Non-Hodgkin's lymphomas (NHL) are a heterogeneous group of disease entities with distinctive clinical, morphological, phenotypic, and genetic characteristics. Most aggressive NHLs occur as primary tumors recognized at diagnosis, but they may also develop secondarily from progression of indolent variants. This transformation is associated with a rapidly progressive clinical course and short survival of the patients. The underlying molecular mechanisms involved in the pathogenesis of aggressive lymphomas and the progression of indolent tumors are not well known.

The TSG101 tumor susceptibility gene is a new putative tumor suppressor gene involved in different types of murine and human tumors. The functional inactivation of the TSG101 gene caused cellular transformation and formation of metastasis in nude mice. Alternatively spliced TSG101 non-coding messages, with loss of coding exons and early termination codons, are frequently detected in many carcinomas without alterations at the DNA level. These non-coding alternatively spliced messages are always coexpressed with a normal transcript in tumor cells, and are also detectable in normal cells at a lower frequency.

The normal TSG101 gene codes for a 46-kDa protein with two well-defined domains. The amino terminus has a three-dimensional structure similar to the catalytic domain of the E2 family of ubiquitin conjugating enzyme inhibitors. The E2 family of inhibitors has a blocking effect on the cell cycle in metaphase. This effect on ubiquitination has been recently demonstrated in the case of two TSG101-related proteins, hMMS2 and CROC-1/UEV-1. Thus alterations in the ubiquitination process might be of relevance for oncogenesis, since cyclin regulation by ubiquitination has been shown to affect the G1 to S transition and the anaphase block in the cell cycle. The TSG101 C-terminus has a coiled-coil domain or leucine zipper1 that permits its binding to transcription factors of the estrogen family of nuclear receptors. Leucine zippers are specific dimerization domains present in many transcription factors. In the case of TSG101, its leucine zipper has been shown to activate the c-fos promoter, and inhibits transcription by nuclear receptors of the steroid family.

All these structural features make the TSG101 protein a candidate to be implicated in oncogenesis. We have recently detected, in 77% of Burkitt's lymphoma cell lines, the presence of a new alternatively spliced message that codes for a 17-kDa protein (isoform B) that lacks the coiled-coil (leucine zipper) domain, but that partially retains the amino terminus with the ubiquitination inhibitor domain. This message has not been detected in normal cells despite the fact that a subpopulation of normal cells is capable of forming alternatively spliced non-coding messages. This isoform B, coded by the new transcript, might have a different function due to the loss of specific protein-protein interactions mediated by the coiled-coil domain (leucine zipper) in the carboxy terminus, and consequent loss of its transcriptional regulation role. Furthermore, the loss of the TSG101 dimerization domain, combined with a partially conserved amino terminal domain, the domain with E2 homology, might determine some differences in the affinity and specificity of the protein interactions mediated by this domain in each of the two isoforms.

However, the function of TSG101 gene and the significance of its different coding transcripts and protein isoforms in the biology of normal cells and in the pathogenesis of human tumors are not yet known. Therefore, in this study we have evaluated the presence of the new TSG101 transcript, coding for isoform B and other non-coding transcripts, in a large series of non-neoplastic lymphoid samples and NHLs. Our results indicate that isoform B is only expressed in tumors, and it may be involved in the pathogenesis of primary aggressive and a subset of transformed human lymphomas.

To determine the role of the different TSG101 coding transcripts in the pathogenesis of non-Hodgkin's lymphomas, we have analyzed the expression of the normal, isoform A, message, the non-coding, and the coding isoform B splicing variants, in 79 cases of Non-Hodgkin lymphomas including 41 indolent and 38 aggressive variants. These transcripts were also examined in normal lymphocytes and reactive tonsils. The different types of messages were determined following the same methodological strategy previously described in Burkitt lymphomas. The message coding for isoform B, when present, is readily detected (Figure 1). Nucleotide sequencing of this new B message confirmed that it joins nucleotide 283, in exon 3, to nucleotide 1055, within exon 9, using a cryptic splice acceptor site, and thus with a different carboxy terminus.

In Table 1, we show the expression of the different TSG101 transcripts in the 79 lymphomas examined. Isoform B was always detected in tumors in which alternatively spliced non-coding transcripts were already expressed, and is always co-expressed with the isoform A message (Figure 1). This observation suggests that expression of non-coding splicing-variants may be an initial process preceding the formation of the coding TSG101 isoform B message. Non-coding messages were observed in 10-25% of the normal lymphoid cell samples, and the isoform B was not detected in these cases. However, non-coding transcripts were found in 42-100% of the tumors (Table 1) ($P < 0.0001$, using the likelihood ratio chi-square method to compare with the distribution in normal lymphoid cells). These transcripts were constantly expressed in primary large B cell lymphomas (100%), but only in 42-55% of chronic lymphocytic leukemias (CLL) and follicular lymphomas (FL) ($P < 0.001$) (Table 1).

Isoform B was detected in 25 to 75% of tumors. All the cases with coding isoform B expression also had alternatively spliced non-coding transcripts. This coding splicing variant affiliations. (TSG101B) was more frequently found in aggressive (53%) than in indolent lymphomas (27%) ($P < 0.03$). However, the expression was variable in different types of tumors. Thus, no differences in the expression were observed between FL or typical MCL and their corresponding aggressive counterparts such as FL grade 3 and blastoid MCL (Table 1). However, it was present in 75% of primary LCL and in 57% of LCL transformed from CLL (Richter's syndrome), but in only 21% of CLL ($P < 0.001$) (Table 1). These findings suggest that the detection of TSG101 isoform B might be associated with more aggressive variants of some subtypes of lymphomas, such as CLL and LCL. These tumors were previously characterized for the expression of p16INK4a, cdc25 and c-myc, but no correlation could be unambiguously established with these parameters.

The higher incidence of alternatively spliced non-coding transcripts and the constant association of the coding isoform B transcript with these non-coding messages suggests a two-step process in the differential expression of these splicing variants in NHL. In an initial step, some tumors would develop an increased rate of alternative-splicing reactions resulting in the heterogeneous formation of non-coding messages. In a second step, a subgroup of tumors would develop the expression of the message coding for the new TSG101 isoform B, lacking the leucine zipper domain. The association of the TSG101 isoform B with more aggressive variants of lymphomas suggests a possible role in the pathogenesis of these tumors. The mechanisms leading to the activation of these alternative-splicing variants and the biological consequences of the expression of coding isoform B lacking the transcriptional repressor domain are not yet known.

The biological effect of TSG101 isoform B is likely to be the result of the modification of its structure, by the loss of its leucine zipper, and modification of its N-terminus. Therefore, the TSG101 isoform B has lost its gene regulation potential. The partially different structure of TSG101B N-terminal domain might also permit different interactions with other cellular proteins by change either in affinity or specificity. However, the presence of this coding transcript in more aggressive variants of lymphomas suggests that it may be involved in the progression of these tumors. The modification of TSG101 terminus might be altering the regulation of ubiquitination reactions that control cyclin levels and might allow replication before DNA integrity has been checked, thus favoring the appearance of additional mutations. Two TSG101-related proteins are able to complement defects regarding cell sensitivity to DNA damaging agents in yeast, and also are capable of trans-activating the c-fos gene expression. These two effects are likely to be caused by each of the two domains of the TSG101-related proteins, and thus the variant isoform B, has retained its potential effect on cell cycle, while losing its gene regulation potential. It is likely that the effect of TSG101 isoforms, and their contribution to the tumor phenotype, is mediated by altering the regulatory properties of the cell cycle.