

1 ***MALAT1* Expression is Associated with Aggressive Behavior in Indolent B-** 2 **Cell Neoplasms**

3 Elena María Fernández-Garnacho¹, Ferran Nadeu^{1,2}, Silvia Martín¹, Pablo Mozas^{1,3},
4 Andrea Rivero^{1,3}, Julio Delgado^{1,2,3}, Eva Giné^{1,2,3}, Armando López-Guillermo^{1,2,3}, Martí
5 Duran-Ferrer¹, Itziar Salaverria^{1,2}, Cristina López^{1,2}, Sílvia Beà^{1,2,3}, Santiago Demajo¹,
6 Pedro Jares^{1,2,3}, Xose S Puente^{2,4}, José Ignacio Martín-Subero^{1,2,5}, Elias Campo^{1,2,3,5}, Lluís
7 Hernández^{1,2*}

8

9 ¹Lymphoid Neoplasm Program, Institut d'Investigacions Biomèdiques August Pi I Sunyer
10 (IDIBAPS), Barcelona, Spain. ²Centro de Investigación Biomédica en Red de Cáncer
11 (CIBERONC), Madrid, Spain. ³Hospital Clínic of Barcelona, Universitat de Barcelona,
12 Barcelona, Spain. ⁴University of Oviedo, Oviedo, Spain. ⁵Institució Catalana de Recerca i
13 Estudis Avançats (ICREA), Barcelona, Spain.

14

15

16 **Abstract**

17 *MALAT1* is a long non-coding RNA with oncogenic roles in cancer but poorly
18 studied in indolent B-cell neoplasms. Here, *MALAT1* expression was analyzed using RNA-
19 seq, microarrays or qRT-PCR in primary samples from various clinico-biological subtypes
20 of chronic lymphocytic leukemia (CLL, n=266) and follicular lymphoma (FL, n=61). In
21 peripheral blood (PB) CLL samples, high *MALAT1* expression was associated with a
22 significantly shorter time to treatment, independently from other known prognostic factors,
23 such as IGHV mutational status. Coding genes whose expression levels were associated
24 with *MALAT1* in CLL were predominantly related to oncogenic pathways stimulated in the
25 lymph node (LN) microenvironment. Further analysis of *MALAT1* expression by
26 microarrays in paired CLL samples from PB/LN showed that its levels were maintained
27 between both anatomical compartments, supporting that the clinical value of *MALAT1*
28 expression found in PB is mirroring expression differences already present in LN.
29 Similarly, high *MALAT1* expression in FL predicted for a shorter progression-free survival,
30 and its correlated expressed genes were associated with pathways promoting FL
31 pathogenesis. In summary, *MALAT1* expression is related to pathophysiology and clinical
32 behavior of indolent B-cell neoplasms. Particularly in CLL its levels could be a surrogate
33 marker of the microenvironment stimulation and may contribute to refine the clinical
34 management of these patients.

35 **Keywords:** long non-coding RNAs, *MALAT1*, Lymphomas, CLL, FL, Biomarker.

36 ***Corresponding Author:** Lluís Hernández, Institut d'Investigacions Biomèdiques August
37 Pi I Sunyer (IDIBAPS), Centre Esther Koplowitz (CEK), Rosselló 153, Barcelona 08036,
38 Spain. Phone: 34-93-2275400, ext.4582; Fax: 34-93-2275717; orcid.org/0000-0002-4854-
39 3069

40 E-mail: hernan@recerca.clinic.cat

41 Introduction

42 Long non-coding RNAs (lncRNAs) regulate the expression of protein-coding genes
43 and are increasingly described as key players in physiological and pathological conditions,
44 most remarkably in cancer^{1,2}. One of the lncRNAs most frequently related to oncogenesis
45 is *MALAT1*, which has been implicated in the regulation of key cellular pathways such as
46 MAPK/ERK, PI3K/AKT, WNT/B-catenin, and NF- κ B³ and, as a consequence, involved in
47 many cancer-associated processes such as cell proliferation, migration, invasion,
48 apoptosis and angiogenesis⁴. There are also several studies supporting *MALAT1*
49 expression as a clinical biomarker mainly associated with a poor prognosis in solid
50 tumors³, although in some neoplasms such as diffuse large B-cell lymphoma (DLBCL),
51 colorectal and breast cancer, high levels of *MALAT1* have been linked to a favorable
52 outcome⁵⁻⁸.

53

54 In B-cell non-Hodgkin lymphomas (B-NHL), we and others have previously shown
55 that lncRNA deregulation is associated with immune cell-related functions and cell
56 proliferation control^{9,10}. In the case of *MALAT1*, its expression has been shown to be
57 upregulated in some lymphoid neoplasms such as DLBCL¹¹, chronic lymphocytic leukemia
58 (CLL)¹², and mantle cell lymphoma (MCL)¹³. These results suggest that *MALAT1*
59 expression could be associated with poor prognosis even in indolent B-NHL, but most of
60 these studies include relatively few cases and the possible clinical relevance of *MALAT1*
61 expression in these lymphoid neoplasms is still not well known. In this study, we have
62 performed a detailed characterization of the biological and clinical impact of *MALAT1*
63 deregulation in two types of indolent B cell lymphomas with different underlying
64 pathobiological mechanisms, namely CLL and follicular lymphoma (FL).

65

67 **Material and methods**

68 ***Description of the transcriptional data and patient cohorts***

69 This study exploited previously published genome wide transcriptional data from
70 CLL and FL. Regarding CLL, we reanalyzed our RNA-seq data from peripheral blood (PB)
71 samples obtained from 266 patients, and 25 additional monoclonal B cell lymphocytosis
72 donors (MBL) of the Spanish ICGC consortium with full clinical annotations¹⁴. We have
73 also analyzed the microarray data from the GSE21029 GEO dataset, including 17 patients
74 with paired PB and lymph node (LN) samples and 7 additional samples from PB¹⁵. For the
75 analysis of *MALAT1* in FL, we extracted RNA from formalin-fixed paraffin-embedded
76 (FFPE) LN samples of 61 FL patients (stages 1 to 3A) from the Hematopathology
77 collection of the Biobank of the Hospital Clínic de Barcelona-IDIBAPS (Spain). In addition,
78 microarray data generated from sorted FL neoplastic cells of 23 patients were obtained
79 from the GSE107367 GEO dataset¹⁶.

80

81 ***RNA extraction and qRT-PCR***

82 For the 61 FFPE FL samples, 3-10 cuts of 10µm each were used per sample for
83 RNA extraction using Allprep DNA/RNA FFPE kit (Qiagen). RNA integrity was analyzed
84 with TS4200 using DV200 index as recommended by the manufacturer (Agilent). In this
85 way, samples with DV200 less than 35% were excluded to avoid excessive degraded
86 material. Reverse transcriptase reaction was performed using High-Capacity Reverse
87 transcription kit (Applied Biosystems) with an RNA input of 90 ng.

88

89 *MALAT1* expression was analyzed by qRT-PCR using a short amplicon specifically
90 designed for FFPE samples. *MALAT1* expression levels were normalized using three
91 reference genes (*ACTB*, *GAPDH* and *YWHAZ*) which have been previously used in B-cell

92 lymphomas, including FL samples^{17,18}. Primer quantities were optimized for each amplicon
93 to reach a high efficiency (range: 90-96%) and their sequences were: 5'-
94 CCCCTTCCCTAGGGGATTTCA-3' (*MALAT1* forward), 5'-
95 AAGCCCACAGGAACAAGTCC-3' (*MALAT1* reverse), 5'- CCAACCGCGAGAAGATGAC-
96 3'(ACTB forward), 5'-TAGCACAGCCTGGATAGCAA-3' (*ACTB* reverse), 5'-
97 AGGTGAAGGTCGGAGTCA-3'(GAPDH forward), 5'
98 CAACAATATCCACTTTACCAGAGTTAA-3' (*GAPDH* reverse), 5'-
99 CAAAGACAGCACGCTAATAATGCA-3' (*YWHAZ* forward), and 5'-
100 TCAGCTTCGTCTCCTTGGGTA-3' (*YWHAZ* reverse). Amplification reactions were carried
101 out using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems, Foster City, CA)
102 following supplier's recommendations in a Step One Plus thermocycler (Applied
103 Biosystems). Transcript expression relative quantification was performed referred to a
104 calibrator sample of universal human reference RNA (Invitrogen).

105

106 ***Bioinformatic analyses and statistics***

107 Microarray data were normalized with RMA-based methodology and used to
108 analyze the expression levels of *MALAT1* transcript. Although the expression microarrays
109 were enriched in probe sets for coding genes, several probe sets that hybridize exclusively
110 to the *MALAT1* transcript were initially included. The specificity of the probe sets was
111 confirmed in the Affymetrix database (NettAffx™). Only one probe set was excluded
112 (223579_s_at) at this stage. The correlations among the different probe sets were checked
113 by strand and another probe set was excluded as an outlier (224559_at). Finally, the mean
114 values of the remaining probe sets were retained as reliable for measuring the expression
115 levels of *MALAT1* (probe sets 1558678_s_at, 223940_x_at, 224558_s_at, 224567_x_at,
116 224568_x_at and 226675_s_at).

117

118 To identify potential target genes, ranked lists of coding genes according to the
119 degree of Pearson's correlation to *MALAT1* were obtained for the different indolent B-cell
120 lymphomas under study using either RNA-seq or microarray data. Only those statistically
121 significant correlations after correction for multiple comparisons (adjusted p-value<0.05)
122 were considered for the downstream analyses. We separately analyzed positive and
123 negative correlations of coding genes with *MALAT1* using the webtool Metascape¹⁹. The
124 Metascape output of enriched (Reactome) pathways were obtained as tables, and
125 summary plots depicting relevant pathways with very high statistical significance (q<0.05)
126 were prepared with Graphpad Prism v7. Furthermore, we used Gene Set Enrichment
127 Analysis (GSEA v3.0 from <http://www.gsea-msigdb.org>) to perform additional pathway
128 enrichment analyses in CLL using the C2 curated gene signatures from MSigDB/GSEA
129 website and several specific signatures previously related to differences between LN and
130 PB in CLL²⁰. Only statistically significant signatures with NES higher than 2 as absolute
131 value were considered as potentially relevant.

132 In the clinical association studies, optimal cut-off points for *MALAT1* expression
133 groups were obtained using the *maxstat* algorithm which optimized the log-rank statistics
134 (*maxstat* package, R Software, Vienna). Cumulative incidence and Kaplan-Meier curves,
135 scatterplots, and box plot graphs were generated using both the R environment and
136 GraphPad Prism v7. Univariate and multivariate Fine-Gray regression models considering
137 *MALAT1* as a continuous variable (as a more stringent statistical approach) were used for
138 measuring its impact on time to treatment (TTT). Univariate analysis of *MALAT1* as a
139 continuous variable was performed with the Cox regression model for overall survival
140 (OS). Both calculations were performed with R v.4.0.3. Student's t-test was used to
141 evaluate the differences in the mean expression of *MALAT1* among subtypes and

142 molecular factors with previously described prognostic value in CLL. Paired t-test was
143 used to evaluate differences in *MALAT1* expression between paired samples of CLL in PB
144 and LN¹⁵. Linear regression on scatter plots was representing the correlation between
145 *MALAT1* expression and the epiCMIT score (epigenetically-determined Cumulative
146 MIToses)²¹.

147

148 **Results**

149 ***MALAT1* expression in CLL**

150 We initially evaluated the expression levels of *MALAT1* in CLL and MBL patients.
151 *MALAT1* expression was significantly higher in CLL than in MBL ($P < 0.001$) (Fig. 1A). We
152 also stratified *MALAT1* levels according to the IGHV mutational status (mutated vs non-
153 mutated) and epigenetic subtypes (naïve-like, memory-like and intermediate) and no
154 significant differences were observed among the different groups (Fig. 1A).

155

156 We next evaluated whether *MALAT1* expression was associated with the clinical
157 behavior of the tumor. Using the *maxstat* algorithm, we segregated CLL patients into high
158 and low expression groups (Fig. S1A). Patients with high *MALAT1* expression (N=65) had
159 a significantly shorter time to treatment (TTT) than the *MALAT1*-low group (N=201)
160 ($p < 0.0001$) (Fig. 1B). This finding was also confirmed in the subset of patients clinically
161 classified as Binet A (N=239, 57 with high and 182 low *MALAT1* levels) ($P = 0.0003$) (Fig.
162 S2A). On the contrary, *MALAT1* levels were not related to the overall survival (OS) of the
163 patients (Fig. S1B and Fig. S2B). The adverse impact of high *MALAT1* expression on TTT
164 was further confirmed in IGHV-mutated CLL and in the three epigenetic CLL subtypes
165 (Figure 1B). Similar findings were observed when the analyses were restricted to patients
166 with Binet A CLL (Fig. S2C-S2D). To evaluate the independent prognostic impact of
167 *MALAT1*, we performed multivariate regression models. First, we checked that *MALAT1*
168 as a continuous isolated variable also had significant prognostic value for TTT (HR=1.32;
169 95%CI: 1.18-1.48); $p < 0.0001$) but not for OS (HR = 1.05; 95%CI: 0.89-1.24; $p = 0.561$).
170 Next, the multivariate analyses confirmed that *MALAT1* expression also had an
171 independent prognostic value for TTT considering Binet stage and the IGHV mutational
172 status ($P = 0.0004$) and Binet stage and the epigenetic subgroups ($P < 0.0001$) (Fig. 2A). We

173 also evaluated the possible association of *MALAT1* levels with other molecular factors
174 previously shown to have prognostic value in CLL, such as mutations in driver genes,
175 number of chromosomal aberrations^{14,22}, IGLV3-21 variant/R110 mutation²³, or the DNA
176 methylation-based epiCMIT score related to the proliferative history of the tumor cells²¹.
177 *MALAT1* expression levels were not related to any of these variables (Fig. S3 and Fig.S4).
178 Overall, these data indicate that *MALAT1* expression has a prognostic value for TTT in
179 CLL irrespective of other known genetic and epigenetic prognostic parameters.

180

181 To determine whether *MALAT1* deregulation could be related to genetic alterations,
182 we analyzed the gene mutational status in the whole-genome sequences of 150 cases
183 from the ICGC Consortium¹⁴. Only five CLL and one MBL revealed mutations in the
184 *MALAT1* locus and were not related to the expression of the gene (Table S1A). No copy
185 number alterations affecting 11q13, where *MALAT1* is located, were found, whereas only 2
186 CLL cases had copy number neutral loss of heterozygosity. None of these alterations were
187 related to the expression of the gene (Table S1B).

188

189 To evaluate the possible functional implications of *MALAT1* expression we
190 searched for coding genes that were significantly correlated, either positively or negatively,
191 with *MALAT1* in the different CLL subgroups (Table S2) and we subsequently performed
192 pathway enrichment analyses. Among the positively correlated genes, the significantly
193 selected pathways were remarkably associated with functions related to activation,
194 proliferation, and survival of CLL cells in the LN microenvironment (Fig. 2B, Fig. S5, Table
195 S3 and Table S4). These pathways included general cell proliferation-related functions
196 such as signaling by PI3K/AKT, MAPK, IL10 and IL4, among others²⁴⁻²⁹. As these results
197 pointed to a possible influence of microenvironment stimuli in *MALAT1* expression we
198 used GSEA to analyze CLL specific signatures and finding, both in U-CLL and M-CLL,

199 significant positive enrichments in a signature of genes previously described as
200 significantly upregulated in LN vs. PB (Fig. 3B)²⁰. Finally, we also analyzed previously
201 published microarray data of paired LN and PB CLL samples¹⁵, where we found that
202 *MALAT1* levels did not change significantly between LN and PB (P=0.097) (Fig. 3C). All
203 these results suggest that the relationship between *MALAT1* expression and shorter TTT
204 that we have observed in PB samples could represent a surrogate biomarker for the
205 degree of stimulation and proliferation of CLL cells in the LN.

206

207 ***MALAT1* expression in Follicular Lymphoma**

208 Based on the results of *MALAT1* in CLL and its potential association with the LN
209 microenvironment, we further explored the role of *MALAT1* in FL, another indolent B cell
210 lymphoid neoplasm with predominant nodal presentation. *MALAT1* expression levels were
211 analyzed by qRT-PCR in a series of 61 grade 1-3A FL. These patients had been
212 homogenously treated with R-CHOP. Using the *maxstat* algorithm, we identified two
213 groups of FL with high (N=11) and low (N=50) *MALAT1* expression. The clinical, biological,
214 and histological characteristics, including histological grade, were similar in both
215 subgroups (Table S5). FL cases with high *MALAT1* expression had a significantly shorter
216 progression-free survival (PFS) than cases with low expression (P=0.017) (Fig. 4A).
217 However, no significant differences were observed in transformation to DLBCL or OS
218 between cases with low and high expression (P=0.088 and P=0.0177) (Fig. S6).

219

220 To explore the possible biological role of *MALAT1* expression in FL we reanalyzed
221 microarray expression data previously generated on purified FL B-cells¹⁶. We observed
222 that the coding genes positively correlating with *MALAT1* expression (Table S6) were
223 associated with cell proliferation and other pathways like VEGF, MAPK, interleukin
224 signaling (including some particular pathways related with IL-3 and IL-5) and BCR

225 signaling, also described as involved in FL oncogenesis (Fig. 4B and Table S7)^{30,31}.

226 Moreover, among the genes positively correlated with *MALAT1* involved in BCR signaling

227 we noticed *PIK3CD* that it has been previously described as predictor of poor prognosis in

228 FL³².

229 Finally, we performed a comparison of all genes involved in pathways significantly

230 enriched in FL and compared to those found in CLL (considering together U-CLL and M-

231 CLL gene lists). A large proportion of those genes were exclusive for each lymphoma type

232 (88% in FL and 99% in CLL) (Fig. S7 and Table S8). Concordantly, cell cycle-related or

233 MAPK pathways that were initially observed in both neoplasms involved different sets of

234 *MALAT1* positively correlated genes (Tables S9 and Table S10). On the other hand, a total

235 of 12 genes were found in common in the comparative analysis, including several involved

236 in cytokine and interleukin signaling (*CAPZA1*), cell cycle (*GAR1*), or both (*HSP90AA1*,

237 *LMNB1* and *LYN*)(Table S11).

238

239 Discussion

240 In this report, we provide novel insights into the clinical and biological role of
241 *MALAT1* in indolent B cell neoplasms. Our study revealed that *MALAT1* upregulation was
242 associated with a detrimental clinical behavior in the different entities studied, associated
243 with a shorter TTT in CLL or shorter PFS in FL. Interestingly, *MALAT1* upregulation was a
244 prognostic factor in CLL independently of the IGHV mutational status, epigenetic
245 subgroups, Binet stage or genetic alterations. In FL, *MALAT1* overexpression was also
246 associated with shorter progression-free survival although the clinical and biological
247 features of patients with low and high expression were similar.

248

249 We also studied the potential causes and consequences of *MALAT1* upregulation
250 that might justify its clinical behavior in indolent B cell neoplasms. In CLL, *MALAT1*
251 mutations were rare and not related to the expression of the gene. This finding is similar to
252 those in other neoplasms such as bladder cancer, head and neck squamous cell cancer,
253 and lung adenocarcinomas, in which single nucleotides variants and indels were found,
254 although considered passenger events caused by a transcription-associated mutational
255 process^{33,34}.

256

257 Therefore, other mechanisms should be related with *MALAT1* overexpression and,
258 in fact, its promoter has been described in solid tumor models to be regulated by
259 transcription factors as *HIF1 α* and *STAT3*, which are induced by microenvironmental
260 factors as hypoxia and cytokines/interleukins, respectively³⁵⁻³⁷. In CLL, microenvironment
261 stimulation occurs in the LN and is a key process for activation, proliferation and survival of
262 these tumoral B cells³⁸. Interestingly, our guilt-by-association analysis on the coding genes
263 correlated with *MALAT1* in PB CLL samples revealed significant enrichments in a gene

264 signature highly expressed in nodal CLL²⁰, as well as for interleukin-related signatures. On
265 top of that, we found that *MALAT1* levels remain similar in PB compared to LN paired CLL
266 samples, supporting that the clinical value of *MALAT1* expression found in PB is mirroring
267 expression differences in LN associated with the degree of the microenvironment
268 stimulation.

269

270 In FL, our study also revealed biological insights that seem to explain the clinical
271 behavior of *MALAT1* upregulation in these neoplasms. In these tumors, coding genes
272 correlated with *MALAT1* expression were enriched in cell cycle-related processes as well
273 as pathways related to cell proliferation, migration and angiogenesis, such as MAPK and
274 VEGF pathways, which have been previously linked to *MALAT1* function in other tumor
275 models³⁹. We also found that cases with high *MALAT1* expression were significantly
276 associated with high levels of genes involved in BCR signaling and interleukin signaling,
277 such as *LYN* and *PIK3CD*, the expression of which promote cell growth and have been
278 associated with poor prognosis in FL^{32,40}. In this regard, these are relevant associations
279 supporting our results about the clinical impact of *MALAT1*, even though our series
280 included a relative limited number of cases.

281

282 Finally, the comparison of *MALAT1* associated pathways and the involved genes
283 showed a low degree of overlapping in the corresponding gene signatures, even in cell
284 cycle-related or MAPK pathways that were initially observed as enriched in both
285 neoplasms. These results suggest that *MALAT1* could be contributing to the regulation of
286 different transcriptional gene sets in both lymphomas. In spite of these differences at the
287 gene level, we also observed that processes as cell proliferation and MAPK pathways
288 were potentially commonly affected by *MALAT1* upregulation, and thus could explain its
289 common clinical impact in both neoplasms.

290

291 In summary, these findings highlight that *MALAT1* overexpression plays a role in
292 the pathobiology and clinical behavior of indolent B cell neoplasms and related with a more
293 aggressive behavior of patients with higher *MALAT1* levels. Particularly in CLL, *MALAT1*
294 could serve as a clinical biomarker that seems to be a surrogate marker of the degree of
295 stimuli the CLL cells receive from the microenvironment. Therefore, *MALAT1* expression
296 could be taken in account in further studies as complementary to other known prognostic
297 factors to improve the clinical management of these patients.

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317 **References**

318

319 1. Huang, Y. *et al.* Regulatory long non-coding RNA and its functions. *J. Physiol. Biochem.* **68**,
320 611–618 (2012).

321 2. Bhan, A., Soleimani, M. & Mandal, S. S. Long Noncoding RNA and Cancer: A New Paradigm.
322 *Cancer Res.* **77**, 3965–3981 (2017).

323 3. Li, Z.-X. *et al.* MALAT1: a potential biomarker in cancer. *Cancer Manag Res* **10**, 6757–6768
324 (2018).

325 4. Zhang, X., Hamblin, M. H. & Yin, K.-J. The long noncoding RNA Malat1: Its physiological and
326 pathophysiological functions. *RNA Biol* **14**, 1705–1714 (2017).

327 5. Wang, Y. *et al.* The Long Noncoding RNA MALAT-1 is A Novel Biomarker in Various Cancers: A
328 Meta-analysis Based on the GEO Database and Literature. *J Cancer* **7**, 991–1001 (2016).

329 6. Lenz, G. *et al.* Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* **359**, 2313–
330 2323 (2008).

331 7. Kwok, Z. H., Roche, V., Chew, X. H., Fadieieva, A. & Tay, Y. A non-canonical tumor suppressive
332 role for the long non-coding RNA MALAT1 in colon and breast cancers. *Int J Cancer* **143**, 668–
333 678 (2018).

334 8. Kim, J. *et al.* Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat Genet* **50**,
335 1705–1715 (2018).

336 9. Zhou, M. *et al.* Discovery and validation of immune-associated long non-coding RNA
337 biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell
338 lymphoma. *Mol. Cancer* **16**, 16 (2017).

- 339 10. Roisman, A. *et al.* Differential expression of long non-coding RNAs are related to proliferation
340 and histological diversity in follicular lymphomas. *Br J Haematol* **184**, 373–383 (2019).
- 341 11. Wang, Q.-M., Lian, G.-Y., Song, Y., Huang, Y.-F. & Gong, Y. LncRNA MALAT1 promotes
342 tumorigenesis and immune escape of diffuse large B cell lymphoma by sponging miR-195. *Life*
343 *Sci* **231**, 116335 (2019).
- 344 12. Ahmadi, A. *et al.* Altered expression of MALAT1 lncRNA in chronic lymphocytic leukemia
345 patients, correlation with cytogenetic findings. *Blood Res* **53**, 320–324 (2018).
- 346 13. Wang, X. *et al.* LncRNA MALAT1 promotes development of mantle cell lymphoma by
347 associating with EZH2. *J Transl Med* **14**, 346 (2016).
- 348 14. Puente, X. S. *et al.* Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*
349 **526**, 519–524 (2015).
- 350 15. Herishanu, Y. *et al.* The lymph node microenvironment promotes B-cell receptor signaling, NF-
351 kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* **117**, 563–
352 574 (2011).
- 353 16. Pangault, C. *et al.* Follicular lymphoma cell niche: identification of a preeminent IL-4-
354 dependent T(FH)-B cell axis. *Leukemia* **24**, 2080–2089 (2010).
- 355 17. Potashnikova, D., Gladkikh, A. & Vorobjev, I. A. Selection of superior reference genes'
356 combination for quantitative real-time PCR in B-cell lymphomas. *Ann Clin Lab Sci* **45**, 64–72
357 (2015).
- 358 18. Yagi, K. *et al.* Expression of multidrug resistance 1 gene in B-cell lymphomas: association with
359 follicular dendritic cells. *Histopathology* **62**, 414–420 (2013).
- 360 19. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-
361 level datasets. *Nat Commun* **10**, 1523 (2019).

- 362 20. Dadashian, E. L. *et al.* TLR Signaling Is Activated in Lymph Node-Resident CLL Cells and Is Only
363 Partially Inhibited by Ibrutinib. *Cancer Res* **79**, 360–371 (2019).
- 364 21. Duran-Ferrer, M. *et al.* The proliferative history shapes the DNA methylome of B-cell tumors
365 and predicts clinical outcome. *Nature Cancer* **1**, 1066–1081 (2020).
- 366 22. Quesada, V. *et al.* The genomic landscape of chronic lymphocytic leukemia: clinical
367 implications. *BMC Med* **11**, 124 (2013).
- 368 23. Nadeu, F. *et al.* IGLV3-21R110 identifies an aggressive biological subtype of chronic
369 lymphocytic leukemia with intermediate epigenetics. *Blood* **137**, 2935–2946 (2020).
- 370 24. Mittal, A. K. *et al.* Chronic lymphocytic leukemia cells in a lymph node microenvironment
371 depict molecular signature associated with an aggressive disease. *Mol Med* **20**, 290–301
372 (2014).
- 373 25. Brocke-Heidrich, K. *et al.* Interleukin-6-dependent gene expression profiles in multiple
374 myeloma INA-6 cells reveal a Bcl-2 family-independent survival pathway closely associated
375 with Stat3 activation. *Blood* **103**, 242–251 (2004).
- 376 26. Kowalska, W. & Bojarska-Junak, A. Monocytic MDSC as a source of immunosuppressive
377 cytokines in chronic lymphocytic leukemia (CLL) microenvironment. *Folia Histochem Cytobiol*
378 **58**, 25–36 (2020).
- 379 27. Mirabilij, S. *et al.* Biological Aspects of mTOR in Leukemia. *Int J Mol Sci* **19**, (2018).
- 380 28. Lezina, L. *et al.* CD40L/IL-4-stimulated CLL demonstrates variation in translational regulation of
381 DNA damage response genes including ATM. *Blood Adv* **2**, 1869–1881 (2018).
- 382 29. Shukla, A., Shukla, V. & Joshi, S. S. Regulation of MAPK signaling and implications in chronic
383 lymphocytic leukemia. *Leuk Lymphoma* **59**, 1565–1573 (2018).
- 384 30. Devan, J., Janikova, A. & Mraz, M. New concepts in follicular lymphoma biology: From BCL2 to
385 epigenetic regulators and non-coding RNAs. *Semin Oncol* **45**, 291–302 (2018).

- 386 31. Duś-Szachniewicz, K., Rymkiewicz, G., Agrawal, A. K., Kołodziej, P. & Wiśniewski, J. R. Large-
387 Scale Proteomic Analysis of Follicular Lymphoma Reveals Extensive Remodeling of Cell
388 Adhesion Pathway and Identifies Hub Proteins Related to the Lymphomagenesis. *Cancers*
389 (*Basel*) **13**, (2021).
- 390 32. Dong, T. *et al.* The Expression of CD9 and PIK3CD is Associated with Prognosis of Follicular
391 Lymphoma. *J Cancer* **6**, 1222–1229 (2015).
- 392 33. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature*
393 **502**, 333–339 (2013).
- 394 34. Rheinbay, E. *et al.* Analyses of non-coding somatic drivers in 2,658 cancer whole genomes.
395 *Nature* **578**, 102–111 (2020).
- 396 35. Lei, L. *et al.* Functions and regulatory mechanisms of metastasis-associated lung
397 adenocarcinoma transcript 1. *J Cell Physiol* **234**, 134–151 (2018).
- 398 36. Wang, Y. *et al.* TGF- β -induced STAT3 overexpression promotes human head and neck
399 squamous cell carcinoma invasion and metastasis through malat1/miR-30a interactions.
400 *Cancer Lett* **436**, 52–62 (2018).
- 401 37. Zheng, T., Ma, G., Tang, M., Li, Z. & Xu, R. IL-8 Secreted from M2 Macrophages Promoted
402 Prostate Tumorigenesis via STAT3/MALAT1 Pathway. *Int J Mol Sci* **20**, (2018).
- 403 38. Stevenson, F. K., Forconi, F. & Kipps, T. J. Exploring the pathways to chronic lymphocytic
404 leukemia. *Blood* **138**, 827–835 (2021).
- 405 39. Arun, G., Aggarwal, D. & Spector, D. L. MALAT1 Long Non-Coding RNA: Functional
406 Implications. *Noncoding RNA* **6**, (2020).
- 407 40. Tauzin, S., Ding, H., Burdevet, D., Borisch, B. & Hoessli, D. C. Membrane-associated signaling in
408 human B-lymphoma lines. *Exp Cell Res* **317**, 151–162 (2011).

409

410

411

412 **Statements & Declarations**

413 ***Acknowledgements***

414 This research was funded by Ministerio de Ciencia e Innovación (MCI), grant
415 numbers RTI2018-094274-B-I00 to EC.; Fundació La Marató de TV3 (“projecte finançat
416 per Fundació La Marató de TV3”) 201920-30 to LH. Suport Grups de Recerca AGAUR
417 2014-SGR-795 and 2017-SGR-736 of the Generalitat de Catalunya to EC and JIM-S.,
418 respectively. AGAUR 2018 FIB00696, Generalitat de Catalunya to EMF-G. EC and JIM-S
419 are Academia Researchers of the "Institució Catalana de Recerca i Estudis Avançats"
420 (ICREA) of the Generalitat de Catalunya. We are indebted to the Genomics core facility of
421 the Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS).

422

423 ***Authors' contributions***

424 Conceived and designed the Study: Lluís Hernández and Elías Campo. Collected
425 and processed clinical/expression data from ICGC, Hospital Clinic de Barcelona and GEO:
426 Elena María Fernández-Garnacho, Ferran Nadeu, Pablo Mozas, Andrea Rivero, Julio
427 Delgado, Eva Giné, Armando López-Guillermo, Martí Duran-Ferrer, Itziar Salaverria,
428 Cristina López, Sílvia Beà, Santiago Demajo, Pedro Jares, Xose S Puente and Lluís
429 Hernández; Processed FFPE material from FL samples for RNA purification: Silvia Martín.
430 Analyzed the RNAseq/microarrays/qRT-PCR data and performed the statistical analysis:
431 Elena María Fernández-Garnacho, Ferran Nadeu, Pablo Mozas, Andrea Rivero and Lluís

432 Hernández wrote the manuscript: Elena María Fernández-Garnacho, Lluís Hernández,
433 Elías Campo and José Ignacio Martín-Subero. All authors read and approved the final
434 manuscript.

435 ***Data availability***

436 The data that support the findings of this study are available from the
437 corresponding author upon reasonable request.

438

439 ***Competing interests***

440 The authors declare that they have no competing interests.

441

442 ***Ethics approval and consent to participate***

443 Samples from FL patients were obtained from the Hematopathology collection of
444 the Biobank of the Hospital Clínic de Barcelona-IDIBAPS (Spain). RNAseq data and full
445 clinical annotations from CLL patients were obtained from the Spanish ICGC consortium,
446 which had an approval of the Institutional Review Board of Hospital Clínic de Barcelona, in
447 accordance with national regulations and the Declaration of Helsinki. All patients provided
448 written informed consent. The remaining data from other samples were obtained from
449 GEO microarray public repository including previously published works with their particular
450 ethical compliances stated in the original articles.

451

452

453

454

455 **Figure legends**

456 **Figure 1. *MALAT1* expression levels in MBL, CLL molecular subgroups, and**
457 **relationship to time to treatment (TTT). (a) *MALAT1* expression was higher in CLL**
458 **patients than in MBL ($P < 0.001$). Dotted line indicates the threshold of *MALAT1* that**
459 **determines differences in outcome. No significant differences were observed in *MALAT1***
460 **expression between CLL subgroups defined by IGHV mutational status (M-CLL versus U-**
461 **CLL) (bottom left panel) or epigenetic subtypes (n-CLL: naïve-like, m-CLL: memory-like**
462 **and i-CLL: intermediate) (bottom right panel). (b) CLL patients with high *MALAT1***
463 **expression showed significantly shorter TTT compared to those with low levels in the**
464 **global cohort (top panel), and almost every one of the different CLL subgroups related to**
465 **the IGHV mutational status (bottom left) and epigenetic subtypes (bottom right).**

466 **Figure 2. *MALAT1* expression is an independent prognostic factor and is related to**
467 **relevant pathways in CLL. (a) Plot for multivariate model analyses, where *MALAT1***
468 **expression as a continuous variable shows its independent prognostic value considered**
469 **together with Binet stage and IGHV mutational status (left panel) or epigenetic subtypes**
470 **(right panel). For epigenetic groups, m-CLL was taken as a reference. (b) Summary of**
471 **most relevant significant pathway enrichments found using Metascape tool regarding**
472 **coding genes positively (top panel) and negatively (bottom panel) correlated with *MALAT1***
473 **expression in CLL subtypes defined by IGHV mutational status. Only statistically**

474 significant Reactome pathways after multiple comparison correction are shown for
475 compact representation.

476 **Figure 3. *MALAT1* levels correlate with many genes that are upregulated in LN**
477 **compared to PB CLL samples, whereas its expression was comparable in both**
478 **anatomical compartments. (a)** GSEA analysis shows that the nodal CLL signature (see
479 material and methods) was significantly enriched in coding genes correlated with *MALAT1*
480 expression in PB CLL samples of M-CLL and U-CLL subtypes. **(b)** No significant
481 differences in *MALAT1* expression between paired PB and LN tissues (GSE21029) were
482 observed.

483 **Figure 4. *MALAT1* expression is associated with a more aggressive behavior and is**
484 **related to pathogenetic pathways in FL. a.** Kaplan-Meier curves for progression-free
485 survival (PFS) according to *MALAT1* expression in FL. Patients with high *MALAT1*
486 expression had a significantly shorter PFS than those with low levels. **b.** Summary of most
487 relevant significant pathway enrichments among coding genes positively correlated with
488 *MALAT1* in FL samples. Only statistically significant Reactome pathways after multiple
489 comparison correction are shown for compact representation.

490

491 **Supplementary Figure Legends**

492 **Fig.S1. (a)** Optimal cut-off established by *maxstat* algorithm regarding *MALAT1*
493 expression and TTT in CLL. **(b)** Optimal cut-off established by *maxstat* algorithm regarding
494 *MALAT1* expression and OS in CLL.

495 **Fig.S2. (a)** Binet A CLL cases with high *MALAT1* expression had a significantly shorter
496 TTT than those with low levels. **(b)** No difference in OS was detected in all CLL cases with

497 high or low *MALAT1* expression neither using the same cutoff defining significant
498 differences for TTT (see Suppl. Fig1b) nor the optimal for OS (data not shown). **(c)** TTT
499 and *MALAT1* expression levels in IGHV-mutated and unmutated Binet A CLL cases. **(d)**
500 TTT and *MALAT1* expression levels in Binet A CLL epigenetic subtypes.

501 **Fig.S3. (a)** Similar *MALAT1* levels in CLL cases with a different number of genetic
502 alterations. **(b)** No significant differences in *MALAT1* expression comparing CLL cases
503 with (mutated) or without (wild-type) individual gene alterations after correction for multiple
504 comparisons. **(c)** No significant differences in *MALAT1* expression comparing CLL cases
505 with (altered) or without (wild-type) individual chromosomal alterations after correction for
506 multiple comparisons.

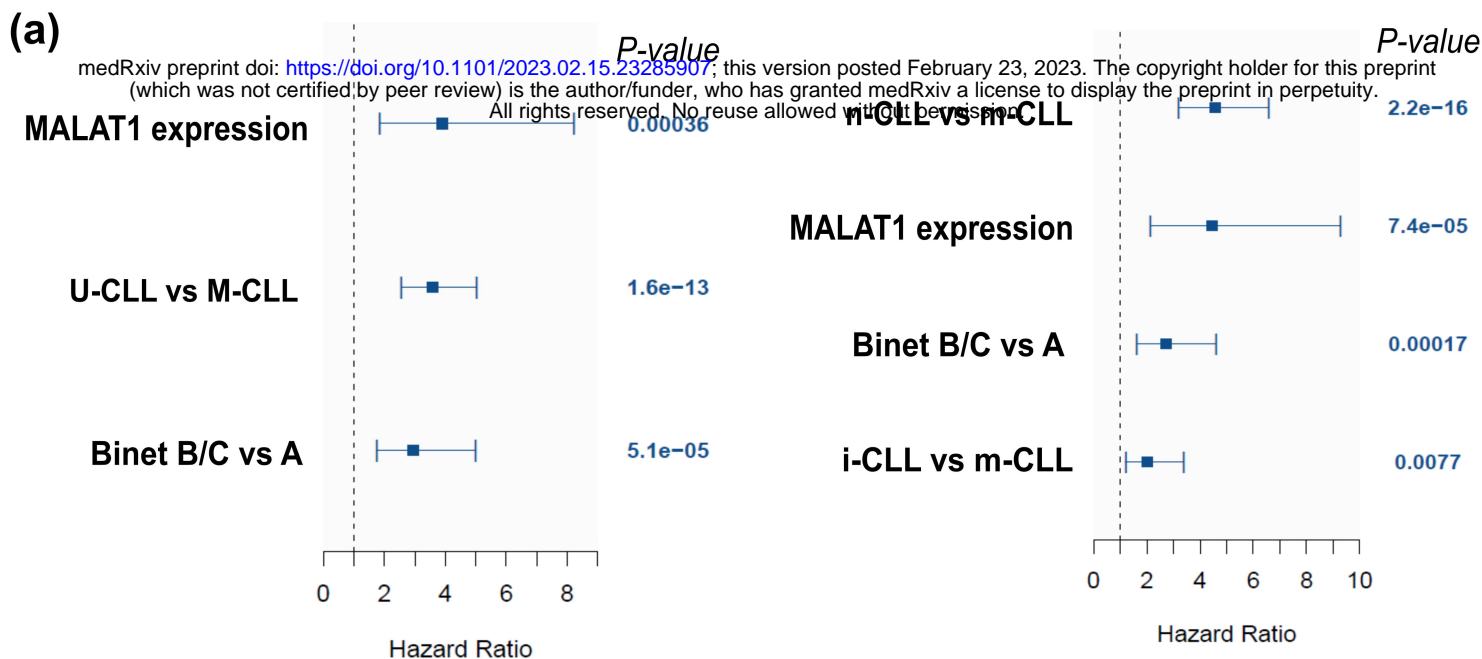
507 **Fig.S4. (a)** Boxplot showing the distribution of the values of the three epiCMT indexes in
508 the three epigenetic groups of CLL cases analyzed according to the *MALAT1* expression
509 categories. No significant differences were found in any comparison. **(b)** Scatterplots
510 showing the linear correlation analyses between the three epiCMT indexes and *MALAT1*
511 expression values, in the three CLL epigenetic subtypes. None of these correlations were
512 statistically significant.

513 **Fig.S5.** Summary of most relevant significant pathway enrichments found using
514 Metascape tool involving signatures previously related to CLL pathogenesis and poor
515 prognosis. Separated analyses were performed involving coding genes positively (top
516 panel) and negatively (bottom panel) correlated with *MALAT1* expression in the different
517 epigenetic CLL subtypes. Only statistically significant Reactome pathways after multiple
518 comparison correction are shown for compact representation.

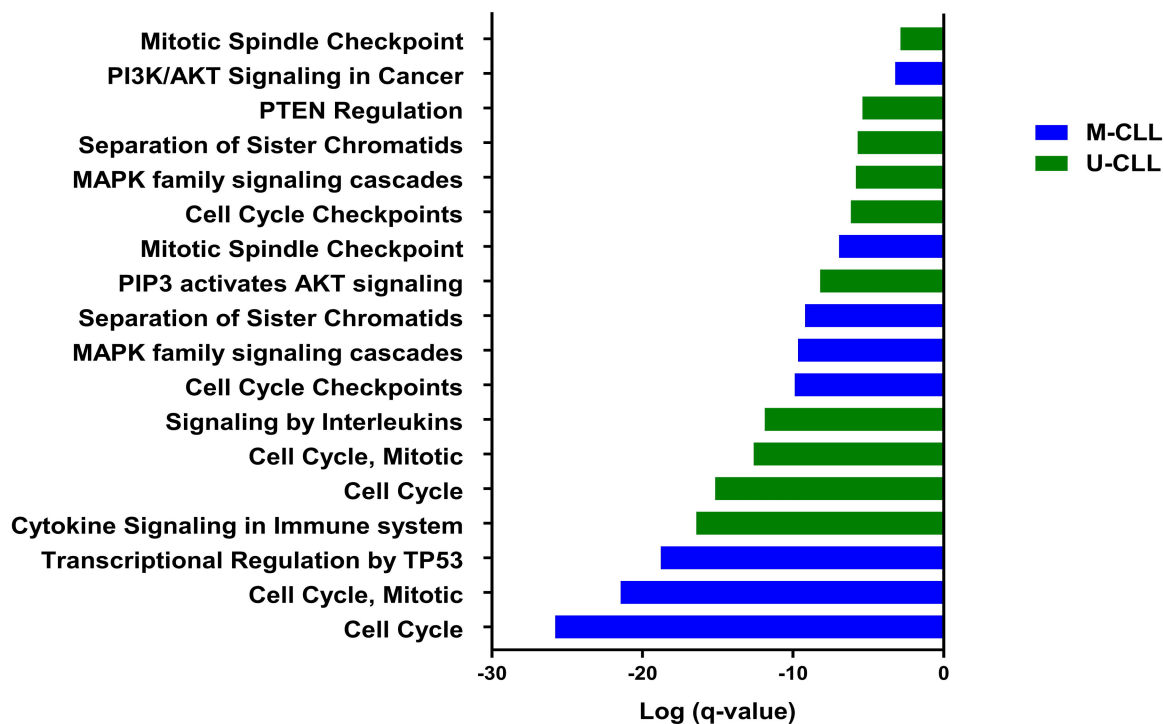
519 **Fig.S6. (a)** Kaplan-Meier curve showing the lack of significant differences in OS of FL
520 cases regarding *MALAT1* expression. **(b)** Cumulative incidence curves showing the lack
521 of significant differences regarding *MALAT1* expression and the risk of transformation.

522 **Fig.S7.** Venn diagram showing the overlapping between FL and CLL of *MALAT1*-
523 correlated genes in pathways found significantly enriched in separated analysis.

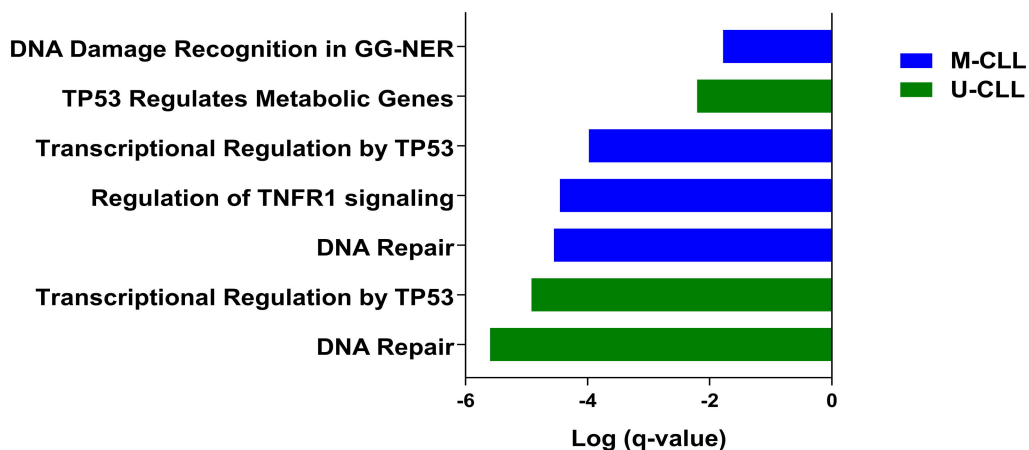
524



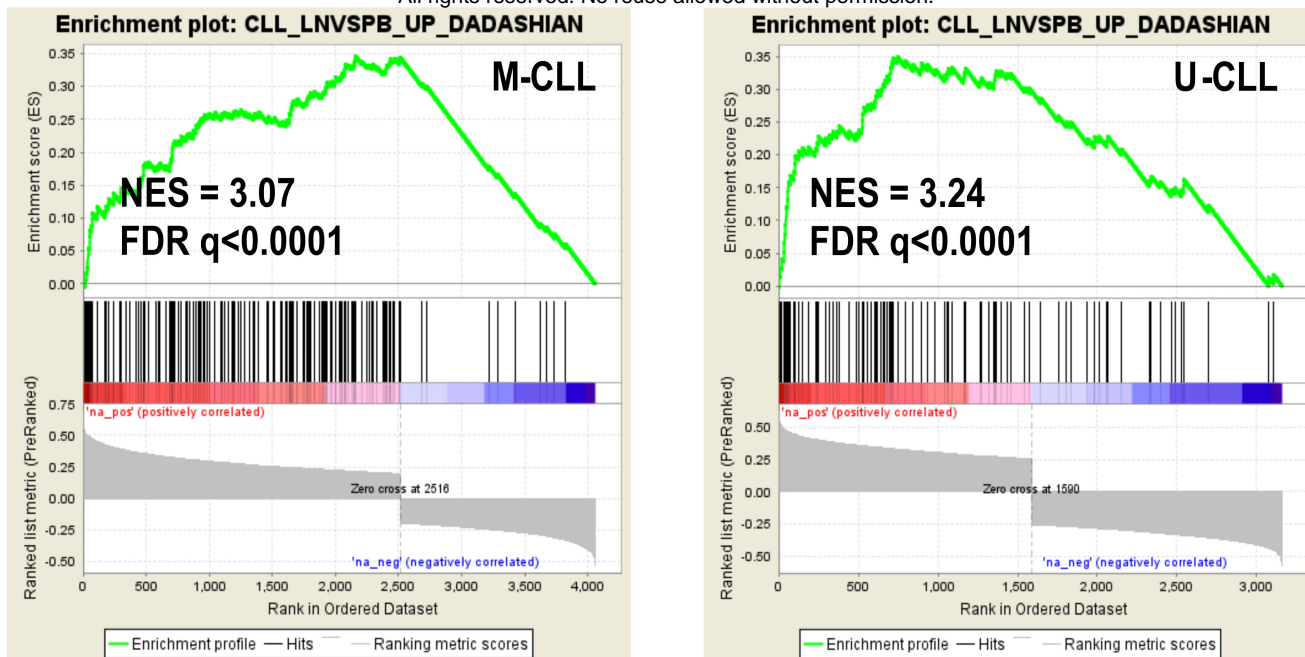
(b) Enriched pathways in positively correlated genes with MALAT1 (CLL)



Enriched pathways in negatively correlated genes with MALAT1 (CLL)



(a) medRxiv preprint doi: <https://doi.org/10.1101/2023.02.15.23285907>; this version posted February 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



(b)

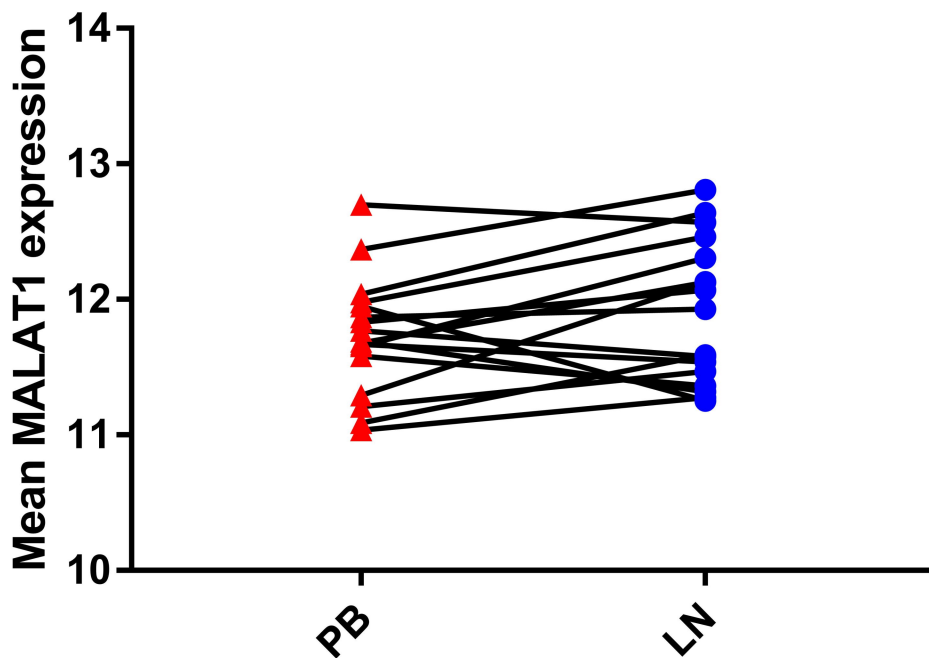
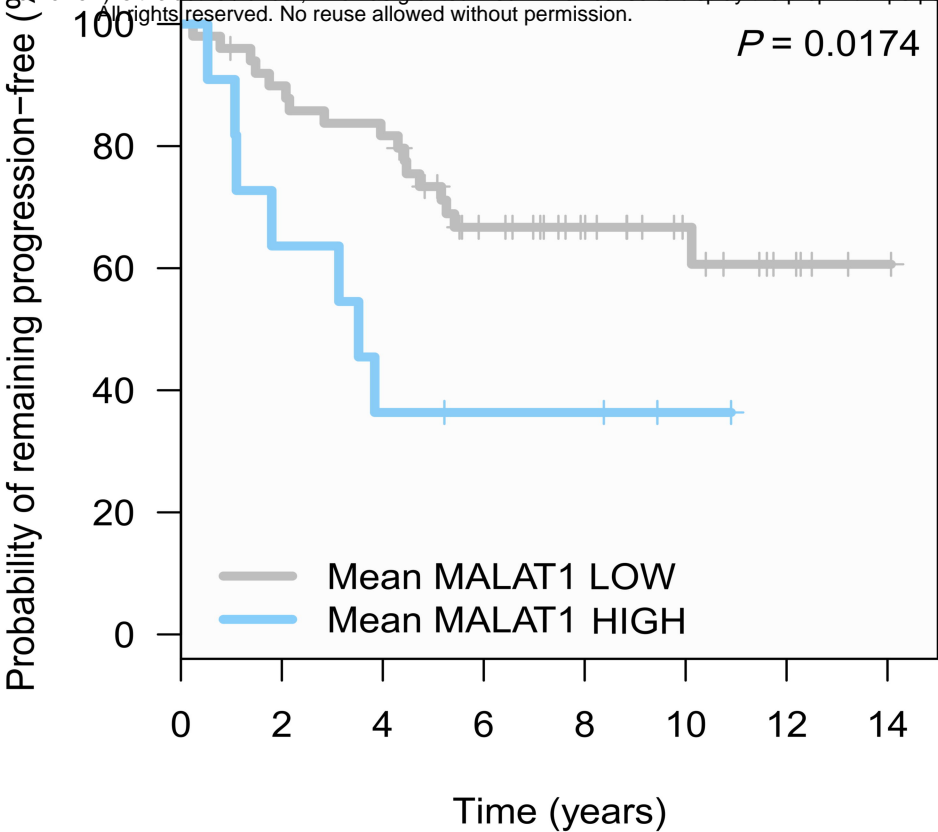


Figure 4

(a) preprint doi: <https://doi.org/10.1101/2023.02.15.23285907>; this version posted February 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



No. at risk:

	0	2	4	6	8	10	12	14
Mean MALAT1 LOW	50	44	40	27	19	11	5	1
Mean MALAT1 HIGH	11	7	4	3	3	1	0	0

(b)

Enriched pathways in positively correlated genes with MALAT1 (FL)

