Micro-RNA networks in T-cell prolymphocytic leukemia reflect T-cell activation and shape DNA damage response and survival pathways

by Till Braun, Markus Glaß, Linus Wahnschaffe, Moritz Otte, Petra Mayer, Marek Franitza, Janine Altmüller, Michael Hallek, Stefan Hüttelmaier, Alexandra Schrader, and Marco Herling

Haematologica 2020 [Epub ahead of print]


Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
**Article**

**Micro-RNA networks in T-cell prolymphocytic leukemia reflect T-cell activation and shape DNA damage response and survival pathways**

Till Braun*1, Markus Glaß2, Linus Wahnschaffe*4, Moritz Otte1, Petra Mayer1, Marek Franitza3, Janine Altmüller3, Michael Hallek1, Stefan Hüttelmaier*2, Alexandra Schrader*1, and Marco Herling*1

1 Department I of Internal Medicine, Center for Integrated Oncology (CIO), Aachen-Bonn-Cologne-Duesseldorf, Excellence Cluster for Cellular Stress Response and Aging-Associated Diseases (CECAD), Center for Molecular Medicine Cologne (CMMC), University of Cologne (UoC), 50937 Cologne, Germany
2 Institute of Molecular Medicine, Section for Molecular Cell Biology, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Charles Tanford protein center, 06120 Halle, Germany
3 Cologne Center for Genomics, Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50937 Cologne, Germany

*equal contribution

Correspondence: marco.herling@uk-koeln.de; Tel.: +49-221-478-51464

**Running title:** T-cell receptor activity shaped miR-omes of T-PLL

**Article Summary:** Using small-RNA and transcriptome sequencing in 46 well-characterized T-PLL, we identified a set of 34 differentially deregulated miRs that resemble those of T-cell receptor activated T-cells and whose targets implicate operative regulatory networks that affect DNA-damage response as well as pro-proliferative and cell survival signaling. Specific miR species discriminate T-PLL patient subsets and allow prognostication in an overall survival score for T-PLL (miROS-T-PLL).

**Word Counts:** Abstract 247; Main text 4726

6 Figures; 2 Tables; 11 Supplementary Figures; 14 Supplementary Tables
Abstract:

T-cell prolymphocytic leukemia (T-PLL) is a poor-prognostic mature T-cell malignancy. It typically presents with exponentially rising lymphocyte counts, splenomegaly, and bone marrow infiltration. Effective treatment options are scarce and a better understanding of T-PLL's pathogenesis is desirable. Activation of the $TCL1$ proto-oncogene and loss-of-function perturbations of the tumor suppressor $ATM$ are T-PLL's genomic hallmarks. The leukemic cell reveals a phenotype of active T-cell receptor (TCR) signaling and aberrant DNA-damage responses. Regulatory networks based on the profile of micro-RNAs (miRs) have not been described for T-PLL. In a combined approach of small-RNA and transcriptome sequencing in 46 clinically and molecularly well-characterized T-PLL, we identified a global T-PLL-specific miR expression profile that involves 34 significantly deregulated miR species. This pattern strikingly resembled miR-ome signatures of TCR-activated T-cells. By integrating these T-PLL miR profiles with transcriptome data, we uncovered regulatory networks associated with cell survival signaling and DNA-damage response pathways. Despite a miR-ome that discerned leukemic from normal T-cells, there were also robust subsets of T-PLL defined by a small set of specific miRs. Most prominently, miR-141 and the miR-200c-cluster separated cases into two major subgroups. Furthermore, increased expression of miR-223-3p as well as reduced expression of miR-21 and the miR-29 cluster were associated with more activated T-cell phenotypes and more aggressive disease presentations. Based on the implicated pathobiological role of these miR deregulations, targeting strategies around their effectors appear worth pursuing. We also established a combinatorial miR-based overall survival score for T-PLL (miROS-T-PLL), that might improve current clinical stratifications.
Introduction

T-cell prolymphocytic leukemia (T-PLL) is a neoplasm of post-thymic T-cells. It represents the most frequent mature T-cell leukemia in Western countries, however, with an incidence of ≈2/million/year, it is still classified as an orphan disease. T-PLL patients typically present at a median age of ≈65 years, with exponentially rising lymphocytosis, marked bone marrow infiltration, and splenomegaly. Phenotypically, T-PLL cells resemble mature, antigen-experienced T-cells. The aggressive growth of T-PLL cells is paralleled by a refractory behavior towards most conventional chemotherapies. The current treatment of choice, the CD52-antibody alemtuzumab, is efficient in inducing initial responses, but nearly all patients relapse within 12-24 months thereafter. Second-line options are even less efficient and the median overall survival (OS) of T-PLL patients is <2 years.

Activating translocation of the *T-cell leukemia/lymphoma 1 (TCL1)* proto-oncogene is the most prevalent genetic aberration in T-PLL. In addition, loss-of-function perturbations of the tumor suppressor *ataxia telangiectasia mutated (ATM)* are reported for >80% of T-PLL patients. Both alterations contribute towards a phenotype of enhanced T-cell receptor signaling (TCR) and an aberrant DNA-damage response, including resistance to p53-mediated cell death, deregulated cell cycle control, and deficient DNA-repair mechanisms. Activating lesions in janus kinase (JAK) and signal transducer and activator of transcription (STAT) molecules as well as epigenetic aberrations have emerged as further hallmarks of T-PLL pathology, resulting in a sustained survival signaling and pro-oncogenic cell cycle deregulation. Despite these recent advances, a better understanding of T-PLL pathobiology is of importance in order to identify novel treatment options.

MicroRNAs (miRs) have increasingly been recognized as relevant in the pathogenesis of hematopoietic and solid tumors. They are small non-coding RNAs with an average length of 22 nucleotides. By targeting specific mRNAs, miRs function as post-transcriptional repressors. Importantly, most miRs regulate a large set of genes, often resulting in a cooperative effect on a given cellular pathway, rather than a specific effect on a single gene. Both onco-miRs and tumor suppressive miRs have been causally implicated in mature B- and T-lymphoid malignancies. As a prominent example, chronic lymphocytic leukemia
(CLL) harbors an unique miR expression signature with miR-181b downregulation as the best investigated miR deregulation.\textsuperscript{20} When overexpressing miR-181b in the \textit{Eµ-TCL1A} CLL mouse model, leukemic expansion is decelerated.\textsuperscript{21} Moreover, miR-181b as well as the miRs-29 and -34b/c were shown to target the proto-oncogene \textit{TCL1A}, reflected by an association of their downregulation with oncogenic TCL1A overexpression in CLL.\textsuperscript{22} Likewise, specific miRs have been identified to be involved in the pathogenesis of mature T-cell tumors such as cutaneous T-cell lymphoma (CTCL; e.g. deregulation of miR-29 and -200)\textsuperscript{17,23} or NK/T-cell lymphoma (downregulated miR-150).\textsuperscript{24} In T-PLL, frequent genomic aberrations of \textit{argonaute RISC catalytic component 2 (AGO2)}, a master regulator of miR processing,\textsuperscript{25} provide first hints for altered miR activity and miR expression signatures (miR-omes).\textsuperscript{1} However, global miR deregulations, likely involved in T-PLL’s pathophysiology, have not been reported.

In the presented study, we performed small-RNA sequencing to investigate the spectrum of differential miR expression in T-PLL. We identify global, T-PLL-specific miR alterations associated with gene signatures affiliated to functional categories of survival signaling and DNA-damage response pathways. In addition, we show that the miR-omes of T-PLL cells and of activated T-cells are remarkably similar. Finally, we identify associations of miR alterations with cellular activation, clinical tumor burden, and patient outcome, all underlining the impact of miR deregulations in T-PLL.
Methods

Patient cohort
Primary isolates of 48 well-annotated T-PLL patients and of T-cells from 6 age-matched healthy donors were studied (banked 2009-2019; patient characteristics in Table 1). The diagnosis of T-PLL was confirmed according to WHO criteria and consensus guidelines. All patients (median age: 68 years) provided informed consent according to the Declaration of Helsinki. Collection and use of the samples have been approved for research purposes by the ethics committee of the University Hospital of Cologne (#11–319). Most samples (82.6%) were collected prior to any first-line treatment (n=38/46).

Sequencing and data processing
RNA from peripheral blood mononuclear cells (PBMCs) of T-PLL patients (median purity 95.4%) and CD3+ pan-T-cells of 6 age-matched healthy controls (median purity 90.2%) was subjected to library preparation and sequenced on the NovaSeq 6000 (n=48 T-PLL) and the HiSeq4000 platform (n=46 T-PLL, Illumina, San Diego, USA) according to manufacturer’s instructions for polyA-RNA and small-RNA sequencing, respectively. Details on cell isolation, stimulation, RNA isolation, library preparation, sequencing, and data processing are given in the Online Supplementary Methods.

Gene Set Enrichment Analysis (GSEA)
GSEAs were performed on pre-ranked lists using the GSEA-software (v3.0) and MSigDB (v7.0) HALLMARK gene sets. For each considered miR, Spearman correlation coefficients for this miR and all protein-coding genes were determined by comparing the respective gene’s Count Per Million (CPM) and Fragments Per Kilobase of Million mapped reads (FPKM). Sorted lists of correlation coefficients were then used as input for the GSEA.

MiR target prediction
To obtain putative mRNA targets for each miR, predicted miR bindings were first determined using the R-package multiMiR (v1.6.0, Database Version 2.3.0), all of 8 prediction databases (diana_microt, elmmo, microcosm, miranda, mirdb, pictar, pita, targetscan), and a 20% default prediction cutoff. All bindings predicted by <2 different databases were removed. From the remaining predicted genes, we chose those as putative miR targets that showed a negative Spearman correlation (rho<0; p<0.05; false discovery rate (FDR)<0.25) of their expression values with the expression of the respective miRs.
**Correlations with clinical data**

To test for associations of miR-223-3p, miR-21, miR-29, and miR-200c/141 with clinical characteristics, cytogenetics, immunophenotypes, and outcome data, cases were divided into groups by the mean or tertiles as cutoffs according to the distribution of expression values within the patient cohort. Further details on statistics are provided in the Online Supplementary Methods.

**Survival score**

To develop a survival score, we (i) randomly divided our cohort into a training set (n=22) and a validation set (n=22). We then (ii) identified miRs that were expressed in at least 80% of T-PLL samples and (iii) that were highly associated with OS in the training set (upper tertile of patients with highest vs. tertile with the lowest expression) using log-rank tests. To the resulting 4 miRs with the strongest single OS associations (lowest p-values) we added several other parameters that had been described to be prognostically relevant in T-PLL (e.g. leukocyte counts, TCL1 mRNA expression)¹,¹¹,³⁰ towards a multi-parameter training model. Next, (iv) we subjected all parameters to a recursive partitioning algorithm using the rpart R package (v.4.1-15) to identify optimum individual cutoffs in our training set. Such deprivatizations finally retained the 4 factors miR-200a-3p, miR-223-3p, miR-424-5p, and TCL1A, for which optimum cutoffs best allowed discriminations of T-PLL patients with shorter vs. longer OS. (v) We then built multivariate scores for all possible combinations of these 4 parameters, adding 1 point to the total score if the respective expression cutoff was passed, and calculated optimum thresholds for these scores. (vi) We selected the score which allowed best discrimination of OS in our training set and (vii) verified our score in the validation set as well as in the total cohort of 44 T-PLL patients.
Results

Global T-PLL-specific miR deregulations highlight differential expression of miR-200c and miR-141 clusters.

To investigate the spectrum of cellular miRs expressed in T-PLL, small-RNA sequencing was performed of peripheral blood (PB)-isolated tumor cells from 46 T-PLL patients and of pan-T-cells of PB from 6 healthy donors. T-PLL patient characteristics are presented in Table 1; sample purities in Online Supplementary Figure S1. As T-PLL cases show a spectrum of (often non-conventional) memory T-cell phenotypes and of small naïve sub-sets,5 we choose age-matched CD3+ pan-T-cells as controls (reflecting a representative mix of populations) in these global profiling analyses.

In total, we identified 2094 miRs, of which 37 miRs displayed a differential expression in T-PLL vs. healthy-donor T-cells (q<0.05, Online Supplementary Table S1). Of these, 14 miR sequences were upregulated (0.7% of all identified miRs) and 23 were downregulated in T-PLL (1.1%, Figure 1A). While miR-6724-5p (fold-change (fc)=0.18, q<0.0001) and miR-206 (fc=0.04, q<0.0001) showed the strongest downregulation, miR-5699-3p (fc=122, q=0.02), miR-200c-3p (fc=38.2, q=0.005), and miR-141-3p (fc=43.2, q=0.005) were the most upregulated. Considering all T-PLL cases, miR-141-3p and miR-21-5p showed the highest absolute abundance while miR-206, miR-651-3p and miR-6774-5p displayed the lowest absolute expression (Online Supplementary Figure S2). Among the 37 deregulated miRs, miR-6724-5p was annotated 4 times due to its expression from 4 different genomic loci. The following analyses are, therefore, based on 34 miRs, containing a sum expression for miR-6724-5p. To validate our results from small-RNA sequencing, we performed qRT-PCR analyses of three highly deregulated miRs (miR-223-3p, miR-200c-3p, miR-141-3p, Online Supplementary Figure S3A-C) in 8 T-PLL and 4 healthy-donor T-cell controls. A strong correlation between the results from small-RNA sequencing and qRT-PCR (r²=0.84, p<0.0001, Pearson, Online Supplementary Figure S3D) confirmed differential (over)expression of these miRs in T-PLL and underlines the robustness of the sequencing data. Using a previously published independent data set from single nucleotide polymorphism (SNP) arrays of 83 T-PLL1 (overlap of 23 cases with the cohort presented
here), we identified only small fractions of cases to carry genomic losses of significantly
downregulated miRs: miR-140-3p, miR-196b-5p (both CN<1.5 in 4.82% of cases),
miR-339-3p, and miR-589-5p (both CN<1.5 in 8.43%, Online Supplementary Table S2).

Unsupervised clustering by Principal Component Analysis (PCA) based on miRs differenti-
tially expressed in T-PLL vs. healthy-donor T-cells indicated a homogeneity across T-PLL
samples and confirmed the global differences in the miR profiles between T-PLL and
healthy controls (Figure 1B). Interestingly, unsupervised hierarchical clustering analysis of
miR expression revealed two clusters of T-PLL cases, which were distinguished by ex-
pression of miR-200c and miR-141 family members (Figure 1C): 23 T-PLL cases showed
low miR-141/-200c expression as compared to CD3+ pan-T-cell controls whereas 23 T-PLL
samples had higher-than-normal-T-cell expression of miR-141 and -200c. Besides higher
serum lactate dehydrogenase (LDH) levels (p=0.03, Mann-Whitney-Wilcoxon test (MWW))
and a lower incidence of TP53 deletions (by FISH, p=0.02, Fisher’s exact test) in the
miR-141/-200c high expressing cohort, we did not find other differences between these two
subsets (Online Supplementary Table S3). Comparing the transcriptomes of cases allo-
cated to these two separate clusters, we identified 356 genes to be differentially expressed
(Online Supplementary Table S4). In line with GSEA based on miR141/200c correlated
genes (see following analyses), we identified the HALLMARK pathways E2F TARGETS (nor-
malized enrichment score (NES)=4.38, q<0.0001) and G2M CHECKPOINT (NES=2.68,
q<0.0001) as significantly altered between the two clusters of T-PLL. Furthermore, global
miR-ome profiles were not associated with distinct cellular immunophenotypes, e.g. neither
with CD45RA/RO expression (i.e. “memory-like” vs. “naïve-like” T-PLL) nor with CD4/8
expression (Online Supplementary Figures S4A, B).

**MiR profiles of T-PLL resemble those of TCR-activated T-cells and form regulatory
networks around nodes of DNA-damage response and pro-survival signaling**

To align the miR-ome data with those of global transcriptome alterations, polyA-RNA se-
quencing was performed on PB-isolated tumor cells from 41 miR-characterized T-PLL pa-
tients and 7 additional T-PLL patients as well as on CD3+ PB pan-T-cells from 6
age-matched healthy donors. In total, we detected 948 protein-encoding mRNAs to be differentially expressed (q<0.05, *Online Supplementary Figure S5A* and *Table S5*). Using this set of deregulated genes, PCA corroborated homogeneity among T-PLL cases and a clear distinction to normal T-cell controls (*Online Supplementary Figure S5B*). In accordance with published data, *TCL1A* (fc=1843, q<0.0001) and *CTLA4* (fc=0.06, q<0.0001) were among the most differentially expressed genes in T-PLL vs. T-cell controls (*Online Supplementary Figure S5C*). In their transcriptome profiles two T-PLL clustered closer to control T-cells than the bulk of cases. However, their miR expression signature (Figure 1C) and clinical presentation did not differ from the overall cohort.

GSEA (*HALLMARK* gene sets²⁷) determined 34 gene sets as upregulated in T-PLL when compared to T-cell controls, of which 19 gene sets were significantly enriched at a FDR of <5%. 16 gene sets were downregulated in T-PLL (11 with FDR <0.05, *Online Supplementary Table S6*). The identified significantly altered pathways associated with cancer and/or immunology are presented in Figure 2A. These included several *HALLMARK* gene sets reflecting dysregulations in DNA-damage response pathways (e.g. DNA_REPAIR normalized enrichment score (NES)=-2.18, q=0.005; *E2F_TAGETS*, NES=2.05, q=0.01) and pro-survival signaling (e.g. *INFLAMMATORY RESPONSE*, NES=3.12, q<0.0001; *TNFA_SIGNALING_VIA_NFKB*, NES=2.44, q=0.001), in line with previously published data.¹

As T-PLL cells generally display a mature, T-cell activated phenotype,¹¹ we investigated whether T-PLL cell miR-omes resemble those of TCR-activated healthy-donor derived T-cells. For that, PBMCs (to avoid direct manipulation of T-cells) of 4 healthy donors were cultured for 72 hours with and without stimulation by anti-CD3/CD28 crosslinking, followed by miR sequencing of CD3⁺ enriched cells. Sample purities and experimental controls are shown in *Online Supplementary Figure S6*. We identified 56 miRs to be differentially expressed in response to TCR activation (q<0.05, *Online Supplementary Figure S7A* and *Table S7*). PCA indicated homogeneity within both groups (T-PLL and normal PBMCs) as well as global differences between their TCR-induced miR profiles (*Online Supplementary Figure S7B*). We identified miRs known to be affected by TCR activation (e.g. *miR-150-5p*)³¹ as well as previously unreported miRs (e.g. *miR-18a-5p*; *Online Supplementary Figure S7C*).
Integrative PCA based on differentially expressed miRs in T-PLL vs. healthy controls (Figure 1C) showed that the stimulated (over unstimulated) T-cells clustered closer to T-PLL (Figure 2B). Fittingly, unsupervised clustering comparing TCR-stimulated T-cells to unstimulated controls confirmed that the miR profiles of T-PLL cells resemble the miR-ome of TCR-activated healthy-donor derived T-cells (Figure 2C).

We next assessed implicated functional relationships by predicted mRNA targets for each deregulated miR in T-PLL. For that, we (i) ranked mRNAs based on their degree of correlation with a specific miR and (ii) performed Hallmark set GSEA on these ranked mRNAs. Pathways reflecting dysregulations of DNA-damage response and pro-survival signaling emerged as predominantly associated with the alterations of miR expression. Exemplary Hallmark pathways are shown in Figure 3A, a full list of gene sets is displayed in Online Supplementary Figure S8. For example, we obtained highly significant NES for the E2F_TARGET and the IL2_STAT5_SIGNALING Hallmark gene sets for (i) the transcriptome of T-PLL as compared to the one of healthy controls (E2F_TARGETS: NES=1.92, q=0.02; IL2_STAT5_SIGNALING: NES=-2.86, q<0.0001, Online Supplementary Table S6) and (ii) for most of the miRs differentially expressed in T-PLL (Figure 3A).

Overall, there was a striking similarity of the miR profiles of T-PLL cells with those of TCR-activated T-cells. By integrating T-PLL’s miR profiles with transcriptome data via GSEA (based on differentially expressed mRNAs and on mRNAs ordered by their correlation to the respective miR), we uncovered prominent regulatory networks around DNA-damage response and pro-survival pathways in T-PLL.

Increased miR-223-3p expression is linked to activated T-cell phenotypes and associates with signatures of altered DNA-damage response and cell-cycle deregulation.

We next focused on phenotypic and clinical associations of miRs (i) differentially expressed in our cohort and (ii) already linked to B- and / or T-cell leukemogenesis. In our sequencing analysis of small-RNAs, miR-223-3p was significantly upregulated in T-PLL over CD3+ pan-T-cells from age-matched healthy donors (fc=9.85, p=0.0002, Figure 3B). GSEA based on mRNAs ranked by their correlation to miR-223-3p expression revealed an association of
miR-223-3p with signatures of altered DNA-damage responses (e.g. **P53_PATHWAY**: NES=2.59, q<0.0001) and deregulated cell-cycle mediators (e.g. **G2M_CHECKPOINT**: NES=-2.34, q=0.003, Figure 3C). We further evaluated potential miR-223-3p target mRNAs, by combining target prediction databases and miR-ome – transcriptome correlations. Defining criteria of mRNAs as putative miR targets in T-PLL are outlined in the Methods section. In total, we identified 8 putative targets of miR-223-3p in T-PLL (Figure 3D). Notably, FOXO1 (rho=-0.41, p=0.005) was identified as one of them and was significantly downregulated in T-PLL (n=32/48 cases with fc<0.5 compared to normal T-cell controls). Tumor suppressive FOXO1 is a prominent regulator of redox balances and DNA-insult mediated cell death.32

When T-PLL cases were divided into three subgroups based on miR-223-3p expression (high, medium, low), we found that miR-223-3p expression correlated significantly with surface expression levels of the TCR-activation markers CD38 (mean surface expression: 65.5% vs. 3.6%, p=0.006, MWW) and CD69 (mean: 8.5% vs. 1.0%, p=0.1, MWW, Figure 3E, **Online Supplementary Table S8** with summary of clinical data).

**Increased levels of miR-200c and miR-141 species are associated with deregulated cell-cycle molecules, activated phenotypes, and more aggressive presentations.**

Another miR family, miR-200c/-141, was significantly upregulated in a subset of 23 T-PLL cases (Figure 4A, 1C). Upregulation for miR-141-3p was 43.2-fold (p<0.0001), for miR-141-5p 29.0-fold (p=0.0001), for miR-200c-3p 38.2-fold (p<0.0001), and for miR-200c-5p 56.6-fold (p=0.003) over all cases. MiR-141-3p showed the highest absolute CPM values among all deregulated miRs in the entire cohort of T-PLL (mean CPM^miR-141-3p^ =26561; mean CPM of all significantly deregulated miRs in T-PLL =1440, fc=18.4; **Online Supplementary Figure S2**). GSEA based on mRNAs ranked by their correlation to miR-200c/-141 family members revealed significant enrichments of the **HALLMARK E2F_TARGET** (NES=3.64, q<0.0001) and **HALLMARK G2M_CHECKPOINT** gene sets (NES=3.05, q<0.0001, both based on miR-141-3p correlated mRNAs, Figure 4B). Additionally, we identified 93 mRNAs as potential targets of miR-200c/-141 (Figure 4C) in
T-PLL. While 9 mRNAs showed an overlap between miR-141-3p and miR-200c-3p target mRNAs, we found only a small set of putative targets which were shared between either miR-141-3p or miR-200c-3p and miR-141-5p. Exemplarily, KAT2B, a known tumor suppressor affecting DNA damage and cell cycle regulation, emerged as a potential target of miR-200c-3p (rho=-0.41, p=0.005, Spearman). Surface expression of the T-cell activation marker CD40L was elevated in cases with high miR-141-3p (mean 16.5% vs. 0.05%, p=0.03, MWW), high miR-200c-3p (mean 16.5% vs. 0.05%, p=0.03, MWW), and high miR-200c-5p expression (mean 20.0% vs. 0.0%, p=0.009, MWW, Figure 4D). Serum LDH levels at the time of sample correlated with increased miR-141-3p (mean 898 U/l vs. 509 U/l, p=0.03, MWW, Figure 4E) and elevated miR-200c-3p expression (mean 917 U/l vs. 509 U/l, p=0.02, MWW, Online Supplementary Table S9 with summary of clinical data).

Reduced miR-21 expression is linked to features of advanced or aggressive disease.
The small-RNA sequencing analysis also revealed a 3.7-fold reduction of miR-21-3p expression (fc=0.27, p<0.0001) and a 3.2-fold reduction of miR-21-5p expression (fc=0.31, p<0.0001) in T-PLL (Online Supplementary Figure S9A). Interestingly, absolute expression (CPM) values of miR-21-5p were the second most altered among all deregulated miRs in T-PLL (mean CPM of all significantly deregulated miRs in T-PLL of 1440 vs. mean CPMmiR-21-5p of 15526, fc=10.8, Online Supplementary Figure S2), suggesting a highly T-PLL-specific loss of expression. GSEA of miR-21 associated mRNAs implicated relevance of this miR in apoptosis and cell-cycle regulation, as gene sets like HALLMARK P53_PATHWAY and G2M_CHECKPOINT were significantly deregulated (NES of APOPTOSIS_PATHWAY considering miR-21-5p associated mRNAs=3.89, q<0.0001; NES of G2M_CHECKPOINT considering miR-21-3p associated mRNAs=2.39, q=0.001, Online Supplementary Figure S9B). Furthermore, we assessed potential target mRNAs of miR-21-3p and miR-21-5p as described above (n=42, e.g. MAP3K1 as a putative target of both miR-21-3p and miR-21-5p, Online Supplementary Figure S9C). In contrast to the current concept of miR-21 being a potent suppressor of cell cycle arrest and apoptosis induction, we did not find negative correlations with mRNAs mediating these published effects (e.g.
PDCD4, rho=0.04, p=0.79; BTG2, rho=0.28, p=0.05; correlations based on miR-21-5p expression, Spearman).\textsuperscript{34} Dichotomized by mean miR-21-5p expression, T-PLL with low miR-21-5p levels revealed higher WBC counts (mean: 154 G/l vs. 90.0 G/l, p=0.02, MWW) and lower platelet counts (mean: 110 G/l vs. 153 G/l, p=0.03, MWW, Online Supplementary Figure S9D) at the time of sampling, indicating a more active growth behavior of T-PLL with low miR-21 expression. Fittingly, serum levels of LDH (mean: 933 U/l vs. 522 U/l, p=0.02, MWW, Online Supplementary Figure S9E) were elevated in patients with low cellular miR-21-5p expression (Online Supplementary Table S10 with summary of clinical data).

**Reduced expression of miR-29 clusters is associated with alterations of survival signaling and cell-cycle regulators reflected in features of a more active disease.**

As analyzed by small-RNA sequencing, the miR-29 family members miR-29a-3p (fc=0.29, p<0.0001), miR-29b-1-5p (fc=0.47, p=0.001), and miR-29c-3p (fc=0.29, p<0.0001) showed a homogenous downregulation in T-PLL over healthy-donor T-cells (Online Supplementary Figure S10A). GSEA of associated mRNAs showed an enrichment of genes of the HALLMARK E2F_TARGET (NES=-3.52, q<0.0001) and of the TNFA_SIGNALING_VIA_NFKB gene sets (NES=-2.08, q=0.01, both miR-29a-3p correlated mRNAs) in low miR-29 expressing T-PLL (Online Supplementary Figure S10B). Furthermore, we identified 79 putative target mRNAs of the miR-29a-3p / miR-29b-1-5p / miR-29c-3p cluster in T-PLL by the above described strategy (Online Supplementary Figure S10C). Indicating an association of reduced miR-29 species with aberrant survival signaling, surface TCR activation markers were elevated in T-PLL with low miR-29a-3p (Online Supplementary Figure S10D), namely CD38 (mean expression: 43.1% vs. 12.9%, p=0.02, MWW) and CD69 (mean: 27.9% vs. 0.89%, p=0.02, MWW). A more active disease state as indicated by lower platelet counts (mean: 110 G/l vs. 186 G/l, p=0.15, MWW) and higher LDH serum levels (mean: 1850 U/l vs. 708 U/l, p=0.06, MWW) at sampling was linked to low miR-29b-1-5p expression (Online Supplementary Figure S10E, F). Lower miR-29c expression tended to be associated with a higher incidence of effusions (6/12 over 1/10 with high miR-29c-3p-expression; p=0.07, Fisher’s exact test, Online Supplementary Figure S11A). In addition, genomic ATM dele-
tions were more frequent in cases with low miR-29b-1-5p expression (9/11 vs. high 1/11 with miR-29b-1-5p; p=0.002, Fisher’s exact test, Online Supplementary Figure S11B, Online Supplementary Table S11 with summary of clinical data).

**A combinatorial miR-based overall survival score for T-PLL (miROS-T-PLL).**

Based on the observed correlation of miR expression with clinical parameters, we aimed to establish a prognostic score that stratifies T-PLL patients according to miR expression levels. To identify best candidates in an unbiased fashion, we first associated miR expression levels with OS for all miRs detected in at least 80% of T-PLL samples (n>36 cases) and compared T-PLL patients with highest expression levels (upper tertile) to those with lowest expression of the respective miR (lower tertile; Online Supplementary Table S12 with summary of clinical data). In this analysis, miR-98-3p (median OS in high vs. low expression: 16.3 months vs. 29.4 months, p=0.0008, log-rank, Figure 5A), miR-200a-3p (52.7 months vs. 19.1 months, p=0.001, log-rank, Figure 5B), miR-223-3p (14.9 months vs. 26.0 months, p=0.001, log-rank, Figure 5C), and miR-424-5p (14.4 months vs. 26.0 months, p=0.0007, log-rank, Figure 5D) were most significantly correlated with OS. We subjected a training model composed of these 4 miRs (miR-98-3p, miR-200a-3p, miR-223-3p, miR-424-5p) and of selected factors that had shown to be of prognostic relevance in T-PLL (e.g. WBC counts, TCL1A mRNA level)\textsuperscript{1,11,30} to parameter shaving by recursive partitioning. This algorithm identified optimum individual thresholds stratifying OS in the randomly created 22-case training set. Using these cutoffs, three miRs (200a-3p, miR-223-3p, miR-424-5p) and TCL1A expression remained as most significant discriminators for OS. Multiple combinatorial scores of these 4 parameters were built and for these scores optimum thresholds for discrimination of OS were calculated (recursive partitioning, see Methods). Best separation was obtained using a miR-exclusive 3-tier score: miR-200a-3p \( fc<2.21 \), miR-223-3p \( fc\geq9.8 \), miR-424-5p \( fc\geq0.91 \); relative to healthy-donor T-cells) with a cutoff of \( \geq2 \) sum points (Table 2). Finally, we verified the miROS-T-PLL score in the 22-case validation set (p=0.0004, log-rank) and in the total cohort of 44 T-PLL (median OS high vs. low miROS-T-PLL: 14.4 months vs. 29.4 months, p<0.0001, log-rank, Figure 5E).
To identify variables underlying (as potential confounders) the miR-based prognostic separation, we associated the expression of miRs, which we used for the score, as well as the miROS-T-PLL score itself, with genomic, mRNA expression, immunophenotypic, and clinical data (Online Supplementary Tables S13 and S14). We did not detect significant differences in the distribution of these parameters between the groups determined by expression of the three miRs or by the miROS-T-PLL score, further validating the newly established score.
Discussion
Here, we report a pilot analysis of cellular miR expression in a cohort of 46 T-PLL patients. We identified 34 miRs to be significantly deregulated in comparison to PB-derived T-cells from age-matched healthy donors. These miRs included those which had already been reported as altered in T-cell acute lymphoblastic leukaemia (T-ALL, e.g. miR-223-3p)\textsuperscript{35} and in CTCL (e.g. miR-29 and -200).\textsuperscript{17,23} They also contained miRs that had not been described in the neoplastic context (e.g. miR-10395-5p). The global profiles of deregulated miRs in T-PLL showed a rather uniform pattern across the analyzed cases. Together with the integrated information from transcriptome sequencing, this set of data allows for the first time insights into miR-based regulatory networks in T-PLL.

It is important to mention, that 4 out of the 34 differentially expressed miRs presented with low CPM values (<1), either being unspecific background in the sequencing technology or representing biological relevant miRs expressed at low levels. In addition, for two of the small-RNAs identified to be differentially expressed in T-PLL (miR-6724-5p, miR-5699-3p), miR-base\textsuperscript{36} assigned questionable confidence in their annotation, although their expression was previously reported in other entities (e.g. bladder cancer).\textsuperscript{37}

In line with T-PLL’s phenotype of augmented TCR activation,\textsuperscript{1,5,11} our comparative profiling revealed a resemblance of T-PLL’s miR-ome to the one of TCR-activated healthy-donor derived T-cells. A specific remodeling of the miR repertoire upon TCR activation had been shown,\textsuperscript{38} however, analysis of full-spectrum miR expression by small RNA sequencing upon TCR activation in healthy-donor derived pan-T-cells had not been reported before. We identified here previously unknown miRs (e.g. upregulation of miR-18a-5p) to be altered upon TCR activation in addition to those that had been described (miR-17-5p ormiR-150-5p).\textsuperscript{31,39} We conclude that constitutive TCR activation shapes the characteristic miR-ome of T-PLL cells.

As hallmarks of T-PLL’s miR-ome, we identified miR-223-3p, miR-21, the miR-29 family, and the miR-200c/-141 cluster as significantly deregulated. These miRs have previously emerged as either onco-miRs or tumor suppressive miRs in other T- or B-cell malignancies.\textsuperscript{23,40–43} We further identified putative target signatures, potentially mediating the postu-
lated effects of pro-survival signaling and aberrant DNA-damage responses (e.g. downregulation of FOXO1 upon miR-223 upregulation). Limiting, the postulated target genes are, although predicted through multiple robust algorithms, based on associations without proven biological effects, which has to be addressed in future experiments. Deregulation of these hallmark miRs in T-PLL was further associated with either a pronounced cellular activation phenotype or more aggressive clinical presentations.

MiR-21 stood out as one of the most abundant miRs in T-PLL. In contrast to the current concept of miR-21 being an onco-miR, we detected significantly downregulated miR-21 levels in T-PLL samples as compared to healthy-donor T-cells. Our integrative correlations further revealed an association of low miR-21 expression with more aggressive disease presentations. We identified SKP2 and MAP3K1 as predicted targets of miR-21 in T-PLL, potentially mediating these features. Notably, we did not find significant negative correlations of miR-21 with those mRNAs that were previously described to mediate the effect of this miR as a potent suppressor of cell cycle inhibition and apoptosis. This indicates a T-cell specific and context-dependent function of miR-21.

Our computational analyses of mRNAs targeted by deregulated miRs in T-PLL suggest a strong impact of altered miR clusters on activation, death resistance, and aberrant DNA-damage responses. Abnormal activity of these pathways, triggered by TCL1A over-expression and damaging ATM aberrations, has emerged as a hallmark of T-PLL pathobiology. We propose that pro-tumorigenic miR networks in their function as post-transcriptional regulators may further enhance the effects of these key genomic lesions, contributing substantially to the pathogenesis of T-PLL. Similar cooperative miR-mRNA networks were postulated for T-ALL and CLL. The causes of the miR deregulations we observed here, remain unknown and are, besides TCR-activation, likely multifactorial. As T-PLL is characterized by a strong genomic instability and high burdens of reactive oxygen species (ROS), mutations or copy number alterations of miR-encoding genes provide possible explanations, in addition to epigenetic mechanisms. Notably, incidences of genomic losses of the downregulated miR-140-3p, miR-196b-5p, miR-339-3p, and miR-589-5p were in the order of 4-9% of T-PLL cases in our cohort.
In summary, we identified a T-PLL-specific miR-ome, with 34 differentially deregulated miRs, that appears instructed by TCR activation. By integrating altered miR expression with the information derived from transcriptome analyses, we postulate that the miR-ome of T-PLL shapes (dys)regulated networks towards apoptotic resistance, cell-cycle abrogation, and defective DNA-damage repair (Figure 6). Highlighting the pathobiological impact of the discovered miR deregulations, we developed the first clinical survival score that predicts T-PLL patients' outcomes, based on expression levels miR-200a-3p, miR-223-3p, and miR-424-5p.

Clinical management of T-PLL remains challenging and current studies mainly focus on the development of targeted treatments directed against anti-apoptotic factors like BLC2\textsuperscript{47} or sustained pro-survival signaling mediated via JAK/STAT pathways.\textsuperscript{48,49} Our study suggests that developing a miR-directed targeting strategy could allow tackling a combination of those pathways and might therefore be a promising approach to successfully eradicate T-PLL cells. However, these concepts have to be tested in subsequent pre-clinical studies. Furthermore, the presented miROS-T-PLL survival score might help in better discriminating rather indolent vs. aggressive T-PLL disease phases at the time of diagnosis.
Acknowledgments
This research was funded by the DFG Research Unit FOR1961 (Control-T; HE3553/4-2), the Köln Fortune program, and the Fritz Thyssen Foundation (10.15.2.034MN). This work was also funded by the EU Transcan-2 consortium ‘ERANET-PLL’ and by the ERAPerMed consortium ‘JAKSTAT-TARGET’. A.S. was supported by a scholarship of the German José Carreras Leukemia Foundation (DJCLS 03 F/2016). We gratefully acknowledge patients with their families for their invaluable contributions.

Conflict of Interests
The authors declare no competing financial interests.
References


**Table 1: Clinical, cytogenetic and immunophenotypic characteristics of analyzed T-PLL (n = 46 cases analyzed by small-RNA sequencing)**

<table>
<thead>
<tr>
<th>Patients’ characteristics</th>
<th>At diagnosis</th>
<th>At sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, years (range)</td>
<td>68 (32-88); n = 46</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Male = 25; Female = 21</td>
<td></td>
</tr>
<tr>
<td>Median OS from diagnosis, months (range)</td>
<td>17.1 (0.4-98.7); n = 44</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical presentation</th>
<th>At diagnosis</th>
<th>At sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median WBC count, x10⁹/L (range)</td>
<td>91.4 (16.7-825.2); n = 38</td>
<td>114.1 (20.8-756.2); n = 43</td>
</tr>
<tr>
<td>Median hemoglobin, g/dL (range)</td>
<td>12.6 (6.6-16.6); n = 34</td>
<td>12.6 (6.2-15.6); n = 36</td>
</tr>
<tr>
<td>Median platelet count, x10⁹/L (range)</td>
<td>129 (33-438); n = 34</td>
<td>118 (48-394); n = 36</td>
</tr>
<tr>
<td>Median LDH, U/L (range)</td>
<td>567 (178-9423); n = 29</td>
<td>795 (226-8634); n = 36</td>
</tr>
<tr>
<td>Splenomegaly (%)</td>
<td>n = 18/29 (62.07)</td>
<td></td>
</tr>
<tr>
<td>Hepatomegaly (%)</td>
<td>n = 5/27 (18.52)</td>
<td></td>
</tr>
<tr>
<td>Lymphadenopathy (%)</td>
<td>n = 15/26 (57.69)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cytogenetic features</th>
<th>n =</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>inv(14)(q11;q32) (%)</td>
<td>27/36;</td>
<td>(75.00)</td>
</tr>
<tr>
<td>t(14;14)(q11;q32) (%)</td>
<td>3/35;</td>
<td>(8.57)</td>
</tr>
<tr>
<td>t(X;14)(q28;q11) (%)</td>
<td>2/36;</td>
<td>(5.56)</td>
</tr>
<tr>
<td>TCR gene rearrangement (%)</td>
<td>37/39;</td>
<td>(94.87)</td>
</tr>
<tr>
<td>MYC amplification (%)</td>
<td>23/27;</td>
<td>(85.19)</td>
</tr>
<tr>
<td>ATM deletion (%)</td>
<td>15/33;</td>
<td>(45.45)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Immunophenotype</th>
<th>n =</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCL1 (%)</td>
<td>35/38;</td>
<td>(92.11)</td>
</tr>
<tr>
<td>CD3⁺ (%)</td>
<td>36/40;</td>
<td>(90.00)</td>
</tr>
<tr>
<td>CD5⁺ (%)</td>
<td>40/40;</td>
<td>(100.00)</td>
</tr>
<tr>
<td>CD7⁺ (%)</td>
<td>39/40;</td>
<td>(97.50)</td>
</tr>
<tr>
<td>CD4⁺/CD8⁺ (%)</td>
<td>27/39;</td>
<td>(69.23)</td>
</tr>
<tr>
<td>CD4⁺/CD8⁻ (%)</td>
<td>7/39;</td>
<td>(17.95)</td>
</tr>
<tr>
<td>CD4⁻/CD8⁺ (%)</td>
<td>5/39;</td>
<td>(12.82)</td>
</tr>
</tbody>
</table>

¹ range reaches from lowest value to highest value in the cohort; ² percentages are out of total cases with sufficient data; ³ evaluated by fluorescence in-situ hybridization
Table 2: Prognostic score (miROS-T-PLL) including miR-200a-3p, miR-223-3p, and miR-424-5p expression levels.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0 Points</th>
<th>1 Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-200a-3p expression$^1$ (FC, rel. to healthy CD3+ pan T-cells)</td>
<td>≥ 2.21</td>
<td>&lt; 2.21</td>
</tr>
<tr>
<td>miR-223-3p expression$^1$ (FC, rel. to healthy CD3+ pan T-cells)</td>
<td>&lt; 9.80</td>
<td>≥ 9.80</td>
</tr>
<tr>
<td>miR-424-5p expression$^1$ (FC, rel. to healthy CD3+ pan T-cells)</td>
<td>&lt; 0.91</td>
<td>≥ 0.91</td>
</tr>
</tbody>
</table>

Prognostic Groups

- Lower Risk: 0-1 points
- Higher Risk: 2-3 points

$^1$ evaluated by Small-RNA sequencing and compared to the mean expression of CD3+ pan T-cells of 6 healthy donors
Figure 1: Global miR deregulations highlight differential expression of miR-141 and miR-200c clusters in T-PLL.

a) MicroRNA (MiR) profiles were evaluated using small-RNA sequencing of primary human peripheral-blood derived T-PLL cells (n=46 cases) and CD3+ healthy-donor derived pan-T-cell controls (n=6 donors). Volcano plot showing fold-changes (fc) and false-discovery rates (FDRs) of all identified miRs (n=2094). Differentially expressed miRs (n=37, Online Supplementary Table S1) are highlighted in blue (downregulation) and red (upregulation). Relative proportions are based on all identified miRs.

b) Principle component analysis (PCA) based on miRs differentially expressed comparing T-PLL cases (n=46) to healthy-donor controls (n=6). Separate clustering of cases and controls indicates global differences in miR expression profiles.

c) Heatmaps of differentially expressed miRs (n=34; FDR<0.05) in T-PLL samples vs. CD3+ pan-T-cells of healthy donors. Left 2-column heatmap: mean counts per million (CPM) values compared between healthy-donor controls (n=6) and T-PLL (n=46; red=higher expression; blue=lower expression). Right heatmap: Colors represent z-scores of respective CPM values calculated for each miR (red=higher z-score; blue=lower z-score). Control samples with slightly lower purities (77% and 85%) and T-PLL cases presenting a rather unique transcriptome, which clusters closer to control T-cells (Figure S4C), are indicated by asterixis (brown: T-PLL cases; green: control cases).

Figure 2: Global alterations of T-PLL miR-ome / transcriptome networks resemble those of TCR-activated T-cells.

a) GSEA of differentially expressed mRNAs in T-PLL using HALLMARK gene sets (n=994 genes, Online Supplementary Table S6 and Online Supplementary Figure S5; total n=48 gene sets). Gene expression profiles were assessed using global mRNA sequencing of primary T-PLL cells (n=48 cases) and healthy donor-derived CD3+ pan-T-cells (n=6 donors). Color code represents assignment to dysregulation to either altered DNA-damage response pathways (black) or pro-survival signaling (grey). Exemplary GSEA plots are presented (DNA_REPAIR: Normalized Enrichment Score (NES) =-2.18, q=0.008, E2F_TARGETS: NES=2.05, q=0.01, INFLAMMATORY RESPONSE: NES=3.12, q<0.0001).

b and c) T-PLL miR-omes resembled those of activated healthy-donor T-cells: age-matched healthy-donor derived PBMCs were isolated via density gradient centrifugation. T-cell activation was achieved via antibody-mediated CD3/CD28 crosslinking. After 72 hours (hrs.), CD3+ primary T-cells were isolated by magnetic-activated cell sorting (MACS) (negative selection; see Method section for details; cell purities are given in Online Supplementary Figure S6A, see Online Supplementary Figure S6B and C for control experiments on stimulation). MiR
profiles were generated using small-RNA sequencing. Unstimulated cultured controls clustered together with directly isolated control samples, indicating that there was a negligible cell culture effect on miR expression profiles. Color code: T-PLL in brown, controls in green colors (light green: CD3⁺ T-cells cultured in vitro for 72 hrs. without stimulation; green: CD3⁺ T-cells submitted to miR-ome sequencing directly after enrichment; dark-green: CD3⁺ T-cell cultures at 72 hrs. subsequent to T-cell receptor (TCR) activation).

b) PCA based on differentially expressed miRs comparing T-PLL to healthy-donor derived T-cells (n=37 miRs, Online Supplementary Table S1). TCR-activated healthy-donor derived T-cells clustered with T-PLL cases.

c) Heatmap and clustering based on differentially expressed miRs comparing healthy-donor derived T-cells: unstimulated controls vs. TCR-activated condition (n=56 miRs, FDR<0.05). Colors represent z-scores of CPM values calculated for each miR (blue=lower z-score; red=higher z-score).

Figure 3: Functional affiliations of predicted miR-targets reflect associations with processes of altered DNA-damage response and pro-survival signaling.

a) GSEA heatmap based on miR associated mRNAs. GSEAs were conducted for all significantly deregulated miRs (n=34 miRs; Online Supplementary Table S1) using ranked correlation indices between mRNA and miR expression in 41 T-PLL and 6 healthy-donor derived T-cell samples. Exemplary HALLMARK pathways are displayed. Full figure is displayed in Online Supplementary Figure S8. Color code summarizes NES scores (blue=negative NES; red=positive NES). Statistical significance is summarized via asterisks (*p<0.05; **p<0.01; ***p<0.001; Kolmogorov-Smirnov-test).

b) Differential expression of miR-223-3p as analyzed by small-RNA sequencing shows significant upregulation in T-PLL (n=46) over normal T-cell controls (n=6); (fc=9.85; p=0.0002).

c) Exemplary GSEA plots of miR-223-3p correlated mRNAs: P53_PATHWAY: NES=2.59, q<0.0001, G2M_CHECKPOINT: NES=-2.34, q=0.003).

d) Predicted targets (by seed sequences, see Methods section for details) that correlated negatively with miR-223-3p expression in all analyzed cases and controls represent regulatory networks involved in DNA-damage response and pro-survival signaling. Font color represents differential expression of mRNAs comparing T-PLL cells (n=48 cases) and healthy-donor derived CD3⁺ pan-T-cells (n=6 donors; for description of global mRNA sequencing results refer to Online Supplementary Figure S5 and Online Supplementary Table S5, blue=lower expression; red=higher expression). Color of highlighted boxes represents assignment of genes to functional groups of DNA-damage response pathways (black) and pro-survival signaling (grey).
e) Groups of low and high miR-233-3p expression were assigned by results of small-RNA sequencing via comparison of the lower vs. upper tertile of cases. Primary T-PLL cases were evaluated for CD38 and CD69 surface expression using flow cytometry (see Online Supplementary Table S8 for the comprehensive dataset). T-PLL with high miR-223-3p expression levels presented with a more activated T-cell phenotype (median CD38 expression: 65.5% vs. 3.6%, p=0.006; median CD69 expression: 8.5% vs. 1.0%, p=0.1; Mann–Whitney–Wilcoxon (MWW) test).

Figure 4: Increased expression of miR-200c and miR-141 clusters is associated with deregulation of cell-cycle regulators reflected in more activated phenotypes and more aggressive disease course.

a) Differential expression of miR-141 and miR-200 family members as analyzed by small-RNA sequencing showed significant upregulation of miR-141-3p (fc=43.2; p<0.0001), miR-141-5p (fc=29.0; p=0.0001), miR-200c-3p (fc=38.2; p<0.0001) and miR-200c-5p (fc=56.6; p=0.003, n=46 T-PLL, n=6 controls).

b) Exemplary GSEA plots of miR-200/-miR-141 correlated mRNAs: E2F_TARGETS: NES=3.64, q<0.0001, G2M_CHECKPOINT: NES=3.05, q<0.0001 (both based on miR-141-3p correlated mRNAs).

c) Predicted targets (by seed sequences, see Methods section for details) correlating negatively with miR-141/-miR-200 expression in all analyzed cases and controls showed regulatory networks involved in DNA-damage response and pro-survival signaling. Font color represents differential expression of mRNAs comparing T-PLL cells (n=48 cases) and healthy-donor derived CD3+ pan-T-cells (n=6 donors; for description of global mRNA sequencing results refer to Online Supplementary Figure S5 and Online Supplementary Table S5, blue=lower expression; red=higher expression). Color of highlighting boxes represents assignment of genes to functional groups of DNA-damage response pathways (black) and pro-survival signaling (grey).

d) Groups of low and high miR-141/-200 expression were assigned by results of small-RNA sequencing: after division into three tertiles, cases of the lower were compared to those of the upper tertile. Cases were evaluated for CD40L surface expression using flow cytometry. T-PLL with higher miR-141-5p, miR-200c-3p, and miR-200c-5p expression presented with a more activated T-cell phenotype (median CD40L expression: 16.5% vs. 0.05%, p=0.03; 16.5% vs. 0.05%, p=0.03, 20.0% vs. 0.0%, p=0.009; MWW).

e) Higher serum LDH levels (see Online Supplementary Table S9 for a summary of clinical data) were associated with high expression of miR-21-3p and miR-200c-3p (median: 898 vs
509 U/l; p=0.03; median: 917 vs 509 G/l; p=0.02; MWW). Groups were divided into three tertiles and the lower was compared against the upper tertile.

**Figure 5: MiROS-T-PLL survival score stratifies T-PLL patients based on miR-200a-3p, miR-223-3p, and miR-424-5p expression.**

a-d) Associations of miR expression with overall survival (OS) from diagnosis (see Online Supplementary Table S12 for the comprehensive dataset): a) miR-98-3p (median OS high vs. low expression: 16.3 months vs. 29.4 months, p=0.0008, log-rank test), b) miR-200a-3p (median OS high vs. low expression: 52.7 months vs. 19.1 months, p=0.001, log-rank test), c) miR-223-3p (median OS high vs. low expression: 14.9 months vs. 26.0 months, p=0.001, log-rank test), and d) miR-424-5p (median OS high vs. low expression: 14.4 months vs. 26 months, p=0.0007, log-rank test) were most significantly associated with OS, when comparing the tertile of T-PLL patients with the highest expression to the tertile with the lowest expression of the respective miR.

e) Analysis via training (n=22) and validation data sets (n=22): Optimum thresholds were calculated by recursive partitioning for all possible combinations miR-200a-3p, miR-223-3p, and miR-424-5p expression levels, adding 1 point to the total score if the respective expression cutoff was passed. Best results were obtained using the following thresholds: miR-200a-3p: fc < 2.21, miR-223-3p: fc ≥ 9.8, miR-424-5p: fc ≥ 0.91 (fold changes relative to CD3+ pan-T-cells derived from healthy donors) and a cutoff of ≥ 2 points. A significant OS association was observed in the validation set (p=0.0004, log-rank test) as well as e) in the total cohort of 44 T-PLL patients (median OS high vs. low expression: 14.4 months vs. 29.4 months, p<0.0001, log-rank test).

**Figure 6: Graphical summary of postulated miR/mRNA-based deregulated networks in T-PLL.**

T-cell activity shaped miR-ome / transcriptome networks are displayed as identified by our combinatorial approach of small-RNA and transcriptome sequencing analyses in 41 clinically well characterized T-PLL cases. MiRs differentially expressed in T-PLL and mRNAs associated with these miRs (p<0.05) are presented. The background color of miRs and mRNAs (dots) indicates the fold-changes of differential expression as compared to age-matched healthy-donor derived CD3+ pan-T-cells (blue=downregulation, red=upregulation). MiR-223-3p, the miR-21 family, the miR-29 family, and the miR-200c/141 family emerged as hallmarks of the T-PLL miR-ome, as they were (i) significantly deregulated among T-PLL cases, (ii) presented putative targets involved in oncogenic pathways of T-PLL’s pathobiology,
(iii) showed associations with prognostic parameters, and (iv) were already described in the leukemogenesis of other B- and T-cell malignancies. Deregradations of these 4 miR-families were associated with cooperative effects on DNA-damage response pathways as well as on pro-proliferative and cell survival signaling (as revealed by GSEAs based on correlated mRNAs). Genes previously described as hallmarks of T-PLL (e.g. *TCL1A*, *CTLA4*, and *MYC*) were found within the network of the identified miR-associated mRNAs.
Figure 2

A

GSEA based on differentially expressed mRNAs comparing T-PLL (n = 50) to healthy-donor derived CD3⁺ T-cells (n = 6)

- Oxidative Phosphorylation
- IL2 STAT5 Signaling
- MYC Targets
- Allograft Rejection
- DNA Repair
- TGF Beta Signaling
- Protein Secretion
- PI3K AKT MTOR Signaling
- Interferon Gamma Response
- Mitotic Spindle
- KRAS Signaling Down
- IL6 JAK STAT3 Signaling
- E2F Targets
- TNFA Signaling Via NFKB
- KRAS Signaling Up
- Inflammatory Response

q-value < 0.05

association with

- black: DNA damage response pathways
- gray: survival signaling
- white: other

Normalized Enrichment Score

B

PCA based on differentially expressed miRs comparing T-PLL to healthy-donor derived CD3⁺ T-cells

C

Heatmap and clustering based on miRs comparing healthy-donor derived T-cells: unstimulated control vs. TCR activated condition

pan CD3⁺ T-cells derived from age-matched healthy donors:

- 72 hrs. in culture w/o CD3/CD28 stim.
- directly isolated after CD3 enrichment
- 72 hrs. in culture with CD3/CD28 stim.
Figure 3

A

B

miR-223-3p expression

$\rho = 0.0002$

C

Enrichment plot: HALLMARK_P53PATHWAY

Enrichment plot: HALLMARK_G2M_CHECKPOINT

D

PURA

SACS

BRF3

TBC1D4

BRWD1

miR-223-3p

FOXO1

MPP7

STYX

E

CD38

CD69

Positive T-cells [%]

$\rho = 0.006$

$\rho = 0.1$

$n = 8$

$n = 9$

$n = 6$

$n = 8$

$n = 14$

$n = 14^*$

$\rho < 0.00$

$FDR < 0.25$
Figure 4

A

Expression of miR-141-5p and miR-200c-3p

miR-141-3p

miR-200c-3p

$p = 0.0001$

$p = 0.0003$

$p < 0.0001$

$p < 0.0001$

CD3+ pan T-cells (n = 6)

T-PLL (n = 46)

B

Enrichment plot: HALLMARK_G2M_CHECKPOINT

Enrichment plot: HALLMARK_E2F_TARGETS

C

miR-141-3p

miR-200c-3p

D

CD40L

E

LDH

miR-141-3p

miR-200c-3p

$p = 0.03$

$p = 0.03$

$p = 0.009$

$p = 0.03$

$p = 0.02$

low expression

medium expression

high expression

 rho < 0.00

p < 0.05

FDR < 0.25

color of highlighting box: DNA repair

n = 12

n = 10

n = 11

n = 9

n = 9

n = 10

n = 8

n = 8

n = 11
Figure 6

Sustained survival signaling
TCL1A
miR-21
miR-223-3p
miR-141/200c family
Impaired DNA damage response
miR-29a-3p/29c-3p
MYC

Color spectrum of mRNA names
log2(FC)
-14 0 14

Color spectrum of miRNA names
log2(FC)
-8 0 8
Supplementary Methods
Supplementary Tables 1-14 as Excel files only

Age-matched healthy-donor derived controls

CD3⁺ pan-T-cells of 6 age-matched healthy donors were used as controls (age older than 55 years). For that peripheral-blood mononuclear cells (PBMCs) were isolated fromuffy coats by density gradient centrifugation (Histopaque, Sigma-Aldrich, St. Louis, Missouri, USA). Enrichment of CD3⁺ T-cells was performed by negative selection from PBMCs using magnetic-bead based cell enrichment (Biolegend, San Diego, California, USA) according to manufacturer’s guidelines. In order to compare T-PLL cells to TCR activated T-cells, PBMCs of 4 healthy donors were stimulated using CD3/CD28 antibody mediated cross-linking (OKT3, Biolegend; 15E8, Biolegend) and kept in culture for 72 hours. As unstimulated controls, PBMCs of the same donor were kept in culture as well. After 72 hours, CD3⁺ T cells were enriched via negative magnetic selection (TCR-activated condition and unstimulated controls).

Total RNA isolation and library preparation

Total RNA was isolated from PBMCs of 46 T-PLL patients and CD3⁺ pan-T-cells of 6 age-matched healthy controls using the mirVana kit (Thermofisher, Waltham, Massachusetts, USA) according to manufacturer’s guidelines (total RNA isolation, no enrichment for small RNAs). In order to remove remaining DNA, DNase treatment of samples was performed by using the DNA-free kit (Invitrogen, Carlsbad, California, USA). RNA quality was assessed using the 4150 TapeStation (Agilent, Santa Clara, California, USA) and samples with RNA integrity number (RIN) < 6 were excluded (median RIN=8.76). For polyA-RNA sequencing, samples were subjected to library preparation using the TruSeq® Stranded mRNA Library Prep kit (Illumina, San Diego, California, USA) and then sequenced on the NovaSeq 6000 platform (Illumina, San Diego, California, USA) according to standard protocols. For small RNA sequencing, library preparation was performed using the Small RNA-Seq Library Prep Kit (Lexogen, Vienna, Austria) including size selection with Lexogen’s Gel extraction module. The barcoded libraries were size restricted between 140 and 170 bp. Samples were then analyzed using the HiSeq4000 platform (Illumina, San Diego, California, USA) according to manufacturer’s instructions. Prior to sequencing, quality of both small RNA and polyA-RNA libraries was accessed using the 4150 TapeStation (Agilent, Santa Clara, California, USA) according to standard protocols.

Small RNA sequencing data processing and analysis

Low-quality read ends as well as remaining parts of sequencing adapters were clipped off using Cutadapt (v1.14). Trimmed reads were aligned to the human genome (UCSC hg38) using Bowtie2 (v2.3.2).¹ Indexing and sorting of mapped reads was performed using
samtools (v1.5). FeatureCounts (v1.5.3) was used for summarizing gene-mapped reads. MiRBase annotations (v22/GRCh 38) were used for assigning mapped reads to known miR genes. Differential gene expression (DGE) was determined using R/edgeR (v3.26.8) and applying Trimmed Mean of M-values (TMM) normalization. Counts Per Million (CPM) transformation was applied to obtain normalized expression values. Hierarchical clustering was performed by using Euclidean distances and Ward's clustering algorithm. Separation of the samples into two clusters was conducted by cutting the clustering tree at the highest level (i.e. visual inspection).

**PolyA-RNA sequencing data processing and analysis**

Reads were aligned to the human genome (UCSC hg38) using Hisat2 (v2.1.0). Indexing and sorting of mapped reads as well as removal of secondary alignments was performed using samtools (v1.5). FeatureCounts (v1.5.3) was used for summarizing gene-mapped reads. Ensembl annotations (GRCh 38.89) were used for assigning mapped reads to known genes. DGE was determined using R/edgeR (v3.26.8) and applying TMM normalization. Fragments Per Kilobase of Million mapped reads (FPKM) transformation was applied to obtain normalized expression values. Hierarchical clustering was performed by using Euclidean distances and Ward's clustering algorithm.

**Quantitative real-time PCR**

Isolated RNA from PBMCs of T-PLL patients (n=11) and CD3+ pan-T-cells of age-matched healthy donors (n=4) was reverse-transcribed into cDNA using the TaqMan™ Advanced miRNA cDNA synthesis kit (Thermofisher, Waltham, Massachusetts, USA). TaqMan™ Advanced miRNA Assays for miR-223-3p, miR-141-3p, miR-200c-3p, and miR-30c-5p (Thermofisher, Waltham, Massachusetts, USA) were used for quantitative real-time PCRs according to the manufactures’ instructions. Thermal cycling and detection were carried out using an ABI 7500 Fast System. MiR-30c-5p presented one of the lowest coefficients of variation among all samples (coefficient of variation=0.27, FC=1.06), accompanied by a low fold-change comparing T-PLL cases to normal T-cell controls. We, therefore, used miR-30c-5p as the endogenous control. Relative quantification was calculated by using the $2^{-ΔΔCT}$ method, with the mean $2^{-ΔΔCT}$ of the 4 healthy-donor derived samples as the reference for relative expression.

**Clinical data analysis**

Detailed information on clinical characteristics, cytogenetics, immunophenotypes, and follow up was accessed for all patients. To detect possible associations between these data and miR expression in 46 T-PLL patients, we performed a screening approach testing for an extended set of parameters. For miR-223-3p, miR-21, miR-29, and miR-200c/141, patients...
were divided in groups using either the tertiles or the mean as cutoffs according to the
distribution of expression values within the patient cohort. We (i) assigned T-PLL patients to
a group according to the expression of the respective miR and (ii) then compared clinical
data between these groups. As clinical data had a few blanks for some parameters in
individual T-PLL patients, group sizes slightly differed between analyses of different miRs
and parameters. Systematic comparison was then performed using Mann-Whitney-Wilcoxon
(MWW) test and Fisher's exact test for continuous and categorical data, respectively.
Associations of miR expression with overall survival (OS) from diagnosis were examined with
log-rank statistics. Blood counts and clinical chemistry indices were evaluated at the time of
sampling. All statistical analyses were carried out using the R software and R packages. A p-
value < 0.05 was considered statistically significant.

References

3. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to
5. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital
**Figure S1:** Primary T-PLL samples submitted to miR-ome sequencing exhibit tumor cell purities higher than 88%.

Bar chart displaying immunophenotypic data showing the proportion of CD3⁺ cells in samples submitted to RNA sequencing. While T-PLL samples were sequenced out of total peripheral blood mononuclear cells (PBMCs, median purity 95.4%; range 88-99%), samples from age-matched healthy donors were enriched for CD3⁺ T-cells by negative selection (median purity 90.2%; range 77-97%; see Methods section for details on cell isolation).

**Figure S2:** CPM values obtained by small-RNA sequencing indicate high absolute expression of significantly deregulated miRs in T-PLL.

Data supplement to Figure 1. Bar chart displaying mean CPM values for a total of 2094 miRs detected by small-RNA sequencing in at least one T-PLL patient (n=46 total cohort). MiRs which were significantly deregulated in 46-T-PLL cases compared to CD3⁺ pan-T-cells of 6 healthy donors are highlighted in red (n=34), miRs not differentially expressed in T-PLL are displayed in grey. MiR-141-3p (mean CPM=26562) and miR-21-5p (mean CPM=15526), both differentially expressed in T-PLL, were among the miRs with highest absolute CPM values.

**Figure S3:** Differential expression of miR-223-3p, miR-200c-3p, and miR-141-3p in T-PLL as analyzed by quantitative real-time PCR

a) and b) Relative expression of miR-223-3p and miR-200c-3p as analyzed by quantitative real-time PCR (qRT-PCR; n=8 T-PLL, n=4 controls). Relative expression was calculated using the 2^ΔΔCT method, with the mean 2^ΔΔCT values of the 4 controls used as the reference. MiR-223-3p showed significant upregulation in T-PLL (fc=9.40; p=0.04, Welch’s test).

c) ΔCT values of miR-141-3p as analyzed by qRT-PCR (n=8 T-PLL, n=4 controls). MiR-141-3p showed no observed expression in CD3⁺ pan-T-cells of age-matched healthy donors, therefore expression is presented as ΔCT values. NE = no expression.

d) Pearson’s correlation comparing small-RNA sequencing data and qRT-PCR results of miR-223-3p (n=8 T-PLL, n=4 CD3⁺ pan-T-cells). Relative expression was calculated using expression of miR-30c-5p as an internal housekeeper control and using the mean expression of the 4 healthy-donor T-cell samples as a biological reference. Results from the qRT-PCR and small-RNA sequencing experiments showed a highly significant correlation (r²=0.84, p<0.0001, Pearson).
Figure S4: Global miR-ome profiles of T-PLL cases are not associated with distinct T-cell phenotypes.

a and b) PCA based on differentially expressed miRs in T-PLL cases (n=46) compared to healthy-donor derived CD3+ pan-T-cells (Online Supplementary Table S1). CD45RA / CD45RO as well as CD4 / CD8 surface expression was detected in primary T-PLL cells via flow cytometry.

Figure S5: T-PLL cells show altered gene expression profiles compared to healthy-donor derived CD3+ pan-T-cells.

Differentially expressed mRNAs (n=948 transcripts; Online Supplementary Table S5) were evaluated using transcriptome sequencing analysis of peripheral-blood derived primary T-PLL cells (n=48 cases) and healthy-donor derived CD3+ pan-T-cells (n=6 donors, isolation via negative selection). a) Volcano plot highlighting differentially expressed mRNAs in blue (lower expression) and red (higher expression) colors. Given percentages indicate relative proportions based on all identified mRNAs (n=18215 totally detected transcripts).

b) PCA and c) heatmaps based on the 100 most differentially expressed mRNAs. Left heatmap: mean Fragments Per Kilobase Million (FPKM) values compared between healthy-donor derived controls (n=6) and T-PLL (n=48; red=higher expression; blue=lower expression). Right heatmap: Colors represent z-scores of respective FPKM values calculated for each mRNA (blue=lower z-score; red=higher z-score).

Figure S6: Primary CD3+ pan-T-cells show elevated activation marker expression upon CD3/CD28 antibody mediated crosslinking.

Data supplementing Figure 2B and C) and Online Supplementary Figure S7: a) Primary CD3+ T-cells were isolated from peripheral blood of healthy donors using PBMC isolation and subsequent MACS enrichment after 72 hrs. (negative selection, see Methods section for details). Sample purities were assessed using flow cytometry for CD3+ positive cells within the isolates (median: 94.8%; range: 78-99%).

b) Stimulation control of TCR-activated healthy-donor derived T-cells (Figure 2B and C): isolated PBMCs were submitted to antibody-mediated CD3/CD28 cross-linking to induce TCR activation. Surface marker expression (CD25, CD38, CD69) was analyzed by flow cytometry at 72 hrs. after such TCR (CD3/CD28) stimulation.

c) Overlay plots of TCR-stimulated condition to unstimulated control (CD25, CD38, CD69 surface expression) of one exemplary healthy-donor sample.
Figure S7: TCR activation induces global changes in miR expression profiles of healthy-donor derived CD3+ pan-T-cells.
Age-matched healthy-donor derived PBMCs were isolated via density gradient centrifugation. TCR-mediated T-cell activation was achieved via continued antibody-based CD3/CD28 cross-linking. After 72 hrs., CD3+ primary human T-cells were isolated by MACS enrichment (negative selection; see Method section for details; cell purities are given in Online Supplementary Figure S6A, see Online Supplementary Figure S6B and C for control experiments). MiR-ome profiles were generated using small-RNA sequencing. a) Volcano plot showing log2 fc and –log10 FDR values of differentially expressed miRs in response to CD3/CD28 stimulation. Given percentages show relative proportions based on all identified miRs (n=56/1587 totally detected miRs; FDR<0.05). b) PCA of differentially expressed miRNAs upon CD3/CD28 cross-linking (light green: unstimulated controls; dark green: stimulated condition). c) Heatmap showing miR expression in unstimulated vs. stimulated healthy-donor derived T-cells (n=52 miRs; FDR<0.05). Colors represent z-scores of respective CPM values calculated for each miR (blue=lower z-score; red=higher z-score).

Figure S8: Functional associations of predicted miR-targets using GSEA highlight processes of DNA-damage responses and pro-survival signaling.
Data supplement Figure 3A: GSEA heatmap based on miR associated mRNAs. GSEAs were conducted for all significantly deregulated miRs (n=34 miRs; Online Supplementary Table S7) using ranked correlation indices between mRNA and miR expression in 41 T-PLL and 6 healthy-donor derived T-cell samples. Color code summarizes NES scores (blue=negative NES; red=positive NES). Statistical significance is summarized via asterisks (*p<0.05; **p<0.01; ***p<0.001; Kolmogorov-Smirnov-test).

Figure S9: Reduced expression of miR-21 is associated with alterations of apoptosis and cell-cycle regulators.
a) Differential expression of miR-21 family members as analyzed by small-RNA sequencing shows significant downregulation of miR-21-3p (fc=0.27 ; p<0.0001) and miR-21-5p (fc=0.31; p<0.0001, n=46 T-PLL cases, n=6 normal T-cell controls). b) Exemplary GSEA plots of miR-21 correlated mRNAs: APOPTOSIS: NES=3.89, q<0.0001 (miR-21-5p correlated mRNAs); G2M_CHECKPOINT: NES=2.39, q=0.001 (miR-21-3p correlated mRNAs).
c) Predicted targets (by seed sequences, see Methods section for details) correlating negatively with miR-21 expression in all analyzed cases and controls implicate regulatory networks involved in DNA-damage response and pro-survival signaling. Font color represents differential expression of mRNAs comparing T-PLL cells (n=48 cases) to healthy-donor derived CD3+ pan-T-cells (n=6 donors; for description of global mRNA sequencing results refer to Figure 2A: blue=lower expression, red=higher expression). Color of highlighted boxes indicates assignment of genes to functional groups of DNA-damage response pathways (black) and pro-survival signaling (grey).

d) Low miR-21 expression was associated with higher white blood cell counts (WBC) (median: 154 vs. 90.0 G/l; p=0.02; MWW) and lower platelet counts (median: 110 vs. 153 G/l; p=0.03; MWW) at the time of sample acquisition. See Online Supplementary Table S10 for summary of clinical data. Groups were divided by the mean miR-21 expression value.

e) Low miR-21 expression was associated with higher serum lactate dehydrogenase (LDH) levels at the time of sample acquisition (median: 933 vs. 522 U/l; p=0.02; MWW). Groups were divided by the mean miR-21 expression.

Figure S10: Reduced expression of miR-29 clusters is associated with alterations of cell survival signaling and cell-cycle regulators reflected in more activated phenotypes and a more aggressive disease course.
a) Differential expression of miR-29 family members as analyzed by small-RNA sequencing showed significant downregulation of miR-29a-3p (fc=0.29; p<0.0001), miR-29b-1-5p (fc=0.47; p=0.001), and miR-29c-3p (fc=0.29; p<0.0001) in 46 T-PLL, as compared to 6 normal T-cell control samples.

b) Exemplary GSEA plots of miR-29-correlated mRNAs: E2F_TARGETS: NES=-3.52, q<0.0001, TNFA_SIGNALING_VIA_NFKB: NES=-2.08, q=0.01 (both based on miR-29a-3p correlated mRNAs).

c) Predicted targets (by seed sequences, see Methods section for details) that negatively correlate with miR-29 expression in all analyzed cases and controls revealed regulatory networks involved in DNA-damage response pathways and pro-survival signaling. Font color represents differential expression of mRNAs comparing T-PLL cells (n=48 cases) and healthy-donor derived CD3+ pan-T-cells (n=6 donors; for description of global mRNA sequencing results refer to Figure 2A and Online Supplementary Table S5, blue=lower expression; red=higher expression). Color of highlighted boxes represents assignment of genes to functional groups of DNA-damage response pathways (black) and pro-survival signaling (grey).
d) Groups of low and high miR-29 expression were assigned by results of small-RNA sequencing: Cases were divided into three tertiles and the lower was compared against the upper tertile. Groups were then evaluated for CD38 and CD69 surface expression using flow cytometry. T-PLL cases with low miR-29 expression levels presented with a more activated T-cell phenotype (median CD38 expression: 43.1% vs. 12.9%, p=0.02; median CD69 expression: 27.9% vs. 0.89%, p=0.02; MWW).

e and f) Platelet counts and serum LDH values of analyzed cases were assessed (see Online Supplementary Table S11 for a summary of clinical data). Groups were divided into tertiles of miR-29 expression and the lower was compared against the upper tertile using MWW. e) Low miR-29 expression was associated with lower platelet counts at the time the sample was taken (median: 110 vs 186 G/l; p=0.15; MWW) and with f) higher serum LDH levels at the time of sample (median: 1850 vs 708 U/l; p=0.06; MWW).

Figure S11: Low miR-29 expression is associated with higher rates of effusions and more prevalent in cases with genomic ATM deletions.

Groups of low, medium, and high miR-29c-3p expression were assigned by results of small-RNA sequencing using tertile-based expression levels. Clinical presentation of cases was recorded for all analyzed T-PLL patients (see Online Supplementary Table S11 for the comprehensive dataset) and the high expression group was compared against the low expression tertile. a) Cases with low miR-29c-3p expression showed a higher proportion of effusions (n=6/12 vs. n=1/10 cases; p=0.07; Fisher’s exact test).

b) Genomic deletions of ATM were assessed using FISH studies probing the 11q22.3 consensus region. Cases with low miR-29b-1-5p expression showed a significantly higher proportion of ATM deletions (n=9/11 vs. n=1/11 cases; p=0.002; Fisher’s exact test).
Figure S1

Purity

% CD3+ T-cells

- pan CD3+ T cells (n = 6 age-matched healthy donors)
- PBMCs isolated from T-PLL patients (n = 46 cases)
Figure S3

A  miR-223-p

B  miR-200c-3p

C  miR-141-3p

D  miR-223-3p: small-RNA sequencing vs. RT-PCR
Figure S4

A

PCA (miR Seq)

T-PLL (n = 46 cases)
- CD45 RO-
- CD45 RO+
- unknown

pan CD3+ T-cells (n = 6 age-matched healthy donors)

B

PCA (miR Seq)

T-PLL (n = 46 cases)
- CD4+
- CD8+
- CD4-/CD8+
- unknown

pan CD3+ T-cells (n = 6 age-matched healthy donors)
**Figure S5**

**A**
Differentially expressed mRNAs in T-PLL

- **log2 fold change**
  - Log2 fold change values are plotted along the x-axis.
  - Log10 FDR values are plotted along the y-axis.

- **Differentially expressed mRNAs**
  - Various genes are highlighted, including CCR8, RNF17, CLEC4E, TCL1A, and TUNAR.

- **CTLA4**
  - Differentially expressed mRNAs are annotated with a log10 FDR value of 2.4% (n = 430).

- **Pan CD3+ T-cells**
  - Differentially expressed mRNAs are annotated with a log10 FDR value of 2.8% (n = 518).

**B**
PCA (mRNA Seq)

- **Dim1 (24.9%)**
  - Principal Component Analysis on mRNA Seq data.

- **Dim2 (11%)**

- **94.8% (n = 17267)**

- **2.8% (n = 518)**

**C**

- **Pan CD3+ T-cells** (6 age-matched healthy donors)
- **T-PLL** (48 cases)

- **mean FPKM**

- **log10(FPKM)**

- **z-score**
Figure S6

A

C}

CD25

CD38

CD69

CD25 expression (gated on CD3+ cells) [%]

CD38 expression (gated on CD3+ cells) [%]

CD69 expression (gated on CD3+ cells) [%]

fluorescence intensity

fluorescence intensity

fluorescence intensity

72 hrs. in culture without anti CD3/CD28 stim.

72 hrs. in culture with anti CD3/CD28 stim.
Figure S7

A. Differentially expressed miRs upon anti-CD3/CD28 stimulation in pan T-cells derived from healthy controls

B. PCA (miR Seq)

C. pan CD3+ T-cells (n = 4 age-matched healthy donors)

72 hrs. in culture without anti CD3/CD28 stim. | 72 hrs. in culture with anti CD3/CD28 stim.
GSEA based on miR-correlated mRNAs for 34 differentially expressed miRs comparing T-PLL to healthy-donor derived controls
**Figure S9**

### A

**miR-21-3p**
- **miR-21-5p**

<table>
<thead>
<tr>
<th>miR expression</th>
<th>CD3+ pan T-cells ($n = 6$)</th>
<th>T-PLL ($n = 46$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>1.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>0.5</td>
<td>$p &lt; 0.0001$</td>
<td>$p &lt; 0.0001$</td>
</tr>
</tbody>
</table>

### B

**Enrichment plot: HALLMARK_APOPTOSIS**
- **Enrichment plot: HALLMARK_G2M_CHECKPOINT**

### C

**miR-21-5p**
- NIPAL1
- LRP6
- MSH2
- ABCD3
- FCHO2
- GID4
- MAP3K1
- BCL7A
- SLC1A10
- TET1
- SLC7A6
- SCML2
- SKP2

**miR-21-3p**
- DDX4
- AAMP
- NFYB

**color spectrum of mRNA names**
- -14
- 0
- 14

**color spectrum of miR names**
- -8
- 0
- 8

**color of high-lighting boxes:**
- DNA repair
- survival signaling

### D

**WBC**
- Absolute no. of cells [G/l]
- $p = 0.02$
- $p = 0.03$
- $n = 31$
- $n = 25$
- $n = 10$

**Platelets**
- Absolute no. of cells [G/l]
- $p = 0.02$
- $p = 0.03$
- $n = 18$
- $n = 9$

**low miR-21-5p expression**
- **high miR-21-5p expression**

### E

**LDH**
- [UI]
- $p = 0.02$
- $p = 0.03$
- $n = 18$
- $n = 9$

**low miR-21-5p expression**
- **high miR-21-5p expression**
**Figure S10**

**A**

<table>
<thead>
<tr>
<th>miR expression</th>
<th>rel. to CD3+ T-cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-29a-3p</td>
<td>$p &lt; 0.0001$</td>
</tr>
<tr>
<td>miR-29b-1-5p</td>
<td>$p = 0.001$</td>
</tr>
<tr>
<td>miR-29c-3p</td>
<td>$p &lt; 0.0001$</td>
</tr>
</tbody>
</table>

CD3+ pan T-cells ($n = 6$)

T-PLL ($n = 46$)

**B**

Enrichment plot.

**C**

miR-29c-3p

miR-29a-3p

miR-29b-1-5p

**D**

<table>
<thead>
<tr>
<th>CD38</th>
<th>CD69</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p = 0.02$</td>
</tr>
</tbody>
</table>

**E**

Platelets

$|p = 0.15|

**F**

LDH

$|p = 0.06|$
Figure S11

A  Effusions

B  ATM deletion
Supplementary Table 1: Differentially expression of miRs comparing T-PLL cases ($n = 46$) to age-matched healthy-donor derived CD3$^+$ pan T-cells ($n = 6$).

Supplementary Table 2: Somatic copy number alterations of significantly deregulated miRs in T-PLL as detected by single nucleotide polymorphism arrays in an independent cohort of T-PLL patients.

Supplementary Table 3: Associations of immunophenotypic, cytogenetic, and clinical data with two distinct T-PLL subgroups, as revealed by unsupervised hierarchical clustering analysis of miR expression.

Supplementary Table 4: Differentially expressed mRNAs comparing two clusters of T-PLL cases (high miR-141/200c expression cluster vs. low miR-141/200c expression cluster) defined by small-RNA sequencing.

Supplementary Table 5: Differentially expressed mRNAs comparing T-PLL cases ($n = 48$) to age-matched healthy-donor derived CD3$^+$ pan T-cells ($n = 6$).

Supplementary Table 6: GSEA (HALLMARK) of differentially expressed mRNAs comparing T-PLL cases ($n = 48$) to age-matched healthy-donor derived CD3$^+$ pan T-cells ($n = 6$).

Supplementary Table 7: Differentially expressed mRNAs comparing healthy-donor derived T-cells: unstimulated control vs. TCR activated condition ($n = 4$).

Supplementary Table 8: Associations of miR-223-3p expression with immunophenotypic, cytogenetic, and clinical data (lower tertile vs. upper tertile of miR-223-3p expression).

Supplementary Table 9: Associations of miR-200c and miR-141 expression with immunophenotypic, cytogenetic, and clinical data (lower tertile vs. upper tertile of the expression of the respective miR).

Supplementary Table 10: Associations of miR-21-3p and miR-21-5p expression with immunophenotypic, cytogenetic, and clinical data (groups were divided by mean miR-21 expression).

Supplementary Table 11: Associations of miR-29c-3p, miR-29b-1-5p, and miR-29c-3p expression with immunophenotypic, cytogenetic, and clinical data (groups were divided be median miR-29 expression).

Supplementary Table 12: Associations of miR expression levels with overall survival for all miRs detectable in at least 80% of samples ($n = 37$), comparing T-PLL patients with highest expression levels (upper tertile) to those with lowest expression of the respective miR (lower tertile).

Supplementary Table 13: Associations of the miROS-T-PLL score and its miRs with incidences of genomic alterations previously described in T-PLL.

Supplementary Table 14: Associations of the miROS-T-PLL survival score with clinical, immunophenotypic and cytogenetic data (miROS-T-PLL <2 vs ≥ 2 points).