The rs78378222 prevalence and the copy loss of the protective allele A in the tumor tissue of diffuse large B-cell lymphoma

Elena N. Voropaeva[1], Yuriy L. Orlov[2,3], Tatiana I. Pospelova[4], Anna A. Gurageva[1], Mikhail I. Voevoda[1], Vladimir N. Maksimov[1], Olga B. Seregina[4] and Maria I. Churkina[4]

[1] Research Institute of Internal and Preventive Medicine, Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] The Digital Health Institute, I.M. Sechenov First Moscow State Medical University, Moscow, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
[4] Novosibirsk State Medical University of the Ministry of Health of the Russian Federation, Novosibirsk, Russia

Corresponding authors
Elena N. Voropaeva,
vena.81@mail.ru
Yuriy L. Orlov, orlov@bionet.nsc.ru

ABSTRACT

Background: Rare single nucleotide polymorphisms (SNPs) are likely to be a crucial genetic factor for human diseases, including cancer. rs78378222 is rare SNP in 3′-untranslated region (UTR) of *TP53* gene leading to disturbance of 3′-end mRNA processing. The frequency of rs78378222 varies in several studied populations. The meta-analysis of 34 genome-wide association studies indicated that rs78378222 was significantly associated with an increased risk of cancer overall. Bioinformatic analysis indicates that somatic loss of the protective A allele of rs78378222 occurs in the tumor tissue of some malignant. The goal of the current study is to document the rs78378222 prevalence and evaluate the copy loss status of the protective allele A in the tumor tissue of patients with diffuse large B-cell lymphoma (DLBCL).
Methods: Total DNA was isolated from FFPE-samples and peripheral blood of patients with DLBCL and comparable in age and sex controls. rs78378222 genotyping was performed by the PCR-RFLP method using restriction endonuclease *Hin*dIII. Direct Sanger's sequencing was used to confirm the presence of C allele of the rs78378222. The search for *TP53* gene mutations was carried out by Sanger's direct sequencing method, according to the IARC protocol.
Results: The result of genotyping of 136 DNA samples from DLBCL tumor tissue suggested that frequency of the rs78378222 was 11/136 (8.1%). Rare allele C frequency was 11/272 (4.2%). A total of 5/11 DLBCL rs78378222 heterozygous samples had the heterozygosity loss in the *TP53* gene. Only one of these cases was combined with *TP53* gene mutations which have proven oncogenic potential—p. Arg196Gln, other four cases have not mutations in the coding regions of gene.
Conclusions: At the stages of DLBCL initiation or progression a loss of the protective allele A of rs78378222 occurs. Further efforts are needed to study possible molecular mechanisms underlying somatic alterations in DLBCL in this region of the *TP53* 3′-UTR as well as functional studies to illustrate how the presents of rs78378222 may affect tumor progression of lymphoma.

INTRODUCTION

TP53 gene (OMIM No. 191117) is located on 17p13.1. The dominant transcript of mRNA has a length 2586 b.p., including eleven exons and 3′- and 5′-untranslated regions (UTRs) (*Diederichs et al., 2016*). *TP53* gene is a central member in the molecular network of constancy and integrity of the genome, which controls cell cycle regulation and DNA repair. Its activation leads to upregulation of genes involved in the intrinsic (*PUMA*, *BAX*) and extrinsic (*TNFRSF10B*, *FAS*) pathways of apoptosis (*Ciardullo et al., 2019*). Consequently, *TP53* is considered a key tumor suppressor gene.

Experimental studies on mice models suggested that the restoration of the gene *TP53* function is sufficient for regression of some cancer types. The studies carried out on *TP53*-knockout mice and rats showed that malignant lymphomas are the prevailing form of neoplasms (*McCoy et al., 2013*).

B-lymphocytes undergoing stress tend to have p53-depended apoptosis, in contrast to other cells that have cell cycle arrest, p53-independent apoptosis or necrosis in stress condition (*Wen et al., 2019*). Besides, *TP53* gene aberrations in B-lymphocytes leads to less effective functioning of intracellular signaling pathways that can arrest cell cycle in G1 and G2 phases; DNA repair defects; more effective cell adaptation to hypoxia and angiogenesis stimulation; reduction control under telomere structure and differentiation inhibition. For these reasons, the disturbance of programed cell death due to dysfunction p53 is a basis for development and progression of lymphoproliferative disorders (*Voropaeva et al., 2017*). Moreover, as the *TP53* gene plays an important role in mediation of chemotherapeutic agents' action and action of target drugs, the deficit of its function leads to formation of multidrug resistance phenotype of lymphoma cells (*Voropaeva et al., 2015*).

The vast majority of studies dedicated to the role of mutations in nucleotide sequences of *TP53* were focused on analysis of 5–8 gene exons, while the UTRs were ignored. Nevertheless, the changes in these regions may have direct biological effect on B cells maturation and may activate lymphomagenesis (*International Cancer Genome Consortium, 2010*).

High-throughput sequencing techniques have allowed us to significantly deepen our understanding of the genetic alterations that may be responsible for tumor growth. In addition to the protein-coding sequences, the major non-coding fraction of the genome can be affected by tumor-promoting mutations (*Diederichs et al., 2016*). Thus, *Li et al. (2013)* first showed that mutations in UTRs of the *TP53* gene are frequently observed in neoplasms. 5′- and 3′-UTR do not act as a matrix for the protein synthesis, but they are a part of highly conservative elements of the gene. According to their data analyses, the majority of DLBCL patients had the mutations in 3′-UTR in *TP53* and almost all identified replacements lied in the location of confirmed earlier or suggested by in silico analysis micro-RNA binding sites.

Recently hundreds of SNPs in *TP53* have been identified, a part of which is being studied as regard as their influence on DLBCL susceptibility (*Voropaeva et al., 2015*, *2014*, *2016*; *Pospelova et al., 2010*). One of the SNPs of *TP53* gene 3′-UTR (rs78378222) leading to change of normal polyadenylation signal sequence (*Ciardullo et al., 2019*).

rs78378222 is considered to be affected by the negative natural selection and malignant neoplasms are thought to be the mechanism of such selection (*Stacey et al., 2011*). It is assumed that this SNP may be a common causal genetic variant for different cancer types in different populations (*Zhou, Yuan & Yang, 2012*). In whole genome studies rs78378222 was described as a risk factor of prostate cancer, glioma, basal cells carcinoma, esophageal cancer, colorectal adenoma (*Wang et al., 2016*). The risk-associated allele is a "C," resulting in an alternative polyadenylation signal (AATACA) replacing the canonical polyadenylation signal (AATAAA) (*Li et al., 2013*). However, in the literature there are reports of fluctuations in the frequency of detection of rs78378222 in various studied populations: from the absence of detection of this marker to its detection with a frequency of 5–6% (*Wang et al., 2016*; *Rao et al., 2014*).

Experiments on cell cultures have shown that the rare C allele compared to the A allele of the rs78378222 leads to both a significant decrease of the p53 protein expression and a decrease of the level of programed cell death induction under the influence of genotoxic factors. (*Li et al., 2013*).

Bioinformatic analysis by *Wang et al. (2015)* suggests how the rare variant rs78378222 may affect the function of p53. It has been shown that the C allele break the correct termination of synthesis and processing of the 3′-end of mRNA. It was shown, that the gene transcripts containing the rare allele C of the rs78378222 are on ~3 kb longer than the gene transcripts containing allele A (*Wang et al., 2015*).

Simultaneously, it was shown that a normal copy of the A allele rs78378222 is lost in the tumor tissue, suggesting a two-stroke theory of carcinogenesis for which both alleles are altered, one at the germline and the second by somatic alteration (*Wang et al., 2015*).

However, *Wang et al. (2015)* have shown that loss of the protective common allele A of rs78378222 occurs in glioma but not in lung cancer. Thus, the loss of protective allele A may not be observed in all tumors. Given the presence of such facts the goal of the current study is to document the rs78378222 prevalence and evaluate the copy loss status of the protective allele A in the tumor tissue of patients with one of the most frequent variant of Non-Hodgkin's lymphomas—diffuse large B-cell lymphoma (DLBCL).

MATERIALS AND METHODS

Total DNA was isolated from FFPE-samples of lymph nodes and extranodal tumor lesions biopsies of 136 patients with DLBCL. We used tissue sections in which at least 50% of the cells were tumor cells. Genomic DNA from blood of independent 150 patients with DLBCL and comparable in age and sex 170 controls was isolated by phenol-chloroform extraction method using proteinase K.

Complete clinical data were available for 280 DLBCL patients: 148 men and 132 women, who were admitted to Novosibirsk Hematological Center during years 2012–2018.

Table 1 Clinical parameters of the DLBCL patients study group (n = 280).

Parameter	Frequency
Age	
<70 years	107 (38.2%)
M 70 years	173 (61.8%)
Sex	
M	148 (52.9%)
F	132 (47.1%)
ECOG score	
0+1	169 (60.4%)
2-3	111 (39.6%)
B-symptoms	
No	113 (40.4%)
Yes	167 (59.6%)
IPI score	
0–2	107 (38.2%)
3–5	173 (61.8%)
Stage	
I–II	42 (15.0%)
III–IV	238 (85.0%)
Extranodal involvement	
No	145 (51.8%)
Yes	135 (48.2%)
Leukemization	
No	199 (71.1%)
Yes	81 (28.9%)
LDH level	
Normal	153 (54.6%)
Elevated	127 (45.4%)

Note:
ECOG, Eastern Cooperative Oncology Group; IPI, International Prognostic Index; LDH, Lactate dehydrogenase.

As many as 85% of these patients had advanced (III–IV) stages of the disease and more than 60% of them had a poor prognosis according to the International Prognostic Index (IPI). More detailed clinical parameters are presented in the Table 1. The study was approved by the Ethics Committee of the Research Institute of Internal and Preventing Medicine (approval number #2019-01). All patients and controls are ethnical Russians. All study participants provided written informed consent to participate in a research study. The study was carried out in accordance with the World Medical Association (WMA) Declaration of Helsinki (2000) and the Protocol to the Convention on Human Rights and Biomedicine (1999). Sensitivity of the PCR-RFLP and sequencing methods in revealing the presence of unequal amounts of alleles was proven earlier. It was shown that the presence of differences even 10–20% of allele in the total template mixture could be detected by restriction digestion analysis and by sequencing (*Siddique et al., 2005*).
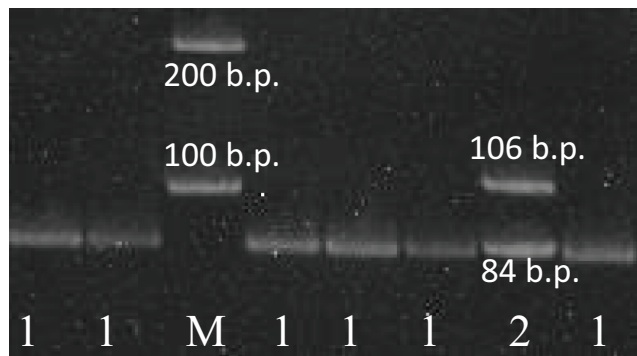
Figure 1 The results of the rs78378222 genotyping. The results of the rs78378222 genotyping by PCR-RFLP method (endonuclease *Hin*dIII). 1–84 b.p. (genotype A/A); 2–106+84 b.p. (genotype A/C); K– negative control; M – 100 b.p. molecular weight marker.

Genotyping of the rs78378222 in *TP53* gene was performed by the PCR-RFLP method, described below. The PCR reaction mixture (25 µl/reaction) contained 75 mM Tris-HCl pH 9.0, 20 mM $(NH_4)_2SO_4$, 0.01% Tween 20, MgCl2 10 mM 1.5 µl, mixture of deoxynucleotidetriphosphates 10 mM 0.5 µl, 1 u.a. SynTaq DNA-polymerase (Syntol, Moscow, Russia) preparations, 0.1 mkM each of the primers (forward 5′-CACACAGG TGGCAGCAAAGCT-3′ and reverse 5′-AGCACATCTGCATTTTCACCCC-3′) and 1 µl 0.1 mkg/ml genome DNA. The reaction was carried out in thermocycler with using the following conditions: initial denaturation stage—5 min at 95 °C; followed by 31 cycles of denaturation stage—30 s at 95 °C, annealing stage—30 s at 60 °C, extension stage—45 s at 72 °C; post extension stage—3 min. The product of PCR was incubated with 10 u.a. restriction endonucleasae *Hin*dIII in thermostat at 37 °C during 12 h. The fragments of hydrolysis were visualized in 5% polyacrylamide gel. The distributions of bands on the electrophoregram and results interpretation are shown in Fig. 1. The PCR conditions were chosen in such a way as to avoid running out of amplicon in excessive amounts. Restriction endonuclease was added to the test tube in excess and kept for a long time. We believe that these conditions allowed us to avoid the possibility of genotyping error due to undigested amplicons.

Direct Sanger's sequencing was used to confirm the presence of C allele of the rs78378222 (Fig. 2).

Search for *TP53* gene mutations was carried out by Sanger's direct sequencing method, according to the IARC protocol (http://p53.iarc.fr/download/tp53_directsequencing_iarc. pdf). The PCR products were cleared using microcolumns with SephadexTM G-50 medium (USA). The sequencing of samples was carried out by the method of capillary electrophoresis on the Hitachi 350 Genetic Analyser (Applied Biosystem, Foster City, CA, USA) using BigDye® Terminator kit v 1.3. and v 1.1. (Applied Biosystems, Foster City, CA, USA) and polymer POP-7. The sequencing results analysis and fragment alignment was conducted by program Chromas, SeqScape v.2.7, Sequence Scanner. The NG_017013.2 sequence of the *TP53* gene was used as a reference.
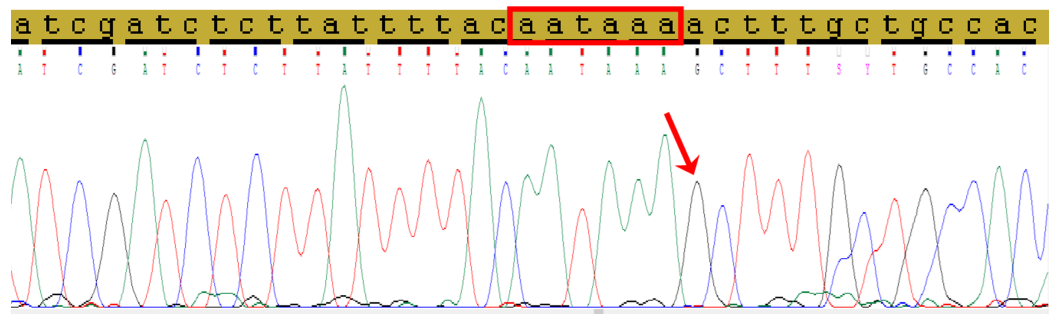
Figure 2 The fragment of 3′-UTR of *TP53* gene containing the rs78378222. The normal poly-adenylation signal is highlighted by the red rectangle, the single nucleotide replacement c.*1175A>G as a result of incorporation the site restriction into the reverse primer for RFLP analysis is shown by red arrow.

Table 2 Frequency distributions of rs78378222 alleles and genotypes between DLBCL patients and controls.

rs78378222	Cases N (%)		Controls N (%)	Cases blood vs. Controls		Cases blood vs. Cases tumor	
	Blood	Tumor		χ^2	p	χ^2	p
Genotype							
A/A	144 (96.0)	125 (91.9)	166 (97.6)	0.71	0.40	2.13	0.15
A/C	6 (4.0)	11 (8.1)	4 (2.4)				
C/C	–	–	–				
Allele							
A	294 (98.0)	261 (95.8)	336 (98.8)	0.70	0.40	2.07	0.15
C	6 (2.0)	11 (4.2)	4 (1.2)				

The comparison of alleles and genotype frequencies of rs78378222 between groups was done by Pearson's χ^2 test in 2 × 2 contingency tables. Differences were considered statistically significant at threshold $p < 0.05$.

RESULTS

The result of genotyping of 136 DNA samples from tumor tissue of patients with DLBCL suggested that genotype A/C of rs78378222 frequency was 11/136 (8.1%). Rare allele C frequency was 11/272 (4.2%). We determined the frequency of polymorphism in healthy tissue patients with lymphoma and controls.

As shown in Table 2, there were no significant differences in distributions of rs78378222 alleles and genotypes between controls and cases of DLBCL ($p = 0.400$) or healthy and tumor tissue ($p = 0.150$) of DLBCL patients. The genotype frequencies of the SNP in both cases and controls were in agreement with the Hardy–Weinberg equilibrium.

rs78378222 minor allele frequency in healthy controls was higher than the in the "1000 genomes project" (13/5008 (0.3%), $p = 0.004$) (*Abecasis et al., 2012*), which can be explained by the different prevalence of the allele in populations of different ethnicities. Thus there were no significant differences in rs78378222 minor allele frequency between

Table 3 rs78378222 minor allele frequency in several studies.

Cohort	N (%)	p*	References
Current study	4/340 (1.2)	–	–
1000 genomes	13/5008 (0.3%)	0.004	Abecasis et al. (2012)
GnomAD	373/31380 (1.2)	1.0	Karczewski et al. (2019)
Estonian	107/4480 (2.4)	0.151	(https://www.ncbi.nlm.nih.gov/snp/rs78378222)
Northern Sweden	11/600 (2.0)	0.440	(https://www.ncbi.nlm.nih.gov/snp/rs78378222)

Note:
* p-value vs. current study.
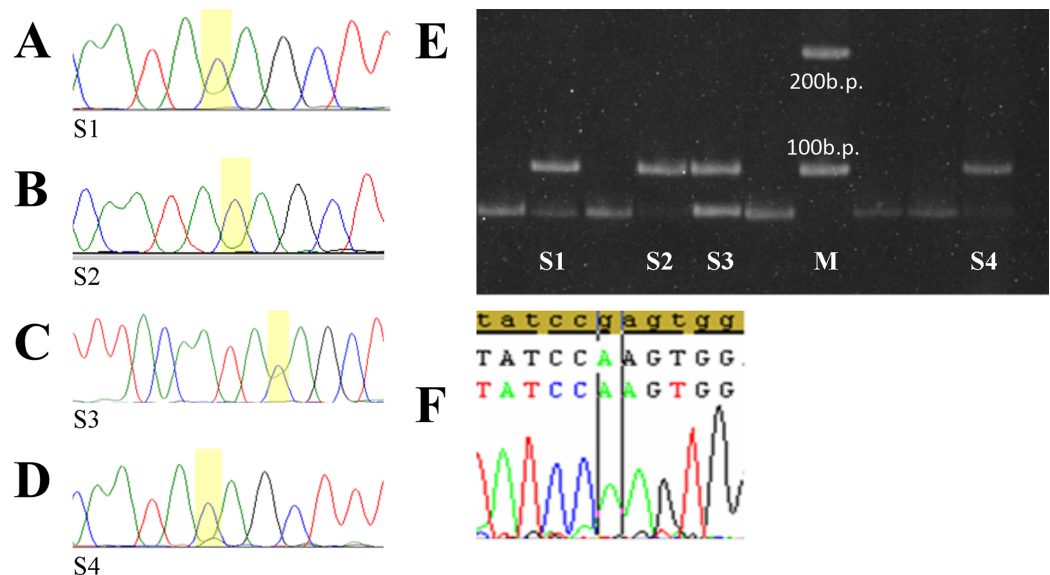
Figure 3 Selected tumor samples with rs783782. Selected tumor samples with rs78378222: (A–D) sequences chromatograms; (E) results of genotyping by PCR-RFLP method; (F) p.R196Q homozygote mutation. S1, 2, 4—tumor samples with C/C genotype rs78378222 (loss of heterozygosity in the TP53), S3—sample with A/C genotype rs78378222, M – 100 b.p. molecular weight marker.

current study controls and Estonian (p = 0.151) (https://www.ncbi.nlm.nih.gov/snp/rs78378222) and Northern Sweden (p = 0.440) (https://www.ncbi.nlm.nih.gov/snp/rs78378222) cohorts (shown in Table 3).

According to design of the rs78378222 genotyping by PCR-RFLP analysis, the presence of rare allele C destroys the site of restriction endonuclease HindIII. It is known that fixing the tissue by formalin and long-term storage of samples in paraffin blocks may results to artificial DNA changes. For this reason we carried out Sanger's direct sequencing of 3′-UTR fragment of TP53 gene, containing the rs78378222 from DNA samples of patients with DLBCL who have minor allele C.

Five of tumor samples had prepotency C allele "dose" on electrophoregrams and sequences chromatograms (Figs. 3A and 3B). Sensitivity of the PCR-RFLP and sequencing methods in revealing the presence of differences even 10–20% of allele in the total template mixture was proven earlier (Siddique et al., 2005). Since the cases of rare homozygous

genotype C/C rs78378222 detection in normal tissue have not been described previously, and DNA extraction was carried out from paraffin blocks contained not less than 50% of tumor tissue, the gained results demonstrated a heterozygosity loss in the *TP53* of the rs78378222 in tumor tissue of patients DLBCL.

Only one of these cases was combined with *TP53* gene mutations which can be considered pathogenic—p.Arg196Gln (Fig. 3C). Codon 196 is phylogenetically conservative. It is a part of the beta-sheets of DNA-binding domain and one of the "hotspots" of the *TP53* gene mutations (*Bouaoun et al., 2016*). Mutation p.Arg196Gln is described in the COSMIC database (Genomic Mutation ID COSV52674826) and the IARC *TP53* mutation database (MUT_ID2461) for various tumors (*Bouaoun et al., 2016*; https://cancer.sanger.ac.uk/cosmic/mutation/overview?id=96071916#references). Analysis with using tools which predicts possible impact of an amino acid substitution on the structure and function of a human protein PolyPhen-2 (Polymorphism Phenotyping v2) and SIFT shows that p.Arg196Gln refers to the number of damaging substitutions. Functional assessment of p53 mutants revealed that this mutation leads to loss of transactivation activity of protein (*Marutani et al., 1999*; *Monti et al., 2003*).

In other four cases rs78378222 with heterozygosity loss were the only aberration in *TP53*. Detailed description of the mutations identified in the Russian DLBCL patient's cohort was described in previous publication (*Voropaeva et al., 2017*).

DISCUSSION

Currently, the international research projects such as Cancer Genome Atlas Research Network (TCGA) and International Cancer Genome Consortium (ICGC) have shown the presence in tumors a large number of driver mutations in the regions of genes that do not encode the protein sequence (*International Cancer Genome Consortium, 2010*; *Esgueva et al., 2012*). Aberrations in 3′-UTR of *TP53* gene may be a universal carcinogenesis mechanism, playing role in pathogenesis of DLBCL as well. In particular mutations in 3′-UTR of *TP53* gene may alter the interactions with regulatory miRNAs or cause the destruction of conservative sequence of polyadenylation signal and the occurrence of additional signals (*Diederichs et al., 2016*; *Li et al., 2013*). *Li et al. (2013)* first reported the prognostic value of aberrations in 3′-UTR of *TP53* gene in DLBCL patients.

Rare variants are more likely to have a functional impact and tend to have a greater effect size than do common variants. Thus, rare variants are likely to be a crucial genetic factor for human diseases, including cancer (*Gorlov, Gorlova & Frazier, 2011*; *Gorlov, Gorlova & Sunyaev, 2008*).

Enciso-Mora et al. (2013) in the whole genome sequencing study described rs78378222. It is rare SNP in 3′-UTR of *TP53* gene leading to disturbance of 3′-end mRNA processing. This polymorphism change canonical sequence of *TP53* gene polyadenylation signal. This signal is required for the detection of sequences by polyadenylation complex, cleavage, polyadenylation and export of mature RNA to cytoplasm (*Li et al., 2013*).

Recently, several studies have investigated the association between a newly reported rs78378222 and cancer risk, but generated inconsistent findings. The meta-analysis on the

basis of high-grade data from 34 Genome-wide association studies indicated that rs78378222 was significantly associated with an increased risk of overall cancer. Furthermore, stratified analyses indicated that rs78378222 increased the risk of cancers: nervous system, skin and others (*Wang et al., 2016*). Among millions of polymorphisms in human genome the rs78378222 has the strongest correlation with basal cell carcinoma (*Diskin et al., 2014*). rs78378222 is considered as one of the mechanism of cancer predisposition in Li-Fraumeni syndrome (LFS) and its variant, Li-Fraumeni-like Syndrome (LFL), autosomal dominant disorders that are characterized by predisposition to multiple early onset cancers (*Macedo et al., 2016*).

There was no purpose to estimate the association of polymorphism with a predisposition to lymphoma. The power of current study is not enough to resolve this purpose. Identification the role of polymorphism in predisposition to development DLBCL requires a multi-center joint study, what can be the basis for our research in the future.

The goal of the current study is to document the rs78378222 prevalence and evaluate the copy loss status of the protective allele A in the tumor tissue of patients with Diffuse large B-cell lymphoma (DLBCL). It is suggested that with at least some malignant neoplasms, the loss of the protective allele A in the tumor may occur. Thus Wang et al. have shown that loss of the allele A occurs in glioblastoma, but not in lung cancer (*Wang et al., 2015*).

The result of genotyping of 136 DNA samples from tumor tissue of patients with DLBCL suggested that frequency of the rs78378222 was 11/136 (8.1%). Rare allele C frequency was 11/272 (4.2%). And it was higher than the rs78378222 minor allele frequency in the "1000 genomes project" ($p < 0.001$), which could be explained by both the different prevalence of the allele in populations of different ethnicities and tumor origin of the samples.

The frequency of rs78378222 varies in different cancer patients studies: from the absence of detection of this SNP in oral, cervical and breast cancer patients from South India to its detection with the frequency 5.4% in cohort LFS/LFL patients (*Wang et al., 2016*; *Rao et al., 2014*). The detailed analysis Genome Aggregation Database (GnomAD) (*Karczewski et al., 2019*) showed that rs78378222 allele C frequency in European population is 1.7% whereas in Latino and African—is 0.4% and 0.2%, in East Asian and Ashkenazi Jewish—is 0%.

Since data on the rs78378222 prevalence in the Russian population is not presented in the literature and specialized databases, we defined the frequency of polymorphism in healthy tissue patients with lymphoma and controls. There were no significant differences in distributions of rs78378222 alleles and genotypes between controls and cases of DLBCL or healthy and tumor tissue DLBCL patients. There were no significant differences in rs78378222 minor allele frequency between current study controls and Estonian (https://www.ncbi.nlm.nih.gov/snp/rs78378222) and Northern Sweden cohorts (shown in Table 2) (https://www.ncbi.nlm.nih.gov/snp/rs78378222).

Remarkably, none of the DNA samples from healthy tissue from cases and controls has rs78378222 in homozygous state. However, the part (5/11) of analyzed DLBCL samples

had the heterozygosity loss in *TP53* gene in carriers of genotype A/C of the rs78378222. One of these cases was combined with p.Arg196Gln *TP53* gene mutation which have proven oncogenic potential—, other four cases have not mutations in the coding regions of gene.

Attention to this SNP is due to the fact that rs78378222 is thought to be the unique SNP of *TP53* gene with the reduced p53 function. The *TP53* mRNA level decrease in the cells having the rs78378222 compared to the cells without this polymorphism but with another closely-spaced marker rs114831472 was confirmed. Allele C of rs78378222, as compared with the A allele, dramatically lowers p53 protein expression and cellular apoptosis (*Li et al., 2013*; *Macedo et al., 2016*). It is also suggested that since the micro-RNA binding site is located on 3′-UTR of gene, SNP of 3'-UTR *TP53* including rs78378222 may affect the gene expression by canceling, weakening or creation of new binding sites (*Gorlov, Gorlova & Sunyaev, 2008*; *Macedo et al., 2016*).

Findings another study suggest that allele C rs78378222 lead to haploinsufficiency of p53. It is situation in which the total level of a gene product produced by the cell is about half of the normal level and that is not sufficient to permit the cell to function normally (*Gorlov, Gorlova & Sunyaev, 2008*).

Thus, rs78378222 presents a rare instance of *TP53* alleles with unambiguous reduced function, in contrast with vast majority of coding sequence mutants which are associated with loss of p53 function, dominant negative effect of mutants and/or gain of uncharacteristic for wild-type p53 functions (*Li et al., 2013*).

Research group *Wang et al. (2015)* performed the integrated analysis of TCGA data and suggested a functional mechanism for germline rare variant in *TP53* rs78378222. Using RNA sequencing data, they observed aberrant transcripts with ~3 kb longer than normal for those individuals. Using exome sequencing data, they further showed that loss of haplotype carrying common protective allele A occurred somatically in tumor tissue. Thus analysis suggests rare risk allele C disrupts mRNA termination, and an allelic loss of a genomic region harboring common protective allele A occurs during initiation or progression some tumors, for example glioblastoma (*Wang et al., 2015*).

CONCLUSIONS

At the stages of DLBCL initiation or progression a loss of the protective allele A of rs78378222 occurs. Further efforts are needed to study possible molecular mechanisms underlying somatic alterations in DLBCL in this region of the *TP53* 3′-UTR as well as functional studies to illustrate how the presents of rs78378222 may affect tumor progression of lymphoma.

Funding

The work was supported by a grant of the President of the Russian Federation to support young Russian scientists - doctors of sciences MD-2706.2019.7. The study was supported

by Budgetary Project No. AAAA-A17-117112850280-2. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
President of the Russian Federation: MD-2706.2019.7.
Budgetary Project: AAAA-A17-117112850280-2.

Competing Interests
Yuriy L. Orlov is an Academic Editor for PeerJ.

Author Contributions
- Elena N. Voropaeva conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Yuriy L. Orlov analyzed the data, prepared figures and/or tables, and approved the final draft.
- Tatiana I. Pospelova conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Anna A. Gurageva conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Mikhail I. Voevoda analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Vladimir N. Maksimov performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Olga B. Seregina performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Maria I. Churkina conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics
The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The Institutional Review Board - Ethics Committee of the Research Institute of Internal and Preventive Medicine - Branch of the Federal State Budget Scientific Institution "The Federal Research Center Institute of Cytology and Genetics" of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia approved this research (approval number #2019-01).

All study participants provided informed consent to participate in a research study. The study was carried out in accordance with the World Medical Association (WMA) Declaration of Helsinki (2000) and the Protocol to the Convention on Human Rights and Biomedicine (1999).
Data Availability

The raw measurements are available in Tables 1–3. Anonymized patient information is available in a Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10335#supplemental-information.

REFERENCES

Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA, 1000 Genomes Project Consortium. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**:56–65 DOI 10.1038/nature11632.

Bouaoun L, Sonkin D, Ardin M, Hollstein M, Byrnes G, Zavadil J, Olivier M. 2016. TP53 variations in human cancers: new lessons from the IARC TP53 database and genomics data. *Human Mutation* **37(9)**:865–876 DOI 10.1002/humu.23035.

Ciardullo C, Aptullahoglu E, Woodhouse L, Lin WY, Wallis JP, Marr H, Marshall S, Bown N, Willmore E, Lunec J. 2019. Non-genotoxic MDM2 inhibition selectively induces pro-apoptotic p53 gene signature in chronic lymphocytic leukemia cells. *Haematologica* **104(12)**:2429–2442 DOI 10.3324/haematol.2018.206631.

Diederichs S, Bartsch L, Berkmann JC, Fröse K, Heitmann J, Hoppe C, Iggena D, Jazmati D, Karschnia P, Linsenmeier M, Maulhardt T, Möhrmann L, Morstein J, Paffenholz SV, Röpenack P, Rückert T, Sandig L, Schell M, Steinmann A, Voss G, Wasmuth J, Weinberger ME, Wullenkord R. 2016. The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Molecular Medicine* **8(5)**:442–457 DOI 10.15252/emmm.201506055.

Diskin SJ, Capasso M, Diamond M, Oldridge DA, Conkrite K, Bosse KR, Russell MR, Iolascon A, Hakonarson H, Devoto M, Maris JM. 2014. Rare variants in TP53 and susceptibility to neuroblastoma. *Journal of the National Cancer Institute* **106(4)**:dju047 DOI 10.1093/jnci/dju047.

Enciso-Mora V, Hosking FJ, Di Stefano AL, Zelenika D, Shete S, Broderick P, Idbaih A, Delattre J-Y, Hoang-Xuan K, Marie Y, Labussière M, Alentorn A, Ciccarino P, Rossetto M, Armstrong G, Liu Y, Gousias K, Schramm J, Lau C, Hepworth SJ, Schoemaker M, Strauch K, Müller-Nurasyid M, Schreiber S, Franke A, Moebus S, Eisele L, Swerdlow A, Simon M, Bondy M, Lathrop M, Sanson M, Houlston RS. 2013. Low penetrance susceptibility to glioma is caused by the TP53 variant rs78378222. *British Journal of Cancer* **108(10)**:2178–2185 DOI 10.1038/bjc.2013.155.

Esgueva R, Park K, Kim R, Kitabayashi N, Barbieri CE, Dorsey PJ, Abraham C, Banerjee S, Leung RA, Tewari AK, Terry S, Shevchuk MM, Rickman DS, Rubin MA. 2012. Next-generation prostate cancer biobanking: toward a processing protocol amenable for the International Cancer Genome Consortium. *Diagnostic Molecular Pathology: American Journal of Surgical Pathology, Part B* **21(2)**:61–68 DOI 10.1097/PDM.0b013e31823b6da6.

Gorlov IP, Gorlova OY, Frazier ML. 2011. Evolutionary evidence of the effect of rare variants on disease etiology. *Clinical Genetics* **79(3)**:199–206 DOI 10.1111/j.1399-0004.2010.01535.x.

Gorlov IP, Gorlova OY, Sunyaev SR. 2008. Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms. *American Journal of Human Genetics* **82(1)**:100–112 DOI 10.1016/j.ajhg.2007.09.006.

International Cancer Genome Consortium. 2010. International network of cancer genome projects. *Nature* **464**:993–998 DOI 10.1038/nature08987.

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, MacArthur DG. 2019. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* DOI 10.1101/531210.

Li Y, Gordon MW, Xu-Monette ZY, Visco C, Tzankov A, Zou D, Qiu L, Montes-Moreno S, Dybkaer K, Orazi A, Zu Y, Bhagat G, Richards KL, Hsi ED, Choi WWL, Van Krieken JH, Huang Q, Ai W, Ponzoni M, Ferreri AJM, Winter JN, Go RS, Piris MA, Møller MB, Wu L, Wang M, Ramos KS, Medeiros LJ, Young KH. 2013. Single nucleotide variation in the TP53 39 untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the International DLBCL Rituximab-CHOP Consortium Program. *Blood* **121(22)**:4529–4540 DOI 10.1182/blood-2012-12-471722.

Macedo GS, Araujo Vieira I, Brandalize AP, Giacomazzi J, Inez Palmero E, Volc S, Rodrigues Paixão-Côrtes V, Caleffi M, Silva AM, Achatz MI, Hainaut P, Ashton-Prolla P. 2016. Rare germline variant (rs78378222) in the TP53 3'UTR: evidence for a new mechanism of cancer predisposition in Li-Fraumeni syndrome. *Cancer Genetics* **209(3)**:97–106 DOI 10.1016/j.cancergen.2015.12.012.

Marutani M, Tonoki H, Tada M, Takahashi M, Kashiwazaki H, Hida Y, Hamada J, Asaka M, Moriuchi T. 1999. Dominant-negative mutations of the tumor suppressor p53 relating to early onset of glioblastoma multiforme. *Cancer Research* **59(19)**:4765–4769.

McCoy A, Besch-Williford CL, Franklin CL, Weinstein EJ, Cui X. 2013. Creation and preliminary characterization of a Tp53 knockout rat. *Disease models & Mechanisms* **6(1)**:269–278 DOI 10.1242/dmm.009704.

Monti P, Campomenosi P, Ciribilli Y, Iannone R, Aprile A, Inga A, Tada M, Menichini P, Abbondandolo A, Fronza G. 2003. Characterization of the p53 mutants ability to inhibit p73 beta transactivation using a yeast-based functional assay. *Oncogene* **22(34)**:5252–5260 DOI 10.1038/sj.onc.1206511.

Pospelova TI, Voropaeva EN, Voevoda MI, Berezina OV. 2010. P53 gene polymorphism as a potential marker of predisposition to the development of non-Hodgkins malignant lymphoma. *Hematology and Transfusiology. Russian Journal* **55(1)**:11–17 [in Russian].

Rao AK, Vinothkumar V, Revathidevi S, Arunkumar G, Manikandan M, Arun K, Rajkumar KS, Ramani R, Ramamurthy R, Munirajan AK. 2014. Absence of TP53 poly-a signal sequence variant rs78378222 in oral, cervical and breast cancers of South India. *Asian Pacific Journal of Cancer Prevention* **15(21)**:9555–9556 DOI 10.7314/APJCP.2014.15.21.9555.

Siddique MM, Balram C, Fiszer-Maliszewska L, Aggarwal A, Tan A, Tan P, Soo KC, Sabapathy K. 2005. Evidence for selective expression of the p53 codon 72 polymorphs: implications in cancer development. *Cancer Epidemiology, Biomarkers & Prevention* **14(9)**:2245–2252 DOI 10.1158/1055-9965.EPI-05-0153.

Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, Magnusson OT, Gudjonsson SA, Sigurgeirsson B, Thorisdottir K. 2011. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nature Genetics* **43(11)**:1098–1103 DOI 10.1038/ng.926.

Voropaeva EN, Pospelova TI, Voevoda MI, Maksimov VN. 2016. The results of comprehensive analysis of TP53 gene status in patients with diffuse large cell lymphoma. *Hematology and Transfusiology. Russian Journal* **61(3)**:138–143 [In Russian] DOI 10.18821/0234-5730-2016-61-3-138-143.

Voropaeva EN, Voevoda MI, Maksimov VN, Pospelova TI. 2014. Linkage disequilibrium and haplotypes of the rs1042522, rs1625895, and rs1787862 markers of TP53 in patients with diffuse large B-cell lymphoma. *Molecular Biology* **48(5)**:664–670 DOI 10.1134/S0026893314050173.

Voropaeva EN, Voevoda MI, Maksimov VN, Pospelova TI. 2015. Prognostic impact of the TP53 rs1625895 polymorphism in DLBCL patients. *British Journal of Haematology* **169(1)**:32–35 DOI 10.1111/bjh.13237.

Voropaeva EN, Voevoda MI, Maksimov VN, Pospelova TI. 2017. Frequency, spectrum, and functional significance of TP53 mutations in patients with diffuse large B-cell lymphoma. *Molecular Biology* **51(1)**:53–60 DOI 10.1134/S0026893316060224.

Wang Z, Rajaraman P, Melin BS, Chung CC, Zhang W, McKean-Cowdin R, Michaud D, Yeager M, Ahlbom A, Albanes D, Andersson U, Freeman LEB, Buring JE, Butler MA, Carreón T, Feychting M, Gapstur SM, Gaziano JM, Giles GG, Hallmans G, Henriksson R, Hoffman-Bolton J, Inskip PD, Kitahara CM, Marchand LL, Linet MS, Li S, Peters U, Purdue MP, Rothman N, Ruder AM, Sesso HD, Severi G, Stampfer M, Stevens VL, Visvanathan K, Wang SS, White E, Zeleniuch-Jacquotte A, Hoover R, Fraumeni JF, Chatterjee N, Hartge P, Chanock SJ. 2015. Further confirmation of germline glioma risk variant rs78378222 in TP53 and its implication in tumor tissues via integrative analysis of TCGA data. *Human Mutation* **36(7)**:684–688 DOI 10.1002/humu.22799.

Wang Y, Wu XS, He J, Ma T, Lei W, Shen ZY. 2016. A novel TP53 variant (rs78378222 A>C) in the polyadenylation signal is associated with increased cancer susceptibility: evidence from a meta-analysis. *Oncotarget* **7(22)**:32854–32865 DOI 10.18632/oncotarget.9056.

Wen P, Wei X, Liang G, Wang Y, Yang Y, Qin L, Pang W, Qin G, Li H, Jiang Y, Wu Q. 2019. Long-term exposure to low level of fluoride induces apoptosis via p53 pathway in lymphocytes of aluminum smelter workers. *Environmental Science and Pollution Research International* **26(3)**:2671–2680 DOI 10.1007/s11356-018-3726-z.

Zhou L, Yuan Q, Yang M. 2012. A functional germline variant in the P53 polyadenylation signal and risk of esophageal squamous cell carcinoma. *Gene* **506(2)**:295–297 DOI 10.1016/j.gene.2012.07.007.