TET proteins regulate Drosha expression and impact microRNAs in iNKT cells

Marianthi Gioulbasani[1,2,†], Tarmo Äijö[1, †], Jair E. Valenzuela[1,3], Julia Buquera Bettes[1], Ageliki Tsagaratou [1,4,5*]

[1] Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[2] School of Biology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[3] Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill and North Carolina State University, Raleigh, NC, USA
[4] Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[5] Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[†]These authors contributed equally to this work and share first co-authorship.

*** Correspondence:**
Corresponding Author
ageliki_tsagaratou@med.unc.edu

Keywords: TET proteins, 5hmC, iNKT, Drosha, microRNAs, PLZF

Abstract

DNA demethylases TET2 and TET3 play a fundamental role in thymic invariant natural killer T (iNKT) cell differentiation by mediating DNA demethylation of genes encoding for lineage specifying factors. Paradoxically, differential gene expression analysis revealed that significant number of genes were upregulated upon TET2 and TET3 loss in iNKT cells. This unexpected finding could be potentially explained if loss of TET proteins was reducing the expression of proteins that suppress gene expression. In this study, we discover that TET2 and TET3 synergistically regulate *Drosha* expression, by generating 5hmC across the gene body and by impacting chromatin accessibility. As DROSHA is involved in microRNA biogenesis, we proceed to investigate the impact of TET2/3 loss on microRNAs in iNKT cells. We report that among the downregulated microRNAs are members of the Let-7 family that downregulate *in vivo* the expression of the iNKT cell lineage specifying factor PLZF. Our data link TET proteins with microRNA expression and reveal an additional layer of TET mediated regulation of gene expression.

Introduction

Ten Eleven Translocation (TET) proteins are enzymes that regulate the process of DNA demethylation by oxidizing 5 methylcytosine (5mC) to generate 5 hydroxymethylcytosine

(5hmC) also known as the sixth base of our genome (1). In addition, TET proteins can oxidize 5hmC to generate additional modified cytosines, namely 5 formylcytosine (5fC) and 5 carboxylcytosine (5caC) (2, 3). The TET family of proteins consists of three members: TET1, that is most highly expressed in embryonic stem cells (ESCs), TET2, which is broadly expressed in various cell types and developmental stages, and TET3 that is more highly expressed as cells differentiate (4). All three TET proteins exert critical roles in shaping the development and function of a vast array of cells (5, 6). We have previously demonstrated that 5hmC is dynamically distributed across the genome of thymic T cell subsets (7). During the process of T cell lineage specification, 5hmC was shown to be increased in the gene body of very highly expressed genes and in active enhancers (7). Similar findings have been reported for murine and human peripheral T cells (7-10), indicating the critical role of TET proteins and 5hmC in regulating gene expression in T cells (5).

To dissect the *in vivo* roles of TET proteins in T cell development we generated *Tet2*-/-*Tet3*flx/flx CD4 cre (*Tet2/3* DKO) mice (11). We focused our analysis on concomitant deletion of TET2 and TET3 since our data indicated redundancy between TET proteins (11). The phenotype of the *Tet2/3* DKO mice was complex, revealing that TET proteins are critical regulators of various T cell types. Specifically, TET2 and TET3 are fundamental for the stability of the transcription factor (TF) FOXP3 and thus the functionality and stability of regulatory T cells (Tregs) (12). In addition, *Tet2/3* DKO mice exhibit a striking expansion of invariant natural killer (iNKT) T cells (11).

iNKT cells are an unconventional type of T cells that express an invariant TCR Vα14 chain and recognize lipids instead of peptides (13). iNKT cells develop in the thymus endowed already with effector properties and they have the ability to generate significant amount of cytokines, immediately upon antigen encounter (14). iNKT cell lineage commitment is orchestrated by the TF Promyelocytic leukaemia zinc finger (PLZF) protein, which endows iNKT cells with effector properties (15, 16). In the thymus, we can discern 3 subsets based on the expression of TFs and their functional properties: NKT2 express GATA3, NKT17 express RORγt and NKT1 express T-bet (17-20). iNKT cells exert important roles in recognition of bacterial pathogens and have been shown to be of clinical value in the context of cancer immunotherapy (14, 21-24). Thus, deciphering the molecular mechanisms that shape their differentiation and functionality is of outmost importance in order to take full advantage of their effector properties (25, 26).

We have previously demonstrated that *Tet2/3* DKO iNKT cells show increased expression of the TF RORγt (11, 27, 28). Integration of our genome wide datasets evaluating gene expression, whole genome methylation, whole genome hydroxymethylation and chromatin accessibility analysis in control and *Tet2/3* DKO iNKT cells revealed that TET2 and TET3, by regulating DNA demethylation, upregulate lineage specifying TFs such as T-bet and Th-POK that are critical for iNKT cell lineage diversification and for suppression of RORγt (11), in a TET2 dependent catalytic manner (29). However, not all the observed differences in the gene expression program of *Tet2/3* DKO iNKT cells can be attributed to gain of methylation in promoters or enhancers of the differentially expressed genes (11). That was particular true in the context of genes that were gaining expression upon loss of TET proteins (11). One possibility is that deletion of TET proteins can result in downregulation of repressors, allowing the upregulation of the targeted genes (30, 31). Repression of genes can occur by small RNAs that target mRNAs and can mediate their degradation (32). DROSHA regulates the generation of precursor miRNAs in the nucleus and then further processing occurs in the cytoplasm by DICER and the ARGONAUTE complex (33, 34). Notably, miRNAs are important for iNKT cell development as indicated by Dicer deficient mice (35, 36). In this study, we report that TET proteins regulate expression of *Drosha* in iNKT cells. We demonstrate that *Tet2/3* DKO iNKT cells show altered expression of precursor and mature miRNAs. Among the identified downregulated miRNAs are members of the Let-7 family that has been demonstrated *in vivo* to target and downregulate the transcription factor PLZF in iNKT cells (37).

Results and Discussion

Analysis of our previously published RNA-seq datasets (11) revealed that *Drosha* was downregulated in *Tet2/3* DKO iNKT cells (**Figure 1A**). To further dissect the molecular mechanisms by which TET2 and TET3 can impact expression of *Drosha* we assessed 5hmC distribution across the gene body. 5hmC upon treatment with bisulfite sequencing is converted to cytosine-5-methylenesulfonate (CMS) (38). Analysis of CMS immunoprecipitation with sequencing (CMS-IP seq) (39, 40) datasets (11) revealed that in wild type iNKT cells 5hmC is distributed across the gene body of *Drosha* (**Figure 1B**). We have previously demonstrated that 5hmC is enriched in the gene body of highly expressed genes, whereas the promoters of these genes are devoid of 5hmC, in conventional T cells

and unconventional iNKT cells (7, 11). Similar findings have been reported for naïve and helper T cell subsets (8-10) as well as for regulatory T cells (41). In addition, we have previously shown that 5hmC correlates with chromatin accessibility in both conventional and unconventional T cells (11, 29). We then investigated how loss of TET proteins may impact chromatin accessibility in the *Drosha* locus. Thus, we compared our datasets (11) for assay for transposase accessible chromatin with sequencing (ATAC-seq) (42) for wild type and *Tet2/3* DKO iNKT cells. We demonstrate that in *Tet2/3* DKO iNKT cells there is reduced accessibility in an intragenic genomic region (mm10: chr15:12,894,551-12,896,829) that has increased accessibility and enrichment of 5hmC in wild type iNKT cells (**Figure 1B**). Due to the low abundance of 5hmC in *Tet2/3* DKO thymic T cell subsets we were not able to perform CMS-IP seq for the *Tet2/3* DKO iNKT cells (11). However, we performed whole genome bisulfite sequencing (WGBS) in order to assess at single-nucleotide resolution the modification status of cytosine. Our analysis revealed a gain of methylation at this intragenic region in the *Drosha* locus at the *Tet2/3* DKO iNKT cells (**Figure 1B**).

We hypothesize that this intragenic region may exert regulatory function to promote the expression of *Drosha*. We have previously shown that 5hmC decorates active enhancers (7). Additional studies have demonstrated a strong correlation of 5hmC with active enhancers in various T cell subsets (8, 41). In many cases these regulatory elements that require 5hmC enrichment in order to be active are intragenic, such as the CNS2 enhancer in the *Foxp3* locus (12, 43, 44), an intragenic enhancer that regulates stable expression of *Cd4* gene in CD4 cells (45) as well as the proximal enhancer of *Zbtb7b* gene that encodes Th-POK (29). We have also demonstrated that 5hmC decorates intragenic cite A at the *Zbtb7b* locus to regulate the accessibility and the binding of the transcription factor GATA3 (29). It has been previously suggested that the binding of GATA3 to cite A promotes Th-POK expression (46).

We have previously discovered a shared gene expression program between *Tet2/3* DKO thymic iNKT cells and CD4 single positive (SP) cells (29). As DROSHA is expressed in both subsets we investigated if its expression was also affected in CD4 SP cells. We report that *Drosha* is downregulated in *Tet2/3* DKO CD4 SP cells (**Supplementary Figure 1**). Interestingly, there is 5hmC enrichment at the same intragenic site of the locus in WT CD4 SP cells (**Figure 1B**). In addition, we looked into our data assessing recruitment of GATA3 (by CUT&RUN) in WT and *Tet2/3* DKO CD4 SP cells (29). We discover that GATA3 binds in this region in WT CD4 SP cells, whereas no binding was detected in *Tet2/3* DKO CD4 SP cells. Moreover, we looked into the binding of Th-POK by using publicly available ChIP-seq

datasets (47) and we demonstrate binding of Th-POK in this potentially regulatory region in CD4 SP cells. Collectively, our findings suggest that TET2 and TET3 generate 5hmC and regulate chromatin accessibility in the *Drosha* locus to promote the expression of the gene (**Figure 1**). Further studies are required to elucidate the precise regulatory elements that control the expression of *Drosha*. However, as TET2 and TET3 deletion results in partial reduction of *Drosha* expression and not complete loss it becomes apparent that additional mechanisms are in place to control the expression of this gene.

As DROSHA is involved in regulating the pathway of microRNAs (miRNAs) we asked whether the reduced expression of *Drosha* has an impact on the miRNAs that are expressed in *Tet2/3* DKO iNKT cells. To identify small RNAs that are impacted we isolated thymic iNKT cells by FACS sorting (**Supplementary Figure 2**) from wild type or *Tet2/3* DKO mice and we performed small RNA-seq (**Figure 2A**). Comparison of precursor and mature miRNAs in the WT and the *Tet2/3* DKO iNKT samples confirmed that samples of the same genotype were similar to each other (**Supplementary Figure 3**). Our analysis compared expression of precursor (**Supplementary Table 1**) and mature miRNAs (**Supplementary Table 2**) and we found that among those that were differentially expressed the majority were downregulated (**Figure 2B, C**). This could be due to the downregulation of *Drosha* expression. We then focused on the affected mature miRNAs (**Figure 2C**). The vast majority of the differentially expressed mature miRNAs were downregulated in the *Tet2/3* DKO iNKT cells (**Figure 2C**). An additional mechanism could be that in the absence of TET proteins at least some miRNAs could gain cytosine methylation, resulting in their downregulated expression. However, when we looked into our previously generated WGBS data (11) we did not notice significant changes in methylation for the vast majority of the miRNAs that were differentially expressed in thymic iNKT cells. We only detected some gain of methylation in *mir199b* and *mir7058* (**Supplementary Figure 4**). Our analysis demonstrated that among the downregulated miRNAs were members of the Let-7 family. Specifically, we observed downregulation of Let-7c, Let-7b and Let-7k (**Figure 2C**). Interestingly, Let-7 miRNAs have been previously shown to target *Zbtb16* mRNA, which encodes for PLZF, for degradation in murine iNKT cells *in vivo* (37). Thus, we hypothesize that the downregulation of some of the members of the Let-7 family could result in increased expression of PLZF. We evaluated PLZF levels in WT and *Tet2/3* DKO iNKT cells by Flow cytometry (**Figure 3A**). Our data indicates that *Tet2/3* DKO iNKT cells exhibit upregulation of PLZF (**Figure 3A, B**). Thus, we propose that in *Tet2/3* DKO

iNKT cells downregulation of some of the Let-7 miRNAs results in reduced targeting for degradation of *Zbtb16* mRNA, resulting in increased expression of PLZF (**Figure 3C**).

Conclusions

In this study, we report that TET2 and TET3 regulate the expression of *Drosha*. We also discover various miRNAs that are differentially expressed including downregulation of Let-7 miRNAs. However, we must emphasize that the NKT17 skewing of the *Tet2/3* DKO iNKT cells can be fully rescued by deletion of ThPOK and partially rescued by deletion of T-bet as we have previously shown (11). Importantly, our unbiased, integrative analysis of genome wide datasets indicated that both ThPOK and Tbet are targets of TET proteins based on 5hmC enrichment and gain of methylation upon concomitant TET2 and TET3 loss (11). Thus, in support of our previous findings that TET proteins exert multifaceted roles in regulating gene expression (30, 31, 48), we propose an additional layer of TET-mediated regulation of lineage specification by affecting expression of miRNAs.

Methods

Mice

Mice were housed in pathogen free conditions in the Genetic Medicine Building at University of North Carolina (UNC) Chapel Hill in a facility managed by the Division of Comparative Medicine at UNC Chapel Hill. All the experiments using mice in this study were performed according to our approved protocol by the UNC Institutional Animal Care and Use Committee (protocol no: 22-252). Age and sex-matched mice were analyzed. Male and female mice were used for our experiments. Control (C57BL/6 (B6), strain number: 000664), RRID: IMSR_JAX: 000664) mice were purchased from Jackson (Jax) laboratories and were bred in our facility at UNC. Tet2-/- Tet3flx/flx CD4 cre mice have been previously described (11, 29). Briefly, Tet2-/- mice (49) (Jax strain no; 023359, RRID: IMSR_JAX:023359) were crossed with Tet3flx/flx (50, 51) (Jax strain no: 031015, RRID: IMSR_JAX:031015) CD4cre mice (52). To determine the genotype of the mice, tissue was isolated and genomic DNA was extracted using Phire Animal Tissue Direct PCR kit (Thermo scientific, cat no F-140WH) following the manufacturer's protocol. Then DNA fragments were amplified by PCR using the Phire DNA polymerase (Thermo scientific, cat no F-140WH) and specific primers using Biorad T100 or Biorad C1000 Touch thermocyclers.

195 **Cell preparation**

196 Thymocytes were isolated from young mice 21-25 days old. Thymocytes were dissociated to
197 prepare single cell suspensions as previously described (53, 54).
198
199 **Flow cytometry**
200 Thymocytes were stained with PBS-57 loaded tetramer PE (dilution 1:400, from NIH tetramer
201 core), TCRβ-PERCP/Cy5.5 (dilution 1:200, Biolegend, clone: H57-597, RRID: AB_1575173)
202 and dead cells were excluded by using a live/dead dye efluor 780 (dilution 1:1000,
203 eBioscience, cat. #65-0865-18) in FACS buffer (2% FBS in PBS) as described (54).
204 Intracellular staining for PLZF AlexaFluor 647(dilution 1:100, BD Pharmingen, clone: R17-
205 809, RRID: AB_2738238) was performed using the Foxp3 Transcription factor staining buffer
206 set (eBioscience, cat. no:00-5523-00)(54). Samples were analyzed in a Novocyte 3005 Flow
207 Cytometer (Agilent) using NovoExpress software (Agilent). Subsequently, the acquired data
208 were analysed and plots were generated using FlowJo (Treestar).
209
210 **FACS sorting**
211 Total thymocytes were stained with biotinylated mouse anti-CD24 (Biolegend, clone:M1/69,
212 RRID: AB_312837). CD24+ cells were depleted using mouse streptavidin magnetic beads
213 (anti-mouse Rapidspheres, cat no 19860, STEMCELL Technologies) following the
214 manufacturer's guidelines as previously described (29, 53). Enriched cells were stained with
215 efluor 780 viability dye (dilution 1:1000, eBioscience, cat. #65-0865-18), aGalactosyl-
216 Ceramide loaded tetramer (conjugated with PE, obtained from NIH tetramer core, dilution
217 1:400), TCRβ-PERCP/Cy5.5 (dilution 1:200, Biolegend, clone: H57-597, RRID:
218 AB_1575173). Live TCRβ$^+$, tetramer$^+$ cells were sorted and used to isolate RNA. The purity
219 of the samples after sorting was >98%. The cells were sorted using either a FACSAria II or a
220 FACSymphony S6 Sorter (Becton Dickinson).
221
222 **Statistical Analysis**
223 For the statistical analysis we used Prism software (Graphpad). We applied
224 unpaired student's *t* test. In the relevant figure legends, we indicated p-values for statistically
225 significant differences ($p < 0.05$). Data are mean ± s.e.m. In the graphs, each dot represents
226 a mouse. Unless otherwise indicated the p-value was not statistically significant. Differences

were considered significant when $p < 0.001$ (***); < 0.0001 (****). Both male and female mice from different litters were evaluated, with reproducible results.

RNA isolation, Library preparation of small RNAs and sequencing

FACS sorted iNKT cells were lysed in RLT plus lysis buffer from the miRNeasy plus kit (Qiagen, cat no: 217084). Total RNA was isolated following the instructions provided by the manufacturer and was quantified using Qubit RNA High Sensitivity assay (Invitrogen) in Qubit 4 Fluorometer (Invitrogen). Total RNA was provided to the UNC High Throughput Sequencing Facility (HTSF). RNA integrity was evaluated with a Tapestation (Agilent) using High Sensitivity RNA ScreenTape (Agilent). RNA with RIN value>9 was used for library preparation. Small RNA libraries were generated using the Revvity NETFLEX small RNA sequencing kit V4. Libraries were pooled and sequenced in an Illumina NextSeq2000 P1 Single End 1x50 to obtain 100 million single end reads. 6 biological replicates for wild type and 7 biological replicates for DKO samples were analyzed. Both male and female mice were evaluated.

Datasets used in this study

Small RNA-seq	GSE267135	This study
CMS-IP-seq in iNKT cells (5hmC mapping)	Series GSE66833 consists of the following 4 samples: GSM1632867: iNKT WT Biological Replicate (BR)1 CMS-Seq IP GSM1632868: iNKT WT BR2 CMS-Seq IP GSM1632869: iNKT WT BR1 CMS-Seq Input GSM1632870: iNKT WT BR2 CMS-Seq Input	(11)

Bulk RNA-seq (gene expression) in iNKT cells	From Series GSE66832 Thymic WT iNKT cells: GSM1632848: BR1 GSM1632849: BR2 GSM1632850: BR3 Thymic *Tet2/3* DKO iNKT cells: GSM1632851: BR1 GSM1632852: BR2 GSM1632852: BR3	(11)
Bulk RNA-seq in CD4 SP cells	GSE190230 WT CD4 SP: GSM5718501 GSM5718502 GSM5718503 *Tet2/3* DKO CD4 SP: GSM5718504 GSM5718505	(29)
ATAC-seq in iNKT cells (chromatin accessibility)	GSE85743 WT thymic iNKT cells: GSM2283399: BR1 GSM2283400: BR2 GSM2283401: BR3 *Tet2/3* DKO thymic iNKT cells:	(11)

	GSM2283402: BR1 GSM2283403: BR2 GSM2283404: BR3	
WGBS in iNKT cells	Series GSE72116 WT thymic iNKT cells GSM1855585: BR1 GSM1855586: BR2 *Tet2/3* DKO thymic iNKT cells GSM1855587: BR1 GSM1855588: BR2	(11)
GATA3 CUT&RUN in CD4 SP	GSE190228 WT CD4 SP: GSM5718477: BR1 GSM5718478: BR2 GSM5718475: BR3 GSM5718476: IgG control *Tet2/3* DKO CD4 SP GSM5718480: BR1 GSM5718481: BR2 GSM5718482: IgG control	(29)
ThPOK ChIP-seq in CD4 SP	GSM4486880: BR1 GSM4486881: BR2 GSM4486882: BR3 GSM4486883: control	(47)

Small RNA-seq data analysis

The small RNA samples were processed using nf-core/smrnaseq (2.2.4) using default parameters (55, 56). The differential expression analysis was done using nf-core/differentialabundance (1.4.0) using default parameters (57).

CMS-seq data analysis

The CMS-IP and input reads from 2 biological replicates of WT iNKT cells were mapped against mm10 using Bismark (0.22.3) (58). The mapping was done using the Bowtie 2 (2.4.1) (59) backend in the paired-end mode with the following parameter values: -I 0 -X 600 -N 0. The coverage tracks were generated using HOMER (4.10) (makeBigWig.pl -norm 1e6) (60).

ATAC-seq data analysis

Adapter trimming and quality filtering of the sequencing libraries (3 biological replicates per genotype, 6 samples in total) was done using fastp (0.21.0)(61) with the default parameters. The sequencing libraries were mapped against mm10 using Bowtie 2 (2.4.1) (–very-sensitive -X 2000)(59). Mitochondrial reads were removed after alignment. Additional filtering was done using samtools (1.12)(62) using the following parameter values: -q 30 -h -b -F 1804 -f 2. Reads with identical sequences were filtered and only one was retained for subsequent analysis. The coverage tracks were generated from the samples obtained by pooling the biological replicates using HOMER (4.10) (makeBigWig.pl -norm 1e6)(60).

WGBS, CUT&RUN and ChIP-seq data analysis has been previously described (29).

Figure Legends

Figure 1. TET2 and TET3 regulate expression of *Drosha* in thymic iNKT cells. A. Gene expression of *Drosha* in WT (*in black*) and *Tet2/3* DKO thymic iNKT cells (*in purple*), evaluated by RNA-seq. 3 biological replicates per genotype were assessed. *** ($p = 0.0004$), unpaired t test. Each dot represents an individual biological replicate. Horizontal lines indicate the mean (s.e.m.). **B.** Portraits of epigenetic regulation (determined by 5hmC, 5mC and chromatin accessibility) in thymic iNKT cells and transcriptional regulation in CD4 SP cells of the *Drosha* locus. Genome browser view of 5hmC distribution (by CMS-IP seq) in the gene body of *Drosha* in WT iNKT cells reveals enrichment of this modification indicating TET activity. 2 biological replicates were analyzed.

Evaluation of 5mC by WGBS in WT and *Tet2/3* DKO iNKT cells. Note the gain of methylation in *Tet2/3* DKO iNKT cells in the indicated region that coincides with enrichment of 5hmC at WT iNKT cells (mm10: chr15:12,894,551-12,896,829). 2 biological replicates of WGBS per genotype were evaluated.

Portraits of chromatin accessibility (assessed by ATAC-seq) of the *Drosha* locus in WT and *Tet2/3* DKO thymic iNKT cells. Peaks indicate accessibility. Note the loss of accessibility in the *Tet2/3* DKO iNKT cells at the indicated site that is enriched for 5hmC in the WT iNKT cells. 3 biological replicates per genotype were evaluated.

5hmC distribution (determined by CMS-IP seq) in the gene body of *Drosha* in WT CD4 SP cells reveals enrichment of this modification in the same intragenic site. 2 biological replicates were evaluated.

GATA3 binds in this potentially regulatory site in WT CD4 SP cells as determined by CUT&RUN, whereas GATA3 binding is not detected in *Tet2/3* DKO CD4 SP cells. 3 biological replicates for WT CD4 cells and 2 biological replicates for *Tet2/3* DKO were analyzed.

ThPOK binding is detected by ChIP-seq in WT CD4 SP cells at the indicated intragenic site. 3 biological replicates were evaluated.

The arrows indicate the direction of transcription.

Figure 2. Differential expression of precursor (hairpin) and mature miRNAs in WT and *Tet2/3* DKO thymic iNKT cells. A. Experimental outline. **B.** Heatmap indicating hairpin miRNAs whose adjusted p-value < 0.05 and absolute log2 fold-change > 2. The z-score normalized expression values are shown. **C.** Heatmap indicating mature miRNAs whose adjusted p-value < 0.05 and absolute log2 fold-change > 2. The z-score normalized expression values are shown. Both male and female mice were used for each genotype. N=6 WT mice and N=7 *Tet2/3* DKO mice were used.

Figure 3. Let-7 miRNAs downregulation in *Tet2/3* DKO thymic iNKT cells contributes in upregulation of PLZF.

A. Representative flow cytometry plots of thymocytes isolated from wild type and *Tet2/3* DKO mice identify iNKT cells as aGalCer-loaded tetramer$^+$ and TCRβ intermediate cells. Representative histogram for the lineage specifying transcription factor PLZF indicates increased expression, determined by intracellular staining and Flow cytometry, in the *Tet2/3* DKO thymic iNKT samples (*in purple*) compared to WT (*in black*) counterparts. **B.** Plot

comparing the median fluorescence intensity (MFI) of PLZF expression in WT (*in black*) and *Tet2/3* DKO (*in purple*) iNKT cells. 4 biological replicates per genotype were assessed. ** (p =0.0004), unpaired t test. Each dot represents an individual biological replicate. Horizontal lines indicate the mean (s.e.m.). **C.** Model for TET mediated regulation of members of the Let-7 miRNA family to impact PLZF expression in thymic iNKT cells.

Supplementary Figure 1: TET2 and TET3 regulate expression of *Drosha* in thymic CD4 SP cells. A. Gene expression of *Drosha* in WT (*in black*) and *Tet2/3* DKO thymic iNKT cells (*in purple*), evaluated by RNA-seq. 3 biological replicates for WT and 2 biological replicates for Tet2/3 DKO CD4 SP cells were assessed. *** (p =0.0003), unpaired t test. Each dot represents an individual biological replicate. Horizontal lines indicate the mean (s.e.m.).

Supplementary Figure 2: **Sorting strategy and purity assessment of FACS sorted iNKT cells. A.** Representative flow cytometry plots indicating the gating selection to exclude doublets and isolate live (LD APC/CY7 negative), wild type iNKT cells (aGalCer loaded tetramer positive, TCRβ intermediate cells) by FACS sorting. **B.** FACS plots indicating purity of a representative sample of wild type iNKT cells after FACS sorting. **C.** As in **A.** for *Tet2/3* DKO iNKT sample. **D.** FACS plots indicating purity of a representative sample of *Tet2/3* DKO iNKT cells after FACS sorting.

Supplementary Figure 3: **Evaluating similarity of RNA samples**. **A.** Dendrograms indicating the clustering of precursor miRNAs and **B.** mature miRNAs samples.

Supplementary Figure 4: **Methylation portraits in mature miRNAs.** Assessing cytosine methylation by WGBS revealed some gain of methylation in *Tet2/3* DKO iNKT cells in two of the mature miRNAs that were downregulated in *Tet2/3* DKO iNKT cells. 5mC distribution in WT and *Tet2/3* DKO iNKT cells for A. *Mir199b* and B. *Mir7058*.

Supplementary Table 1: Results of differential gene expression analysis of hairpin miRNAs in wild type and *Tet2/3* DKO iNKT cells.

Supplementary Table 2: Results of differential gene expression analysis of mature miRNAs in wild type and *Tet2/3* DKO iNKT cells.

Data availability statement

The datasets generated for this study have been deposited in the Gene Expression Omnibus (GEO) public repository with the accession code: GSE267135.

Author Contributions

MG performed cell isolation, Flow cytometry, prepared samples for FACS sorting, isolated RNA and contributed in maintaining relevant mouse colonies. TÄ analyzed RNA-seq, CMS-IP seq, WGBS, CUT&RUN, ChIP-seq and small RNA-seq data, generated figures and wrote the relevant method sections. JEV and JBB assisted in Flow cytometry experiments and mouse colony management. AT conceived the study, performed experiments, analyzed data, supervised research, acquired funding and wrote the manuscript. All authors agree with the content of the manuscript.

Funding

This study was supported by NIH grant (R35-GM138289), Supplement 3R35-GM138289-02S1 from National Institute of General Medicinal Sciences (NIGMS), and UNC Lineberger Comprehensive Cancer Center Startup funds (to AT). Research reported in this publication and related to FACS sorting was supported in part by the North Carolina Biotech Center Institutional Support Grant 2012-IDG-1006. UNC Flow Cytometry Core and UNC High Throughput Sequencing core (HTSF) are affiliated to UNC Lineberger Comprehensive Cancer Center and are supported in part by P30 CA016086 Cancer Center Core Support Grant to the UNC Lineberger Comprehensive Cancer Center.

Acknowledgments

We acknowledge Ms. Theresa Hegarty (UNC DCM Colony Management Core) for excellent mouse colony management. We thank Ms. Janet Dow, Mr. Roman Bandy and Ms. Ayrianna Woody of the UNC Flow Cytometry Core for FACS sorting. We thank the UNC HTSF for preparing small RNA-seq libraries and for performing sequencing. We are grateful to the NIH tetramer core for generously providing PBS-57 loaded mouse CD1d tetramers. Some of the images were generated with Biorender.

Conflict of Interest

Dr. Tarmo Äijo is a Director of Data Science at Covera Health. No funding from Covera Health was received for this study.

References

1. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-5.

2. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011;333(6047):1300-3.

3. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011;333(6047):1303-7.

4. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol. 2013;14(6):341-56.

5. Tsiouplis NJ, Bailey DW, Chiou LF, Wissink FJ, Tsagaratou A. TET-Mediated Epigenetic Regulation in Immune Cell Development and Disease. Front Cell Dev Biol. 2020;8:623948.

6. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet. 2017;18(9):517-34.

7. Tsagaratou A, Aijo T, Lio CW, Yue X, Huang Y, Jacobsen SE, et al. Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. Proc Natl Acad Sci U S A. 2014;111(32):E3306-15.

8. Ichiyama K, Chen T, Wang X, Yan X, Kim BS, Tanaka S, et al. The methylcytosine dioxygenase Tet2 promotes DNA demethylation and activation of cytokine gene expression in T cells. Immunity. 2015;42(4):613-26.

9. Nestor CE, Lentini A, Hagg Nilsson C, Gawel DR, Gustafsson M, Mattson L, et al. 5-Hydroxymethylcytosine Remodeling Precedes Lineage Specification during Differentiation of Human CD4(+) T Cells. Cell Rep. 2016;16(2):559-70.

10. Vincenzetti L, Leoni C, Chirichella M, Kwee I, Monticelli S. The contribution of active and passive mechanisms of 5mC and 5hmC removal in human T lymphocytes is differentiation- and activation-dependent. Eur J Immunol. 2019;49(4):611-25.

11. Tsagaratou A, Gonzalez-Avalos E, Rautio S, Scott-Browne JP, Togher S, Pastor WA, et al. TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. Nat Immunol. 2017;18(1):45-53.

12. Yue X, Trifari S, Aijo T, Tsagaratou A, Pastor WA, Zepeda-Martinez JA, et al. Control of Foxp3 stability through modulation of TET activity. J Exp Med. 2016;213(3):377-97.

13. Bendelac A, Savage PB, Teyton L. The biology of NKT cells. Annu Rev Immunol. 2007;25:297-336.

14. Crosby CM, Kronenberg M. Tissue-specific functions of invariant natural killer T cells. Nat Rev Immunol. 2018;18(9):559-74.

15. Savage AK, Constantinides MG, Han J, Picard D, Martin E, Li B, et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. Immunity. 2008;29(3):391-403.

16. Kovalovsky D, Uche OU, Eladad S, Hobbs RM, Yi W, Alonzo E, et al. The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. Nat Immunol. 2008;9(9):1055-64.

17. Constantinides MG, Bendelac A. Transcriptional regulation of the NKT cell lineage. Curr Opin Immunol. 2013;25(2):161-7.

18. Lee YJ, Holzapfel KL, Zhu J, Jameson SC, Hogquist KA. Steady-state production of IL-4 modulates immunity in mouse strains and is determined by lineage diversity of iNKT cells. Nat Immunol. 2013;14(11):1146-54.

19. Wang H, Hogquist KA. How Lipid-Specific T Cells Become Effectors: The Differentiation of iNKT Subsets. Front Immunol. 2018;9:1450.

20. Krovi SH, Gapin L. Invariant Natural Killer T Cell Subsets-More Than Just Developmental Intermediates. Front Immunol. 2018;9:1393.

21. Heczey A, Liu D, Tian G, Courtney AN, Wei J, Marinova E, et al. Invariant NKT cells with chimeric antigen receptor provide a novel platform for safe and effective cancer immunotherapy. Blood. 2014;124(18):2824-33.

22. Rotolo A, Caputo VS, Holubova M, Baxan N, Dubois O, Chaudhry MS, et al. Enhanced Anti-lymphoma Activity of CAR19-iNKT Cells Underpinned by Dual CD19 and CD1d Targeting. Cancer Cell. 2018;34(4):596-610 e11.

23. Delfanti G, Cortesi F, Perini A, Antonini G, Azzimonti L, de Lalla C, et al. TCR-engineered iNKT cells induce robust antitumor response by dual targeting cancer and suppressive myeloid cells. Sci Immunol. 2022;7(74):eabn6563.

24. Cortesi F, Delfanti G, Grilli A, Calcinotto A, Gorini F, Pucci F, et al. Bimodal CD40/Fas-Dependent Crosstalk between iNKT Cells and Tumor-Associated Macrophages Impairs Prostate Cancer Progression. Cell Rep. 2018;22(11):3006-20.

25. Morgan RC, Kee BL. Genomic and Transcriptional Mechanisms Governing Innate-like T Lymphocyte Development. J Immunol. 2022;209(2):208-16.

26. Verykokakis M, Kee BL. Transcriptional and epigenetic regulation of innate-like T lymphocyte development. Curr Opin Immunol. 2018;51:39-45.

27. Tsagaratou A. TET mediated epigenetic regulation of iNKT cell lineage fate choice and function. Mol Immunol. 2018;101:564-73.

28. Tsagaratou A. Unveiling the regulation of NKT17 cell differentiation and function. Mol Immunol. 2019;105:55-61.

29. Aijo T, Theofilatos D, Cheng M, Smith MD, Xiong Y, Baldwin AS, et al. TET proteins regulate T cell and iNKT cell lineage specification in a TET2 catalytic dependent manner. Front Immunol. 2022;13:940995.

30. Tsagaratou A. Deciphering the multifaceted roles of TET proteins in T-cell lineage specification and malignant transformation. Immunol Rev. 2021;300(1):22-36.

31. Tsagaratou A. TET Proteins in the Spotlight: Emerging Concepts of Epigenetic Regulation in T Cell Biology. Immunohorizons. 2023;7(1):106-15.

32. Wilson RC, Doudna JA. Molecular mechanisms of RNA interference. Annu Rev Biophys. 2013;42:217-39.

33. Yates LA, Norbury CJ, Gilbert RJ. The long and short of microRNA. Cell. 2013;153(3):516-9.

34. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol. 2013;14(8):475-88.

35. Fedeli M, Napolitano A, Wong MP, Marcais A, de Lalla C, Colucci F, et al. Dicer-dependent microRNA pathway controls invariant NKT cell development. J Immunol. 2009;183(4):2506-12.

36. Zhou L, Seo KH, He HZ, Pacholczyk R, Meng DM, Li CG, et al. Tie2cre-induced inactivation of the miRNA-processing enzyme Dicer disrupts invariant NKT cell development. Proc Natl Acad Sci U S A. 2009;106(25):10266-71.

37. Pobezinsky LA, Etzensperger R, Jeurling S, Alag A, Kadakia T, McCaughtry TM, et al. Let-7 microRNAs target the lineage-specific transcription factor PLZF to regulate terminal NKT cell differentiation and effector function. Nat Immunol. 2015;16(5):517-24.

38. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One. 2010;5(1):e8888.

39. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature. 2011;473(7347):394-7.

40. Huang Y, Pastor WA, Zepeda-Martinez JA, Rao A. The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. Nat Protoc. 2012;7(10):1897-908.

41. Yue X, Samaniego-Castruita D, Gonzalez-Avalos E, Li X, Barwick BG, Rao A. Whole-genome analysis of TET dioxygenase function in regulatory T cells. EMBO Rep. 2021;22(8):e52716.

42. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213-8.

43. Yang R, Qu C, Zhou Y, Konkel JE, Shi S, Liu Y, et al. Hydrogen Sulfide Promotes Tet1- and Tet2-Mediated Foxp3 Demethylation to Drive Regulatory T Cell Differentiation and Maintain Immune Homeostasis. Immunity. 2015;43(2):251-63.

44. Nair VS, Song MH, Ko M, Oh KI. DNA Demethylation of the Foxp3 Enhancer Is Maintained through Modulation of Ten-Eleven-Translocation and DNA Methyltransferases. Mol Cells. 2016;39(12):888-97.

45. Issuree PD, Day K, Au C, Raviram R, Zappile P, Skok JA, et al. Stage-specific epigenetic regulation of CD4 expression by coordinated enhancer elements during T cell development. Nat Commun. 2018;9(1):3594.

46. Wang L, Wildt KF, Zhu J, Zhang X, Feigenbaum L, Tessarollo L, et al. Distinct functions for the transcription factors GATA-3 and ThPOK during intrathymic differentiation of CD4(+) T cells. Nat Immunol. 2008;9(10):1122-30.

47. Chopp LB, Gopalan V, Ciucci T, Ruchinskas A, Rae Z, Lagarde M, et al. An Integrated Epigenomic and Transcriptomic Map of Mouse and Human alphabeta T Cell Development. Immunity. 2020;53(6):1182-201 e8.

48. Theofilatos D, Ho T, Waitt G, Aijo T, Schiapparelli LM, Soderblom EJ, et al. Deciphering the TET3 interactome in primary thymic developing T cells. iScience. 2024;27(5):109782.

49. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc Natl Acad Sci U S A. 2011;108(35):14566-71.

50. Kang J, Lienhard M, Pastor WA, Chawla A, Novotny M, Tsagaratou A, et al. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci U S A. 2015;112(31):E4236-45.

51. Ko M, An J, Pastor WA, Koralov SB, Rajewsky K, Rao A. TET proteins and 5-methylcytosine oxidation in hematological cancers. Immunol Rev. 2015;263(1):6-21.

52. Lee PP, Fitzpatrick DR, Beard C, Jessup HK, Lehar S, Makar KW, et al. A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. Immunity. 2001;15(5):763-74.

53. Theofilatos D, Aijo T, Tsagaratou A. Protocol to isolate mature thymic T cell subsets using fluorescence-activated cell sorting for assessing gene expression by RNA-seq and transcription factor binding across the genome by CUT&RUN. STAR Protoc. 2022;3(4):101839.

54. Gioulbasani M, Tsagaratou A. Defining iNKT Cell Subsets and Their Function by Flow Cytometry. Curr Protoc. 2023;3(7):e838.

55. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol. 2020;38(3):276-8.

56. Alexander Peltzer LP, Phil Ewels, Chuan Wang, Jose Espinosa-Carrasco, Kevin Menden, nf-core bot, Christopher Mohr, Harshil Patel, Gregor Sturm, CKComputomics, Lluc Cabús, Kevin L. Keys, Sébastien Guizard, Maxime U Garcia, Robert Syme, Adam Talbot, Erik Danielsson, Konrad Stawiski MD PhD, … Paolo Di Tommaso. nf-core/smrnaseq: v2.3.0 - 2024-02-23 - Gray Zinc Dalmatian (2.3.0). : Zenodo; 2024 [

57. WackerO JM, Azedine Zoufir, nf-core bot, Alexander Peltzer, Cristina Tuñí i Domínguez, Dave Carlson, Steffen Möller, Marcel Ribeiro-Dantas, Harshil Patel, & James A. Fellows Yates. nf-core/differentialabundance: v1.4.0 - 2023-11-27 (1.4.0). Zenodo; 2023 [

58. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27(11):1571-2.

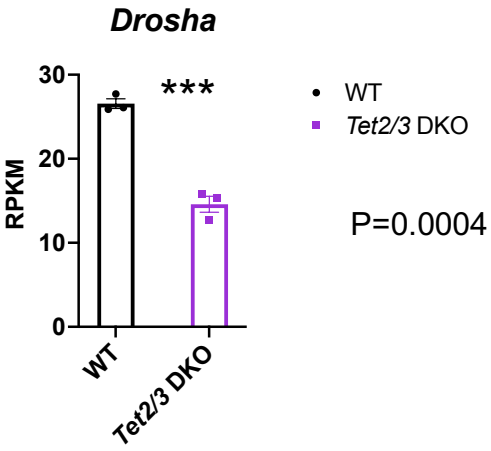
59. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.

60. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.

61. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.

62. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

A. Total iNKTs



B.

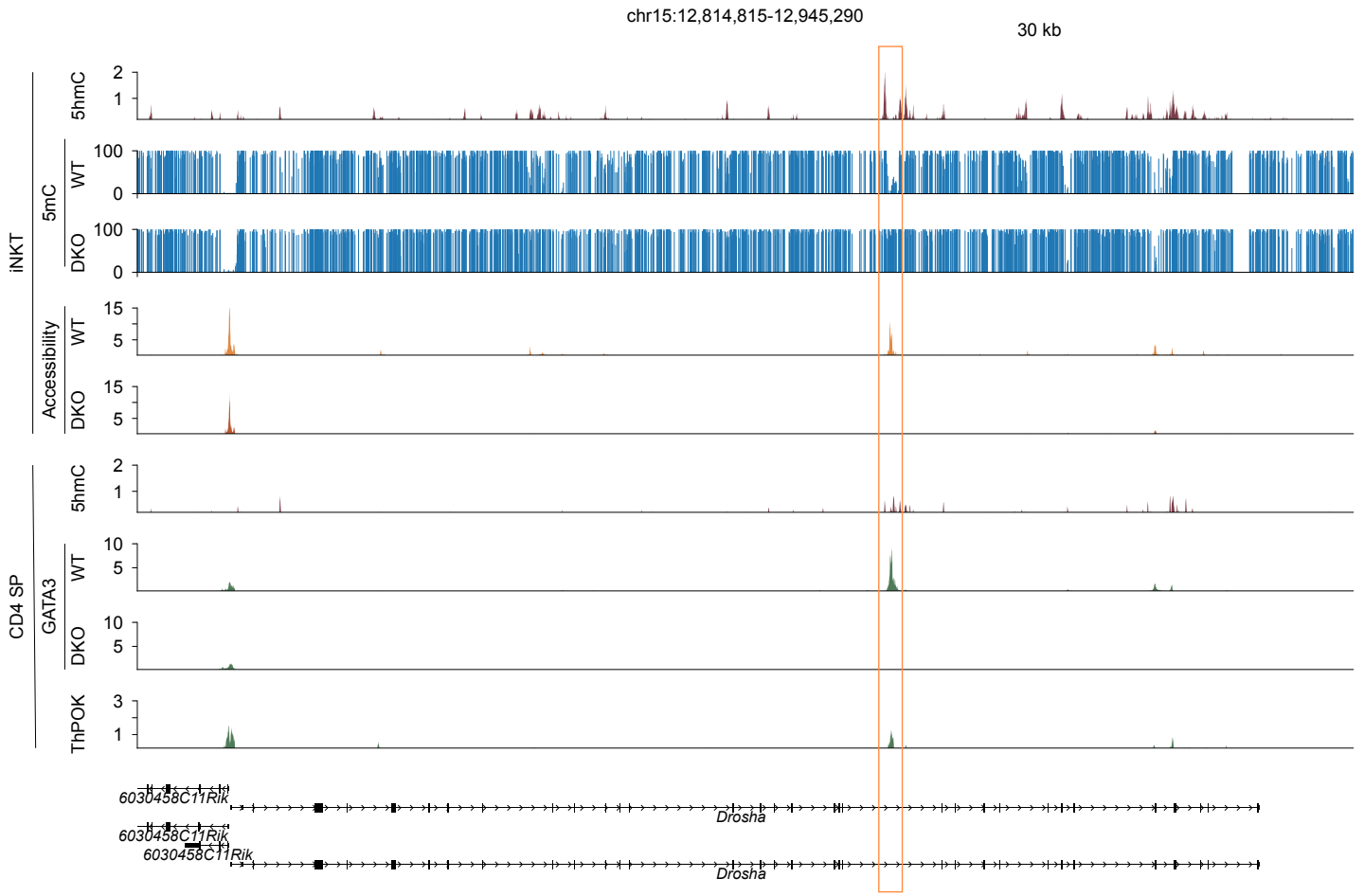
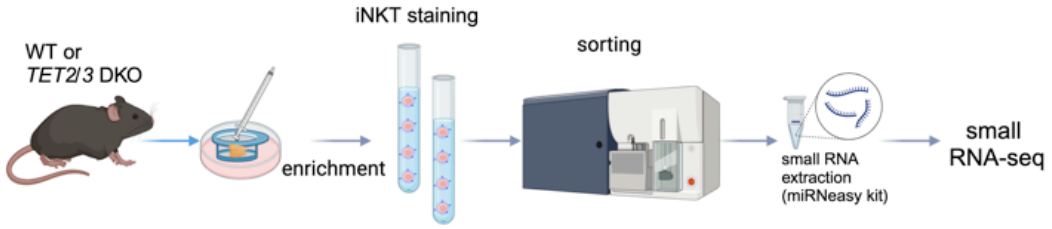
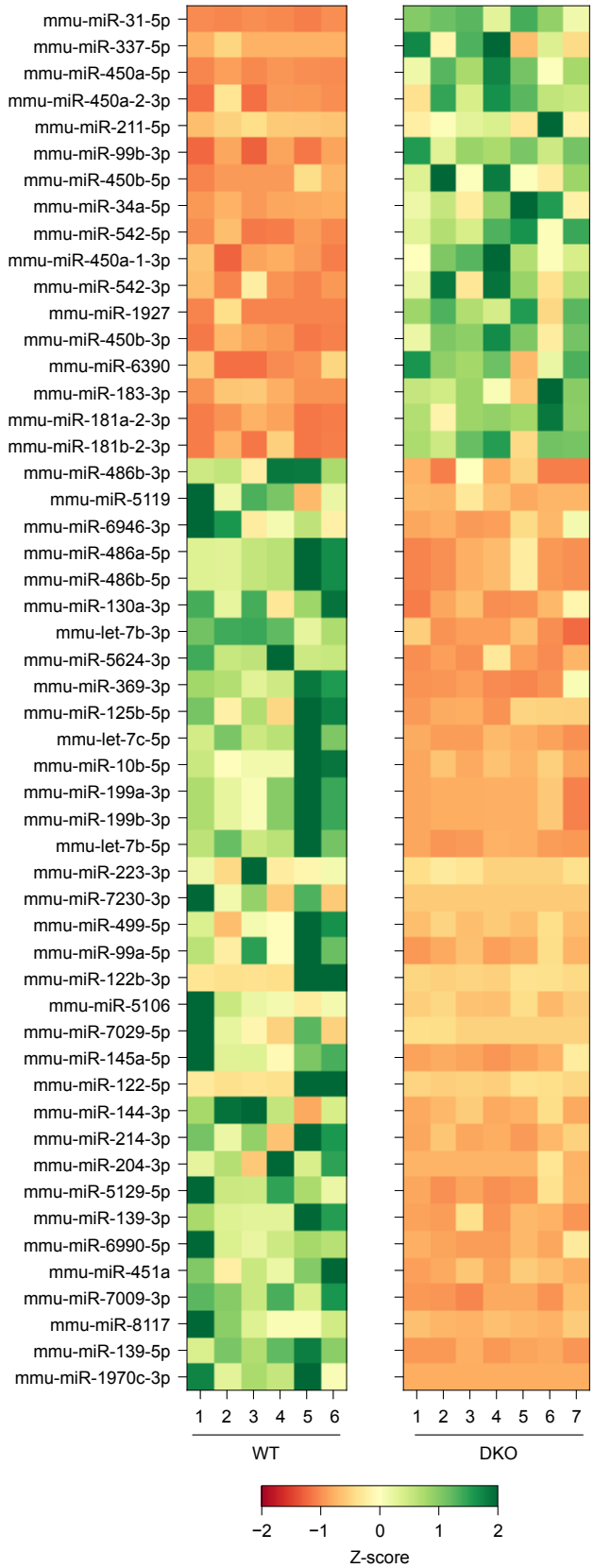


Figure 1

A. Experimental outline



B. Pre-miRNA



C. Mature miRNAs

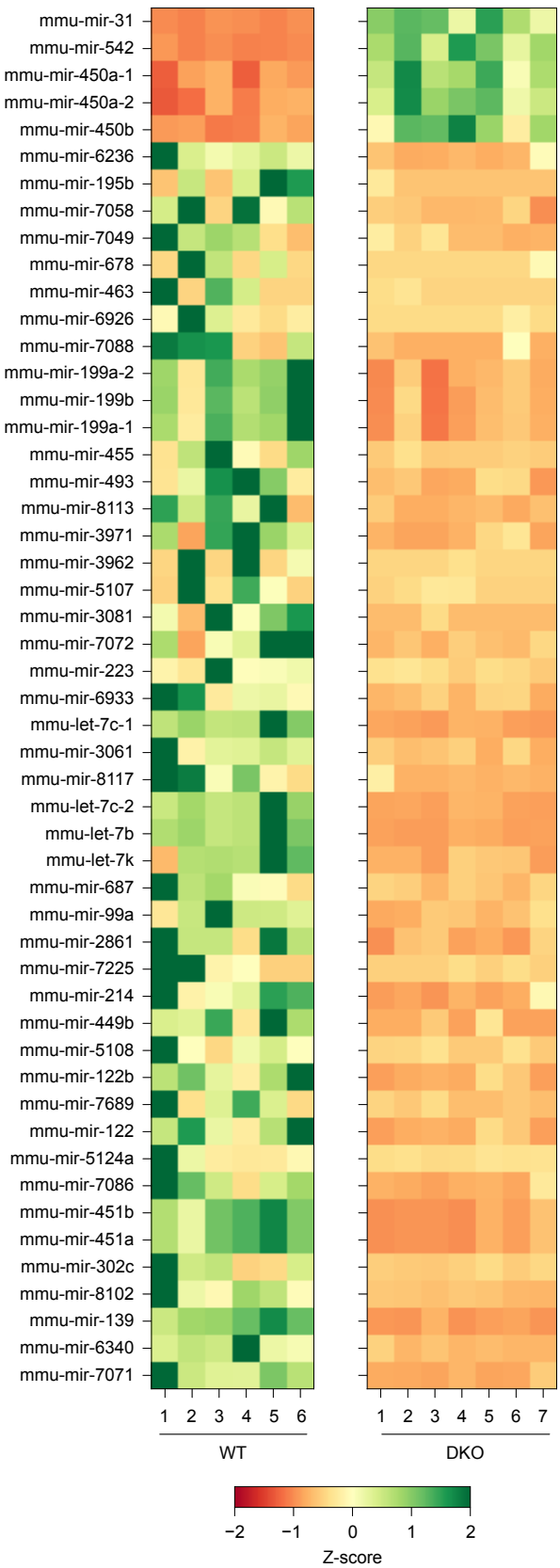
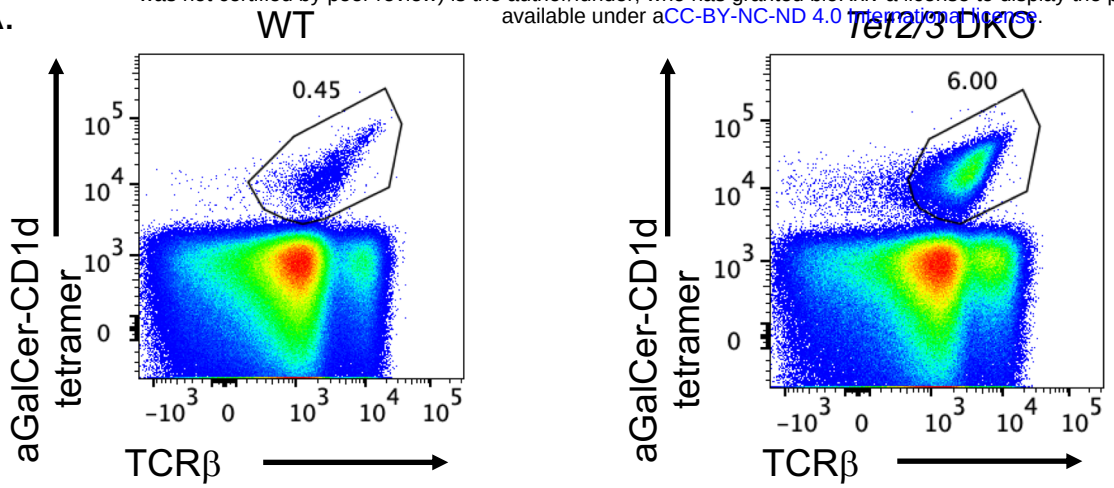
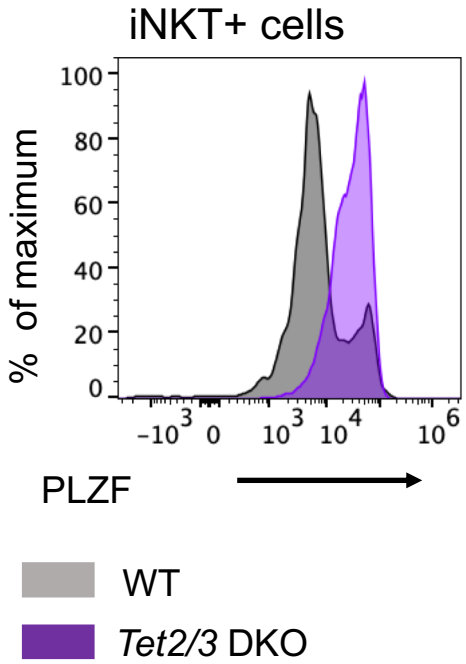


Figure 2

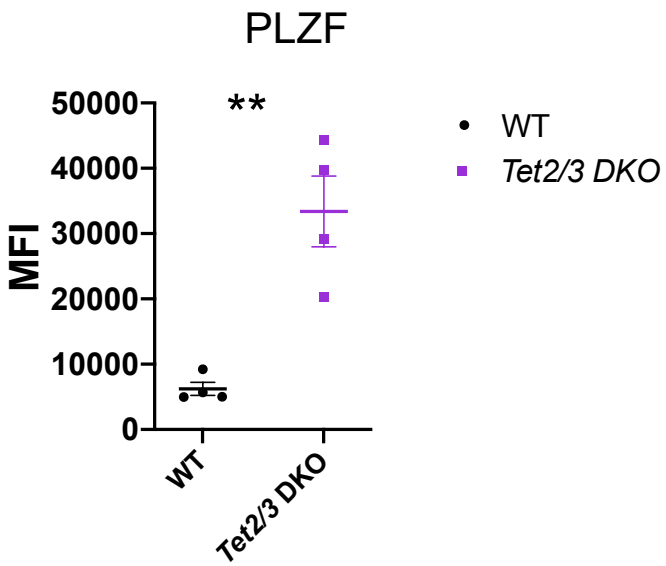
A.



B.



C.



D.

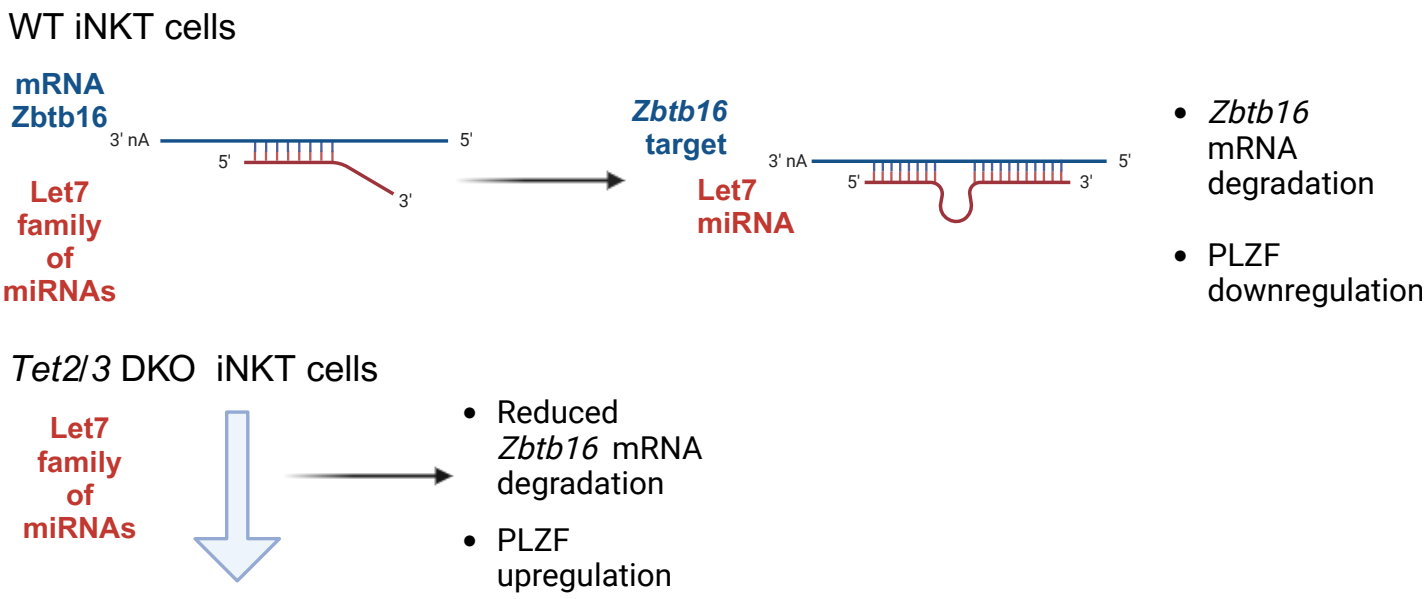


Figure 3