, paraffin embedded, sections of splenic tissue. The antibody and IHC information is provided in (Supplementary Table 5), inclusive of antibody, company of origin, catalog number, antigen retrieval time in minutes, optimized dilution, and antibody staining time in minutes. Each condition was optimized for staining 5-micron splenic sections. Optimization controls for *EME2* and *USP24* were conducted on kidney and liver, respectively. Hematoxylin counter-staining, dehydration, and re-application of cover slip was done using the Tissue-Tek Prisma platform.

Scoring, Protein Comparison, and Prognostication

All slide scans were taken at 40X magnification before being analyzed using VENTANA Image Viewer. Slides were analyzed by three researchers independently. The slide order was randomized, and identifiers were blinded to eliminate researcher bias. Researchers were shown five separate screen-shots at 20X magnification for each separate slide and scored staining

intensity from 1 (no staining) to 5 (diffuse, complete tissue staining) for each image. This scoring was then averaged between the three researchers for each image and staining intensity scores (SIS) were assigned to each image. The five images per sample were then averaged, and each sample was given an overall staining intensity score (OSIS). Each lymphoma diagnosis was then validated one final time by a trained pathologist prior to cohort comparison. Data comparisons between CSP, TBCL, and SMZL sample groups was done using Student's *t*-test.

RESULTS

SSGES Delineation

Gene expression has been utilized as an effective technology for deciphering diagnostic and prognostic subgroups among B-cell lymphomas (23, 36). Despite previous efforts to evaluate the gene expression of SMZL cases (22, 37), definitive biomarkers have not been well-established. We analyzed 437 gene expression profiles, comparing control spleen (CSP), other secondary lymphoid organ B-cell lymphoma (TBCL), and SMZL samples. TBCL samples included diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), mantle cell lymphoma (MCL), nodal marginal zone lymphoma (NMZL), and extranodal marginal zone lymphoma (EMZL), for comparison.

In order to evaluate whether SMZL possessed a unique genomic profile, unsupervised clustering was conducted at a significance level of 0.001, on all 437 samples across the 16,272 available genes. SMZL cases clustered together when analyzed, suggesting there is a unique, delineating molecular signature by which SMZL can be more accurately diagnosed, if deciphered (Figure 1A).

In order to identify a unique SMZL genetic signature, univariate class comparisons were completed. The comparison between SMZL and CSP produced 5,707 gene probes that possessed a significant difference in expression, and SMZL GEPs compared to the TBCLs returned 6,477 gene probes. All probes highly significantly, differentially expressed ($p < 1e-7$) in each respective class comparison were then selected. The comparison between SMZL and CSP produced 1,056 highly differentially expressed genes, while the greater significance between SMZL and TBCL returned 2,374. These two lists then underwent cross-comparison to identify genes that were differentially expressed in SMZL from both datasets. This resulted in 164 genes that were differentially expressed in SMZL when compared with both CSP and TBCLs.

SAM was then conducted using the 164 gene signature to reduce the FDR by conducting gene-specific *t*-tests independently, as opposed to the class comparison *t*-test across the full dataset, further corroborating that the identified differentially expressed gene probes are not selected randomly. Using SAM, 147 and 148 genes were differentially expressed in SMZL when compared with CSP and TBCLs, respectively. These lists were cross-compared one final time, and 135 genes were identified to be uniquely, differentially expressed in SMZL against both, CSP and TBCL comparisons. This collection of 135 genes were coined the "SMZL specific Gene Expression Signature (SSGES)" (Supplementary Table 2, Figure 1B).

Predictive and Functional Analysis

Multiple methods for predictive analysis were conducted to determine the fidelity of our diagnostic tool. This was evaluated for SMZL against CSP and TBCL separately to assess the predictive value of the 135 gene SSGES. For each method of predictive analysis to delineate SMZL from CSP was correctly classified 100% of the time. And, when this same SSGES signature is used to predict diagnosis between SMZL from other B-cell lymphomas (TBCL), it was correct 98% of the time (Supplementary Table 6).

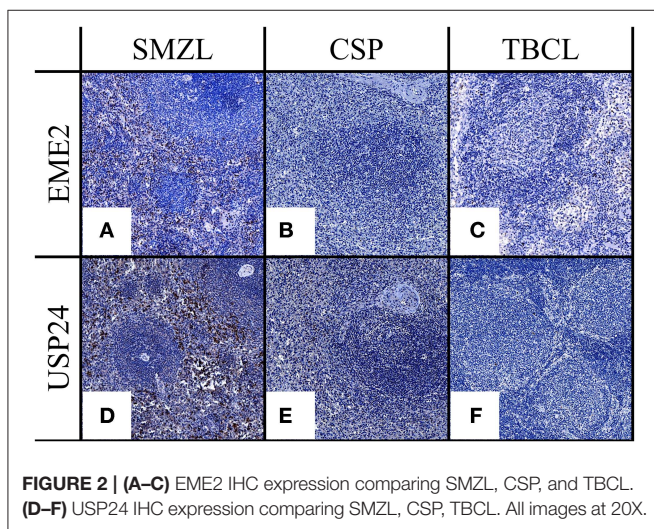
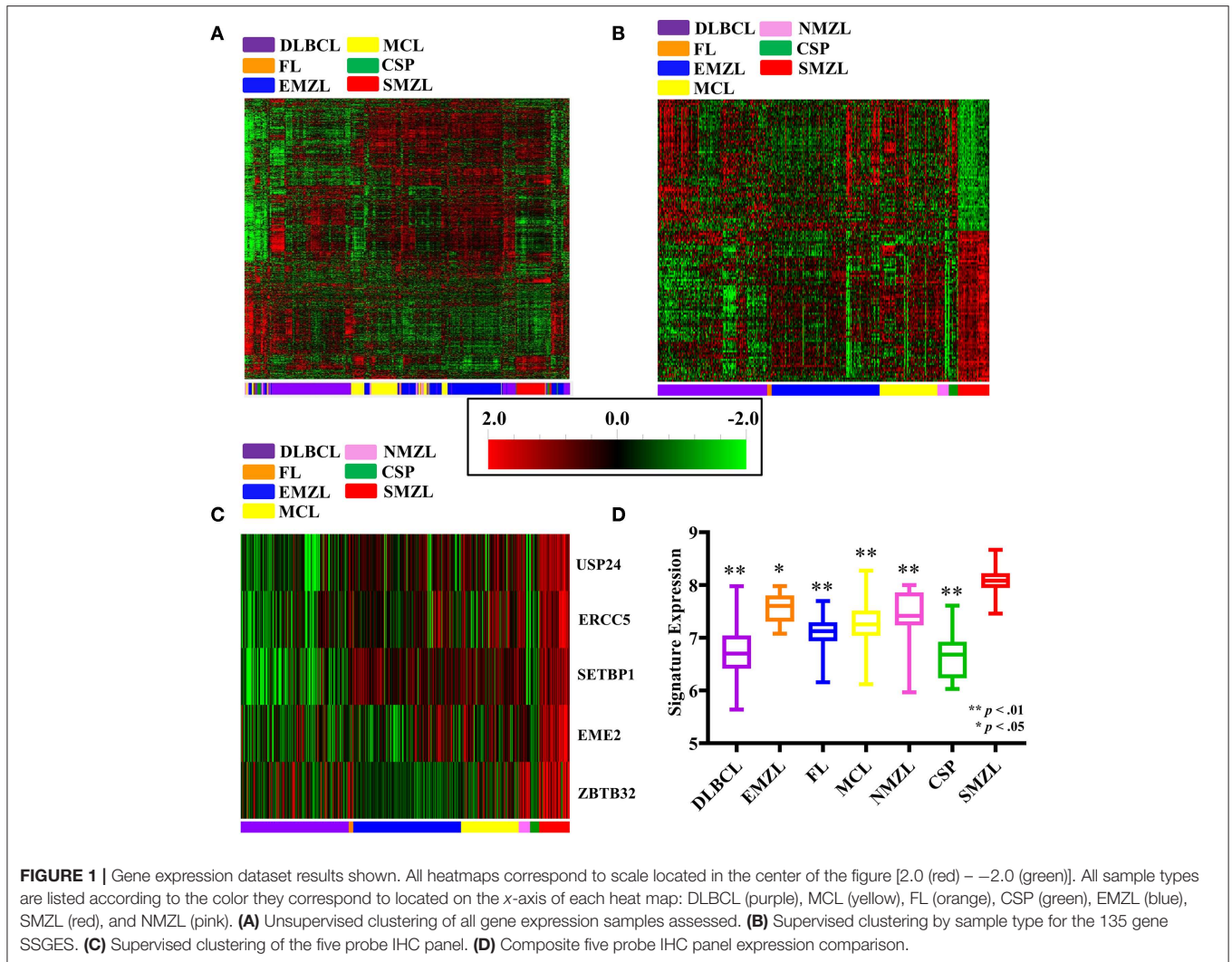
Gene Expression Replication Set: Ensuring Fidelity and Reliability of SSGES

A gene expression replication set containing 705 samples was organized independently of the original SSGES delineating cohort (Supplementary Table 3). The replication cohort contained secondary lymphoid organ B-cell lymphoma samples from datasets not used in the original analysis set as well as splenic controls and the original 43 SMZL samples. Due to the limited sample number of SMZL samples, all samples were used in the original and replication cohort, but new comparison samples were utilized. While this limits the ability to delineate new signatures, the replication cohort demonstrated the strong predictability of SMZL samples against other B-cell lymphoma samples and controls not originally used in signature delineation. Multiple methods predictive analysis was used, and with similar effectiveness to the original analysis, SMZL samples could on average be correctly differentiated from CSP and TBCL samples 98 and 95%, respectively (Supplementary Table 6).

SSGES Protein Evaluation

To provide validation of the gene expression signature and enhance the impact of the study, five markers were selected to provide clinical utility as a protein assay for diagnostic purposes and to investigate for SMZL-specific treatment targets at the protein level. These markers were also evaluated alongside patient data to assess the prognostic significance of their increased expression. The five markers selected to be evaluated using IHC: *EME2*, *ZBTB32*, *ERCC5*, *USP24*, and *SETBP1*. Multiple methods predictive analysis was conducted using this five-gene cohort, with correct classification reaching ~95% in the original cohort, and ~90% when compared to the validation cohorts, indicating the potential for this much smaller panel to demonstrate diagnostic utility (Supplementary Table 6, Figure 1C). Lastly, these five markers demonstrated significantly higher expression in SMZL cases when compared against each malignant cohort individually, further demonstrating their selectivity for SMZL (Figure 1D).

Following staining and respective scoring, expression of each marker was compared. *EME2* was shown to have significantly higher expression in SMZL samples when compared with CSP ($p = 0.05$) or TBCL ($p = 0.01$) cohorts (Figures 2A–C), and *USP24* demonstrated significantly higher expression in SMZL when compared to CSP ($p = 0.02$; Figures 2D,E). *USP24* expression also appeared increased in SMZL compared to TBCL samples ($p = 0.10$) but was not statistically significant (Figures 2D,F). For both *USP24* and *EME2*, increased expression



was consistently observed within the interfollicular zones of SMZL cases (**Figure 3**), but marginal zone staining was only recognized sporadically.

The remaining markers did not present with statistically significant differential expression (**Supplementary Figure 2**), but ZBTB32, SETBP1, and ERCC5 were each indicative of possible prognostic value for SMZL cases. ZBTB32 showed significantly higher expression in SMZL cases with lymphoma symptoms present ($p = 0.04$; **Figure 4A**). Expression scoring values for ERCC5 ($r = 0.76$), SETBP1 ($r = 0.44$), and ZBTB32 ($r = 0.78$) in SMZL cases correlated with the patients' staging upon tissue extraction. (**Figure 4B**), and USP24 increased expression also correlated with increased staging of SMZL ($r = 0.93$) (**Figure 4B**). Further, SETBP1 expression also correlated with increased staging of the TBCL samples ($r = 0.66$), indicating its potential as B-cell lymphoma marker in multiple subtypes. Each of these markers' increased expression correspondence with severity of symptoms among the SMZL cases, points to their

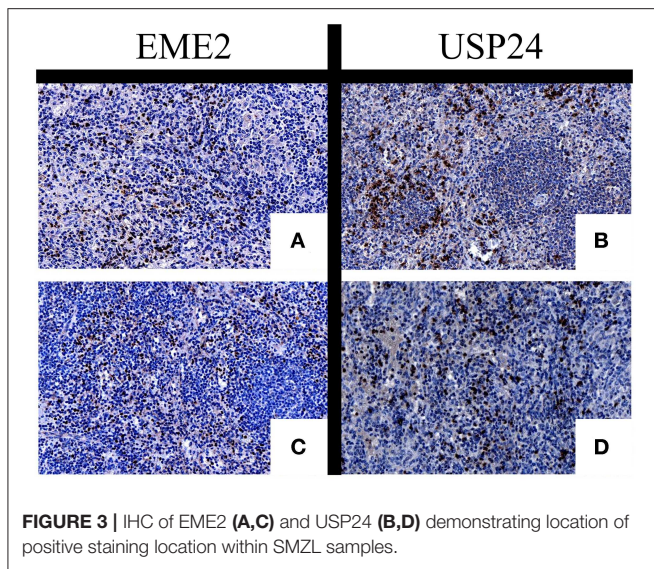


FIGURE 3 | IHC of EME2 (A,C) and USP24 (B,D) demonstrating location of positive staining location within SMZL samples.

potential use as prognostic markers. Thus, they warrant further investigation among expanded patient cohorts to corroborate these preliminary findings.

DISCUSSION

In this study, a 135-gene signature coined the SSGES that demonstrates clinical potential for identification of SMZL was identified. Even more importantly, novel biomarkers have been identified via IHC that could provide diagnostic, as well as prognostic utility for SMZL patients. Given the considerable dearth of information for both diagnosis and prognosis of SMZL in the literature, it's expected that these findings will significantly contribute to patient outcome. Importantly, when we replicated the SSGES against new data sets and tested their predictability and fidelity, they yielded 90–100% accuracy of identifying SMZL, well-above current clinical markers and methods. These results provide preliminary findings that must be further validated in expanded investigations and should guide future inquiries into understanding SMZL pathogenesis.

Many treatment options have been proposed and are being further developed in marginal zone lymphomas, with significant improvement in SMZL relapse and progression free survival rates (38–41). Currently, many of these drug therapies will be tried in SMZL cases prior to resorting to a splenectomy as a treatment option, but there is still limited genetic information driving these decisions (20). In order to force a switch in treatment regimens and recommendations, a genetic panel must be produced that provides diagnostic and prognostic value for identifying SMZL. This study has identified such a panel, through meta-analysis of gene expression profiles among SMZL cases against other B-cell lymphomas and splenic controls. This panel must be further clinically validated and assessed for potential usage systematically, but the high predictive success of the panel against each comparison set is promising for the prospect of future implementation.

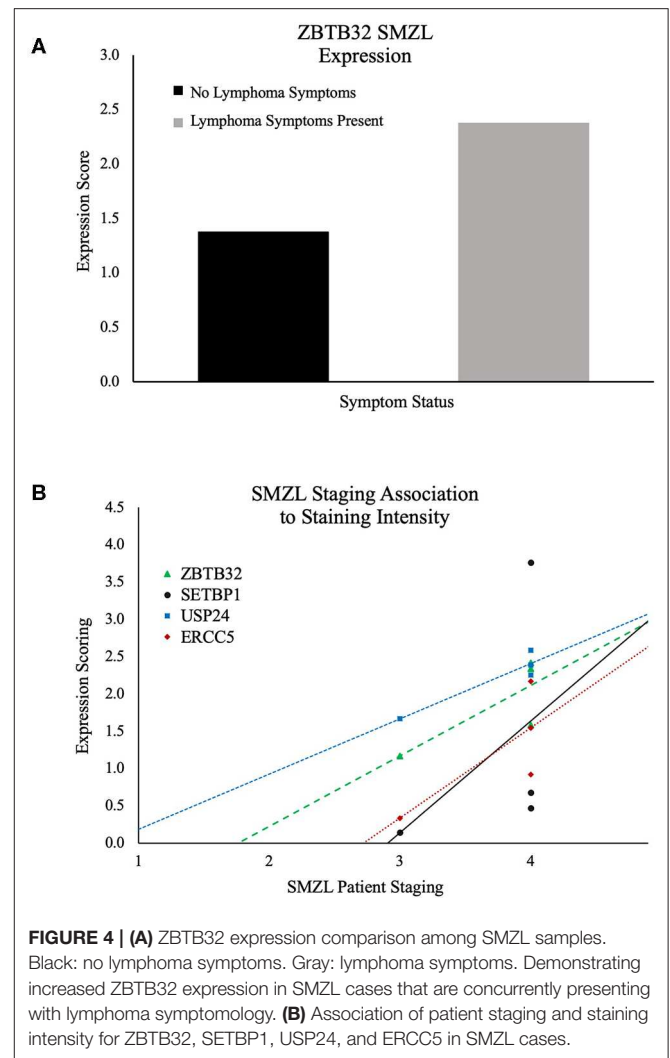


FIGURE 4 | (A) ZBTB32 expression comparison among SMZL samples. Black: no lymphoma symptoms. Gray: lymphoma symptoms. Demonstrating increased ZBTB32 expression in SMZL cases that are concurrently presenting with lymphoma symptomatology. (B) Association of patient staging and staining intensity for ZBTB32, SETBP1, USP24, and ERCC5 in SMZL cases.

While each of the markers investigated in the present investigation are understudied, they each display mechanistic potential as drivers of SMZL carcinogenesis. EME2 in conjunction with MUS81 initiate replication fork restart (42), indicating its increased expression could be exploited by cancer cells for enhanced DNA replication. USP24 has been shown to play a role in anti-apoptotic tumoral activity in myeloma cells and displayed increased expression in cells displaying increased drug resistance (43). USP24 has also been shown to be a regulator of tumor microenvironment signaling and has been proposed as a potential therapeutic target due to this role (44). The mechanisms collectively demonstrate that its increased expression could play a role in SMZL pathogenesis. Increased ZBTB32 expression has been identified among memory B-cells (45), and thus, its increased expression among symptomatic patients may be due to an increased presence of the memory B-cell based SMZL cells, a commonly proposed cell-type of primary SMZL tumor cells. Polymorphisms at *ERCC5* and disrupted expression have been reported to increase solid tumor carcinogenesis due to the

TABLE 1 | Proposed biological impact of markers selected for IHC panel.

| Gene | NCBI gene ID | Biological function (citation) |
|---------------|--------------|--|
| <i>EME2</i> | 197342 | Replication fork restart (42) |
| <i>ERCC5</i> | 2073 | DNA excision disruption (46, 47) |
| <i>SETBP1</i> | 26040 | Cellular proliferation and development promotion (50), tumor suppressor suppression (51) |
| <i>USP24</i> | 23358 | Cell survival (43); Tumor microenvironment regulation (44) |
| <i>ZBTB32</i> | 27033 | Memory B-cell expression (45) |

disruption of DNA excision pathways (46, 47), and in an IHC investigation study, *ERCC5* encoded XPG was shown to have positive expression in breast, ovarian, and sarcoma samples (48). Lastly, *SETBP1* activation has been shown to play a role in leukemic progression and development. Missense mutations of *SETBP1* have been shown to increase the stability of the *SETBP1* protein and result in its subsequent over expression in many myeloid neoplasms (49). Further, *SETBP1* overexpression has been associated with shorter overall survival in leukemic patients (50). It is proposed that *SETBP1* executes its oncogenic properties through suppression of tumor suppressor genes and by inducing increased myeloid cell development and proliferation (50, 51). Each of these mechanisms provide evidence that the demonstrated over expression of *SETBP1* in this study may play a role in oncogenesis and should also be investigated further for its mechanistic value among SMZL patients. A more concise record of the potential biological impact that these markers provided is listed within **Table 1**.

Upon evaluation of IHC, a significant amount of the respective protein expression occurred within the interfollicular zones of the spleens, rather than exclusively within the tumor cells, indicating the markers' potential role as tumor environment effectors. The importance of understanding the tumor environment has become increasingly relevant, as demonstrated by the pursuit of therapies that can be utilized to circumvent its impact on lymphoma progression (52). Despite the markers' inconsistent expression as identifiers of SMZL tumoral cells themselves, this increased interfollicular expression corroborates the increased gene expression observed in the splenic tissue SMZL samples. This increased expression of these proteins within the tumor environment should be considered for their possible role in lymphomagenesis.

This study identified markers that provide potential clinical significance in delineating SMZL from other B-cell lymphomas and improving the biological understanding of SMZL. Future studies much be conducted to validate and expand the information reported in this manuscript. The intention of this study was to identify markers that could assist in diagnosis of SMZL and provide empirical data for weighing treatment options and prognostic outlook. Thus, while this study identified potential markers that may play a role

in SMZL pathogenesis and even progression, significantly more work must be completed to fulfill the goal of clinical utility. Mechanistic studies utilizing *in vitro* B-cell lymphoma models could assess the functional impact of increasing the expression of each of the protein markers discussed above. Additionally, evaluating the expression of these markers in expanded SMZL as well as other lymphoma cohorts not utilized here would further validate these results. Increased prognostic value would also come from evaluating survival time, treatment outcomes, and treatment reactivity along with SSGES expression.

Several limitations arose upon completion of the study. First, the authors were limited to the samples and respective sample information that was available for use within the *GEO* datasets. Ideally, the signature could be derived against additional other splenic derived lymphoma subtypes such as hairy cell leukemia variant, hairy cell leukemia, hepatosplenic T-cell lymphoma, macrophage-rich large B-cell lymphoma, and more DLBCL cases presented in the spleen (53, 54), but these data were unavailable at the time of study completion. Additionally, there was no available information regarding the status and source of control spleens used in this study for gene expression as well as protein expression evaluation. In order to necessitate a splenectomy there must be critical circumstances at hand, and thus, while the spleens did not present with another lymphoma, it is difficult to know the true molecular profile they present due to the variable circumstances that resulted in their availability. The authors were also limited by the number of samples utilized for the IHC investigations. Due to the rarity and indolent presentation of SMZL, sample availability is extremely limited. Hence, the prognostic and diagnostic implications of these results are disclosed as preliminary until further validation in additional sample cohorts. Lastly, the gene expression analysis used herein was limited to exclusive use of Affymetrix probes. Thus, additional gene expression and genomic sequencing data executed using other platforms were not considered for analysis. This limited the study sample size provided and the depth of information that could have been gleaned from a wider array of analysis platforms. Further investigations utilizing data across multiple platforms would provide validation to these results and would further enhance clinical utility of the disclosed SMZL signature. Each of these limitations were recognized throughout the completion of this study. In an attempt to combat them, multiple levels of replication were utilized, samples originating from numerous study sites and regions were included, highly significant statistical filtering was conducted, and the SSGES was analyzed against an independent validation gene expression cohort for its accuracy. These efforts and subsequent results suggest that the SSGES is a reliable tool for diagnosis of SMZL when compared against other marginal zone lymphomas and control spleen samples.

In conclusion, this study has identified a diagnostic gene expression signature that presents many potential markers for diagnosis of SMZL patients from a splenic biopsy. It has also identified five novel protein markers, *EME2*,

ERCC5, SETBP1, USP24, and ZBTB32, that appear to play a role in SMZL pathogenesis. Each of these results warrant further investigation for their broader roles in B-cell lymphoma biology.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI Gene Expression Omnibus (GSE12195, GSE16024, GSE2109, GSE23501, GSE35348, GSE35426, GSE53820, GSE55267, GSE57520, GSE7307, GSE93291, GSE146814).

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: JR, JI, and CC. Performed the experiments: JR and AB. Analyzed the data: JR and CC. Reviewed and validated pathology: TG. Wrote the paper: JR and CC. Reviewed the manuscript: JR, TG, AB, JI, and CC.

REFERENCES

- Armitage JO, Weisenburger DD. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *J Clin Oncol.* (1998) 16:2780–95. doi: 10.1200/JCO.1998.16.8.2780
- Liu L, Wang H, Chen Y, Rustveld L, Liu G, Du XL. Splenic marginal zone lymphoma: a population-based study on the 2001–2008 incidence and survival in the United States. *Leuk Lymphoma.* (2013) 54:1380–6. doi: 10.3109/10428194.2012.743655
- Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood.* (2016) 127:2375–90. doi: 10.1182/blood-2016-01-643569
- Santos TSD, Tavares RS, Farias DLC. Splenic marginal zone lymphoma: a literature review of diagnostic and therapeutic challenges. *Rev Bras Hematol Hemoter.* (2017) 39:146–54. doi: 10.1016/j.bjh.2016.09.014
- Thieblemont C, Davi F, Noguera M-E, Briere J. Non-MALT marginal zone lymphoma. *Curr Opin Hematol.* (2011) 18:273–9. doi: 10.1097/MOH.0b013e3283477815
- Chacón JI, Mollejo M, Muñoz E, Algara P, Mateo M, Lopez L, et al. Splenic marginal zone lymphoma: clinical characteristics and prognostic factors in a series of 60 patients. *Blood.* (2002) 100:1648–54. doi: 10.1182/blood.V100.5.1648.h81702001648_1648_1654
- Arcaini L, Lazzarino M, Colombo N, Burcheri S, Boveri E, Paulli M, et al. Splenic marginal zone lymphoma: a prognostic model for clinical use. *Blood.* (2006) 107:4643–9. doi: 10.1182/blood-2005-11-4659
- Olszewski AJ, Castillo JJ. Survival of patients with marginal zone lymphoma: analysis of the Surveillance, Epidemiology, and End Results database. *Cancer.* (2013) 119:629–38. doi: 10.1002/cncr.27773
- Franco V, Florena AM, Iannitto E. Splenic marginal zone lymphoma. *Blood.* (2003) 101:2464–72. doi: 10.1182/blood-2002-07-2216
- Matutes E. Clinical and biological diversity of splenic marginal zone lymphoma. *Expert Rev Anticancer Ther.* (2009) 9:1185–9. doi: 10.1586/era.09.91
- Lenglet J, Traullé C, Mounier N, Benet C, Munoz-Bongrand N, Amorin S, et al. Long-term follow-up analysis of 100 patients with splenic marginal zone lymphoma treated with splenectomy as first-line treatment. *Leuk Lymphoma.* (2014) 55:1854–60. doi: 10.3109/10428194.2013.861067
- Conconi A, Franceschetti S, Aprile von Hohenstaufen K, Margiotta-Casaluci G, Stathis A, Moccia AA, et al. Histologic transformation in marginal zone lymphomas[†]. *Ann Oncol.* (2015) 26:2329–35. doi: 10.1093/annonc/mdv368

FUNDING

Thanks to the Department of Biology, University of Nebraska at Omaha and the University of Nebraska Foundation for funding this project.

ACKNOWLEDGMENTS

The authors acknowledge the Tissue Science Facility at the University of Nebraska Medical Center for their assistance with tissue staining, imaging, and training services for the Discovery platform. Additionally, the authors thank Dr. George Haddix, the University of Nebraska Foundation, the University of Nebraska at Omaha Department of Biology, and the University of Nebraska Omaha College of Arts and Sciences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.00640/full#supplementary-material>

- Arcaini L, Rossi D, Paulli M. Splenic marginal zone lymphoma: from genetics to management. *Blood.* (2016) 127:2072–81. doi: 10.1182/blood-2015-11-624312
- Iannitto E, Tripodo C. How I diagnose and treat splenic lymphomas. *Blood.* (2011) 117:2585–95. doi: 10.1182/blood-2010-09-271437
- Oliveira LC, Fardilha C, Louro M, Pinheiro C, Sousa A, Marques H, et al. Palliative splenic irradiation for symptomatic splenomegaly in non-Hodgkin lymphoma. *Ecancermedicalscience.* (2018) 12:887. doi: 10.3332/ecancer.2018.887
- Zaorsky NG, Williams GR, Barta SK, Esnaola NF, Kropf PL, Hayes SB, et al. Splenic irradiation for splenomegaly: a systematic review. *Cancer Treat Rev.* (2017) 53:47–52. doi: 10.1016/j.ctrv.2016.11.016
- Rossi D, Trifonov V, Fangazio M, Brusca G, Rasi S, Spina V, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med.* (2012) 209:1537–51. doi: 10.1084/jem.20120904
- Martínez N, Almaraz C, Vaqué JP, Varela I, Derdak S, Beltrán S, et al. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. *Leukemia.* (2014) 28:1334–40. doi: 10.1038/leu.2013.365
- Clipson A, Wang M, de Leval L, Ashton-Key M, Wotherspoon A, Vassiliou G, et al. KLF2 mutation is the most frequent somatic change in splenic marginal zone lymphoma and identifies a subset with distinct genotype. *Leukemia.* (2015) 29:1177–85. doi: 10.1038/leu.2014.330
- Thieblemont C. Improved biological insight and influence on management in indolent lymphoma. Talk 3: update on nodal and splenic marginal zone lymphoma. *Hematology Am Soc Hematol Educ Program.* (2017) 2017:371–8. doi: 10.1182/asheducation-2017.1.371
- Robinson JE, Cutucache CE. Deciphering splenic marginal zone lymphoma pathogenesis: the proposed role of microRNA. *Oncotarget.* (2018) 9:30005–22. doi: 10.18632/oncotarget.25487
- Thieblemont C, Nasser V, Felman P, Leroy K, Gazzo S, Callet-Bauchu E, et al. Small lymphocytic lymphoma, marginal zone B-cell lymphoma, and mantle cell lymphoma exhibit distinct gene-expression profiles allowing molecular diagnosis. *Blood.* (2004) 103:2727–37. doi: 10.1182/blood-2003-06-2160
- Fernández V, Salameo O, Espinet B, Solé F, Royo C, Navarro A, et al. Genomic and gene expression profiling defines indolent forms of mantle cell lymphoma. *Cancer Res.* (2010) 70:1408–18. doi: 10.1158/0008-5472.CAN-09-3419
- Jiang H-B, Yang T-J, Lu P, Ma Y-J. Gene expression profiling of gastric cancer. *Eur Rev Med Pharmacol Sci.* (2014) 18:2109–15.

25. Glas AM, Floore A, Delahaye LJM, Witteveen AT, Pover RCF, Bakx N, et al. Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics*. (2006) 7:278. doi: 10.1186/1471-2164-7-278
26. Sakhinia E, Glennie C, Hoyland JA, Menasce LP, Brady G, Miller C, et al. Clinical quantitation of diagnostic and predictive gene expression levels in follicular and diffuse large B-cell lymphoma by RT-PCR gene expression profiling. *Blood*. (2007) 109:3922–8. doi: 10.1182/blood-2006-09-046391
27. Cadili A, de Gara C. Complications of splenectomy. *Am J Med*. (2008) 121:371–5. doi: 10.1016/j.amjmed.2008.02.014
28. Kristinsson SY, Gridley G, Hoover RN, Check D, Landgren O. Long-term risks after splenectomy among 8,149 cancer-free American veterans: a cohort study with up to 27 years follow-up. *Haematologica*. (2014) 99:392–8. doi: 10.3324/haematol.2013.092460
29. Weledji EP. Benefits and risks of splenectomy. *Int J Surg*. (2014) 12:113–9. doi: 10.1016/j.ijsu.2013.11.017
30. Sammon J, Twomey M, Crush L, Maher MM, O'Connor OJ. Image-guided percutaneous splenic biopsy and drainage. *Semin Intervent Radiol*. (2012) 29:301–10. doi: 10.1055/s-0032-1330064
31. Herek TA, Shew TD, Spurgin HN, Cutucache CE. Conserved molecular underpinnings and characterization of a role for caveolin-1 in the tumor microenvironment of mature T-cell lymphomas. *PLoS ONE*. (2015) 10:e0142682. doi: 10.1371/journal.pone.0142682
32. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. (2001) 98:5116–21. doi: 10.1073/pnas.091062498
33. Shen Y, Iqbal J, Xiao L, Lynch RC, Rosenwald A, Staudt LM, et al. Distinct gene expression profiles in different B-cell compartments in human peripheral lymphoid organs. *BMC Immunol*. (2004) 5:20. doi: 10.1186/1471-2172-5-20
34. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. (2014) 30:523–30. doi: 10.1093/bioinformatics/btt703
35. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science*. (2017) 357:eaan2507. doi: 10.1126/science.aan2507
36. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. (2000) 403:503–11. doi: 10.1038/35000501
37. Watkins AJ, Hamoudi RA, Zeng N, Yan Q, Huang Y, Liu H, et al. An integrated genomic and expression analysis of 7q deletion in splenic marginal zone lymphoma. *PLoS ONE*. (2012) 7:e44997. doi: 10.1371/journal.pone.0044997
38. Cervetti G, Ghio F, Cecconi N, Morganti R, Galimberti S, Petrini M. How to treat splenic marginal zone lymphoma (SMZL) in patients unfit for surgery or more aggressive therapies: experience in 30 cases. *J Chemother*. (2017) 29:126–9. doi: 10.1080/1120009X.2015.1114222
39. Andorsky DJ. Insights into lymphoma from the 2018 American Society of Clinical Oncology annual meeting: commentary. *Clin Adv Hematol Oncol*. (2018) 16(Suppl 16):13–5.
40. Avivi I, Arcaini L, Ferretti VV, Boumendil A, Finel H, Milone G, et al. High-dose therapy and autologous stem cell transplantation in marginal zone lymphomas: a retrospective study by the EBMT Lymphoma Working Party and FIL-GITMO. *Br J Haematol*. (2018) 182:807–15. doi: 10.1111/bjh.15454
41. Ayyappan S, William BM. Marginal zone lymphoma: clinicopathologic variations and approaches to therapy. *Curr Oncol Rep*. (2018) 20:33. doi: 10.1007/s11912-018-0687-9
42. Pepe A, West SC. MUS81-EME2 promotes replication fork restart. *Cell Rep*. (2014) 7:1048–55. doi: 10.1016/j.celrep.2014.04.007
43. Peterson LF, Sun H, Liu Y, Potu H, Kandarpa M, Ermann M, et al. Targeting deubiquitinase activity with a novel small-molecule inhibitor as therapy for B-cell malignancies. *Blood*. (2015) 125:3588–97. doi: 10.1182/blood-2014-10-605584
44. Wang Y-C, Wu Y-S, Hung C-Y, Wang S-A, Young M-J, Hsu T-I, et al. USP24 induces IL-6 in tumor-associated microenvironment by stabilizing p300 and β -TrCP and promotes cancer malignancy. *Nat Commun*. (2018) 9:3996. doi: 10.1038/s41467-018-06178-1
45. Jash A, Wang Y, Weisel FJ, Scharer CD, Boss JM, Shlomchik MJ, et al. ZBTB32 restricts the duration of memory B cell recall responses. *J Immunol*. (2016) 197:1159–68. doi: 10.4049/jimmunol.1600882
46. Walsh CS, Ogawa S, Karahashi H, Scoles DR, Pavelka JC, Tran H, et al. ERCC5 is a novel biomarker of ovarian cancer prognosis. *J Clin Oncol*. (2008) 26:2952–8. doi: 10.1200/JCO.2007.13.5806
47. Liu Y, Hu Y, Zhang M, Jiang R, Liang C. Polymorphisms in ERCC2 and ERCC5 and risk of prostate cancer: a meta-analysis and systematic review. *J Cancer*. (2018) 9:2786–94. doi: 10.7150/jca.25356
48. Aracil M, Dauffenbach LM, Diez MM, Richeh R, Moneo V, Leal JFM, et al. Expression of XPG protein in human normal and tumor tissues. *Int J Clin Exp Pathol*. (2013) 6:199–211.
49. Makishima H, Yoshida K, Nguyen N, Przychodzen B, Sanada M, Okuno Y, et al. Somatic SETBP1 mutations in myeloid malignancies. *Nat Genet*. (2013) 45:942–6. doi: 10.1038/ng.2696
50. Cristobal I, Blanco FJ, Garcia-Orti L, Marcotegui N, Vicente C, Rifon J, et al. SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood*. (2010) 115:615–25. doi: 10.1182/blood-2009-06-227363
51. Nguyen N, Vishwakarma BA, Oakley K, Han Y, Przychodzen B, Maciejewski JP, et al. Myb expression is critical for myeloid leukemia development induced by Setbp1 activation. *Oncotarget*. (2016) 7:86300–12. doi: 10.18632/oncotarget.13383
52. Puente XS, Jares P, Campo E. Chronic lymphocytic leukemia and mantle cell lymphoma: crossroads of genetic and microenvironment interactions. *Blood*. (2018) 131:2283–96. doi: 10.1182/blood-2017-10-764373
53. Wilkins BS. Lymphomas involving the spleen. *Diagn Histopathol*. (2010) 16:116–24. doi: 10.1016/j.mpdhp.2009.12.007
54. Ingle SB, Hinge (Ingle) CR. Primary splenic lymphoma: current diagnostic trends. *World J Clin Cases*. (2016) 4:385–9. doi: 10.12998/wjcc.v4.i12.385

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with several of the authors, JI and AB.

Copyright © 2020 Robinson, Greiner, Bouska, Iqbal and Cutucache. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.