LYMPHOID NEOPLASIA

Single nucleotide variation in the TP53 3′ untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the International DLBCL Rituximab-CHOP Consortium Program

Yong Li,1,2 Michael W. Gordon,1 Zijun Y. Xu-Monette,2 Carlo Visco,3 Alexander Tzankov,4 Dehui Zou,5 Lugui Qiu,6 Santiago Montes-Moreno,6 Karen Dybkaer,7 Attilio Orazi,7 Youli Zu,9 Govind Bhagat,10 Kristy L. Richards,11 Eric D. Hsi,12 William W. L. Choi,13 J. Han van Krieken,14 Qin Huang,15 Weiyun Ai,16 Maurilio Ponzoni,17 Andrés J. M. Ferreri,17 Jane N. Winter,18 Ronald S. Go,19 Miguel A. Pins,19 Michael B. Müller,20 Lin Wu,21 Michael Wang,22 Kenneth S. Ramos,1 L. Jeffrey Medeiros,2 and Ken H. Young2

1Department of Hematopathology and Blood and Marrow Transplantation, The University of Texas MD Anderson Cancer Center, Houston, TX; 2Department of Pathology, Columbia University School of Medicine, New York, NY; 3Department of Pathology, University Hospital, Basel, Switzerland; 4Institute of Hematology, Chinese Academy of Medical Science and Peking Union Medical College, Tianjin, China; 5Department of Pathology, Hospital Universitario Marqués de Valdecilla/IFIMAV, Santander, Spain; 6Department of Hematology, Aarhus University Hospital, Aalborg, Denmark; 7Department of Pathology, Weill Medical College of Cornell University, New York, NY; 8Department of Pathology, The Methodist Hospital, Houston, TX; 9Department of Pathology, Columbia University School of Medicine, New York, NY; 10Department of Hematology, University of North Carolina School of Medicine, Chapel Hill, NC; 11Department of Pathology, Cleveland Clinic, Cleveland, OH; 12Department of Pathology, University of Hong Kong Li Ka Shing Faculty of Medicine, Hong Kong, China; 13Department of Pathology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 14Department of Pathology, University of California San Francisco School of Medicine, San Francisco, CA; 15Departments of Hematology and Pathology, San Raffaele H. Scientific Institute, Milan, Italy; 16Department of Hematology, Feinberg School of Medicine, Northwestern University, Chicago, IL; 17Department of Hematology, Gundersen Lutheran Health System, La Crosse, WI; 18Department of Pathology, Odense University Hospital, Odense, Denmark; 19Department of Genomic Oncology, Roche Molecular Systems Inc., Pleasanton, CA; and 20Department of Lymphoma and Myeloma, The University of Texas MD Anderson Cancer Center, Houston, TX

We identified multiple single nucleotide variants (SNVs) in the TP53 3′ untranslated region (3′UTR) in tumor specimens from 244 diffuse large B-cell lymphoma (DLBCL) patients with a wild-type TP53 coding sequence (CDS) and 1 or more 3′UTR SNVs had a better 5-year survival rate than patients carrying a wild-type CDS and the reference 3′UTR, yet there is no statistically significant difference in overall survival (OS). In contrast, 3′UTR variation predicted poorer OS for patients with a mutant TP53 CDS. We then sequenced TP53 3′UTR in 247 additional DLBCL patients as a validation set. Altogether, we identified 187 novel SNVs; 36 occurred at least twice. Most of the newly identified 3′UTR SNVs were located at sites that are complementary to seed sequences of microRNAs (miRNAs) that are predicted or experimentally known to target TP53. Three SNVs disrupt the seed match between miR-125b and the TP53 3′UTR, thereby reducing suppression of p53 by this miRNA. In addition, a germline SNP (rs78378222) located in the TP53 polyadenylation signal resulted in downregulation of both p53 messenger RNA and protein levels and reduction of cellular apoptosis. This study is the first to demonstrate the prognostic value of the TP53 3′UTR in cancer. (Blood. 2013;121(22):4529-4540)

Introduction

The tumor suppressor p53, encoded by TP53, is recognized as the guardian of the human genome. p53 responds to diverse stressors by regulating many downstream target genes that promote cellular DNA repair, apoptosis, cell cycle arrest, and senescence. Restoring p53 function alone is sufficient to cause regression of several types of cancer in mice.1,2 This characteristic suggests that p53 itself or the p53 pathway must be derailed for tumor development, genetically or otherwise. Three distinct sets of nucleotide alterations can occur in the human TP53 gene: tumor-associated (somatic) mutations, germline Li-Fraumeni mutations, and germline TP53 polymorphisms.3,4 Both tumor-associated mutations and Li-Fraumeni mutations occur in the TP53 coding sequence (CDS) and produce mutant p53 proteins that lack most or all of the normal p53 functions and confer oncogenic properties.3 For instance, a missense base substitution at codon 273 of human TP53 (R273H) is frequently observed in somatic tumors and in tumors from Li-Fraumeni syndrome patients. Mouse models demonstrate that this mutant p53 allele has enhanced oncogenic potential that cannot be explained by simple loss of p53 function.5,6

In general, p53 CDS mutations contribute to oncogenesis by 2 distinct mechanisms: (1) mutant alleles can have dominant negative effects on wild-type (WT) p53 function; and (2) mutant p53 protein can actively promote tumorigenesis independently of WT p53...
function. It is accepted that the WT TP53 CDS produces native p53 tumor suppressor, whereas most TP53 CDS mutations produce oncogenic mutant p53. By contrast, only a few of the >200 naturally occurring germline variants of TP53 in human populations cause measurable perturbation of p53 function.\(^3\) More than 90% of these p53 variants occur in 10 introns alleged to have no cancer-related biological consequences. None of the 20 variants in the TP53 CDS have been linked reproducibly to cancer predisposition.\(^4\) The best-studied TP53 CDS polymorphism is codon 72 (R72P), owing to an earlier finding that human papillomavirus causes more efficient degradation of p53\(^9\)\(^{12}\) than p53\(^{12}\). However, few associations between this polymorphism and various types of cancer have been found.\(^3\) Sequence variation in the untranslated regions (UTRs) of the TP53 gene has not been extensively studied.

Diffuse large B-cell lymphoma (DLBCL) is the most common type of lymphoma and is clinically heterogeneous with subsets of patients with better or worse prognosis.\(^9\)\(^{10}\) Biologically, DLBCL has been known to be heterogeneous with germinal center B-cell (GCB) and activated B-cell (ABC) subtypes. Recent advancements in next-generation sequencing have further broadened our understanding of genetic lesions in DLBCL. Three DLBCL-sequencing projects have provided a wealth of information on whole genomes, exomes, and/or genetic lesions in DLBCL. Three DLBCL-sequencing projects have been known to be heterogeneous with germinal center B-cell (GCB) or the ABC subtype, and that polyallelic deletion and loss of TP53\(^9\)\(^{10}\) has been found to be associated with lymphoma subtypes. Recent advancements in next-generation sequencing have further broadened our understanding of genetic lesions in DLBCL. Three DLBCL-sequencing projects have provided a wealth of information on whole genomes, exomes, and/or genetic lesions in DLBCL.

In a previous study, we found that TP53 CDS mutations predicted poorer survival in 506 n-embedded DLBCL patients, none of whom were included in the 506 DLBCL patients in the previous p53 CDS study.\(^17\) All cases were reviewed by a group of hematopathologists (S.M.M., M.A.P., M.B.M., A.T., and K.H.Y.), and the diagnoses confirmed by World Health Organization classification criteria. The current study and material transfer agreement were approved by the institutional review board of each participating center. The overall collaboration was approved by the Institutional Review Board of The University of Texas MD Anderson Cancer Center, Houston, Texas. This study was conducted in accordance with the Declaration of Helsinki.

All of these patients have undergone molecular and phenotypic classification with available clinical data. There are no statistical differences in regard to clinical parameters (age, gender, stage, lactate dehydrogenase level, tumor size, performance status, chemotherapy regimen, survival time, and recurrence rate), GCB/ABC classification and distribution, and survival among 3 groups (244 cases in the training set, 247 cases in the validation set, and unselected remaining 262 cases).

**DNA sequencing and cell biology**

Genomic DNA was extracted from formalin-fixed, paraffin-embedded tissue samples using the HighPure paraffin DNA extraction kit (Qiagen). The TP53 UTRs were sequenced bidirectionally using the Sanger method at Polymeric DNA Inc. (Alameda, CA). All sequences were compared with the TP53 reference sequence (NM_000546.5). Single nucleotide variants (SNVs) in the 3' UTR were grouped into single nucleotide polymorphisms (SNPs) and novel SNVs (nSNVs). Any nucleotide change listed in the dbSNP database was considered an SNP; the remainder (none occurring in $>$1% of patients) were considered nSNVs.

The p53-null human lung adenocarcinoma cancer cell line H1299 was obtained from the American Type Culture Collection (Manassas, VA). For mutational analysis, the p53 expression vector pCMV-p53 (Clontech, Mountain View, CA) was modified to contain TP53 3' and 3'UTRs as well as the 2.5-kg genomic DNA downstream of the 3'UTR. The resulting constructs were designated as TP53Ext, TP53ExtC1170 T, and TP53ExtA1175 C, respectively. MicroRNA (miRNA) expression vectors and Western blotting analysis were described previously\(^18\) and are included in the supplemental Materials and methods (see the Blood Web site).

**Statistical analysis**

All statistical calculations were performed using GraphPad Prism 5. Overall survival (OS) was calculated from time of diagnosis to time of death from any cause or last follow-up. Progression-free survival (PFS) was calculated from time of diagnosis to time of progression or death from any cause. The actuarial probabilities of OS and PFS\(^17\) were determined using the Kaplan-Meier method, and differences were compared using the log-rank test. The 5-year survival and 5-year PFS rates describe the percentage of patients that are alive 5 years after their cancer is diagnosed; they were compared using Fisher’s exact test. All differences with $P < .05$ were considered statistically significant.

**Results**

**Distribution and frequency of SNVs in TP53 UTRs in 244 de novo DLBCL patients**

In the initial group of 244 DLBCL samples, we identified 128 SNVs in the UTRs of 107 patients, with 14 in the 5'UTR and 114 in the 3'UTR (Figure 1A; supplemental Table 1). Six 3'UTR SNVs have been reported as SNPs in dbSNP and occurred 37 times (Table 1). We designated the newly found noncoding SNVs in the TP53 gene as nSNVs. A total of 108 nSNVs were identified with 121 occurrences

---

**Materials and methods**

**Patients**

For the initial study group, we randomly selected 244 patients (training set) with DLBCL from a total of 506 de novo DLBCL patients treated with R-CHOP, whose TP53 CDSs had been sequenced in a previous study by our consortium.\(^17\) For the validation set, tumor specimens were also obtained from 247 DLBCL patients receiving R-CHOP treatment in the consortium program, none of whom were included in the 506 DLBCL patients in the previous p53 CDS study.\(^17\)
in the 3’UTR in 76 patients. More patients had at least 1 nSNV in the 3’UTR (n = 76, 31%) than they had nonsynonymous mutations in the CDS (n = 49, 20%). In the 76 patients with 3’UTR nSNVs, 32 (42%) had multiple lesions, compared with 2 of 49 (4%) patients with 2 or more CDS mutations. Ten of the TP53 3’UTR nSNVs occurred twice, and 1 occurred 3 times (Figure 1A). The 5’UTR SNVs occurred at a much lower frequency (n = 15 patients, 6%; supplemental Table 1; none of them reported in dbSNP) and were not further analyzed in the current study.

Most TP53 CDS somatic mutations are clustered within exons 5 to 8.19 Although the nucleotide length (544 bp) of exons 5 to 8 accounts for only ~46% of the coding exons (1182 bp), mutations in these exons account for 94% of all CDS somatic mutations.19 Our analysis of TP53 CDS in the 506 DLBCL patients confirmed the uneven codon distribution of single nucleotide missense substitutions; mutations in exons 5 to 8 account for 90% of all mutations in coding exons (exons 3-10 plus parts of exon 2 and exon 11).17 In the initial study group of 244 patients, 92% of all CDS mutations were found in exons 5 to 8 (45 of 49 cases, Figure 1B). In contrast, 3’UTR nSNVs were evenly dispersed across the 1207-bp 3’UTR in the 76 patients with nSNVs (Figure 1A). The distribution of SNVs in the 5’UTR in the 244 cases was similar to that in the 3’UTR (Figure 1C). These data suggest that UTR variants in the TP53 gene are distributed more evenly than are CDS mutations.

**Survival in relation to clinical characteristics and TP53 3’UTR variation**

The clinical and biological characteristics of the 244 DLBCL patients and their impact on OS and PFS are summarized in supplemental Table 2. We first verified that no adverse clinical variables were associated with variation in the 3’UTR. When patients were divided based on the presence of TP53 3’UTR lesions (SNPs only, nSNVs only, or any SNVs [either SNP or nSNV]), no significant difference in OS or PFS was found (Figure 2). TP53 CDS mutations (MUT-CDS) predicted poorer OS (P = .0028) and PFS (P = .0086) in this 244-patient cohort (Figure 3A-B). We then divided patients into 2 groups according to the p53 CDS status and analyzed survival in relation to SNVs. We found that SNPs in the TP53 3’UTR had no impact on
patient OS or PFS (supplemental Figure 1). Of the patients with WT-
CDS only, those without nSNVs (n = 136) seemed to have poorer
OS and PFS than those with nSNVs (n = 59), but the difference was
not statistically significant (P = .15 for OS; P = .17 for PFS,
respectively; Figure 3C-D). When 5-year survival rates were
calculated, nSNVs conferred better survival in patients with WT-
CDS (78% vs 59%, P = .025), and 5-year PFS also showed a trend
(76% vs 58%, P = .079). In contrast, patients with MUT-CDS and
nSNVs (n = 17) had poorer OS and PFS than did patients with
MUT-CDS and no nSNVs (n = 32; P = .019 for OS; P = .018 for
PFS; Figure 3E-F). Indeed, all patients with MUT-CDS and nSNVs
died within 76 months after diagnosis (Figure 3E). We also
combined SNPs and nSNVs as total SNVs; there were similar
survival trends, but they were not statistically significant (supple-
mental Figure 2). After the cohort was divided into ABC- and GCB-
DLBCL subsets, the prognostic impact of nSNVs in OS seemed to
favor patients with WT-CDS and not favor patients with MUT-CDS
in the GCB subset. Yet none were statistically significantly, most
due to a smaller cohort (supplemental Figure 3). Again, nSNVs
conferred better 5-year survival in GCB patients with WT-
CDS (P = .0006). We did not find statistical differences in
frequencies of subtypes, CDS mutation, and TP53 monoallelic
deletion between patients with nSNVs and those without.

Validation of TP53 3'UTR variants
To determine whether nSNVs are common in DLBCL, we sequenced
the TP53 3'UTR in tumor samples from an additional 247 DLBCL
patients. We found 106 nSNVs in 70 (28%) patients with 5 occurring
twice. All 6 SNPs identified in the initial set were found in this 247
patient group (Table 1; supplemental Table 3). For the combined 491
patients, 187 nSNVs were identified in the TP53 3'UTR: 151
occurred once, 31 occurred twice, and 5 occurred 3 times (sup-
plemental Table 4). These data indicate that nSNVs in the
TP53 3'UTR occur frequently in DLBCL. In addition, nSNVs were dis-
persed across the whole TP53 3'UTR (Figure 4A), providing
additional evidence that 3'UTR nSNVs are distributed more evenly
than are CDS mutations.

3'UTR variants in miRNA-binding sites
Eleven miRNAs are reported to target the human TP53 gene: miR-25,
miR-30d,26 miR-125b,21 miR-504,22 miR-380,25 miR-92a, miR-
141,24 miR-200a,25 miR-1285,26 miR-15a, and miR-16.27 Among
these miRNAs, 3 pairs have the same seed sequence: miR-25 and miR-
92a, miR-141 and miR-200a, and miR-15a and miR-16. In addition,
miR-504 and miR-1285 target 2 sites each. Therefore, the TP53 3'UTR
has a total of 10 experimentally validated miRNA-targeting sites.

Table 1. Frequency of known TP53 3'UTR SNPs in DLBCL cohorts

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Position†</th>
<th>MAF‡</th>
<th>Heterozygous frequency</th>
<th>Validation§</th>
<th>First group (244)</th>
<th>Second group (247)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs8065799</td>
<td>A&gt;C</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1042526</td>
<td>A&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs3519705</td>
<td>C&gt;T</td>
<td>0.010</td>
<td>0.015</td>
<td>Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1695680</td>
<td>G&gt;A</td>
<td>0.032</td>
<td>0.024</td>
<td>Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs7951638</td>
<td>G&gt;A</td>
<td>N/A</td>
<td>0.500</td>
<td>Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs3449624</td>
<td>C&gt;A</td>
<td>N/A</td>
<td>0.020</td>
<td>Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1788136</td>
<td>G&gt;A</td>
<td>0.026</td>
<td>0.010</td>
<td>Cluster, Frequency, 1000Genome</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs4968187</td>
<td>G&gt;A</td>
<td>0.035</td>
<td>N/A</td>
<td>Cluster, Frequency, 1000Genome Precious</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>rs916131</td>
<td>G&gt;A</td>
<td>N/A</td>
<td>N/A</td>
<td>Precious</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1041327</td>
<td>T&gt;C</td>
<td>N/A</td>
<td>N/A</td>
<td>Precious</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1811095</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>Cluster</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs35944977</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs11483147</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs17886358</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>0.435</td>
<td>Cluster, Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs34607064</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs6759322</td>
<td>A&gt;C</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs17879353</td>
<td>C&gt;A</td>
<td>N/A</td>
<td>0.007</td>
<td>Cluster, Frequency</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>rs1042535</td>
<td>C&gt;G</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1794291</td>
<td>A&gt;G</td>
<td>N/A</td>
<td>N/A</td>
<td>Cluster</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs34182553</td>
<td>C&gt;A</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs77011241</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>0.389</td>
<td>Frequency, 1000Genome</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs53404853</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs17884306</td>
<td>G&gt;A</td>
<td>0.044</td>
<td>N/A</td>
<td>Cluster, Frequency, 1000Genome Precious</td>
<td>17</td>
<td>24</td>
</tr>
<tr>
<td>rs17884586</td>
<td>C&gt;T</td>
<td>N/A</td>
<td>0.012</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs35659787</td>
<td>G&gt;A</td>
<td>N/A</td>
<td>0.10</td>
<td>Frequency</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs55817367</td>
<td>C&gt;T</td>
<td>0.007</td>
<td>N/A</td>
<td>1000Genome</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1794292</td>
<td>G&gt;C</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs1794293</td>
<td>G&gt;C</td>
<td>N/A</td>
<td>N/A</td>
<td>—</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs114831472</td>
<td>C&gt;G</td>
<td>0.032</td>
<td>0.065</td>
<td>1000Genome</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>rs78378222</td>
<td>A&gt;C</td>
<td>0.010</td>
<td>N/A</td>
<td>Cluster, Frequency, 1000Genome</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

*Minor allele found in the initiation set (244 cases) or the validation set (247 cases). Some patients possess more than 1 SNP. Underlining in column 2 indicates SNPs found in the cohort.
†Counted from first nucleotide of the TP53 3'UTR: “del,” deletion.
‡MAF, minor allele frequency; N/A, not available.
§Validation: “Cluster,” multiple submissions mapped to the same location and clustered into 1 refSNP cluster and assigned a reference SNP ID number; “Frequency,” validated by both population frequency data and genotype data: “1000Genome,” validated by the 1000 Genomes Project; “Precious,” cited as clinically relevant in the literature and belonging to a reserved set of clinically associated SNPs in dbSNP.
Except for the miR-141 and miR-380-5p sites, 8 of 10 miRNA sites had 1 or more nSNVs occurring at least once in 491 patients (Table 2). We analyzed the potential of nSNVs to affect putative miRNA binding as predicted by TargetScan,28,29 in which 8 nucleotides in the target 3’UTR are critical to miRNA suppression: a sequence complementary to the 6-nucleotide seed sequence of the miRNAs and 2 flanking nucleotides (A1 and m8). Of the 146 patients carrying 1 or more nSNVs in the 3’UTR, 95 (65%) had

Figure 2. Kaplan-Meier estimates of OS and PFS in DLBCL patients with SNVs in the TP53 3’UTR. (A) OS of patients with or without SNPs. (B) PFS of patients with or without SNPs. (C) OS of patients with or without nSNVs. (D) PFS of patients with or without nSNV. (E) OS of patients with or without SNVs (either SNP or nSNV). (F) PFS of patients with or without SNVs.
nSNVs that disrupt and/or generate miRNA-binding sites as predicted by TargetScan; 78 of 95 patients carried nSNVs disrupting putative miRNA binding (Table 3; Figure 4B).

p53 expression analysis of 3’UTR nSNVs

To test whether 3’UTR variation impacts p53 expression, we introduced gene expression cassettes carrying human TP53 complementary DNA (cDNA) with 3’UTR variants in the miR-125b–binding site (Figure 5A) into p53-null H1299 cells. As shown in Figure 5B, TP53 3’UTR nSNVs did not alter p53 expression compared with WT. However, when cells were cotransfected with miR-125b, 3 nSNVs (733U, 737A, and 739A) disrupting the seed match between miR-125b and the TP53 3’UTR released the negative regulation of p53 by miR-125b (Figure 5C). One variant (740A) that converts the seed match from 7mer to 8mer, which is predicted to have a greater effect on p53 downregulation, did not enhance p53 suppression by miR-125b. These results demonstrate that miRNAs and nSNVs in the TP53 3’UTR significantly alter p53 protein levels in vitro. We next analyzed the relationship between p53 protein expression with a 50% cutoff17 and 3’UTR nSNVs in the initiation set. We found that there was no correlation between nSNVs and p53 expression in

Figure 3. Kaplan-Meier estimates of OS and PFS in DLBCL patients classified according to TP53 CDS status. (A) OS of patients with or without an MUT-CDS. (B) PFS of patients with or without an MUT-CDS. (C) OS of patients with a WT-CDS and nSNV. (D) PFS of patients with a WT-CDS and nSNVs. (E) OS of patients with an MUT-CDS and nSNV. (F) PFS of patients with MUT-CDS and nSNVs.
the entire cohort, the WT-CDS group, or the MUT-CDS group (supplemental Table 5). There are at least 2 possibilities: (1) nSNVs may have either inhibitory or stimulatory effects on p53 regulation; and (2) the standard assay of p53 expression in tumors (percentage of cells with positive staining rather the signal strength)\textsuperscript{17} cannot detect the subtle modulation of p53 protein levels by nSNVs.

**Functional significance of a germline polymorphism in the TP53 3′UTR**

SNP rs78378222 (A\textsuperscript{1175}C) is located in the polyadenylation signal (PAS) of TP53 and is strongly associated with basal cell carcinoma (BCC).\textsuperscript{30} This germline variant was found in 7 patients (6 A/C and 1 C/C). We predicted a decrease in p53 expression when the TP53 gene contains the PAS variant (AATAAA\textsuperscript{3}AATACA). We introduced gene expression cassettes carrying human TP53 cDNA with this PAS variant (TP53ExtA\textsuperscript{1175}C), or a control SNP rs114831472 (TP53ExtC\textsuperscript{1170}T, not associated with BCC and other cancers) located 1 nucleotide upstream of the PAS, and the 2.5-kb genomic sequence downstream of the 3′UTR (Figure 6A) into p53-null H1299 cells. We found significantly lower p53 protein levels in cells with TP53ExtA\textsuperscript{1175}C than in cells with WT control or TP53ExtC\textsuperscript{1170}T (Figure 6B). Northern blot analysis confirmed that p53 messenger RNA (mRNA) levels were downregulated in cells with the p53 PAS variant (Figure 6C). We assessed the effect of the TP53 PAS variant on apoptosis by using the terminal deoxynucleotidyl transferase-mediated dUTP nick end labeling (TUNEL) assay. In cells harboring the p53 PAS minor allele, cellular apoptosis was significantly lower than that in cells expressing the WT TP53 gene (Figure 6D). These data indicate that the TP53 variant rs78378222 hinders p53 expression and its downstream functions in apoptosis.

**Discussion**

In this work, we found many previously unreported SNVs in the TP53 3′UTR in DLBCL tumor DNA samples. 3′UTR shortening\textsuperscript{31}
‡The 7mer-m8 site comprises the seed match supplemented by a Watson-Crick match to miRNA nucleotide 8. The 7mer-A1 site comprises the seed match supplemented by an A across from miRNA nucleotide 1. The 8mer site comprises the seed match supplemented by both m8 and A1. “NO” indicates that the seed match was abolished. “?” indicates unconventional interaction.

*Number of patients carrying SNVs affecting putative miRNA binding. Unique SNVs are indicated in parentheses. SNVs affecting TP53:miRNA complementarity other than the seed match are not included. Underlining in the column 2 indicates the altered nucleotides.

†Nucleotides in the hsa-miR-3’re UTR are numbered starting from the one after the stop codon. Nucleotides complementary to the seed sequence (positions 2-7) of miRNAs are indicated. Position (pos.) numbers correspond to nucleotides that are complementary to the miRNA seed sequence (in bold).

‡SNPs 1070C>T. One miR-504 site overlaps with the miR-125b site. “x” indicates that base pairing is disrupted.

§The 7mer-m8 site comprises the seed match supplemented by a Watson-Crick match to miRNA nucleotide 8. The 7mer-A1 site comprises the seed match supplemented by an A across from miRNA nucleotide 1. The 8mer site comprises the seed match supplemented by both m8 and A1. “NO” indicates that the seed match was abolished. “?” indicates unconventional interaction.
and dysregulation of 3′UTR-binding proteins or miRNAs in cancer have been reported. This study, however, is the first to demonstrate that 3′UTR variation is widespread in the TP53 gene in cancer patients. Most DLBCLs originate from B cells in the germinal center, which have a high proliferation rate and are prone to genomic alterations. Germinal center B cells expand clonally and undergo physiological somatic hypermutation of their immunoglobulin variable-region genes and class-switch recombination to alter the immunoglobulin subtype mediated by activation-induced cytidine deaminase as part of the immune response. BCL2 point mutations in DLBCL are suspected to be caused by this somatic hypermutation process. TP53 3′UTR nSNVs found in this study are predominantly G→A and C→T substitutions, and the frequency ratio of transition over transversion is 16.7, significantly higher than 2.1 (for all human variants in dbSNP), 2.0 (TP53 CDS mutation in DLBCL), or 7.3 (known SNPs of TP53 3′UTR). However, only a small number of nSNVs are located in WRCY/RGYW sites (IUPAC degenerated codes; supplemental Table 6), hot spots for physiological hypermutations mediated by activation-induced cytidine deaminase. In addition, these nSNVs are not found in many projects that clearly sequence the full TP53 3′UTR from hundreds of individuals. Therefore, most nSNVs found in this study are likely cancer-specific mutations rather than germline polymorphisms or a result of a normally operative somatic hypermutation process. We acknowledge that the absence of patient-matched noncancerous tissue samples prevents us from completely ruling out the possibility of rare private germline variants in these patients. Counting CDS mutations and 3′UTR nSNVs, ~45% of DLBCL patients (109 of 244) had alterations in TP53, making TP53 the most frequently mutated gene identified in DLBCL. A recent study reported an R-CHOP treatment-associated pattern of p53 and cell cycle dysregulation in DLBCL, further emphasizing the importance of p53 in DLBCL pathogenesis and prognosis.

Most of the TP53 3′UTR nSNVs are located in validated or putative miRNA-binding sites. Each specific mammalian cell expresses ~70 miRNAs at an appreciable level yet normal B-cell development and the initiation of malignancy and tumor progression are ordered and dynamic processes in which hundreds of miRNAs are differentially expressed in more than a dozen distinct cell populations. Thus, 3′UTR variation may have a direct biological impact on potential interactions between miRNAs and p53 in B-cell maturation or DLBCL pathogenesis. Because miRNAs repress gene expression, 3′UTR alterations that are predicted to disrupt miRNA-binding sites, as occurred in 78 of 146 patients with 1 or more nSNVs in this study (Table 3), could reduce the efficiency of miRNA suppression and subsequently increase p53 expression. This has been experimentally validated for a few nSNVs located at the miR-125b site. We reasoned that when the TP53 CDS is WT, the increased levels of tumor-suppressive p53 protein confer better 5-year survival in patients with 3′UTR nSNVs, as WT p53 is advantageous for patients undergoing R-CHOP chemotherapy; in contrast, when the TP53 CDS is mutated concomitantly, the elevated levels of oncogenic p53 will worsen survival in patients with 3′UTR nSNVs. Our findings do show that variation in the TP53 3′UTR as a prognostic indicator of survival in DLBCL patients is dependent on the status of the CDS, yet there is no correlation between nSNV and p53 protein expression in the entire cohort, the WT-CDS group, or the MUT-CDS group. It should be noted that 3′UTR variation may also disrupt interactions between p53 mRNA and RNA-binding proteins, although the exact nucleotides have not been pinpointed. Additionally, 3′UTR variation could affect mRNA stability or translation efficiency due to the variation in the seed sequence of the miRNA plus A1 and m8. Some patients had both mutations generating miRNA binding and mutations disrupting it. Three SNPs listed in dbSNP that putatively affect miRNA binding are WT p53 is advantageous for patients undergoing R-CHOP chemotherapy; in contrast, when the TP53 CDS is mutated concomitantly, the elevated levels of oncogenic p53 will worsen survival in patients with 3′UTR nSNVs. Our findings do show that variation in the TP53 3′UTR as a prognostic indicator of survival in DLBCL patients is dependent on the status of the CDS, yet there is no correlation between nSNV and p53 protein expression in the entire cohort, the WT-CDS group, or the MUT-CDS group. It should be noted that 3′UTR variation may also disrupt interactions between p53 mRNA and RNA-binding proteins, although the exact nucleotides have not been pinpointed. Additionally, 3′UTR variation could affect mRNA stability or translation efficiency due to the variation in the seed sequence of the miRNA plus A1 and m8. Some patients had both mutations generating miRNA binding and mutations disrupting it. Three SNPs listed in dbSNP that putatively affect miRNA binding are

### Table 3: Occurrence of nSNVs in the TP53 3′UTR potentially affecting miRNA binding in DLBCL

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Training Set</th>
<th>Validation Set</th>
<th>Total 146 of 491</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TP53 CDS status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MUT-CDS</td>
<td>16 (21)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>WT-CDS</td>
<td>60 (79)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>3′UTR nSNV affecting putative miRNA binding</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generating only</td>
<td>12</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>Disrupting only</td>
<td>31</td>
<td>23</td>
<td>54</td>
</tr>
<tr>
<td>Generating and disrupting</td>
<td>15</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Others</td>
<td>18</td>
<td>33</td>
<td>51</td>
</tr>
<tr>
<td><strong>No. of 3′UTR nucleotides altered</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42 (59.5)</td>
<td>42 (60)</td>
<td>84 (57.5)</td>
</tr>
<tr>
<td>2</td>
<td>19 (23.8)</td>
<td>18 (25.7)</td>
<td>37 (25.3)</td>
</tr>
<tr>
<td>3</td>
<td>8 (10.5)</td>
<td>7 (10)</td>
<td>15 (10.3)</td>
</tr>
<tr>
<td>4</td>
<td>5 (6.4)</td>
<td>3 (4.3)</td>
<td>8 (5.5)</td>
</tr>
<tr>
<td>5</td>
<td>1 (1.2)</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>6</td>
<td>1 (1.2)</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
</tbody>
</table>

Others, outside seed match or not in miRNA sites; NA, not applicable.

*nSNVs predicted by TargetScan 5.0 to be targeted by miRNAs (based on the seed sequence of the miRNA plus A1 and m8). Some patients had both mutations generating miRNA binding and mutations disrupting it. Three SNPs listed in dbSNP that putatively affect miRNA binding.

![Figure 5](https://www.bloodjournal.org/figure5.png)

**Figure 5.** Expression analyses of TP53 3′UTR variation affecting miR-125b binding in p53-null cells. (A) Schematic representation of the p53 cDNA construct and 3′UTR variants. (B) Western blot of p53 expression in p53-null H1299 cells transfected with p53 constructs. (C) Western blot of p53 expression in p53-null H1299 cells transfected with or without miR-125b. β-actin was used as a control.
to other mechanisms. Further investigation is needed to ascertain the role of the nSNVs in p53 regulation and DLBCL prognosis.

The disappearance of the survival advantage at \( \sim 10 \) years for DLBCL patients with a WT-CDS and nSNVs is noteworthy. The frequency of \( TP53 \) CDS mutations may be higher in relapsed DLBCL patients, and this may partially override the impact of 3'UTR variation on survival beyond 5 years. In patients with multiple myeloma, the frequency of \( TP53 \) CDS mutations in de novo disease is only \( \sim 3\% \), yet this number rises up to 30% in patients with relapsed disease.\(^{45}\) Relapsed DLBCL is generally diagnosed by imaging and limited sampling such as fine-needle aspiration biopsy rather than by excisional biopsy; the lack of adequate specimens for genetic analysis precluded our testing for \( TP53 \) CDS mutations in DLBCL patients at relapse and is a limitation of this study.

\( TP53 \) germline variants have been investigated for cancer susceptibility using genome-wide association studies, in which a type I error rate (\( P \)-value threshold for significance) of \( 10^{-7} \) is generally regarded as indicating genome-wide significance.\(^{3,46}\) Until 2011, no genome-wide association study reported a significant association between any germline \( TP53 \) polymorphism (other than Li-Fraumeni Syndrome variants) and any cancer with a \( P \) value \( \leq 10^{-7} \). In 2011, however, rs7837822 was found to be strongly associated with BCC, with an odds ratio (OR) = 2.16 (\( P = 2.2 \times 10^{-20} \)).\(^{36}\) rs7837822 is located in the fifth nucleotide of the \( TP53 \) PAS that is required for

Figure 6. p53 expression and function for the TP53 PAS variant in p53-null cells. (A) Schematic representation of the TP53 constructs. Each carries a WT 5'UTR, WT CDS, a WT or mutant 3'UTR, and a 2.5-kb native genomic sequence downstream of the TP53 PAS. (B) Western blot of p53 expression in p53-null H1299 cells transfected with p53 constructs. \( \beta \)-actin was used as a control. OE, overexposure. (C) Northern blot of p53 mRNAs in H1299 cells carrying exogenous p53 genes. A housekeeping RNA (H1RNA, the RNA subunit of human ribonuclease P) was used as a control. (D) Apoptosis of cells with p53 constructs; representative TUNEL images (left) and percentage of TUNEL-positive cells (right). \( *P \leq 0.05 \).
polyadenylation machinery recognition and subsequent cleavage, polyadenylation, and export of mature mRNAs to the cytoplasm. The risk-associated allele is a “C,” resulting in an alternative PAS (AATACA) replacing the canonical PAS (AATAAA). Among millions of SNPs in the human genome, rs78378222 has the strongest association with BCC. This SNP is also associated with an increased risk of prostate cancer (OR = 1.44; P = 2.4 × 10^{-6}), glioma (OR = 2.35; P = 1.0 × 10^{-5}), and colorectal adenoma (OR = 1.39; P = 1.6 × 10^{-5}), as well as esophageal squamous cell carcinoma (OR = 3.22; P = 1.34 × 10^{-4}).47 In an independent confirmation study, rs78378222 was found to be associated with a 3.5 times higher risk of glioma (P = 1.34 × 10^{-4}).48 Stefansson and colleagues showed that rs78378222[A/C] heterozygotes expressed “somewhat less” p53 transcript levels than WT homozygotes in human blood specimens (P = 0.04).49 In the present study, we tested the role of the alternative PAS in p53 regulation using a p53-null cell line with exogenous p53 constructs. We found that the C allele, as compared with the A allele, dramatically lowers p53 mRNA levels and results in reduced p53 protein expression and cellular apoptosis. Our data are in agreement with a previous report that simian virus 40 mRNA with AATACA as PAS exhibits an eightfold decrease in cleavage efficiency compared with the canonical site.49,50 These findings provide direct evidence that rs78378222 negatively impacts p53 expression and function. p53 stabilization and function rely on various posttranslational modifications. As rs78378222 occurs in the PAS signal of TP53, none of the downstream p53 protein modifications are affected. Thus, this variant presents a unique instance of TP53 alleles with unambiguous reduced function, in contrast with most CDS mutants that are associated with a “loss of function” and/or a “gain of function.”51

In summary, in this study we have shown that sequence variation in the TP53 3’UTR is frequent in DLBCL and that the role of this sequence variation in patient survival depends on the status of the TP53 CDS. To our knowledge, this study is the first to demonstrate the prognostic importance of the TP53 3’UTR in cancer or any other disease. The full extent of the pathobiology of 3’UTR variants remains to be determined.

Acknowledgments

The authors thank the consortium program team of pathologists, hematologists, and clinicians; each of the contributing center principal physicians; their patients; and former and current hematopathology and hematology/oncology fellows for their support.

References


This work was supported by grants from the National Cancer Institute/National Institutes of Health (R01CA138688, IR1CA146299, P50CA136411, and P50CA142509 to Y.L. and K.H.Y.); the Shannon Timmins Leukemia Fellowship Award at The University of Texas MD Anderson Cancer Center (Z.Y.X.-M.); and The University of Texas MD Anderson Cancer Center Institutional Research and Development Fund, Institutional Research Grant Award, an MD Anderson Cancer Center Lymphoma Specialized Programs of Research Excellence Research Development Program Award, an MD Anderson Cancer Center Myeloma Specialized Programs of Research Excellence Research Development Program Award, a Gundersen Lutheran Medical Foundation Award, MD Anderson Cancer Center Collaborative Funds with Roche Molecular System, HTG Molecular Diagnostics, and Daiichi Sankyo Pharmaceutical (K.H.Y.).

Authorship


Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Yong Li, Department of Biochemistry and Molecular Biology, School of Medicine, University of Louisville, 319 Abraham Flexner Way, Louisville, KY 40202; e-mail: yong.li@louisville.edu; and Ken H. Young, Department of Hematopathology, Unit 72, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77230-1439; e-mail: khyoung@mdanderson.org.

From www.bloodjournal.org by guest on September 17, 2017. For personal use only.


Single nucleotide variation in the $TP53$’ untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the International DLBCL Rituximab-CHOP Consortium Program

Yong Li, Michael W. Gordon, Zijun Y. Xu-Monette, Carlo Visco, Alexander Tzankov, Dehui Zou, Lugui Qiu, Santiago Montes-Moreno, Karen Dybkaer, Attilio Orazi, Youli Zu, Govind Bhagat, Kristy L. Richards, Eric D. Hsi, William W. L. Choi, J. Han van Krieken, Qin Huang, Weiyun Ai, Maurilio Ponzoni, Andrés J. M. Ferreri, Jane N. Winter, Ronald S. Go, Miguel A. Piris, Michael B. Møller, Lin Wu, Michael Wang, Kenneth S. Ramos, L. Jeffrey Medeiros and Ken H. Young