Members Of The Glutathione And ABC-Transporter Families Are Associated With Clinical Outcome In Patients With Diffuse Large B-Cell Lymphoma

Short Title: GSH family and ABC-transporter expression in DLBCL

Charalambos Andreadis 1,2, Phyllis A. Gimotty 2, Peter Wahl 2, Rachel Hammond 2, Jane Houldsworth 3, Stephen J. Schuster 1, and Timothy R. Rebbeck 2

1Abramson Cancer Center and 2Center for Clinical Epidemiology and Biostatistics of the University of Pennsylvania, Philadelphia, PA; 3Cell Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY.

CA, PAG, SJS, and TRR designed and contributed to the research; CA and PAG performed the research; PW and RH assisted with the analysis; JH contributed the independent dataset; and CA wrote the paper.

This publication was supported by grant number 5-K12-CA076931 from the National Cancer Institute (Bethesda, MD). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

Correspondence:
Charalambos Andreadis, M.D., M.S.C.E.
Abramson Cancer Center
University of Pennsylvania
16 Penn Tower
3400 Spruce St.
Philadelphia, PA 19104
Tel. 215-662-7909
Fax. 215-662-4064
babis.andreadis@uphs.upenn.edu

Scientific Heading: Neoplasia

Word Count: 4195; Abstract Count: 189
Abstract

Standard chemotherapy fails in 40-50% of patients with diffuse large B-cell lymphoma (DLBCL). Some of these failures can be salvaged with high-dose regimens, suggesting a role for drug resistance in this disease. We examined the expression of genes in the glutathione (GSH) and ATP-dependent transporter (ABC) families in two independent tissue-based expression microarray datasets obtained prior to therapy from patients with DLBCL. Among genes in the GSH family, glutathione peroxidase 1 (GPx1) had the most significant adverse effect on disease-specific overall survival (dOS) in the primary dataset (n=130) (HR 1.68, 95% CI: 1.26 – 2.22, p <0.001). This effect remained statistically significant after controlling for biological signature, LLMPP cell of origin signature, and IPI score, and was confirmed in the validation dataset (n=39) (HR 1.7, 95% CI: 1.05 – 2.8, p=0.033). Recursive partitioning identified a group of patients with low-level expression of GPx1 and multi-drug resistance 1 (MDR1; ABCB1) without early treatment failures and with superior dOS (p<0.001). Overall, our findings suggest an important association of oxidative-stress defense and drug elimination with treatment failure in DLBCL and identify GPx1 and ABCB1 as potentially powerful biomarkers of early failure and disease-specific survival.
Introduction

Anthracycline and alkylator combination chemotherapy, such as CHOP (cyclophosphamide, adriamycin, vincristine, and prednisone), has remained the standard first line treatment for patients with diffuse large B-cell lymphoma (DLBCL) for almost three decades, with the recent addition of rituximab. However, only about 50 – 60% of patients are cured with this approach. An additional 10-20% of patients can be salvaged with high-dose chemotherapy followed by stem cell transplantation, implicating drug resistance as a significant cause of treatment failure in this disease. Several mechanisms of drug resistance, mostly arising from work in epithelial cancer cell lines, have been put forward to account for this variability.

Anthracyclines are lipophillic compounds that enter the cell by free diffusion across the plasma membrane. ATP-dependent drug efflux via transmembrane proteins (ABC transporters) is a major route of anthracycline elimination and thus, a major determinant of intracellular drug levels (Figure 1). Since the identification of MDR1 glycoprotein (ABCB1) in multi-drug resistant cell lines, several additional ABC transporters have been identified, among them MRP1 (ABCC1), MRP2 (ABCC2), and breast cancer-resistance protein (BCRP or ABCG2). Although several lines of work suggest an association between ABC transporter expression and inferior clinical outcome in various malignancies, studies in previously untreated patients with lymphomas, and DLBCL in particular, have produced controversial findings.

Once inside the cell, anthracyclines exert their cytoidal action by several mechanisms, including DNA intercalation, topoisomerase II inhibition, and reactive-oxygen-species (ROS) formation that can cause widespread damage to DNA, phospholipids, and thiol-containing transport proteins. Several anthracycline-resistant cancer cell lines exhibit changes in critical aspects of the ROS defense system. Glutathione (GSH) is the most abundant intracellular
non-protein thiol\textsuperscript{13} and a key contributor to ROS defense. Its intracellular levels are primarily regulated by the enzyme glutamate-cysteine ligase (GCL), which catalyzes the rate-limiting step in overall GSH biosynthesis (Figure 1). The major downstream effector of GSH-mediated ROS defense is the enzyme glutathione peroxidase 1 (GPx1) which reduces hydrogen peroxide at the expense of oxidizing GSH to its disulfide form GSSG. Overexpression of GPx1 and other antioxidant enzymes, as well as depletion of intracellular GSH, have been shown to suppress anthracycline-induced apoptosis in several experimental systems, including cancer cell lines and tumor xenografts.\textsuperscript{14-18} Previous clinical studies have examined GPx1 function\textsuperscript{19} and genetic polymorphisms\textsuperscript{20} as they relate to susceptibility to lymphoma but not response to therapy. A cDNA microarray-based study of tumor samples from a patient population with DLBCL,\textsuperscript{21} found increased expression of antioxidant defense enzymes, including GPx1, to be associated with a better prognosis in this disease.

Alkylating agents act through the covalent binding of alkyl groups to various cellular molecules.\textsuperscript{22} The glutathione family has also been implicated in alkylator resistance in cell lines as well as primary patient tissues (Figure 1).\textsuperscript{13,23-27} Specifically, glutathione-S-transferases (GSTs) comprise a family of related proteins that can enzymatically conjugate GSH to electrophillic chemotherapeutic agents and even facilitate their elimination through the ABC transporter MRP1 (\textit{ABCC1}).\textsuperscript{28} Besides GSH conjugation activity, GSTs have organic peroxidase activity and may also play a role in protection from ROS induced by anticancer drugs.\textsuperscript{13} Moreover, recent studies have demonstrated that GST subclasses \(\pi\) and \(\mu\) have a regulatory role in the cell by binding to key members of the survival and death-signaling pathway, such as c-Jun N-terminal kinase 1 (JNK1) and apoptosis signal-regulating kinase 1 (ASK1).\textsuperscript{29,30} Overexpression of GSTs at the protein and mRNA level has been noted in
lymphomas. Though all classes have been studied, GSTP1 seems to be the one most consistently associated with inferior clinical outcomes.\textsuperscript{31,32} More recently, Ribrag \textit{et al} showed that \textgreater=50\% immunostaining of DLBCL cells for GSTP1 in 69 previously untreated patients was associated with inferior complete response rates and progression-free survival, but not overall survival.\textsuperscript{33}

Based on the previous experimental observations implicating members of the ABC-transporter family and GSH pathway in resistance to anthracyclines and alkylating agents in several malignancies, we initiated this study in DLBCL. We thought that the complex interaction between these gene families at the molecular level could best be delineated with an approach that measures coordinate tissue expression of critical genes. Our experimental hypothesis was that elevated expression of members of the ABC transporter and GSH families would be associated with inferior clinical outcomes in previously untreated patients with DLBCL. Additionally, coordinate expression of members of these pathways assessed via classification and regression tree analysis (CART) could identify subgroups with divergent risk for early treatment failure.

**Methods**

**Datasets and Data**

We used the publicly available dataset reported on by Monti \textit{et al}\textsuperscript{34} for our primary analysis (http://www.broad.mit.edu/cgi-bin/cancer/datasets.cgi). This set includes pre-treatment patient-level tumor-specimen expression data from the Affymetrix HG-U133 A and B microarray set (Affymetrix, Santa Clara, CA). Tumor specimens were nodal biopsies from newly diagnosed, previously untreated patients with DLBCL seen at Dana Farber / Partners Cancer Care and confirmed by expert hematopathology review. Expression signals were
obtained using Affymetrix Microarray Suite 5.0 (MAS 5.0) software with the scaling factor (SF) and normalization factor (NF) set to 1, indicating no scaling or normalization.

Patients in this dataset (N=130) received full-dose CHOP-based therapy (e.g. 3-4 cycles +XRT for localized disease or a minimum of 6 cycles for advanced disease) and had long-term clinical follow-up. Baseline clinical characteristics have been previously reported in Table 1 of the supplementary information. The primary clinical outcome in this expression dataset is consistent with disease-specific overall survival (dOS). DOS was measured from the date of diagnosis to the date of death or last follow-up without disease.

Based on gene expression in this dataset and using consensus clustering, Monti et al have identified three distinct biological subtypes of DLBCL: Oxidative phosphorylation (OxPhos), B-cell receptor / proliferation (BCR), and host response (HR). These subtypes have distinct genetic features and association with the tumor microenvironment but have not been found to be predictive of response to conventional chemotherapy. Since they provide important biological information, they were used to stratify patients in our analysis.

Additionally, the “cell of origin” classification method (activated B-cell like, Germinal center B-cell like, and Type 3) developed by the Lymphoma / Leukemia Molecular Profiling Project (LLMPP) and later defined by the expression of 27 genes, was calculated for this dataset and used as a predictor/stratifying variable in our analysis. Finally, the international prognostic index (IPI), a four-level index derived from baseline clinical variables, was made available on 116 patients and was included as a predictor/stratifying variable in this study.

An independent dataset previously reported on by Houldsworth et al, was updated and used to validate findings from our primary analysis. This set includes pre-treatment tumor-specimen expression data from 39 patients with untreated DLBCL from Memorial Sloan-
Kettering Cancer Center. Expression-level data were generated on the Affymetrix HG-U95Av2 microarray (Affymetrix, Santa Clara, CA) and re-analyzed through MAS 5.0, with SF and NF set to 1. The survival endpoint evaluated in this dataset is disease-specific overall survival after anthracycline-based chemotherapy.

**Derivation of Expression Level**

Genes of interest in the ABC transporter family and GSH pathway were mapped to unique probe sets in the primary dataset using the Netaffx tool from Affymetrix. When several probe sets were available for the same target gene, the signals across these sets were averaged, as long as there was a high correlation between them (Pearson coefficient >0.8). When correlation was poor, a representative probe set was selected for each target gene based on Affymetrix probe set suffix designation (#_at preferred over #_s_at, #_s_at preferred over #_x_at) and target sequence characteristics (exemplar sequence preferred over consensus sequence, mRNA support preferred over EST support). The absolute expression signal of each gene was adjusted by the median expression signal of each subject along a set of 1798 one-to-one match genes across all Affymetrix platform generations. This set of matches across platforms (HU 6800, HG-U95Av2, and HG-U133) was derived using the Affymetrix good match spreadsheets and eliminating 1-to-many and many-to-1 associations. Unless otherwise noted, all expression values reported are median-adjusted.

Gene expression in the secondary dataset was derived by mapping the selected probe sets from the HG-U133 chip to probe sets from the HG-U95Av2 chip using the Affymetrix best match spreadsheets. Previous research suggests that correlation coefficients for replicates analyzed with different Affymetrix platforms are uniformly over 0.8 for best-match probe sets.
with median adjustment. The absolute expression signal of each gene was again adjusted by the median expression signal of each subject along the set of 1798 one-to-one match genes across all Affymetrix generations.

The “redox signature score” recently described by Tome et al. was calculated in the primary dataset as follows: The probe sets corresponding to the genes in the signature were identified through Netaffx. Specifically, these genes include: Thioredoxin interacting protein, superoxide dismutase 1, 2, 3, glutathione peroxidase 1, 3, 4, thioredoxin reductase 1, 2, catalase, GSTAs, GSTOs, thioredoxin, and microsomal GSTs. Absolute expression values were derived as detailed above and adjusted by the median expression signal of the 1798 one-to-one match genes. The adjusted signals were then combined as described by the authors and the signature was generated.

**Statistical Analysis**

The correlation between genes was calculated using Pearson’s correlation coefficient. Comparisons between categorical and ordinal variables were performed using Fisher’s exact test. The non-parametric Kruskal-Wallis test was used to compare GPx1 levels across stratifying variables. All rates were estimated from survival distributions that were computed using the Kaplan-Meier method. Univariate hazard ratios for dOS were derived from the Cox proportional hazards (CPH) model and p-values were calculated based on the Wald test to assess the significance of factors in the model. The stricter p-value cutoff of 0.01 was used to account for multiple comparisons. The proportional hazards assumption was tested using “log-log plots” (i.e. ln(-ln(survival)) versus ln(time)) for each statistically significant variable. Interaction terms
were incorporated in the regression model for each IPI, LLMPP, and biological signature subgroup before generating stratified estimates of effect size.

Recursive partitioning using CART (Salford Systems, San Diego, CA) was used to develop a classification tree for the 2-year disease-specific overall survival rate (dOS2) based on all the genes in Table 1. This algorithm sequentially divides a group of patients into subgroups that become progressively more homogeneous with respect to clinical outcome than the original group. At the first step, all factors entered are examined at every possible cutoff (number of patients minus one) to select the factor that best splits the entire group of patients into two subgroups. At the second, as well as subsequent steps, a factor is selected (again from all potential factors and cutoffs), using the same process to subdivide the current group. The process stops when no further improvement in homogeneity can be made by splitting using the available factors. To minimize overfitting the data, groups with fewer than 5 patients were not further split. The tree was then pruned to develop final risk groups, between which the 2-year dOS rates differed significantly (p<0.01). The 95% confidence intervals for dOS2 were calculated using the binomial distribution. All p-values reported are two tailed. Statistical analysis was performed with the STATA software (STATA Corporation v.9, Research Park, Texas, 2005).

Results

The median-adjusted expression of genes in the ABC transporter and GSH families in previously untreated patients with DLBCL from the primary dataset was calculated and is displayed in Table 1. In univariate analysis, none of the ABC transporters evaluated was associated with dOS. Among members of the glutathione pathway, glutathione peroxidase 1
(GPx1) was found to have a highly significant association with dOS. The hazard ratio for a
tenfold increase in GPx1 expression was 1.68 (95% CI: 1.26 – 2.22, p <0.001).

**Classification and Regression Tree (CART) Analysis**

In order to study early resistance to treatment rather than late relapse, we evaluated the 31
failures (59% of total) occurring in the first 2 years and developed a classification tree for the
rate of dOS at 2 years based on the median-adjusted expression of genes in Table 1. CART
analysis identified the most powerful predictor of the two-year dOS rate to be GPx1 expression
with a cutoff at 8.9 (Figure 2). Specifically, 36/39 subjects (92%, 95%CI: 80% – 97%) with low
GPx1 expression (median 6.1, range: 1.7 – 8.9) were alive and disease-free at 2 years, compared
to 58/91 (64%, 95% CI: 53% – 73%) with high GPx1 expression (median 14.5, range: 8.9 –
50.6) (p=0.001). This analysis was also performed with IPI included for the 116 patients with a
known IPI value. Interestingly, CART identified GPx1 and not IPI as the best predictor of
outcome in that approach.

In the low GPx1 subgroup, MDR1 (ABCB1) emerged as the best predictor of outcome,
at an expression cutoff of 1. ABCB1 expression above the cutoff (median 1.4, range: 1.0 – 2.0)
was thus associated with a two-year dOS rate of 40% in 5 patients (95% CI: 12% – 78%). In
contrast, low expression of ABCB1 (median 0.6, range: 0.1 – 1.0) identified a subgroup of 34
patients with no failures (100% 95% CI: 90% – 100%). The difference in two-year dOS between
the two subgroups defined by ABCB1 was statistically significant (p=0.001). With an additional
median follow-up of 3.1 years, only 7 late relapses occurred in the subgroup (N=34)
characterized by low levels of expression of both ABCB1 and GPx1 (Figure 3). Thus, dOS was
highly statistically significantly different with long-term follow up in this patient cohort
compared to the others (p<0.001). Moreover, this cohort with excellent prognosis was comprised of patients in all IPI and LLMPP signature subgroups (Table 2).

In the high GPx1 expression subgroup, the best predictor of two-year dOS was expression of glutathione-s-transferase GSTA1 with a cutoff of 2.2. The group with low GSTA1 expression (n=63) had a poor outcome (dOS2=54%, 95% CI: 42% – 66%, p=0.004). In contrast, high levels of GSTA1 were favorably associated with outcome, with 24/28 subjects (86%, 95% CI: 68% – 94%) being alive and disease-free at 2 years. In this subgroup 7 additional failures occurred (11 total), with an additional median follow up of 2.8 years (Table 2). There was no statistically significant difference in long-term outcome in this subgroup compared to the others (p=0.30).

**Glutathione Peroxidase 1**

As noted, GPx1 expression was found to be significantly associated with both early failure by recursive partitioning and dOS by survival analysis in the primary dataset. We thus chose to characterize GPx1 expression further in this patient cohort. Median GPx1 expression was 12.3 (range: 1.7 – 50.6). We observed no correlation between GPx1 expression and IPI score (Kruskal-Wallis test for heterogeneity p=0.15) or LLMPP subgroup (p=0.24). Interestingly, we observed a significant association of GPx1 expression with the BCR biological subtype identified by Monti et al\textsuperscript{34} (median 8.9) compared to the OxPhos (median 15) and HR (median 16) subgroups (p<0.0001).

As reported above, GPx1 was significantly associated with dOS in the primary dataset, with a hazard ratio for a tenfold increase in GPx1 expression of 1.68 (95% CI: 1.26 – 2.22, p <0.001). The association of GPx1 expression and dOS persisted when adjusted for IPI score,
with a stratified hazard ratio of 1.72 (95% CI: 1.26 – 2.34, p=0.001) or biological subtype (stratified hazard ratio: 1.63, 95% CI: 1.17 – 2.28, p=0.004). When adjusted for LLMPP signature subgroup, GPx1 expression also remained significantly associated with dOS, albeit with an effect modification (p=0.022) (Figure 4). Specifically, in the activated B-cell-like subgroup (N=24) high GPx1 expression had a pronounced adverse effect on dOS, with a hazard ratio of 7.7 (95%CI: 1.9 – 31, p=0.004). The association with non-ABC subgroups was still significant but at a lower hazard ratio of 1.79 (95% CI: 1.30 – 2.47, p<0.001). These results suggest that there may be a synergistic association of GPx1 expression and the activated B-cell-like phenotype with clinical outcome.

A recent cDNA microarray-based study in patients with DLBCL generated a “redox signature score”\textsuperscript{21} that was found to correlate with overall survival in this lymphoma. Given the significant role of the GSH pathway, and GPx1 in particular, in ROS defense, we wanted to test whether the “redox signature score” had any prognostic significance independently of GPx1 in our dataset. We found no significant association of the redox score with dOS (hazard ratio for a tenfold change: 1.1, 95% CI: 0.99 – 1.2, p=0.070). There was a borderline significant difference in dOS between the highest and lowest quartiles of the score (hazard ratio 2.1, 95% CI: 0.99 – 4.6, p=0.05) with the highest level of the score being associated with a poor clinical outcome, in contrast to the observation by Tome \textit{et al.}\textsuperscript{21} This result is consistent with the association we found, since elevated levels of GPx1 are associated with an elevated redox score and with a poor clinical outcome. In fact, GPx1 expression appears to be the most significant component of that score. When GPx1 expression was removed from the redox signature, no association of the latter with clinical outcome was observed.
Evaluation of GPx1 in an independent dataset

In the independent dataset there are 39 subjects with diffuse large B-cell lymphoma that were evaluated prior to the initiation of chemotherapy. The median dOS was similar in the primary and validation datasets (95 months and 116 months respectively, p=0.22, Figure 5a). In order to compare the two different patient populations, we examined the distribution of subjects across the IPI score (for the 116 patients with available data in the primary dataset) and we found them to be equivalent (p=0.12 for heterogeneity). GPx1 expression in this dataset was also found to be adversely associated with dOS with a hazard ratio for a tenfold change in expression at 1.7 (95%CI: 1.05 – 2.8, p=0.033). Moreover, the survival outcome among subjects by tertile of GPx1 expression was comparable between the two datasets (Figure 5b,c,d). Due to the limited number of patients in the second dataset, we could not study the low-risk group identified by low GPx1 and ABCB1 expression in this patient cohort (no subjects met the cutoff criteria).

Discussion

In an effort to identify the molecular basis for the differential response to anthracycline- and alkylator-based chemotherapy among patients with DLBCL we have analyzed the expression of genes in the ABC transporter family and the GSH pathway in tumor tissue samples obtained prior to therapy. Our study demonstrates that high expression of GPx1, the major effector of GSH-dependent ROS defense, is adversely associated with early failure as well as disease-specific overall survival. This effect is observed even after adjusting for biological subtype, IPI score, and LLMPP cell of origin signature, the most widely used prognostic tools in this disease. Moreover, using CART analysis, we have identified a subgroup of patients with
low levels of GPx1 expression that has superior short-term and long-term clinical outcomes. GPx1 expression was also significantly associated with time to treatment failure in an independent dataset comprised of 39 DLBCL patients from a different institution. We have made the assumption that microarray-derived expression levels reflect the expression of these genes in cancer cells. They may however, also reflect gene expression in the organism as a whole, or expression in tumor infiltrating cells, such as macrophages and T-cells. To that effect, it is interesting that GPx1 levels were not significantly different in the HR biological subtype which is characterized by infiltrating host cells, compared to the others.

Assuming that the derived expression values primarily reflect mRNA levels in the tumor cells, our results suggest that augmented ROS- and oxidative-stress defense may protect DLBCL cells from chemotherapy-induced death. This hypothesis is supported by several lines of investigation. Oxidative stress can induce apoptosis through numerous mechanisms. Several anthracycline- and alkylator-resistant cancer cell lines exhibit increased levels of enzymes within the free-radical defense system, including increases in GSH, GCL and GPx1. L-Buthionine-S-sulphoximine (BSO), a GCL inhibitor, which depletes intracellular GSH levels, has been shown to sensitize cells to treatment with alkylators and anthracyclines. Overexpression of GPx1 and other antioxidant enzymes has been shown to suppress apoptosis in cancer cell lines and tumor xenograft models, including in the setting of doxorubicin treatment. In lymphoma, the oncogene bcl2, overexpressed in 50% of DLBCL, is known to inhibit apoptosis, and has been shown to do so in part through the antioxidant pathway and GPx1. To our knowledge, only one prior study has suggested that GPx1 overexpression did not have a significant protective effect on doxorubicin-induced apoptosis.
Our results challenge the findings of a previous cDNA microarray-based study of tumor samples from a patient population with DLBCL \textsuperscript{21}, which found increased expression of antioxidant defense enzymes, including GPx1, to be associated with better prognosis. When we actually applied the “redox signature score” generated by these authors to our primary dataset, we found no significant association with dOS, except for the lowest and highest quartiles. Even then, the association was in the opposite direction of what was previously described, and in favor of our hypothesis that a high oxidative-stress defense, especially through high levels of GPx1, is associated with inferior clinical outcome. This adverse association was confirmed by us in an independent oligonucleotide dataset. There are several key differences in our approach, which can account for the different conclusions. The dataset on which the Tome \textit{et al}\textsuperscript{21} study is based, includes patients seen across decades at the NCI and treated with a variety of anthracycline-based chemotherapy regimens. The dataset we have employed is comprised of a more homogeneous group of patients from a single institution treated with CHOP chemotherapy. Most importantly, the study by Tome at al\textsuperscript{21} is based on a cDNA platform, while our dataset is derived from an oligonucleotide platform. Comparative microarray studies have shown generally moderate correlation between expression levels among these two inherently different platforms.\textsuperscript{43,44} Moreover, oligonucleotide-based arrays exhibit higher probe set specificity and are considered more reliable for assessing differential gene expression among samples.\textsuperscript{45,46}

In previous reports, several ABC transporters have been implicated in drug resistance in cancer cell lines and patient samples from a variety of malignancies. The association between ABC transporter expression and clinical outcomes in patients with lymphoma and DLBCL has been controversial. The ABC transporter MDR1 (ABCB1) is present by IHC in as many as 49% of newly diagnosed lymphomas.\textsuperscript{47} Several studies have documented inferior complete response
rates among patients with MDR1 overexpression at the protein or mRNA level.\textsuperscript{5,7-8} However, a recently reported large IHC-based series of 207 patients showed no such association.\textsuperscript{6} Similarly, MRP1 (ABCC1) is present by IHC in 44\% to 63\% of previously untreated lymphomas.\textsuperscript{5,48} MRP1 had no impact on response to chemotherapy or survival in 48 patients examined by Filipits \textit{et al}.\textsuperscript{48}, but when co-expressed with MDR1 it was shown to correlate with inferior response rates but not overall survival in a similar patient population.\textsuperscript{5} In our study, we observed a significant association between \textit{MDR1} (ABCB1) expression and outcome by CART analysis but not by the univariate linear-effects Cox model. ABCB1 expression was associated with outcome only in the subgroup of patients with low-level GPx1 expression. The nature of this association (confounded by GPx1 and possibly other untested genes) may explain why it is seen in certain lymphoma patient populations and not in others. We did not observe an association between BCRP (ABCG2), MRP1 (ABCC1), or MRP2 (ABCC2) with outcomes by either analysis, confirming the prior IHC based studies. Our study is based on gene expression rather than protein detection by IHC. Even though mRNA profiling does not account for post-translational modification and protein localization, it allows a quantitative comparison between patient samples compared to IHC, which is semi-quantitative.

The glutathione-S-transferases (GSTs) have been associated with drug resistance to several chemotherapeutic agents in cells lines as well as primary patient tissues\textsuperscript{25}, especially in association with low GSH levels or inhibition of GSH synthesis.\textsuperscript{13,24} Besides GSH conjugation activity, GSTs have organic peroxidase activity\textsuperscript{49} and may also regulate key components of the cell survival and cell death signaling pathways.\textsuperscript{29,30} In our study, GSTA1 was favorably associated with outcome in the subgroup of patients with high GPx1 levels. However, the effect of GSTA1 did not persist with longer follow up.
In conclusion, we have used two independent expression microarray datasets derived from untreated patients with diffuse large B-cell lymphoma to evaluate expression of genes in the ABC transporter and GSH families and how it relates to clinical outcome after CHOP chemotherapy. In a disease such as DLBCL, with a well-established role of drug resistance in treatment failure \textit{in vivo}, we have identified several members of the GSH pathway, and most notably GPx1, to be significantly adversely associated with early failure as well as disease-specific overall survival, in accordance with most published \textit{in vitro} studies. We demonstrated that GPx1 expression had a robust association with outcome across clinically and biologically defined risk groups, as well as across datasets. Moreover, we were able to identify a group of patients with low-level expression of GPx1 and MDR1 (ABCB1) that had superior short-term and long-term clinical outcomes. The generalizability of our observations needs to be assessed in ongoing studies utilizing other anthracycline-containing regimens or chemotherapy and monoclonal antibody treatment combinations in patients with DLBCL. Nevertheless, our findings provide the rationale for prospective intervention-based studies that target patients with a high likelihood of chemotherapy resistance for alternative treatment strategies.

\textbf{Acknowledgements}

We thank Drs. Margaret A. Shipp and Donna Neuberg for providing supplemental patient information on the primary dataset. We also thank Dr. Georgia Hatzivassiliou for her insightful comments and editorial support.
References


Table 1: Association of gene expression with disease-specific overall survival in the primary dataset

<table>
<thead>
<tr>
<th>Gene</th>
<th>Probe set ID(^a)</th>
<th>IQR Range(^b)</th>
<th>Coefficient(^c)</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCB1</td>
<td>209993_at</td>
<td>0.50 – 0.88</td>
<td>0.01</td>
<td>-0.7 – 0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>MDR1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCC1</td>
<td>202804_at</td>
<td>1.2 – 2.3</td>
<td>0.01</td>
<td>-0.2 – 0.3</td>
<td>0.9</td>
</tr>
<tr>
<td>MRP1</td>
<td>202805_s_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCC2</td>
<td>206155_at</td>
<td>0.07 – 0.24</td>
<td>0.34</td>
<td>-0.9 – 1.5</td>
<td>0.6</td>
</tr>
<tr>
<td>MRP2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCG2</td>
<td>209735_at</td>
<td>0.7 – 1.2</td>
<td>-0.11</td>
<td>-0.7 – 0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>BCRP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCLM</td>
<td>203925_at</td>
<td>0.61 – 0.99</td>
<td>0.31</td>
<td>-0.4 – 1.0</td>
<td>0.4</td>
</tr>
<tr>
<td>GCLC</td>
<td>202922_at</td>
<td>1.3 – 2.0</td>
<td>0.24</td>
<td>0.008 – 0.5</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>202923_s_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPx1</td>
<td>200736_s_at</td>
<td>8.5 – 17</td>
<td>0.05</td>
<td>0.02 – 0.08</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GPx2</td>
<td>202831_at</td>
<td>0.94 – 1.6</td>
<td>0.20</td>
<td>-0.1 – 0.7</td>
<td>0.1</td>
</tr>
<tr>
<td>GPx3</td>
<td>201348_at</td>
<td>1.7 – 3.5</td>
<td>0.07</td>
<td>-0.02 – 0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>214091_s_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPx4</td>
<td>201106_at</td>
<td>7.4 – 13</td>
<td>-0.02</td>
<td>-0.07 – 0.04</td>
<td>0.5</td>
</tr>
<tr>
<td>GSTA1</td>
<td>203924_at</td>
<td>0.17 – 0.55</td>
<td>-0.74</td>
<td>-1.6 – 0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>GSTA2</td>
<td>202478_at</td>
<td>0.57 – 1.1</td>
<td>0.09</td>
<td>0.4 – 0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>GSTA3</td>
<td>222102_at</td>
<td>0.11 – 0.24</td>
<td>-0.28</td>
<td>-1.8 – 1.2</td>
<td>0.7</td>
</tr>
<tr>
<td>GSTA4</td>
<td>202967_at</td>
<td>1.4 – 2.2</td>
<td>-0.34</td>
<td>-0.7 – 0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>GSTM1</td>
<td>204550_x_at</td>
<td>1.8 – 2.5</td>
<td>-0.02</td>
<td>-0.1 – 0.06</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>215333_x_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTM4</td>
<td>204149_s_at</td>
<td>0.05 – 0.19</td>
<td>-0.02</td>
<td>-0.4 – 0.4</td>
<td>0.9</td>
</tr>
<tr>
<td>GSTM5</td>
<td>205752_s_at</td>
<td>1.4 – 2.1</td>
<td>-0.30</td>
<td>-0.7 – 0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>GSTP1</td>
<td>200824_at</td>
<td>5.9 – 14</td>
<td>-0.007</td>
<td>-0.04 – 0.02</td>
<td>0.6</td>
</tr>
<tr>
<td>GSTT1</td>
<td>203815_at</td>
<td>0.09 – 0.18</td>
<td>-0.08</td>
<td>-1.7 – 1.5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

IQR: inter-quartile range
a: ID from the Affymetrix HG-U133 A and B microarray set (Affymetrix, Santa Clara, CA).  
b: median-adjusted expression signal.  c: $\beta_x$ coefficient from Cox proportional hazards model.
Table 2: Comparative distribution of IPI and LLMPP signature score among patients in the low- and high-risk groups identified by CART

<table>
<thead>
<tr>
<th></th>
<th>Study Cohort</th>
<th>Low GPx1/ABCB1</th>
<th>High GPx1/GSTA1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td><strong>IPI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45 (39)</td>
<td>11 (37)</td>
<td>13 (52)</td>
</tr>
<tr>
<td>2</td>
<td>27 (23)</td>
<td>8 (27)</td>
<td>6 (24)</td>
</tr>
<tr>
<td>3</td>
<td>32 (28)</td>
<td>8 (27)</td>
<td>2 (8)*</td>
</tr>
<tr>
<td>4</td>
<td>12 (10)</td>
<td>3 (10)</td>
<td>4 (16)</td>
</tr>
<tr>
<td><strong>LLMPP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCB</td>
<td>64 (49)</td>
<td>20 (59)</td>
<td>12 (43)</td>
</tr>
<tr>
<td>ABC</td>
<td>24 (18)</td>
<td>7 (21)</td>
<td>4 (14)</td>
</tr>
<tr>
<td>Type 3</td>
<td>42 (32)</td>
<td>7 (21)</td>
<td>12 (43)</td>
</tr>
</tbody>
</table>

Data available on 116 subjects for IPI and 130 subjects for LLMPP signature. Percentages reflect distribution of patients in each cohort. * Significantly different from entire cohort; Fisher’s exact for heterogeneity (p=0.048)
Figure Legends

Figure 1: Role of the glutathione pathway and ABC-transporter family in drug elimination and defense from ROS-mediated oxidative stress. Shaded diamonds represent intracellular drug and drug metabolites. Glutathione (GSH) is represented by clear pentagons. The rate-limiting step in its de novo biosynthesis is the enzyme glutamate-cysteine ligase (GCL). The balance of drug/drug-metabolites and intracellular GSH is a significant determinant of drug levels and ROS stress. Glutathione peroxidase 1 (GPx1) reduces hydrogen peroxide generated after exposure to certain drugs, by oxidizing GSH to its disulfide form GSSG. Glutathione s-transferases (GST) conjugate GSH to drugs and drug metabolites, thus facilitating drug-conjugate elimination and efflux. GSTs also function as peroxidases to a lesser extend than GPx1 (not shown for simplicity). The ABC transporters ABCB1 (MDR1), ABCC1 (MRP1), and ABCC2 (MRP2) among others are involved in ATP-dependent efflux of drugs or drug-GS conjugates.

Figure 2: Classification and regression tree analysis (CART) for the dOS rate at 2 years. The cutoffs represent median-adjusted expression values of the genes indicated. Circles represent branch points and squares are terminal nodes. The p-values at each split are calculated for the binominal distribution.

Figure 3: Long-term outcome by CART classification. The group with low GPx1 and low ABCB1 expression levels was defined by the 2-year dOS rate and has no early failures. Tick marks reflect censoring.
**Figure 4: Effect of GPx1 expression and the LLMPP signature score on dOS.** Patients are grouped in quartiles by GPx1 expression, with quartile 1 having the lowest and quartile 4 the highest level. The association with dOS is presented for the entire cohort (a), patients with the activated B-cell (ABC) subtype (b), and patients with non-ABC disease subtypes (c). Censoring not shown for simplicity.

**Figure 5: Clinical outcome among patients in the primary and validation datasets.** DOS is presented for the primary and validation datasets (a). DOS is compared among patients in the lowest (b), middle (c), and highest (d) tertile of GPx1 expression in each dataset. Tick marks reflect censoring.
Figure 2

For personal use only.on October 23, 2017. For personal use only.
Figure 3
Figure 4

(a) [Graph showing survival rates for different quartiles over months.]  
(b) [Graph showing survival rates for different quartiles over months.]  
(c) [Graph showing survival rates for different quartiles over months.]
Figure 5
Members of the glutathione and ABC-transporter families are associated with clinical outcome in patients with diffuse large B-cell lymphoma

Charalambos Andreadis, Phyllis A Gimotty, Peter Wahl, Rachel Hammond, Jane Houldsworth, Stephen J Schuster and Timothy R Rebbeck