

Oncorequisite role of an aldehyde dehydrogenase in the pathogenesis of T-cell acute lymphoblastic leukemia

Chujing Zhang,¹ Stella Amanda,¹ Cheng Wang,² Tze King Tan,¹ Muhammad Zulfakar Ali,¹ Wei Zhong Leong,¹ Ley Moy Ng,¹ Shojiro Kitajima,³ Zhenhua Li,⁴ Allen Eng Juh Yeoh,^{1,4} Shi Hao Tan¹ and Takaomi Sanda^{1,5}

¹Cancer Science Institute of Singapore, National University of Singapore, Singapore; ²Department of Anatomy, National University of Singapore, Singapore; ³Institute for Advanced Biosciences, Keio University, Tsuruoka, Japan; ⁴VIVA-NUS CenTRAL, Department of Pediatrics, National University of Singapore, Singapore and ⁵Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.



Haematologica 2021
Volume 106(6):1545-1558

ABSTRACT

Aldehyde dehydrogenases (ALDH) are overexpressed in various types of cancers. One of the ALDH family genes, *ALDH1A2*, is aberrantly expressed in more than 50% of cases of T-cell acute lymphoblastic leukemia (T-ALL). However, its molecular function and role in the pathogenesis of T-ALL are largely unknown. Chromatin immunoprecipitation-sequencing and RNA-sequencing analyses showed that the oncogenic transcription factor *TAL1* and its regulatory partners bind to the intronic regulatory element of the *ALDH1A2* gene, directly inducing a T-ALL-specific isoform with enzymatic activity. *ALDH1A2* was preferentially expressed in the *TAL1*-positive T-ALL subgroup. In T-ALL cell lines, depletion of *ALDH1A2* inhibited cell viability and induced apoptosis. Interestingly, gene expression and metabolomic profiling revealed that *ALDH1A2* supported glycolysis and the tricarboxylic acid cycle, accompanied by NADH production, by affecting multiple metabolic enzymes to promote ATP production. Depletion of *ALDH1A2* increased the levels of reactive oxygen species, while the levels were reduced by *ALDH1A2* overexpression both *in vitro* and *in vivo*. Overexpression of *ALDH1A2* accelerated tumor onset and increased tumor penetrance in a zebrafish model of T-ALL. Taken together, our results indicate that *ALDH1A2* protects against intracellular stress and promotes T-ALL cell metabolism and survival. *ALDH1A2* overexpression enables leukemic clones to sustain a hyper-proliferative state driven by oncogenes.

Introduction

T-cell acute lymphoblastic leukemia (T-ALL) is a hematologic malignancy that arises from immature T-cell precursors.^{1,2} This type of leukemia occurs mostly in children but is also found in adults. A number of chromosomal and genetic abnormalities have been identified in T-ALL, many of which affect genes encoding transcription factors.³ Gene expression and mutational profiles have demonstrated that T-ALL cases can be classified into mutually exclusive subgroups based on the expression of several transcription factors (*TAL1*, *TAL2*, *LYL1*, *LMO1*, *LMO2*, *TLX1/HOX11*, *TLX3/HOX11L2*, *HOXA* and *NKX3-1*).^{4,6} This suggests that each transcription factor drives a distinct cellular program that may work in concert with other molecular pathways in T-ALL cells.

TAL1-positive T-ALL constitutes the largest subgroup, accounting for 40-60% of all primary cases.^{4,7-11} One of the known downstream targets of *TAL1* in T-ALL cells is *ALDH1A2*,^{12,13} a member of the aldehyde dehydrogenase (*ALDH*) family of genes that encode oxidoreductases. *ALDH* proteins convert a variety of aldehydes into carboxylic acids.^{14,15} They detoxify endogenous aldehydes, generated during the

Correspondence:

TAKAOMI SANDA
takaomi_sanda@nus.edu.sg

Received: December 18, 2019.

Accepted: May 14, 2020.

Pre-published: May 15, 2020.

<https://doi.org/10.3324/haematol.2019.245639>

©2021 Ferrata Storti Foundation

Material published in *Haematologica* is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: <https://creativecommons.org/licenses/by-nc/4.0/legalcode>, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.



metabolism of membrane lipids, amino acids, carbohydrates and steroids.¹⁶ The reaction is NAD(P)⁺-dependent and generates an important component of the cellular antioxidant system, NAD(P)H, which also acts as an essential coenzyme in several metabolic pathways, such as glycolysis and the tricarboxylic acid (TCA) cycle.¹⁴ Retinaldehyde dehydrogenases, a subfamily of *ALDH* genes, including *ALDH1A2*, are also capable of converting retinaldehyde into retinoic acid. ALDH activity has also been implicated as a cancer stem cell marker in various solid tumors.^{17,18} The Aldefluor assay is commonly used to isolate an ALDH-positive population to refine the cancer stem cell population.^{17,19} Overexpression of *ALDH* genes confers drug resistance to cancer cells.²⁰ These findings suggest that high levels of ALDH activity may be advantageous and, even required, for cancer cell maintenance.

In this study, we elucidated the molecular function and role of *ALDH1A2* in the pathogenesis of T-ALL. We demonstrated that *ALDH1A2* is directly activated by TAL1 and protects against intracellular stress, supporting leukemia cell metabolism and promoting leukemia cell survival.

Methods

Cell samples

Human leukemia cell lines were cultured in RPMI-1640 medium (BioWest) supplemented with 10% fetal bovine serum (BioWest). A T-ALL patient-derived xenograft sample was provided by Alejandro Gutierrez (Boston Children's Hospital) and expanded in NSG mice. Mouse protocols were approved by the Institutional Animal Care and Use Committee.

Knockdown and overexpression experiments

For knockdown experiments, short-hairpin RNA (shRNA) was inserted into a pLKO.1-puro lentivirus vector. For overexpression experiments, *ALDH1A2* cDNA cloned from Jurkat cells was inserted into a MSCV-IRES-GFP retrovirus vector. For inhibition of the regulatory element, single-guide RNA (sgRNA) was inserted into a FgH1UTG lentivirus vector (Addgene plasmid #70183). Virus was produced by co-transfecting the construct with the packaging and envelope plasmids into 293T cells using FuGENE6 (Roche).

Quantitative reverse transcription polymerase chain reaction

Total RNA was extracted from cells using a NucleoSpin RNA kit (Macherey-Nagel) and reverse-transcribed using a QuantiTect kit (Qiagen). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed on a QuantStudio3 real-time PCR system (Thermo Fisher Scientific) using Power SYBR Green PCR Master Mix (Roche). The primer sequences are described in the *Online Supplementary Information*.

Western blot analysis

Equivalent amounts of protein were resolved on a sodium dodecylsulfate polyacrylamide gel electrophoresis gel, transferred onto polyvinylidene difluoride membranes (Bio-Rad), and subjected to immunoblot analysis with specific antibodies against *ALDH1A2* (Abcam), TAL1 (Millipore), PARP, caspase-3, α -tubulin, and β -actin (Cell Signaling Technology). The proteins were detected using an enhanced chemiluminescence reagent (Thermo).

Confocal microscopic imaging

Images of EGFP-fused *ALDH1A2* protein localization in live cells were captured using an Olympus FV1000 TIRF confocal

microscope (C4). Colocalization of signals was analyzed using IMARIS 9.5 software after staining with Mito Tracker Red CMXRos (Thermo) and Hoechst.

Intracellular reactive oxygen species assay

The cells were treated with freshly prepared CellROX Deep Red reagent (Thermo Fisher Scientific). Fluorescence signals were analyzed with a BD SR II flow cytometer using BD FACSDiva™ software and FlowJo software.

Metabolomic profiling

The extracted metabolites were used for a capillary electrophoresis time-of-flight mass spectrometry basic scan, conducted by Human Metabolome Technologies (see details in the *Online Supplementary Information*). The number of metabolites was normalized to cell volume detected by a Scepter 2.0 cell counter (Millipore).

Extracellular flux analysis

The cells were resuspended in Seahorse assay medium (Agilent) and seeded into XF24 plates. Extracellular acidification rate and the oxygen consumption rate were measured by Seahorse XF24 (Agilent). The details are provided in the *Online Supplementary Information*.

Chromatin immunoprecipitation sequencing and RNA-sequencing

Chromatin immunoprecipitation (ChIP)-sequencing analysis was done in our previous study.²¹ For RNA-sequencing, RNA was extracted using an miRNeasy kit (Qiagen) followed by DNase treatment (Ambion). Construction of the strand-specific library and sequencing of the single-end 100-bp-long reads by a BGISEQ sequencer were conducted at BGI Biotech (Hong Kong). The details are described in the *Online Supplementary Information*.

Zebrafish transgenesis

Zebrafish protocols were approved by the Institutional Animal Care and Use Committee. The human *ALDH1A2* or *mCherry* gene was cloned under the zebrafish *rag2* promoter and injected into zebrafish embryos to create *Tg(rag2: ALDH1A2)* or a control *Tg(rag2: mCherry)* line. The *rag2-myr-mAKT2* construct was kindly provided by Alejandro Gutierrez.²²

Results

The *ALDH1A2* gene is directly activated by the TAL1 complex in T-cell acute lymphoblastic leukemia

To identify the targets directly regulated by TAL1 in T-ALL cells, we referred to a previously performed ChIP-sequencing dataset for TAL1 and other members of the transcriptional complex in a T-ALL cell line (Jurkat).¹⁵ We integrated this result with the RNA-sequencing dataset after knocking down each factor in the same cell line, which was generated in our recent study,²¹ to select genes that were positively regulated by the TAL1 complex. We further filtered genes that were associated with a high level of an active histone mark (H3K27ac) in T-ALL cells but not in normal thymus cells (*Online Supplementary Figures S1A and B*), representing high-confidence TAL1 targets that were aberrantly activated in T-ALL cells.

ALDH1A2 was among the top hits. An early study showed that the TAL1/GATA3/LMO complex binds at the intronic region of the *ALDH1A2* gene and induces its expression in T-ALL cells.¹² Consistently, our ChIP-

sequencing analysis demonstrated that TAL1 and its binding partners (HEB, GATA3, RUNX1 and MYB) bind within the intron with the co-activator CBP and RNA polymerase II in Jurkat cells (Figure 1A). Knocking down each of these factors resulted in downregulation of *ALDH1A2* mRNA expression in the same cell line, as determined by RNA-sequencing (Figure 1B, *Online Supplementary Figure S1B*). The result was validated at the protein level (Figure 1C). Importantly, the TAL1-bound region was associated with a high level of H3K27ac signals in multiple *TAL1*-positive

T-ALL cell lines but not in *TAL1*-negative cell lines (Figure 1A). This status corresponded to an *ALDH1A2* mRNA expression pattern that was detected only in *TAL1*-positive T-ALL cell lines (Figure 1D). Similarly, in the primary T-ALL cells from three independent cohorts,^{4,23,24} *ALDH1A2* expression was significantly higher in the *TAL1*/2- and *LMO1*/2-subgroups than in the *TLX*, *LYL/LMO* and *HOXA*-positive subgroups (Figure 1E, *Online Supplementary Figure S1C*, *Online Supplementary Table S1*). Furthermore, the results from the analysis of the

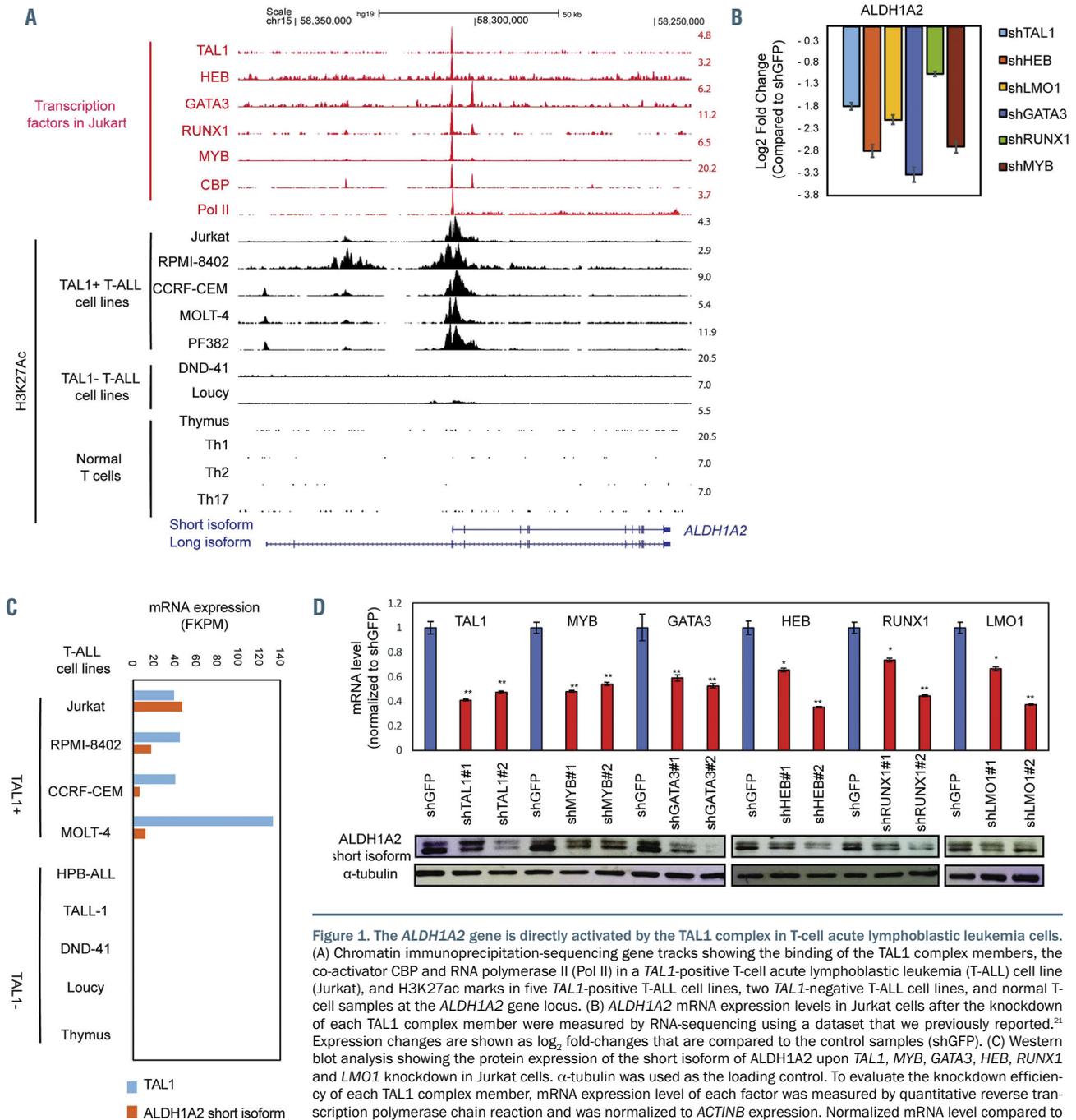


Figure 1. The *ALDH1A2* gene is directly activated by the TAL1 complex in T-cell acute lymphoblastic leukemia cells. (A) Chromatin immunoprecipitation-sequencing gene tracks showing the binding of the TAL1 complex members, the co-activator CBP and RNA polymerase II in a *TAL1*-positive T-cell acute lymphoblastic leukemia (T-ALL) cell line (Jurkat), and H3K27ac marks in five *TAL1*-positive T-ALL cell lines, two *TAL1*-negative T-ALL cell lines, and normal T-cell samples at the *ALDH1A2* gene locus. (B) *ALDH1A2* mRNA expression levels in Jurkat cells after the knockdown of each TAL1 complex member were measured by RNA-sequencing using a dataset that we previously reported.²¹ Expression changes are shown as log₂ fold-changes that are compared to the control samples (shGFP). (C) Western blot analysis showing the protein expression of the short isoform of *ALDH1A2* upon *TAL1*, *MYB*, *GATA3*, *HEB*, *RUNX1* and *LMO1* knockdown in Jurkat cells. α-tubulin was used as the loading control. To evaluate the knockdown efficiency of each TAL1 complex member, mRNA expression level was measured by quantitative reverse transcription polymerase chain reaction and was normalized to *ACT1NB* expression. Normalized mRNA level compared to control (shGFP) is shown. Error bars represent the standard deviation (SD) for technical replicates. **P*<0.05, ***P*<0.01 using a two-tailed Student *t* test. (D) The mRNA expression levels of *ALDH1A2* and *TAL1* in four *TAL1*-positive T-ALL cell lines, four *TAL1*-negative T-ALL cell lines and cells from one normal thymus sample were analyzed by RNA-sequencing. Expression values are shown as fragments per kilobase million (FPKM).

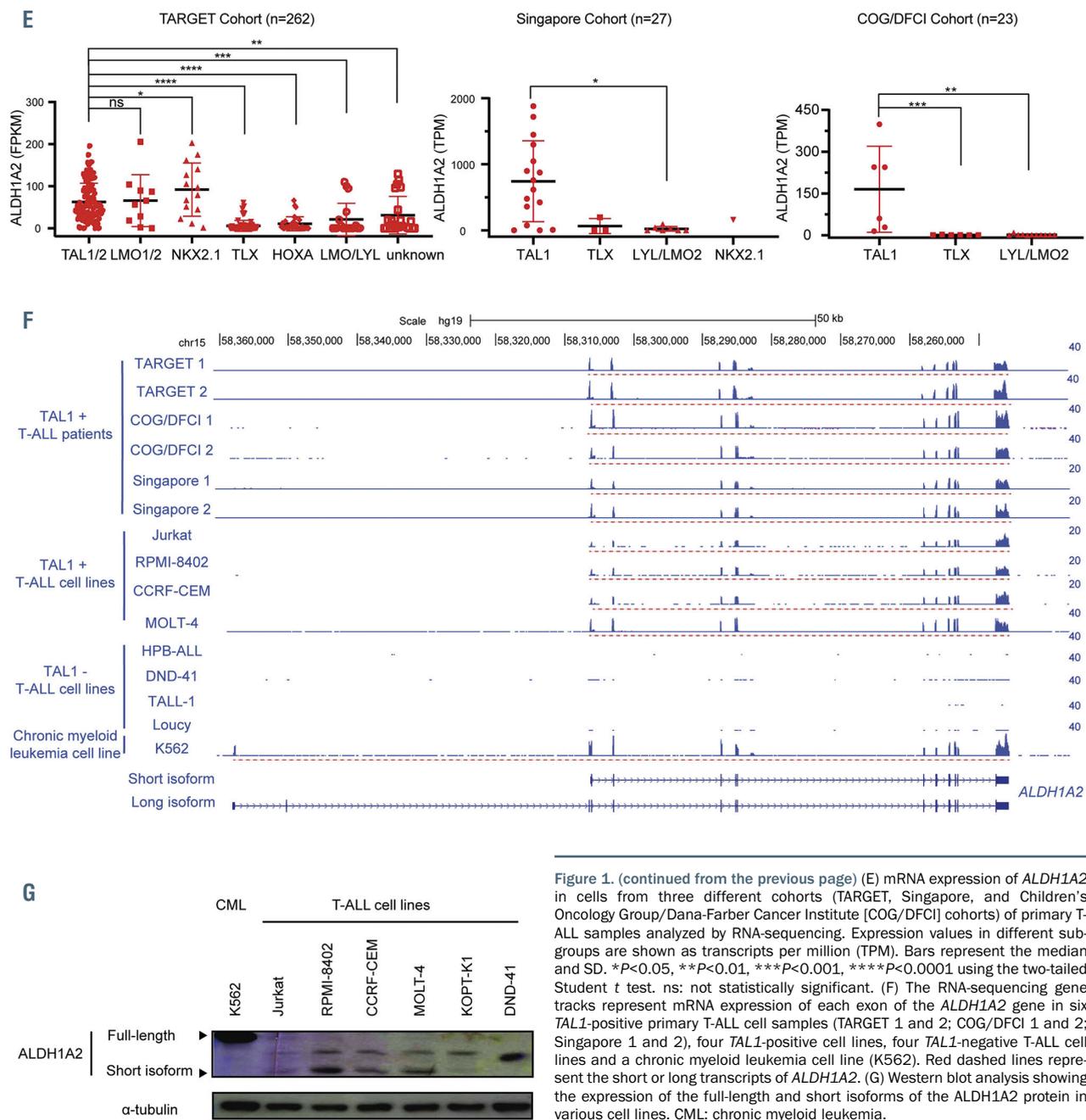


Figure 1. (continued from the previous page) (E) mRNA expression of *ALDH1A2* in cells from three different cohorts (TARGET, Singapore, and Children's Oncology Group/Dana-Farber Cancer Institute [COG/DFCI] cohorts) of primary T-ALL samples analyzed by RNA-sequencing. Expression values in different subgroups are shown as transcripts per million (TPM). Bars represent the median and SD. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ using the two-tailed Student *t* test. ns: not statistically significant. (F) The RNA-sequencing gene tracks represent mRNA expression of each exon of the *ALDH1A2* gene in six *TAL1*-positive primary T-ALL cell samples (TARGET 1 and 2; COG/DFCI 1 and 2; Singapore 1 and 2), four *TAL1*-positive cell lines, four *TAL1*-negative T-ALL cell lines and a chronic myeloid leukemia cell line (K562). Red dashed lines represent the short or long transcripts of *ALDH1A2*. (G) Western blot analysis showing the expression of the full-length and short isoforms of the *ALDH1A2* protein in various cell lines. CML: chronic myeloid leukemia.

Cancer Cell Line Encyclopedia dataset showed the highest expression of *ALDH1A2* mRNA in T-ALL samples among 40 cancer types²⁵ (Online Supplementary Figure S1D).

Notably, RNA-sequencing analysis demonstrated that only the short isoform, which lacks the first three exons, but not the full-length (long) isoform of *ALDH1A2*, was expressed in the *TAL1*-positive T-ALL cells, including primary samples and cell lines (Figure 1F). This result was consistent with that from an earlier study.¹² In contrast, K562 cells, a chronic myeloid leukemia cell line, expressed the long isoform of *ALDH1A2* but not the short isoform (Figure 1F, bottom). The difference in protein size was confirmed by western blot (Figure 1G).

It is noteworthy that the H3K27ac marks were not observed in normal T cells (Figure 1A). Consistent with this finding, in an analysis of a public dataset of normal human

hematopoietic cells,²⁶ *ALDH1A2* expression was not detected in T cells (Online Supplementary Figure S1E). Similarly, the mouse *Aldh1a2* gene was not expressed in hematopoietic cells, while it was detected in bone marrow stroma cells (Online Supplementary Figure S1F). Furthermore, recent single-cell RNA-sequencing studies confirmed that *ALDH1A2* is expressed in thymic stroma cells but not in any T-cell subpopulations both in mice and human cells.^{27,28} These results further indicated that *ALDH1A2* is aberrantly activated in cases of *TAL1*-positive T-ALL.

The short isoform of *ALDH1A2* possesses enzymatic activity and is localized in the cytoplasm

We next investigated the function of the short isoform of the *ALDH1A2* protein. The analysis of the amino acid sequence indicated that the short isoform retains function-

ally critical residues, such as the NAD binding site and the enzymatic catalytic residues, but lacks 96 amino acids in the N-terminus (Figure 2A) that are a part of the homotrimeric interface (Figure 2B). Hence, we postulated that the short isoform might have attenuated enzymatic activity or a different subcellular localization compared to the long isoform.

To investigate this, we first produced each isoform of ALDH1A2 as a purified recombinant His-tagged protein (Online Supplementary Figure S2A) and performed an *in vitro* enzymatic assay. We incubated each isoform with a substrate (retinaldehyde) and a coenzyme (NAD⁺) and measured the amount of NADH produced over 20 min (Figure 2C). This analysis revealed that the short isoform had enzymatic ability, although slightly lower than that of the long isoform (Figure 2C, Online Supplementary Figure S2B). Next, we analyzed the effect of ALDH1A2 on the production of retinoic acid using a reporter system. We cloned a trimerized retinoic acid responsive element, which can be activated by the retinoic acid receptor complex, into the luciferase plasmid and then established Jurkat cells that stably express this construct³⁹ (Online Supplementary Figure S2C). In this setting, the luciferase activity is dependent on the amount of retinoic acid produced internally in the Jurkat cells. We then knocked down the endogenous ALDH1A2 by lentiviral shRNA transduction. The luciferase activity was significantly downregulated after ALDH1A2 had been depleted (Online Supplementary Figure S2C), indicating that the short isoform of ALDH1A2 can mediate the production of retinoic acid in T-ALL cells.

We next analyzed the subcellular localization of the short isoform of the endogenous ALDH1A2 protein. We used the CRISPR/Cas9-mediated method to introduce an EGFP gene fused to the 3' end of the endogenous ALDH1A2 gene into both Jurkat and K562 cells (Online Supplementary Figures S2D and E). We then analyzed the localization of the EGFP signal in the nucleus (with Hoechst staining), mitochondria (with Mitotracker staining) or cytoplasm (without staining). The results demonstrated that the short isoform of ALDH1A2 in Jurkat cells was localized in the cytoplasm with no co-localization in mitochondria or the nucleus (Figure 2D, Online Supplementary Figure S2F), which was similar to the pattern observed for the K562 cells that expressed the long isoform. These results indicate that the short isoform of ALDH1A2 possesses intact enzymatic functions and the same localization pattern as observed for the long isoform.

ALDH1A2 supports T-cell acute lymphoblastic leukemia cell survival and viability

We next analyzed whether ALDH1A2 expression confers any functional advantage to T-ALL cells. We first evaluated the phenotype after depletion of ALDH1A2 by blocking the regulatory element. We designed two independent sgRNA (#1 and #2), which targeted the TAL1-bound region (Figure 3A). We transduced each of them under a doxycycline-inducible system together with a catalytically inactive Cas9 (dCas9) protein fused with the KRAB repressor, thereby epigenetically silencing the transcriptional activity at the TAL1-bound region. We observed successful downregulation of ALDH1A2 protein after the induction of each sgRNA (Figure 3B). Importantly, apoptotic cell death was induced after 72 h of induction, as evidenced by the cleavage of caspase-3 and PARP (markers of apoptosis) (Figure 3B).

To independently validate this result, we utilized a lentiviral shRNA system to knock down ALDH1A2 and analyzed the number of annexin-V-positive cells, which is another means to detect apoptosis. The result confirmed that depletion of ALDH1A2 induced apoptosis in multiple TAL1/ALDH1A2-positive T-ALL cell lines but not in the negative cell lines (Figure 3C). Consistently, cell viability was reduced after the shRNA knockdown of ALDH1A2 in an ALDH1A2-positive cell line (Jurkat) (Figure 3D). Conversely, forced expression of ALDH1A2 in an ALDH1A2-negative T-ALL cell line (DND-41) increased cell viability (Figure 3E).

WIN 18,446 is reported to be a pan-ALDH1A inhibitor.³⁰ It has been shown to strongly inhibit the enzymatic activity of ALDH1A2 both *in vitro* and *in vivo*.^{31,32} Since ALDH1A2 is the only member of the ALDH1A family of genes expressed in T-ALL (Online Supplementary Figure S3A), we tested the effect of WIN 18,446 on cell viability of several T-ALL cell lines. Strikingly, two TAL1/ALDH1A2-positive cell lines (Jurkat and RPMI-8402) were more sensitive to this small-molecule inhibitor than two TAL1/ALDH1A2-negative cell lines (KOPT-K1 and DND-41) (Figure 3F, Online Supplementary Figure S3B). These results indicate that the expression of ALDH1A2 is associated with the viability and survival of T-ALL cells. This phenotype was not rescued by the addition of all-trans retinoic acid (Online Supplementary Figure S3C), thus suggesting that this mechanism is likely independent of the amount of retinoic acid produced.

ALDH1A2 affects metabolic pathways in T-cell acute lymphoblastic leukemia cells

We next investigated the molecular mechanisms by which ALDH1A2 supports cell viability and prevents apoptosis. We performed global gene expression profiling by RNA-sequencing after sgRNA-mediated ALDH1A2 depletion. We selected genes that were significantly downregulated (n=96) or upregulated (n=19) in the ALDH1A2-depleted samples compared to the control samples (Figure 4A, Online Supplementary Figure S4A, Online Supplementary Table S2). Interestingly, the downregulated genes included a number of metabolic enzymes and transporters involved in the glycolysis pathway (yellow, Figure 4A and B, Online Supplementary Figure 4B and C). In contrast, several enzymes involved in amino acid metabolism, such as ASS1 and ASNS, were upregulated after ALDH1A2 depletion (Figure 4A, Online Supplementary Figure S4A). The result was validated for each enzyme individually by qRT-PCR (Online Supplementary Figure S4D).

This finding prompted us to analyze the metabolic state of the T-ALL cells. We performed a capillary electrophoresis time-of-flight mass spectrometry analysis to measure the relative levels of 200 metabolites involved in the major metabolomics pathways after depletion of ALDH1A2 in Jurkat cells (Figure 4C). Strikingly, ALDH1A2 depletion resulted in a reduction in the intermediate metabolites involved in the glycolysis pathway (pink). Suppressed glucose metabolism was accompanied by a reduction in acetyl-CoA (yellow), which is one of the key carbon sources that drive the TCA cycle. We also observed a reduction in citric acid, *cis*-aconitic acid and isocitric acid (yellow), likely due to the decreased integration of glycolysis-derived acetyl-CoA into the TCA cycle.

It is noteworthy that the levels of metabolites derived

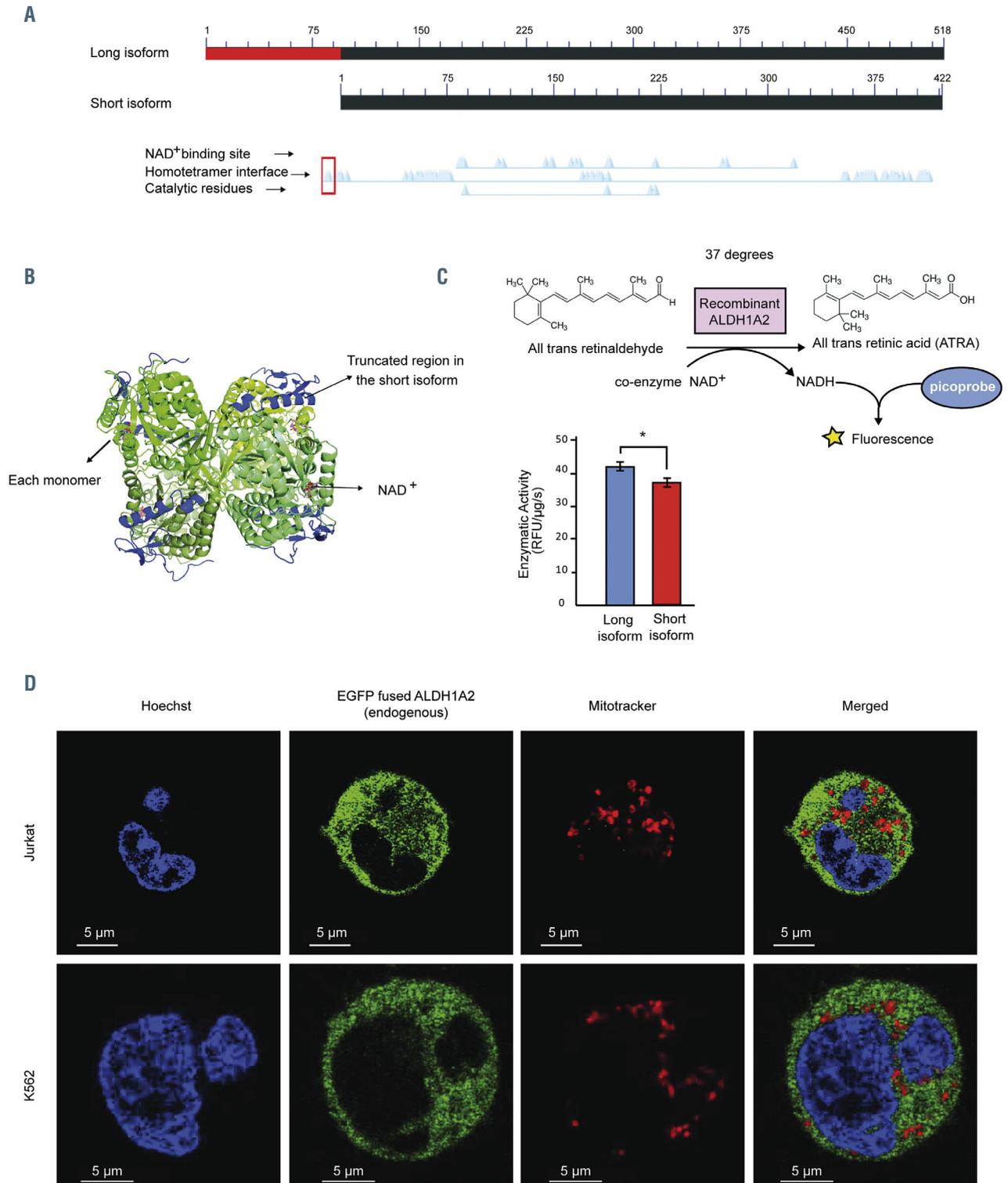


Figure 2. The short isoform of ALDH1A2 has enzymatic activity and is localized in the cytoplasm. (A) Protein domains of the ALDH1A2. NAD⁺ binding site, homotetramer interface and catalytic residues are indicated individually by blue triangles. The truncated short isoform lacks 96 amino acids at the N-terminus. The deleted tetramer interface of the short isoform is marked with a red block. (B) The crystal structure of the homotetrameric ALDH1A2 long isoform (PDB: 4X2Q) is presented as a cartoon model generated by PyMOL. Each monomer is colored in a different shade of green. Within a monomer, the region absent in the short isoform is colored blue. (C) A scheme showing the *in vitro* enzymatic assay. The enzymatic activity of each isoform was evaluated by the velocity of the kinetics curve (see *Online Supplementary Figure S3B*). Error bars represent the standard deviation for technical replicates. **P* < 0.05 using the two-tailed Student *t* test. (D) Confocal images showing the cellular localization of the endogenous EGFP-fused ALDH1A2 proteins in Jurkat and K562 cells. Hoechst and Mitotracker staining evidence the nucleus and mitochondria, respectively.

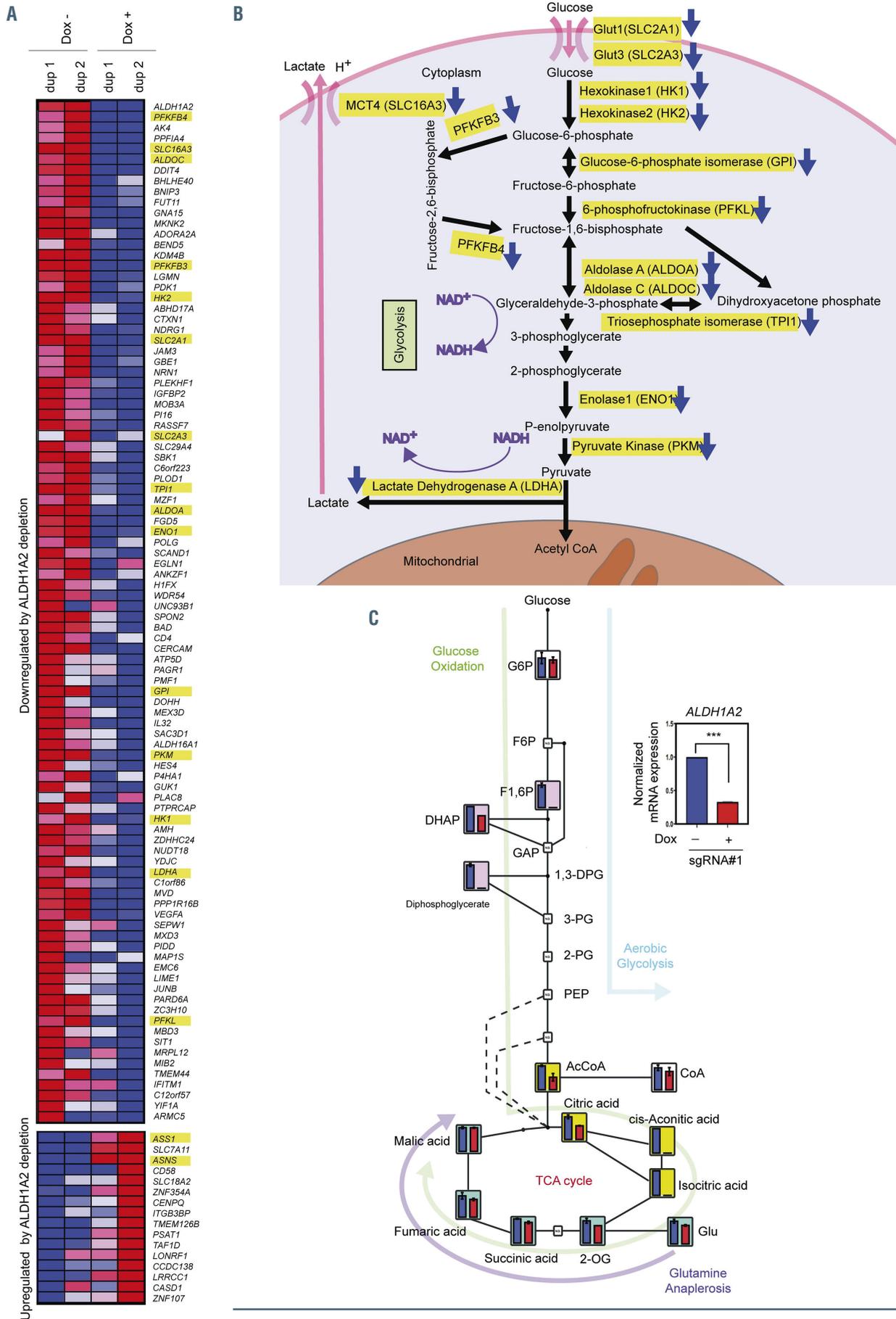


Figure 4. Legend on following page.

Figure 4. ALDH1A2 affects metabolic pathways in T-cell acute lymphoblastic leukemia cells. (A) sgRNA #1 was introduced into Jurkat cells by doxycycline (Dox) treatment for 72 h in duplicate (dup 1 and dup 2). Differentially expressed genes were selected based on the following criteria: $P < 0.05$, \log_2 fold change < -0.4 or > 0.4 , and mean transcript per million (TPM) value of the control > 250 . Heatmap image showing relative mRNA expression levels of the differentially expressed genes in the control (Dox-) and ALDH1A2-depleted cells (Dox+). The genes involved in the glycolysis pathway are highlighted. (B) A scheme showing the metabolic enzymes involved in the glycolysis pathway. The genes significantly downregulated by ALDH1A2 depletion are highlighted. (C) ALDH1A2 expression was depleted in Jurkat cells by Dox-inducible sgRNA-mediated transcriptional repression in biological duplicates. Downregulation of ALDH1A2 mRNA expression was confirmed by quantitative reverse transcription polymerase chain reaction. $***P < 0.001$ using the two-tailed Student *t* test. Relative amounts of the metabolites in the glycolysis pathway and the tricarboxylic acid cycle were measured by capillary electrophoresis time-of-flight mass spectrometry in the control cells (blue) and the ALDH1A2-depleted cell samples (red). Error bars represent the standard deviation for biological replicates.

from 2-oxoglutarate in the TCA cycle were not affected (blue). This finding suggests that the TCA cycle, suppressed by the loss of ALDH1A2, might be supplemented by glutaminolysis through which external glutamine is converted into 2-oxoglutarate. Glutamine has been reported to be the major carbon source for T-ALL cells with activated NOTCH1.^{35,34} Indeed, depletion of glutamine in culture media induced apoptosis, which was further increased by ALDH1A2 depletion after 24 h of doxycycline treatment (Online Supplementary Figure S4E and F). In the presence of glutamine, ALDH1A2 depletion alone did not induce apoptosis before 48 h of induction (Figure 3B, Online Supplementary Figure S4E and F). Thus, these results suggest that external glutamine might be used to replenish downstream metabolites to maintain the TCA cycle in the absence of ALDH1A2, and that ALDH1A2 and glutaminolysis pathways may compensate each other.

ALDH1A2 supports cellular aerobic glycolysis and energy production in T-cell acute lymphoblastic leukemia cells

Our results suggest that ALDH1A2 directly or indirectly affects major metabolic pathways. In particular, aerobic glycolysis is a hallmark of cancer metabolism and has also been reported to be activated in T-ALL cells.³⁵ Because metabolic profiling can represent only the static state, we then analyzed the dynamic state of these pathways after ALDH1A2 depletion using the Seahorse XF24 platform.

We first measured the extracellular acidification rate (ECAR), which is an indicator of lactic acid fermentation and thus reflects the activity of glycolysis. We incubated the cells under conditions of glucose starvation, added glucose, and then measured the ECAR in both the control and ALDH1A2-depleted cells. Strikingly, ALDH1A2 depletion inhibited aerobic glycolysis, as shown by the reduction of the baseline ECAR (“glycolysis” in Figure 5A). This trend was more significant upon the addition of oligomycin, which is an inhibitor of complex V in the electron transport chain and thus maximizes cellular aerobic glycolysis by disturbing mitochondrial respiration (“glycolytic capacity”). Treatment with 2-deoxy-D-glucose, which competitively inhibits the production of glucose-6-phosphate from glucose, completely abolished these activities in both control and ALDH1A2-depleted cells (“glycolytic reserve”), indicating that the effect of ALDH1A2 is glucose-dependent. These results demonstrated that the expression of ALDH1A2 promotes cellular glycolysis and contributes to the maintenance of metabolome plasticity by increasing cellular glycolytic capacity in T-ALL cells.

Because ALDH1A2 affects glycolysis and the TCA cycle, which are the major sources of energy production, we also analyzed the effect of ALDH1A2 depletion on oxidative phosphorylation in mitochondria. We measured the oxygen consumption rate to determine changes in the level of oxidative phosphorylation-dependent ATP generation. As expected, the basal respiration level was

decreased after ALDH1A2 depletion (Figure 5B). To further support this finding, we analyzed the ratio of NAD^+/NADH and the amount of ATP in the same setting. Consistently, depletion of ALDH1A2 increased the ratio of NAD^+/NADH , thus preventing NADH production (Figure 5C). ATP production normalized by the number of cells was also decreased by ALDH1A2 depletion (Figure 5D). Furthermore, in the cells cultured in glucose-free medium, ALDH1A2 depletion inhibited ATP production more strongly than it did in the control cells (Online Supplementary Figure S5A). These results indicate that ALDH1A2 supports energy production, which explains the cell phenotype after ALDH1A2 depletion.

ALDH1A2 reduces the level of reactive oxygen species in T-cell acute lymphoblastic leukemia cells

A high level of oxidative phosphorylation in mitochondria has been known to produce oxidative stress such as reactive oxygen species (ROS). On the other hand, previous studies provided evidence that ALDH family proteins help to alleviate intracellular ROS.³⁶⁻³⁸ The underlying mechanism is attributable to the ability of these proteins to clear cellular aldehydes, which are known ROS inducers.³⁹⁻⁴¹ Thus, we next measured the total level of intracellular ROS after depletion of ALDH1A2 in Jurkat cells using three different settings (sgRNA, shRNA and a small-molecule inhibitor).

Strikingly, genetic inhibition with shRNA (Figure 6A) or sgRNA (Figure 6B) resulted in significant increases in ROS levels. Similarly, treatment with WIN 18,446 increased the level of ROS in Jurkat cells (Figure 6C) as well as in primary leukemia cells that were harvested from a patient-derived xenograft mouse model (Figure 6D), both of which expressed only the short isoform of ALDH1A2 (Online Supplementary Figure S6A). Conversely, overexpression of the short isoform of ALDH1A2 reduced the level of ROS (Figure 6E). Furthermore, treatment with N-acetyl cysteine, an antioxidant, was able to reduce the ROS level after ALDH1A2 depletion in Jurkat cells (Online Supplementary Figure S6B). Importantly, in the setting of sgRNA knockout, ROS was increased at 48 h after doxycycline treatment (Figure 6B) before the induction of apoptosis was observed (Figure 3B). This suggests that accumulation of ROS could be a cause of apoptosis but not a consequence of cell death. These results indicate that although ALDH1A2 supports energy production, which potentially increases oxidative stress, it more predominantly plays a role protecting against the production of ROS and thus supports cell survival, which also explains the phenotype after ALDH1A2 depletion. Of note, N-acetyl cysteine treatment did not restore cell viability after ALDH1A2 depletion (*data not shown*), suggesting that the cell viability phenotype was mainly attributable to the maintenance of glycolysis and energy production.

ALDH1A2 overexpression accelerates tumorigenesis in a zebrafish model of T-cell acute lymphoblastic leukemia

Finally, we investigated the effect of *ALDH1A2* overexpression on tumorigenesis *in vivo* using a zebrafish model. We overexpressed the short isoform of the human *ALDH1A2* gene with a fluorescent marker (mCherry) under the *rag2* promoter in lymphocytes (Online Supplementary Figure S7A). We confirmed that transgenes were successfully integrated into the genome (Online Supplementary Figure S7B) and mCherry was expressed in the thymus (Online Supplementary Figure S7C). We then sorted mCherry-positive cells from the *ALDH1A2*-transgenic and control fish and measured the ROS levels. Strikingly, the ROS level was significantly lower in the

ALDH1A2-transgenic fish than in the control fish (Figure 7A), supporting our results in cell lines.

We next monitored tumor development using offsprings from one of the established founder lines. However, we did not observe any spontaneous tumor development in the *ALDH1A2* single transgenic fish (Online Supplementary Figure S7C). Hence, we then crossbred this line with a transgenic line overexpressing a myristoylated, constitutively active mouse *Akt2* gene (*myr-mAkt2*) which can cause T-ALL²² (Online Supplementary Figure S7B). Interestingly, overall penetrance was significantly increased up to 60% and tumor onset was slightly accelerated in the double transgenic animals as compared to the *myr-mAkt2* single transgenic animals (Figure 7B,

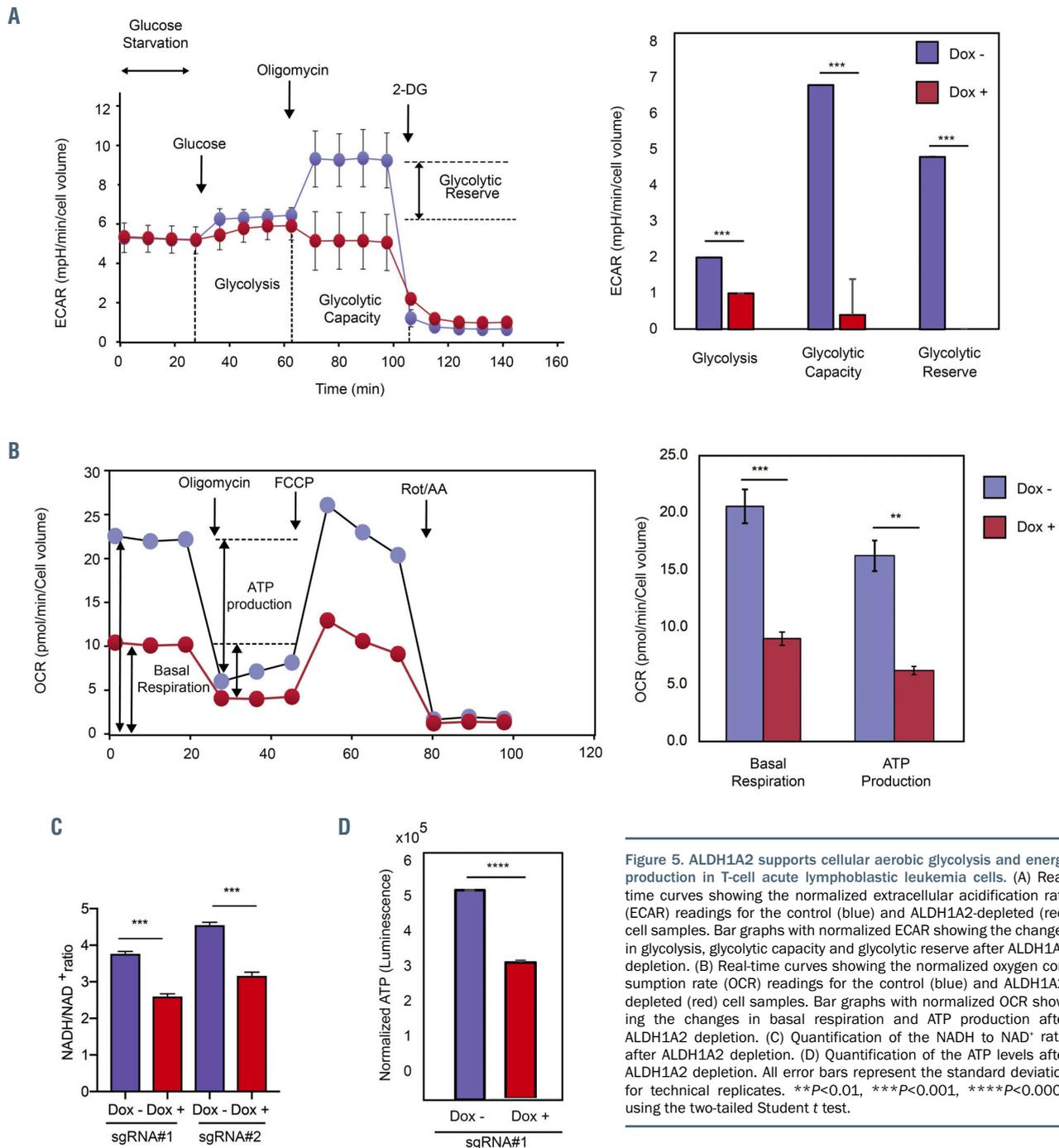


Figure 5. *ALDH1A2* supports cellular aerobic glycolysis and energy production in T-cell acute lymphoblastic leukemia cells. (A) Real-time curves showing the normalized extracellular acidification rate (ECAR) readings for the control (blue) and *ALDH1A2*-depleted (red) cell samples. Bar graphs with normalized ECAR showing the changes in glycolysis, glycolytic capacity and glycolytic reserve after *ALDH1A2* depletion. (B) Real-time curves showing the normalized oxygen consumption rate (OCR) readings for the control (blue) and *ALDH1A2*-depleted (red) cell samples. Bar graphs with normalized OCR showing the changes in basal respiration and ATP production after *ALDH1A2* depletion. (C) Quantification of the NADH to NAD⁺ ratio after *ALDH1A2* depletion. (D) Quantification of the ATP levels after *ALDH1A2* depletion. All error bars represent the standard deviation for technical replicates. ***P*<0.01, ****P*<0.001, *****P*<0.0001 using the two-tailed Student *t* test.

Online Supplementary Figure S7D). These results indicate that although overexpression of *ALDH1A2* alone does not have oncogenic capability, it can promote T-cell leukemogenesis induced by a driver oncogene.

Discussion

ALDH1A2 is one of the first reported downstream targets of TAL1 in T-ALL cells. Using the subtractive PCR method,

Ono *et al.* isolated *ALDH1A2* in T-ALL cell lines that co-expressed TAL1, LMO and GATA3.¹² Our ChIP-sequencing and RNA-sequencing experiments also demonstrated that *ALDH1A2* is directly activated by the TAL1 complex via an intronic regulatory element in T-ALL cells. The short isoform of *ALDH1A2* is specifically expressed in T-ALL cells. Importantly, *ALDH1A2* expression is highly specific to T-ALL cells, mostly to the *TAL1*-positive subgroup. Thus, using a novel approach, our studies reconfirmed that *ALDH1A2* is a signature gene of *TAL1*-positive T-ALL.

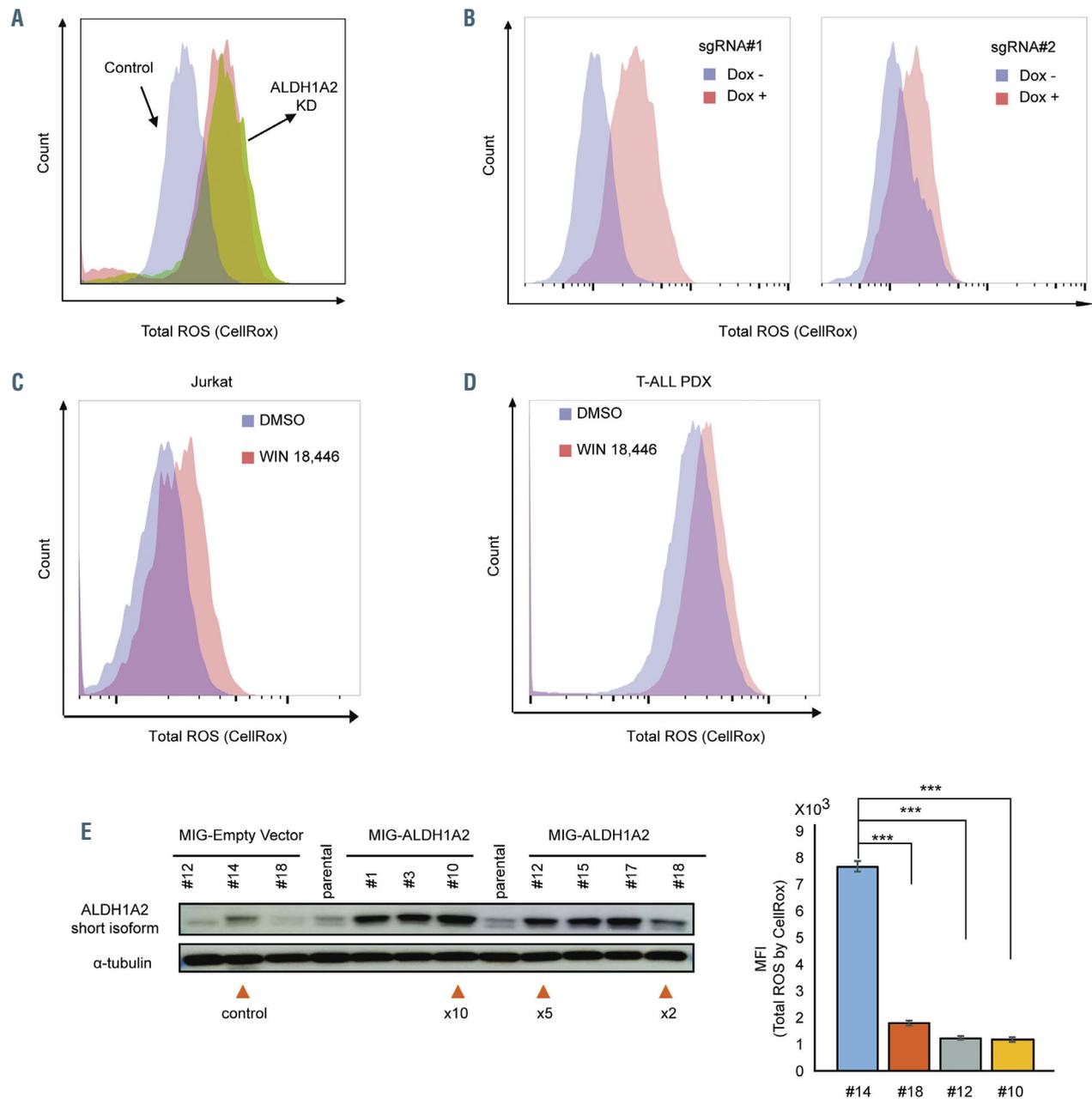


Figure 6. *ALDH1A2* reduces the levels of reactive oxygen species in T-cell acute lymphoblastic leukemia cells. (A, B) Total reactive oxygen species (ROS) levels were measured by flow cytometry analysis using CellROX staining at 72 h after shRNA-mediated *ALDH1A2* depletion (A) or at 48 h after sgRNA-mediated *ALDH1A2* depletion (B) in Jurkat cells. (C, D) Total ROS levels were measured by flow cytometry analysis using CellROX staining at 72 h after treatment with WIN 18,446 of Jurkat cells (C) or primary T-cell acute lymphoblastic leukemia cells harvested from a patient-derived xenograft model (D). (E) Jurkat cells were transduced with *ALDH1A2* cDNA or an empty vector by MSCV retrovirus infection. The level of expression of the short isoform of *ALDH1A2* in established Jurkat single clones was measured by western blot. MIG-ALDH1A2 #10, #12 and #18 were selected as the overexpressing clones, and they expressed different levels of *ALDH1A2*. MIG-EV #14 was selected as the control. Total ROS levels were measured by flow cytometry analysis using CellROX staining and are reported as mean fluorescence intensity (MFI). Error bars represent the standard deviation for biological replicates. ****P*<0.001 using the two-tailed Student *t* test.

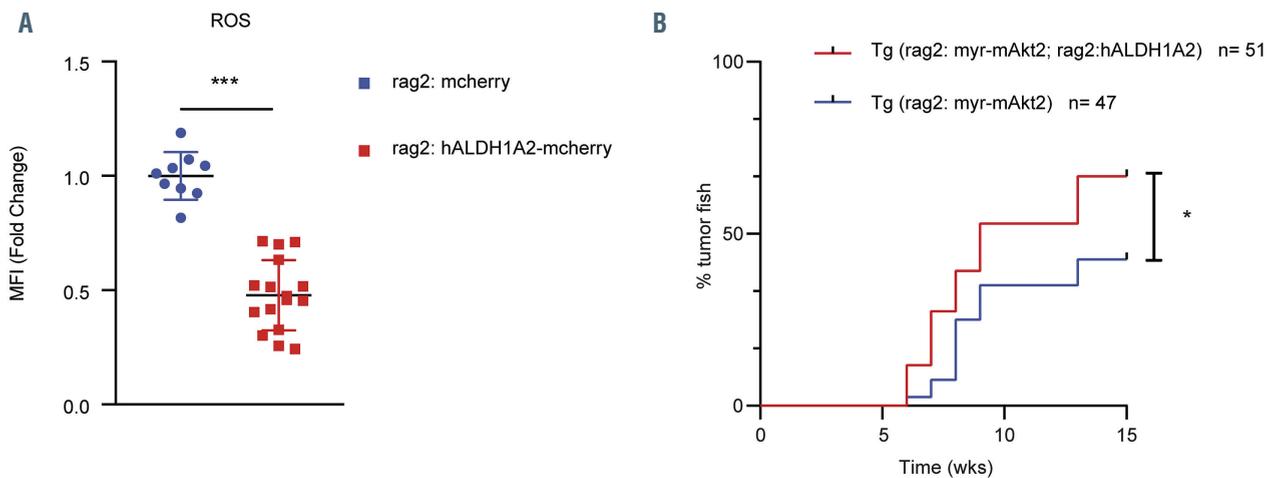


Figure 7. *ALDH1A2* overexpression accelerates tumorigenesis in a zebrafish model. (A) Total reactive oxygen species levels in the mCherry-positive cells from the control and *ALDH1A2*-transgenic zebrafish were measured by flow cytometry and shown by mean fluorescent intensity as the fold-change compared to control cell samples. (B) Tumor development (lymphoma-like and leukemia-like phenotypes) in the single and double transgenic zebrafish was recorded according to the criteria defined by Langenau *et al.*⁵⁰ Tumor onset and penetrance were evaluated by Kaplan-Meier curve analysis. * $P < 0.05$ using the Gehan-Breslow-Wilcoxon test.

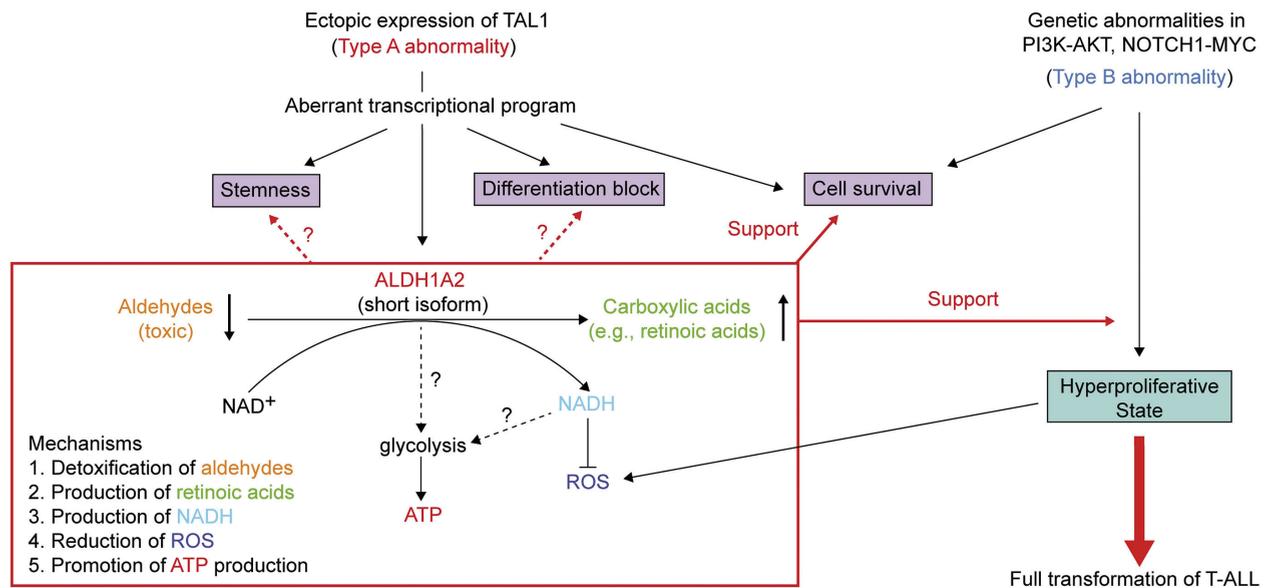


Figure 8. Scheme showing the oncorequisite role of *ALDH1A2* for full transformation of T-cell acute lymphoblastic leukemia.

However, the molecular functions and roles of *ALDH1A2* in the pathogenesis of T-ALL had not been elucidated previously. Here, we first experimentally proved that the T-ALL-specific isoform possesses an enzymatic activity that catalyzes retinaldehyde to retinoic acid with the production of NADH. Thus, the main mechanism involving this protein is characterized by ectopic expression through the activation of an alternative promoter bound by TAL1 rather than the expression of a dominant-negative protein or a loss of function. Additionally, we showed that this protein can reduce ROS levels and support energy production in T-ALL cells. Although the direct mechanism by which *ALDH1A2* promotes the glycolysis pathway remains unelucidated, one possibility is that the high levels of NADH caused by *ALDH1A2* overexpression might lead to pseudo-hypoxic

conditions that upregulate the expression of metabolic enzymes. Alternatively, the reduction in ROS level might lead to upregulation of metabolic enzymes. Further investigation is needed.

In malignant cells, regulation of ROS and metabolic state is crucial to maintain cell proliferation and survival. Although it has been reported that ROS can promote the proliferation and survival of T-ALL cells,⁴² it is maintained at low levels in the T-ALL leukemia-initiating cells in a mouse model.⁴³ It is also noteworthy that, in more than 50% of the *TAL1*-positive T-ALL cases, genetic abnormalities in the PI3K-AKT pathway and the NOTCH1-MYC pathway have been observed,^{4,8-11,44} and these pathways can promote several metabolic processes.⁴⁵ In particular, MYC has been known to promote glutaminolysis. Metabolic dependence on the TCA cycle

has also been reported in the *Myc*-induced zebrafish T-ALL model.⁴⁶ However, hypermetabolic and hyperproliferative states often induce ROS production through the upregulation of mitochondrial bioenergetics. Hence, malignant cells may need to balance the ROS level depending on their state of proliferation. Thus, ALDH1A2 expression may be ideal for leukemic or pre-leukemic clones to protect cells from ROS production and further promote metabolism (Figure 8).

Genetic abnormalities in T-ALL can be classified into “type A” and “type B” abnormalities. The former refer to abnormal expression of transcription factor genes, such as *TAL1*, that delineate distinct molecular pathways, which are more prevalent in one subgroup than others.^{4,47-49} The latter refer to oncogenic pathways that are commonly observed across different subgroups of T-ALL, for example, PI3K-AKT and NOTCH1-MYC.^{2,48,49} Considering the high occurrence of *ALDH1A2* expression with PI3K-AKT and NOTCH1-MYC abnormalities in the *TAL1*-positive subgroup, these pathways may compensate for each other and synergize, which would lead to further increases in cell proliferation and metabolism. *ALDH1A2* expression may be advantageous as a preceding event that serves as a “requisite” before the cells activate other oncogenic pathways, making it an “oncorequisite”.

Disclosures

No conflicts of interest to disclose.

References

- Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet*. 2008;371(9617):1030-1043.
- Aifantis I, Raetz E, Buonamici S. Molecular pathogenesis of T-cell leukaemia and lymphoma. *Nat Rev Immunol*. 2008;8(5):380-390.
- Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. *J Clin Oncol*. 2005;23(26):6306-6315.
- Liu Y, Easton J, Shao Y, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet*. 2017;49(8):1211-1218.
- Ferrando AA, Neuberg DS, Staunton J, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell*. 2002;1(1):75-87.
- Soulier J, Clappier E, Cayuela JM, et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). *Blood*. 2005;106(1):274-286.
- O'Neil J, Shank J, Cusson N, Murre C, Kelliher M. TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB. *Cancer Cell*. 2004;5(6):587-596.
- O'Neil J, Calvo J, McKenna K, et al. Activating Notch1 mutations in mouse models of T-ALL. *Blood*. 2006;107(2):781-785.
- Tremblay M, Tremblay CS, Herblot S, et al. Modeling T-cell acute lymphoblastic leukemia induced by the SCL and LMO1 oncogenes. *Genes Dev*. 2010;24(11):1093-1105.
- Gutierrez A, Sanda T, Grebliunaite R, et al. High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia. *Blood*. 2009;114(3):647-650.
- Bornschein S, Demeyer S, Stirparo R, et al. Defining the molecular basis of oncogenic cooperation between TAL1 expression and Pten deletion in T-ALL using a novel pro-T-cell model system. *Leukemia*. 2018;32(4):941-951.
- Ono Y, Fukuhara N, Yoshie O. TAL1 and LIM-only proteins synergistically induce retinaldehyde dehydrogenase 2 expression in T-cell acute lymphoblastic leukemia by acting as cofactors for GATA3. *Mol Cell Biol*. 1998;18(12):6939-6950.
- Sanda T, Lawton LN, Barrasa MI, et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell*. 2012;22(2):209-221.
- Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert Opin Drug Metab Toxicol*. 2008;4(6):697-720.
- Black W, Vasiliou V. The aldehyde dehydrogenase gene superfamily resource center. *Hum Genomics*. 2009;4(2):136-142.
- Vasiliou V, Pappa A, Estey T. Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism. *Drug Metab Rev*. 2004;36(2):279-299.
- Marcato P, Dean CA, Pan D, et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells*. 2011;29(1):32-45.
- Moreb JS. Aldehyde dehydrogenase as a marker for stem cells. *Curr Stem Cell Res Ther*. 2008;3(4):237-246.
- Marcato P, Dean CA, Giacomantonio CA, Lee PW. Aldehyde dehydrogenase: its role as a cancer stem cell marker comes down to the specific isoform. *Cell Cycle*. 2011;10(9):1378-1384.
- Moreb JS, Ucar D, Han S, et al. The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. *Chem Biol Interact*. 2012;195(1):52-60.
- Leong WZ, Tan SH, Ngoc PCT, et al. ARID5B as a critical downstream target of the TAL1 complex that activates the oncogenic transcriptional program and promotes T-cell leukemogenesis. *Genes Dev*. 2017;31(23-24):2343-2360.
- Gutierrez A, Grebliunaite R, Feng H, et al. Pten mediates Myc oncogene dependence in a conditional zebrafish model of T cell acute lymphoblastic leukemia. *J Exp Med*. 2011;208(8):1595-1603.
- Qian M, Zhang H, Kham SK, et al. Whole-transcriptome sequencing identifies a distinct subtype of acute lymphoblastic leukemia with predominant genomic abnormalities of EP300 and CREBBP. *Genome Res*. 2017;27(2):185-195.
- Aries IM, Bodaar K, Karim SA, et al. PRC2 loss induces chemoresistance by repressing apoptosis in T cell acute lymphoblastic leukemia. *J Exp Med*. 2018;215(12):3094-3114.
- Ghandi M, Huang FW, Jane-Valbuena J, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature*. 2019;569(7757):503-508.
- Casero D, Sandoval S, Seet CS, et al. Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. *Nat Immunol*. 2015;16(12):1282-1291.

27. Park JE, Botting RA, Dominguez Conde C, et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science*. 2020;367(6480):eaay3224.
28. Kernfeld EM, Genga RMJ, Neherin K, et al. A single-cell transcriptomic atlas of thymus organogenesis resolves cell types and developmental maturation. *Immunity*. 2018;48(6):1258-1270.e6.
29. Balkan W, Colbert M, Bock C, Linney E. Transgenic indicator mice for studying activated retinoic acid receptors during development. *Proc Natl Acad Sci U S A*. 1992;89(8):3347-3351.
30. Arnold SL, Kent T, Hogarth CA, et al. Pharmacological inhibition of ALDH1A in mice decreases all-trans retinoic acid concentrations in a tissue specific manner. *Biochem Pharmacol*. 2015;95(3):177-192.
31. Paik J, Haenisch M, Muller CH, et al. Inhibition of retinoic acid biosynthesis by the bisdichloroacetyldiamine WIN 18,446 markedly suppresses spermatogenesis and alters retinoid metabolism in mice. *J Biol Chem*. 2014;289(21):15104-15117.
32. Amory JK, Muller CH, Shimshoni JA, et al. Suppression of spermatogenesis by bisdichloroacetyldiamines is mediated by inhibition of testicular retinoic acid biosynthesis. *J Androl*. 2011;32(1):111-119.
33. O'Neil J, Grim J, Strack P, et al. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *J Exp Med*. 2007;204(8):1813-1824.
34. Herranz D, Ambesi-Impiombato A, Sudderth J, et al. Metabolic reprogramming induces resistance to anti-NOTCH1 therapies in T cell acute lymphoblastic leukemia. *Nat Med*. 2015;21(10):1182-1189.
35. Kishton RJ, Barnes CE, Nichols AG, et al. AMPK is essential to balance glycolysis and mitochondrial metabolism to control T-ALL cell stress and survival. *Cell Metab*. 2016;23(4):649-662.
36. Singh S, Brocker C, Koppaka V, et al. Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. *Free Radic Biol Med*. 2013;56:89-101.
37. Sun X, Zhu H, Dong Z, et al. Mitochondrial aldehyde dehydrogenase-2 deficiency compromises therapeutic effect of ALDH bright cell on peripheral ischemia. *Redox Biol*. 2017;13:196-206.
38. Kim J, Chen CH, Yang J, Mochly-Rosen D. Aldehyde dehydrogenase 2² knock-in mice show increased reactive oxygen species production in response to cisplatin treatment. *J Biomed Sci*. 2017;24(1):33.
39. Yang G, Ibuki Y. α,β -unsaturated aldehyde-induced delays in nucleotide excision repair and the contribution of reactive oxygen species. *Chem Res Toxicol*. 2018;31(2):145-155.
40. Novitskiy G, Traore K, Wang L, Trush MA, Mezey E. Effects of ethanol and acetaldehyde on reactive oxygen species production in rat hepatic stellate cells. *Alcohol Clin Exp Res*. 2006;30(8):1429-1435.
41. Tamura M, Ito H, Matsui H, Hyodo I. Acetaldehyde is an oxidative stressor for gastric epithelial cells. *J Clin Biochem Nutr*. 2014;55(1):26-31.
42. Silva A, Girio A, Cebola I, et al. Intracellular reactive oxygen species are essential for PI3K/Akt/mTOR-dependent IL-7-mediated viability of T-cell acute lymphoblastic leukemia cells. *Leukemia*. 2011;25(6):960-967.
43. Giambra V, Jenkins CR, Wang H, et al. NOTCH1 promotes T cell leukemia-initiating activity by RUNX-mediated regulation of PKC- θ and reactive oxygen species. *Nat Med*. 2012;18(11):1693-1698.
44. Hsieh AL, Walton ZE, Altman BJ, Stine ZE, Dang CV. MYC and metabolism on the path to cancer. *Semin Cell Dev Biol*. 2015;43:11-21.
45. Lien EC, Lyssiotis CA, Cantley LC. Metabolic reprogramming by the PI3K-Akt-mTOR pathway in cancer. *Recent Results Cancer Res*. 2016;207:39-72.
46. Anderson NM, Li D, Peng HL, et al. The TCA cycle transferase DLST is important for MYC-mediated leukemogenesis. *Leukemia*. 2016;30(6):1365-1374.
47. Seki M, Kimura S, Isobe T, et al. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat Genet*. 2017;49(8):1274-1281.
48. Belver L, Ferrando A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nat Rev Cancer*. 2016;16(8):494-507.
49. Iacobucci I, Mullighan CG. Genetic basis of acute lymphoblastic leukemia. *J Clin Oncol*. 2017;35(9):975-983.
50. Langenau DM, Traver D, Ferrando AA, et al. Myc-induced T cell leukemia in transgenic zebrafish. *Science*. 2003;299(5608):887-890.