

## Supplementary Information

### Dynamic changes in tRNA modifications and abundance during T-cell activation - supplementary materials

Roni Rak<sup>1,2</sup>, Michal Polonsky<sup>3,4</sup>, Inbal Eizenberg-Magar<sup>3</sup>, Yufeng Mo<sup>5</sup>, Yuriko Sakaguchi<sup>5</sup>, Orel Mizrahi<sup>1</sup>, Aharon Nachshon<sup>1</sup>, Shlomit Reich-Zeliger<sup>3</sup>, Noam Stern-Ginossar<sup>1</sup>, Orna Dahan<sup>1</sup>, Tsutomu Suzuki<sup>5</sup>, Nir Friedman<sup>3\*</sup>, Yitzhak Pilpel<sup>1\*</sup>

<sup>1</sup>Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.

<sup>2</sup>Institute of Animal Science, Agricultural Research Organization, The Volcani Center, Bet Dagan, Israel

<sup>3</sup>Department of Immunology, Weizmann Institute of Science, Rehovot, Israel

<sup>4</sup>Division of Biology and Biological Engineering, California Institute of Technology, USA.

<sup>5</sup>Department of Chemistry and Biotechnology, the University of Tokyo, Japan.

\*corresponding authors

#### Corresponding Authors:

Yitzhak Pilpel, Ph.D.

Weizmann Institute of Science  
Herzl st. Rehovot 76100, ISRAEL

Email: [pilpel@weizmann.ac.il](mailto:pilpel@weizmann.ac.il)

Nir Friedman, Ph.D.

Weizmann Institute of Science  
Herzl st. Rehovot 76100, ISRAEL

Email: [nir.friedman@weizmann.ac.il](mailto:nir.friedman@weizmann.ac.il)

#### Classification:

BIOLOGICAL SCIENCES; cell biology

#### This PDF file includes:

- Extended Material and Methods
- 27 supplementary Figures
- Two supplementary tables
- Four datasets legends

## Extended Material and methods

### T-Cell isolation and FACS sorting

Two independent biological repeats were done as follows. Naïve cells were extracted from 6 and 21 spleens of B6 mice using the StemCell CD62L<sup>+</sup> CD44<sup>+</sup> kit (female, seven weeks, 6-9% naive). Each two (out of the 6) or seven (out of the 21) spleens were pooled to a total of 3 samples extracted separately. Naïve cells were activated by adding anti-CD3/anti-CD28 activation beads (1:1 bead: cell ratio, Thermo Fischer Scientific Cat#11131D). Cells were grown at  $0.25 \times 10^6$  cell/ml in 24 wells and were collected at 20h, 48h, and 72h after activation.

A portion from the naïve state was taken to FACS sorting for live cells for the first biological repeat, followed by RNA extraction. At 20h, only live cells were sorted. At 48 and 72 h, cells were sorted into CD62L<sup>+</sup> CD44<sup>+</sup> and CD62L<sup>-</sup> CD44<sup>+</sup>. The cells were stained as follows and analyzed using ARIA III FACS:

| Antibody | fluorophore  |
|----------|--------------|
| CD4      | Pacific-Blue |
| CD62L    | PE/Cy7       |
| CD44     | FITC         |
| CD69     | PE           |
| CD25     | APC          |

### RNA extraction

Sorted cell and unsorted culture cells were washed twice with PBS. The supernatant was removed, and cells were re-suspended in TRIzol (Thermo Fisher Scientific). RNA was extracted according to manufacture protocol.

### Cell culture

HeLa cells (ATCC; CRL-3216) were grown in DMEM high glucose medium (Biological Industries; 01-052-1A) supplemented with 10% FBS, 1% penicillin/ streptomycin (P/S) and 1% L-Glutamine. Transfections were done using JetPEI transfection reagent (polyplus transfection) using standard manufacture protocol.

### Generation of *TYW1* knock-out

HeLa cells were seeded onto 10cm plates and grown for 24h until reaching a confluence of approximately 70% (seeding  $\sim 2 \times 10^6$  cells). 5 $\mu$ g of LentiCRISPR V2 vector (Addgene #52961) containing either one of two sgRNA for *tyw1* gene or non-targeting control sgRNA was transfected to cells using JetPei transfection reagent. Guide RNA sequences were chosen from Predesigned Alt-R<sup>®</sup> (IDT).

Guide sequences are:

**TWT1 guide 1:** 5'-AACCACTCTGCACTTTCAGT,-3'

**TWT1 guide 2:** 5'-TCTTCACTGGAGCTCTCAAA-3'.

**Non-targeting guide:** 5'- TTTCGTGCCGATGTAACAC-3'.

Four hours after transfection, the medium was replaced with a fresh medium. Two days after the transfection, the medium was replaced with a selection medium (DMEM containing 2 µg/ml puromycin) and refreshed every day for three days until non-transfected cells reached 100% death (no adherent cells remained in the dish). Transfected cells were diluted in 96 wells plate to a concentration of a single cell per well and grown for two weeks with a selection medium to generate single clones. Knock-out of *tyw1* gene in each clone was validated using sanger sequencing with the following primers:

For *TYW1*- guide 1 validation:

Fwr: 5'-GGTGCCCTTTGTTCTTTGAGATG-3'

Rev: 5'-GTGTGCGTGTGTGCACGTG-3'

For *TYW1*- guide 2 validation:

Fwr: 5'- GAGAGCTTGCCTTTGCCCC -3'

Rev: 5'- CTTTCTTTCTTCACATGATCC -3'

We arbitrarily selected clones that show editing in the PAM position (Fig S4-3B) for downstream analysis.

#### Frameshift assay

Control and *TYW1*-KO (knock-out) HeLa cells were seeded onto 6-well plates such that cell confluence will be approximately 70% the next day (~150,000 cells). Cells were transiently transfected with frameshift reporter plasmid. The plasmid express mCherry, followed by HIV gag-pol frameshift signal, followed by GFP out of frame (-1). As a control, we used in-frame GFP and out of frame (+1) GFP in the same plasmid context (Fig 4C and S4-3A). The plasmids were a generous gift from Martin Mikl, Weizmann Institute, Israel<sup>1</sup>. The cells were analyzed using ATTUNE FACS 48h after transfection. Frameshift rate was calculated as GFP/mCherry ratio, in the linear range of fluorescent slope (Fig. S4-3).

#### mRNA sequencing

mRNA sequencing for the first biological repeat was performed using MARS-Seq protocol as described<sup>2</sup>. Libraries were sequenced using a 75bp single read output run on the Illumina NEXTseq platform.

Reads were trimmed using cutadapt<sup>3</sup> and mapped to mm10 mouse genome using STAR<sup>4</sup> v2.4.2a (default parameters). The pipeline quantifies the genes annotated in RefSeq (that have expanded with 1000 bases toward 5' edge and 100 bases toward 3' bases): /home/labs/bioservices/services/ngs/support/gtf/mm10.genes.3utr.gtf. Counting was done using htseq-count<sup>5</sup> (union mode). Further analysis is done for genes having a minimum of 5 reads in at least one sample. Normalization of the counts and differential expression analysis was performed using DESeq2<sup>6</sup> with the following parameters: betaPrior=True, cooksCutoff=FALSE, independentFiltering=FALSE. Raw P values were adjusted for multiple hypothesis testing using the procedure of Benjamini and Hochberg. The pipeline was constructed using Snakemake<sup>7</sup>.

mRNA-seq for the second biological repeat was done in parallel to ribosome footprint sequencing by a tailored mRNA-sequencing protocol as previously described<sup>8</sup>. Cells were harvested with Tri-Reagent (Sigma-Aldrich), total RNA was extracted, and poly-A selection was performed using Dynabeads mRNA DIRECT Purification Kit (Invitrogen). mRNA samples were subjected to DNaseI treatment and 3' dephosphorylation using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) and T4 PNK (NEB) followed by 3' adaptor ligation using T4 ligase (NEB). The ligated products were used for reverse transcription with SSIII (Invitrogen) for first-strand cDNA synthesis. The cDNA

products were 3' ligated with a second adaptor using T4 ligase and amplified for eight cycles in a PCR for final library products of 200-300bp. Sequencing reads were aligned as previously described<sup>9</sup>.

### Ribosome footprint sequencing

For Ribo-seq libraries, cells were treated with 100µg/ml cycloheximide (CHX) for 1 minute. Cells were then placed on ice, washed twice with PBS containing 100µg/ml CHX, and lysed with lysis buffer (1% triton in 20mM Tris 7.5, 150mM NaCl, 5mM MgCl<sub>2</sub>, 1mM dithiothreitol supplemented with 10 U/ml Turbo DNase and 100µg/ml cycloheximide). After lysis, samples were flash-frozen and stored at -80°C until further processing. Subsequently, Ribo-seq library generation was performed as previously described<sup>8</sup>. Sequencing reads were aligned as previously described<sup>9</sup>. To compare transcript expression level, mRNA and footprint counts were normalized to units of RPKM to normalize for gene length and sequencing depth, based on the total number of reads. Translation efficiency was calculated as the ratio between read counts (RPKM) of ribo-seq protected RNA fragments and mRNA sequencing. A and P site position was determined using ribowaltz<sup>10</sup>

### tRNA sequencing and read analysis

tRNA sequencing protocol was adapted from Zheng *et al.*<sup>11</sup>, with minor modification. The protocol involves the use of a highly persistent and thermo-stable reverse transcriptase. The primers are DNA-RNA hybrids with an overhang of thymidine, which enables the reverse transcription of adenosine ending RNAs. This leads to the enrichment of mature tRNAs due to their shared CCA tail. We isolated small RNAs of (<200bp) from the total RNA using SPRI-beads (Agencourt AMPure XP, Beckman Coulter; A63881), using dual side size-selection protocol. First, RNA and beads were mixed at a 1:1.8 ratio and the supernatant was collected. The small RNA was isolated by mixing the clear supernatant with beads at a 1:0.8 ratio, with X1.34 isopropanol. Reverse transcription was done using TGIRT™-III Enzyme (InGex, LLC) with the indicated primers. 3' adaptor was ligated to the cDNA using T4 ligase (NEB; M0202S). The cDNA was purified using Dynabeads myOne SILANE (life Technologies; 37002D) after each step. The library was amplified using NEBNext PCR mix and cleaned using SPRI-beads. Samples were pooled and sequenced using a 75bp single read output run on MiniSeq high output reagent kit.

| Primer name                      | Primer sequence                               |
|----------------------------------|---|
| Reverse-transcription DNA primer | 5'- CACGACGCTCTCCGATCTT -3'                   |
| Reverse-transcription RNA primer | 5'-rArGrArUrCrGrGrArArGrArGrCrGrUrCrGrUrG -3' |
| 3'-ligation adaptor              | 5'- AGATCGGAAGAGCACCA-3'                      |

Read were trimmed using homerTool<sup>12</sup>. Alignment to the genome and mature tRNAs gene sequence was done using Bowtie2<sup>13</sup> with parameters --very-sensitive-local. Reads aligned with equal alignment score to the genome and mature tRNAs were annotated as mature tRNAs. Reads aligned to multiple tRNA genes were randomly assigned when mapping to identical anti-codons and discarded from the analysis if aligned to different anti-codons. Read count was done using BedTools-coverage count<sup>14</sup>. Variant calling for mutation detection and indels was done using samtools<sup>15</sup> command "mpileup" with the parameters: "-A -q1 -d100000".

Using this tRNA sequencing protocol, 45%-60% (depending on the sample) out of the reads are mapped to mature tRNAs, the cytosolic or mitochondrial. An additional 1%-6% of the reads were aligned to pre-mature tRNAs. Of 273 unique mature tRNA sequences in the mouse genome, 120-156 tRNA genes get sufficiently high read counts that allow their quantification. On the other hand, 116-141 tRNAs did not get sufficient reads count (under 20 reads). All the tRNA genes with low scores ( below 50, tRNA-Scan SE<sup>16</sup>) have low read counts and can be considered pseudo-tRNA (Fig. S5-1A). The low read count indicates that those RNAs are either not expressed or cannot pass through maturation. Individual tRNAs were next grouped by anti-codon. We calculated the gene copy number for each anti-

codon based on the number of tRNA genes in the genome, with a prediction score above 50. We found that although they exist in the genome, tRNAs of certain anti-codons, with low gene copy numbers, have a very low expression (Fig. S5-1B). For all of those anti-codons, the matching codons can be translated by wobble interactions with a different tRNA group. These results are similar to measurements done in human HEK293 cells and the brain<sup>17</sup>, and they agree with the prediction that approximately a third of the tRNA genes in mice are inactive in all tissues, and around 30% are tissue specific<sup>18</sup>.

#### qRT-PCR for tRNA

small RNAs of (<200bp) were isolated from the total RNA using SPRI-beads (Agencourt AMPure XP, Beckman Coulter; A63881), by dual side size-selection protocol. Reverse transcription was done using TGIRT™-III Enzyme (InGex, LLC), with the indicated primers as in tRNA-sequencing protocol. Expression levels of tRNA<sup>Val</sup>(AAC), tRNA<sup>Gly</sup>(GCC) and tRNA<sup>iMet</sup>, was normalized to tRNA<sup>His</sup>(GTG) (constant expression based on tRNA sequencing). Levels were determined using quantitative RT-PCR with light cycler 480 SYBR I master kit (Roche Applied Science) and the LightCycler 480 system (Roche Applied Science), according to the manufacturer’s instructions.

The primers used were:

| Primer name          | Primer sequence                 |
|----------------------|---------------------------------|
| Mus-HisGTG2 Forward  | 5'- GCCGTGATCGTATAGTGGTTAG -3'  |
| Mus-HisGTG2 Reverse  | 5'- GTGCCGTGACTCGGATT -3'       |
| Mus- AlaAGC3 Forward | 5'- GGGATGTAGCTCAGTGGTAGA -3'   |
| Mus- AlaAGC3 Reverse | 5'- TGGTGGAGATGCTGGGGATC -3'    |
| Mus-ValAAC2 Forward  | 5'- GTTTCCGTAGTGTAGTGGTCATC -3' |
| Mus-ValAAC2 Reverse  | 5'- TGTTTCCGCCCGGTTTC -3'       |
| Mus-iMet Forward     | 5'- CGCAGCGGAAGCGTG -3'         |
| Mus-iMet Reverse     | 5'- GCAGAGGATGGTTTCGATCC -3'    |

#### Codon usage calculation

Codon usage and codon bias were calculated based on mRNA expression. We summed the usage of each codon in each gene multiplied by the expression of the gene.

$$\{CodonUsage\} = \sum_{i=1}^n \text{codonFrequency in gene } i \times \text{mRNA expression of gene } i$$

#### Estimating tAI in based on tRNA expression

We generated a measure of modified tAI (tRNA apadtation index, implemented in Fig. 2). Our measure is calculated similarly to the canonical tAI calculation<sup>19</sup>, with one major change- we determine tRNA availability based on the

read aligned to the indicated tRNA, as done in our previous work<sup>20</sup>. As such, the new measure can be computed for every condition. We defined the tRNA availability of each tRNA type (anti-codon) by the sum reads aligned to its tRNAs. Then, we determined the translation efficiency of an individual codon by the extent of reads of tRNAs that serve in translating it, incorporating both the fully matched tRNA as well as tRNAs that contribute to translation through wobble rules<sup>21</sup> (W-value).

Since there are biases in the tRNA sequencing methods we employed this method as a comparative approach- we report fold-changes between conditions and not the actual value.

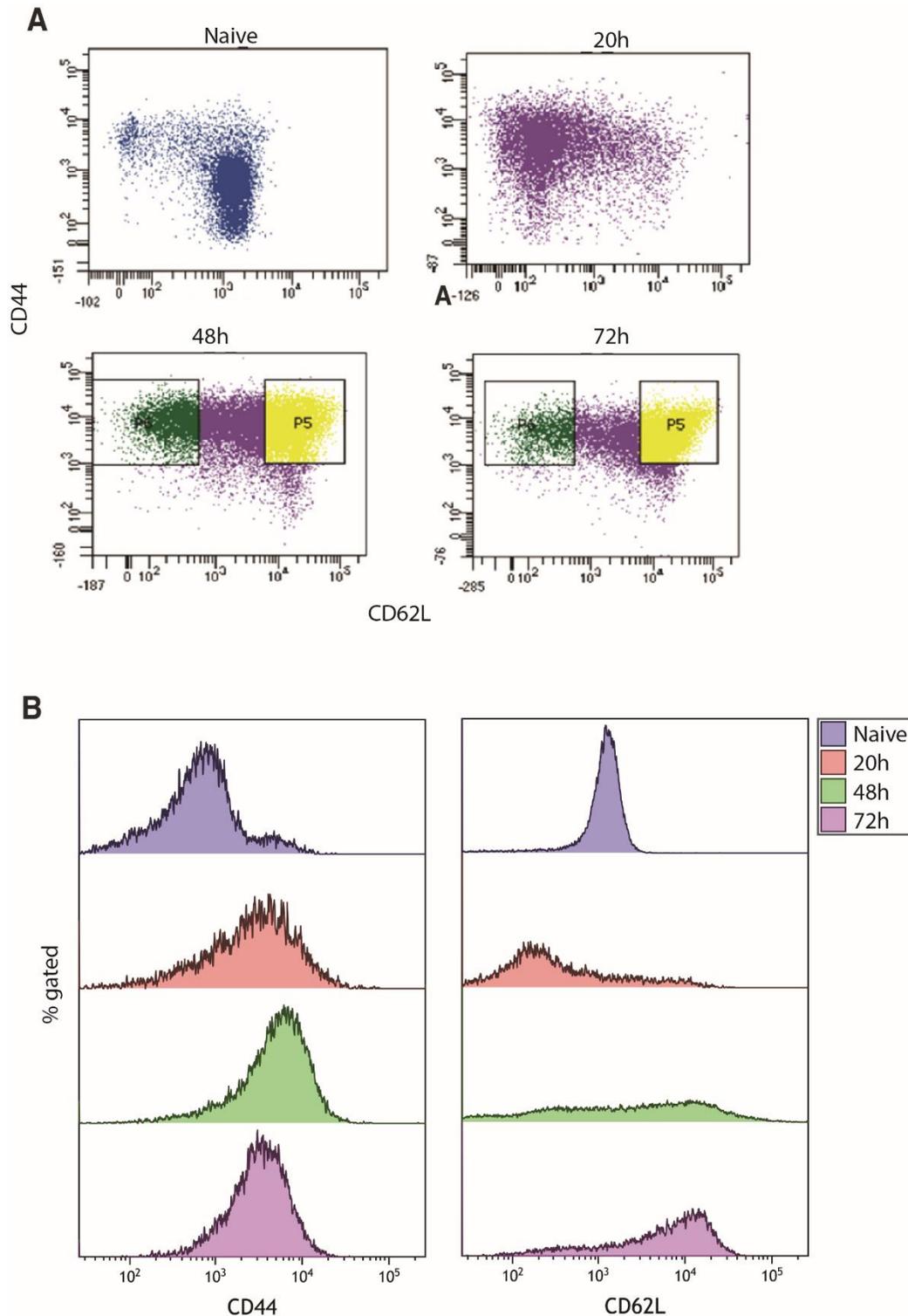
#### LC/MS shotgun analyses of tRNAs

Capillary liquid chromatography/mass spectrometry (LC/MS) shotgun analyses of tRNA fraction were conducted as described<sup>22,23</sup>. Total RNAs extracted from the culture cells were resolved by 10% PAGE with 7M urea. The gel bands corresponding to class I tRNA fraction were excised, followed by elution, the tRNA fraction was precipitated by ethanol. Two pmol of class I tRNA fraction was digested with 50 units RNase T1 (Thermo Scientific) in 20 mM NH<sub>4</sub>OAc (pH 5.3) at 37°C for one hour, followed by the addition of an equal volume of 0.1 M triethylamine acetate (TEAA; pH 7.0). The digests were subjected to the L-column2 (C18, 5µm, ID 0.3× 5 mm, CERI) for desalting and chromatographed by HiQ sil C18W-3 nanospray column (C18, 3µm, 120 Å pore size, ID 0.1 × 100 mm, Techno Alpha) with solvent system consisted of 0.4 M 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) (pH 7.0) (solvent A) and 0.4 M HFIP (pH 7.0) in 50% methanol (solvent B) at a flow rate of 300 nl per min with a linear gradient of 5–100% B solvent over 35 min with a splitless nano HPLC system (DiNa, Techno Alpha). The eluent was ionized by ESI source and introduced into an ion trap-orbitrap hybrid mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific). Deprotonated ions of RNA fragments were detected with a negative polarity mode and scanned over an *m/z* range of 600–2000 throughout the separation. Extracted ion chromatograms (XICs) were plotted according to the theoretical *m/z* of each fragment.

#### Illustrations

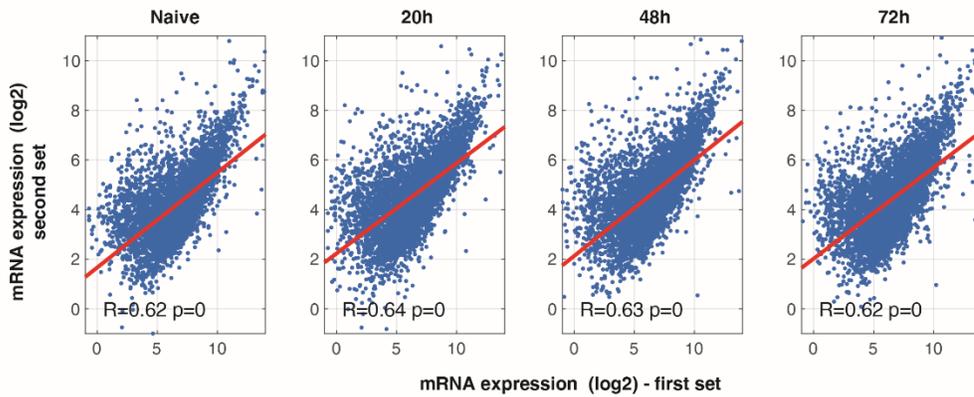
Illustrations in Figure 1 and Figure 5 were created with Biorender

Supplementary Figures

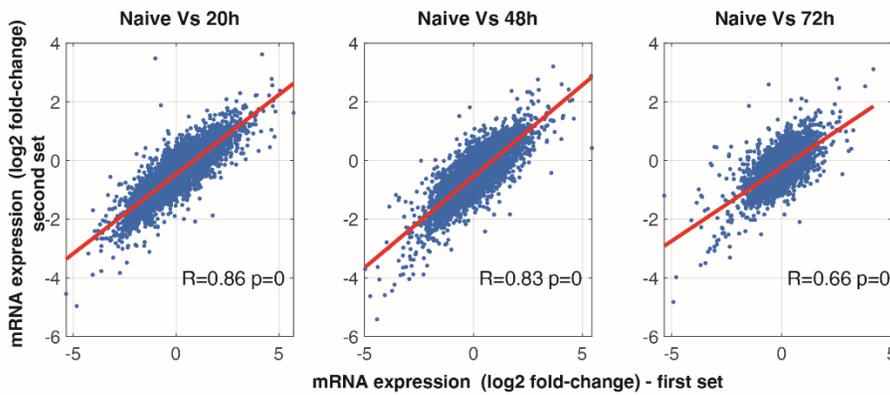


**Sup Figure 1-1- FACS analysis of T-cells activation.** CD4<sup>+</sup> T-cells were isolated from mice spleen and activated in vitro. Naïve and activated T-cells at indicated time-points were collected and sorted for CD62L<sup>+</sup> or CD62L<sup>-</sup> at later time points. A. FACS analysis of CD44 and CD62L levels. Marked in green is the CD62L<sup>-</sup> population collected, marked in yellow is the CD62L<sup>+</sup> population collected. B. Histogram representation of CD44 and CD62L levels in the indicated time point, showing elevation of CD44 levels in the activated cells and dynamics of CD62L levels.

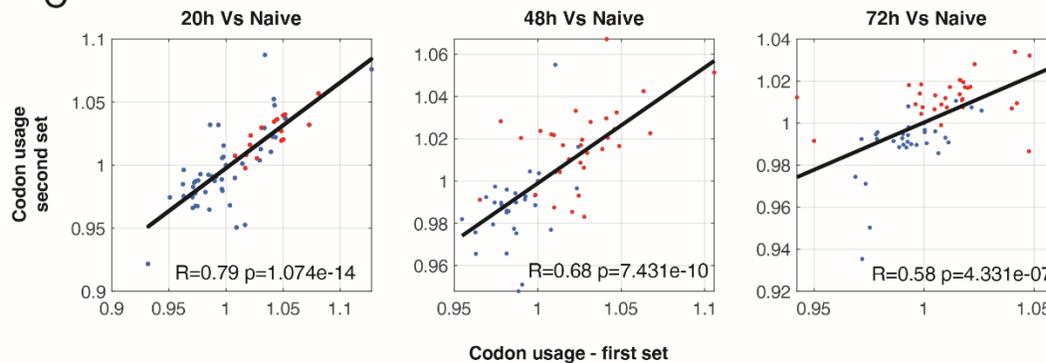
A



B



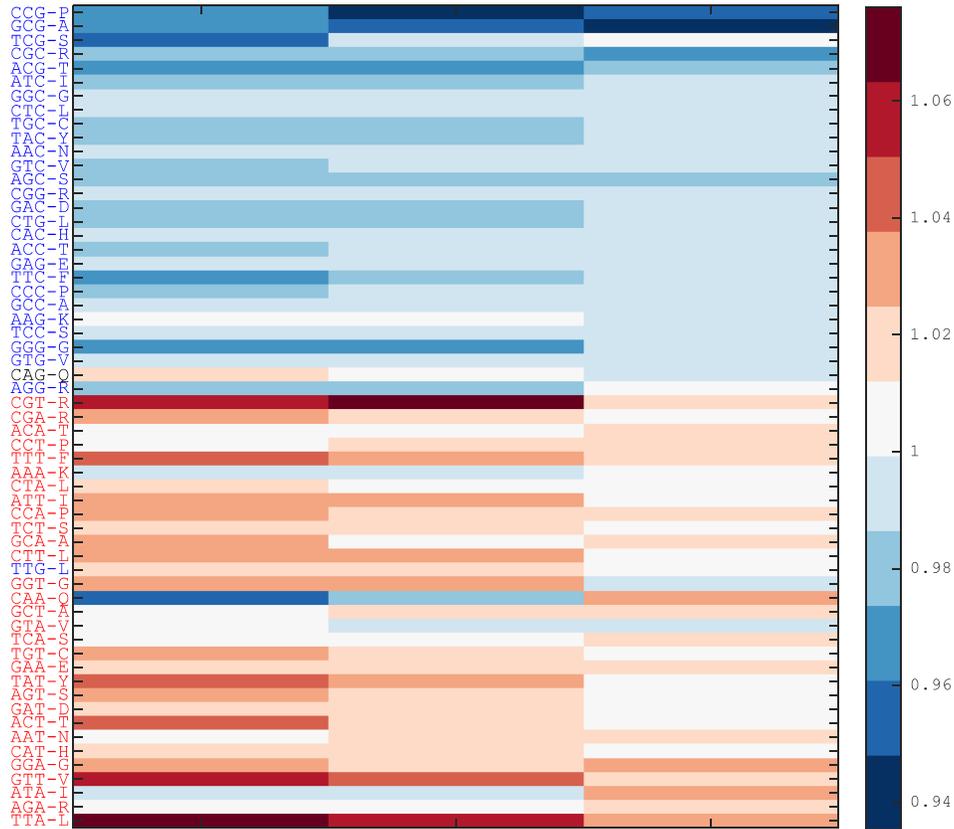
C



**Sup Figure 1-2- Comparisons of mRNA sequencing results of two biologically independent experiments.** The first set is composed of six biological conditions (naïve, 20h, 48h CD62L<sup>+</sup>, 48h CD62L<sup>-</sup>, 72h CD62L<sup>+</sup>, 72h CD62L<sup>-</sup>) with three biological repeats for each condition, and the second from 4 biological conditions (naïve, 20h, 48h, 72h) with two biological repeats for each condition. For comparison between the sets CD62L<sup>+</sup> and CD62L<sup>-</sup> in the first set were averaged. A. Comparison of read counts (RPKM) for each mRNA. B. mRNA expression fold-changes (log2) between the indicated samples. C. Changes in codon usage, calculated based on mRNA expression (see M&M). marked in blue are G/C ending codons, marked in red are A/T ending codons.

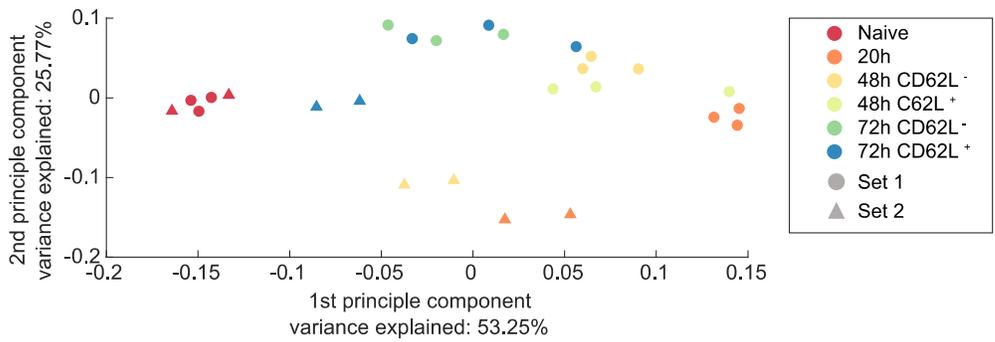
**A**

Differentiation  
codons

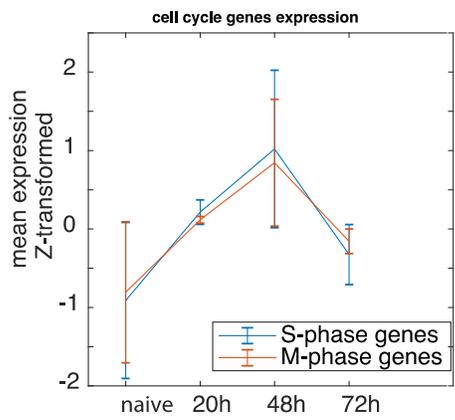


A/T ending codons  
G/C ending codons

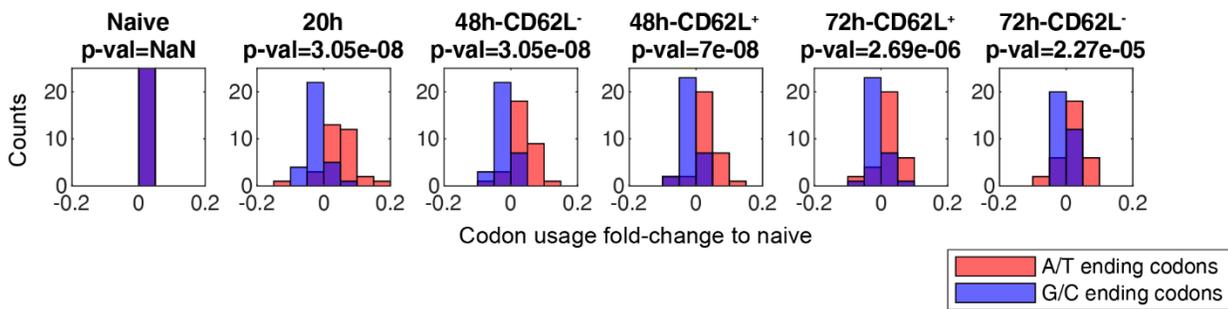
**B**



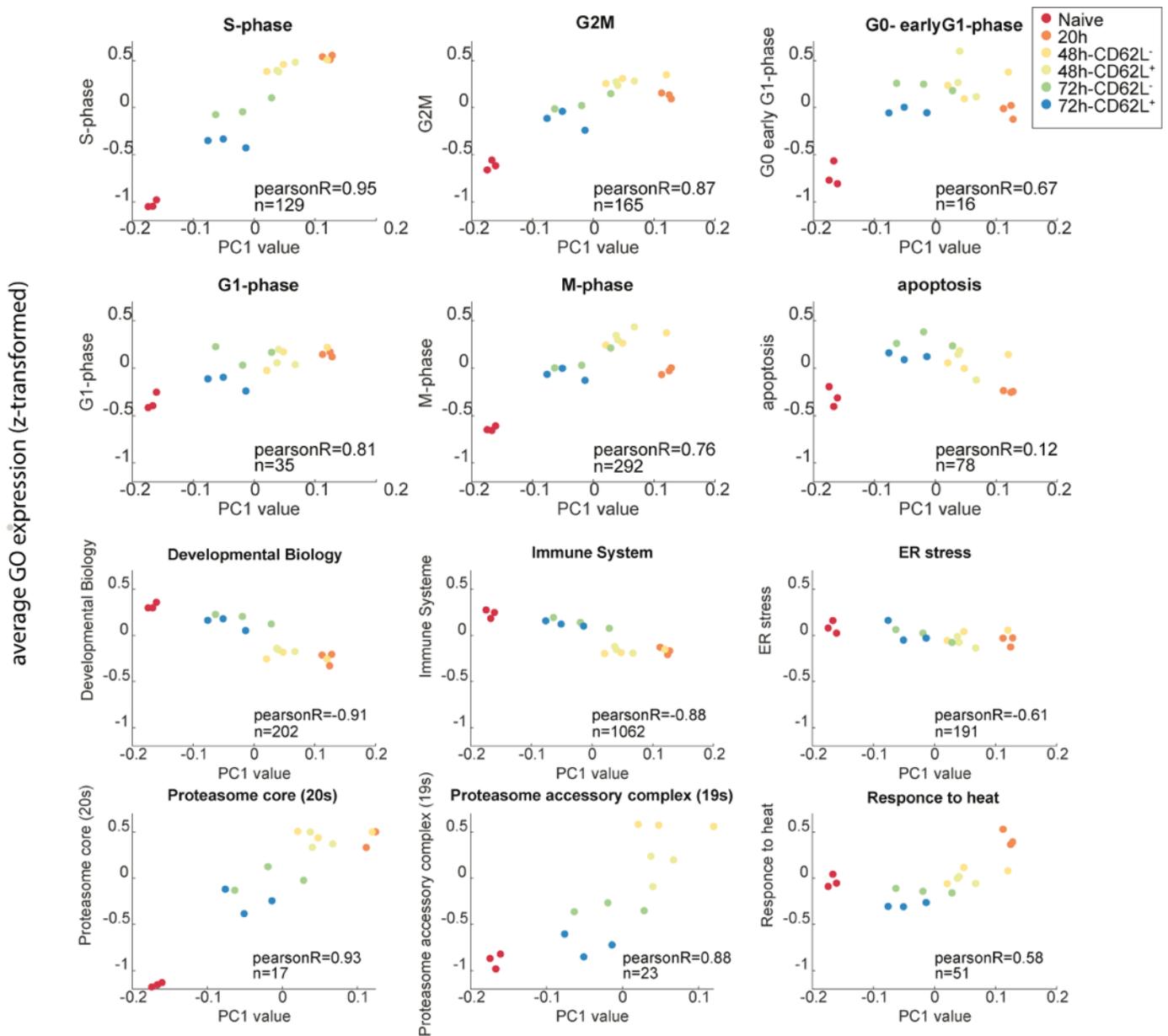
**C**



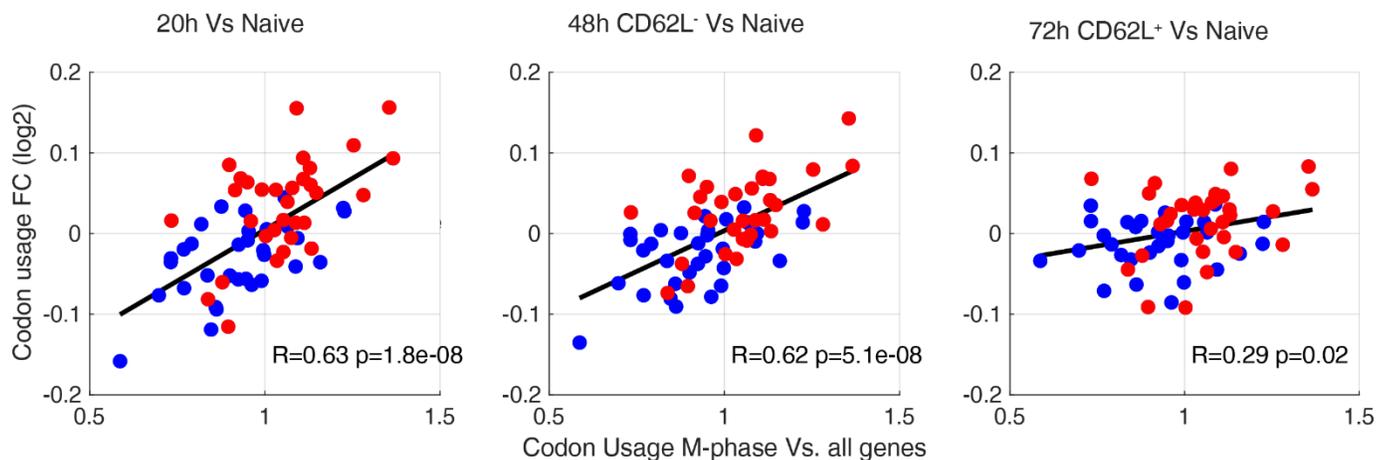
**Sup Figure 1-3- Changes in codon bias during T-cells activation. A.** Analysis of codon usage changes based on mRNA expression (normalized to amino-acids, i.e., “codon-bias”) as measured by RNA-seq. The analysis is similar to the one shown in Figure 1C, performed on an independent biological experiment. Codons are sorted based on the ratio of average codon usage of all genes from the “G2/M phase” GO category to the “pattern specification process” GO category. The separation between A/T (red) and G/C (blue) ending codons is significant in all comparisons ( $p < 10^{-5}$ , rank-sum test). **B.** Codon bias differentiates proliferative and arrested T-cells. Shown here is a principal component analysis of the samples, based on their codon bias, as presented in panel A and Figure 1C (similar to Figure 1D yet, presented here are the two biological repeats). The first component (53% variance explained) best separates the naïve and 20h samples. Circles represent the first biological repeat, and triangles are the second biological repeat. **C.** Expression of S and M – phase genes in naïve and activated T-cells from the second biological experiment (Z-transformed average of normalized read count of S-phase pathway (147 genes) and M phase pathway (326 genes),  $n=2$ , mean  $\pm$  std).



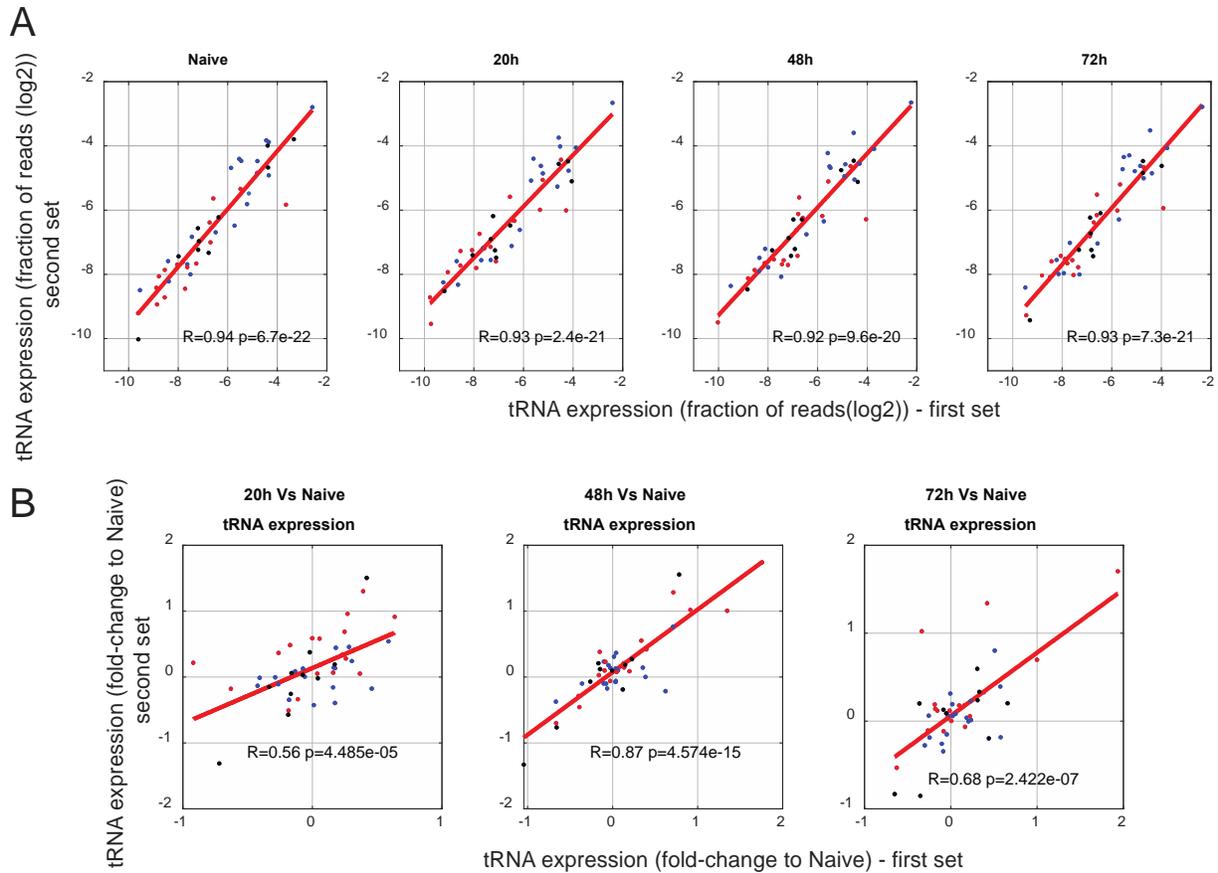
**Sup Figure 1-4- Significant changes between A/T vs. G/C ending codons.** Each histogram depicts the fold-changes (log<sub>2</sub>FC) in codon representation between the indicated sample to the naïve sample. P-values for differences between the groups are shown in the title, calculated by the Mann-Whitney U test.



**Sup Figure 1-5 – Expression of GO-related genes.** Scatter plots of position in PC1 (x-axis, calculated based on expressed codon usage, as in Figure 1C) plotted with average expression of the genes in the indicated GO category (y-axis, Z-transformed average of n genes). Each dot represents a sample. Each plot represents the indicated GO category.



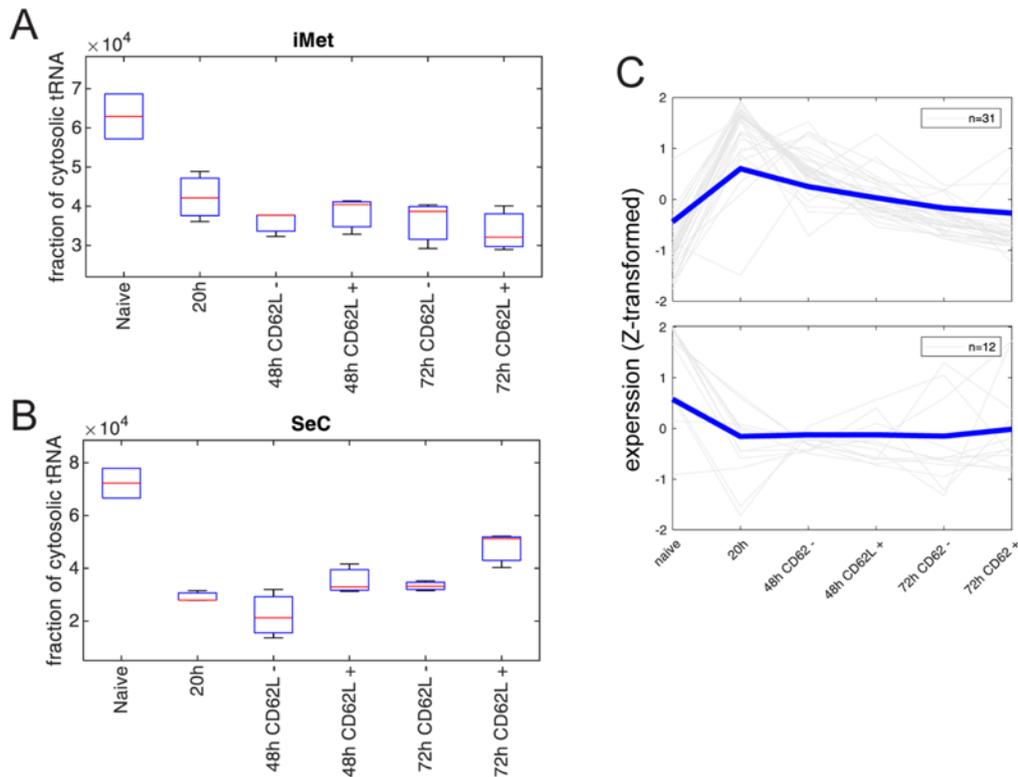
**Sup Figure 1-6 – Changes in codon demand in correlation with M-phase codon usage.** Correlation of changes in codon usage of the indicated samples compared to naïve, and the codon representation of M-phase genes compared to all genes in the genome. A/T ending codons are shown as red dots, G/C ending codons are shown as blue dots (R, p = Pearson correlation, trend-line in black).



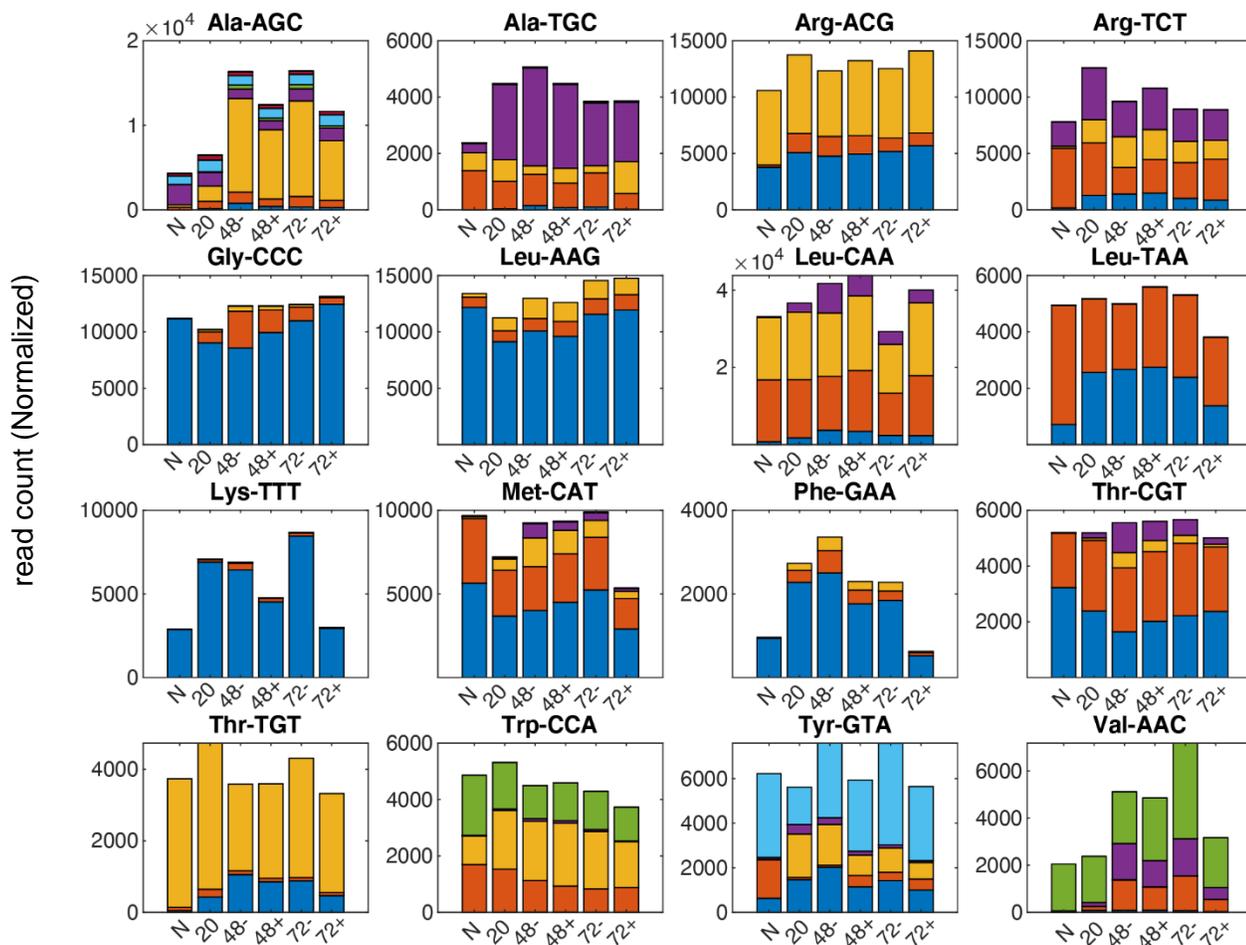
**Sup Figure 2-1- Comparisons of tRNA sequencing results of two biologically independent experiments.** The first set is composed of six biological conditions (naïve, 20h, 48h CD62L<sup>+</sup>, 48h CD62L<sup>-</sup>, 72h CD62L<sup>+</sup>, 72h CD62L<sup>-</sup>) with three biological repeats for each condition, and the second from 4 biological conditions (naïve, 20h, 48h, 72h) with three biological repeats for each condition. For comparison between the sets, CD62L<sup>+</sup> and CD62L<sup>-</sup> in the first set were averaged. A. Averaged tRNA expression summarized by anti-codon (fraction of reads) of set 1 (x-axis) vs. set 2 (y-axis). (R and P for Pearson correlation, trend-line in red) B. tRNA expression fold-changes (log2) between the indicated samples matched in the two experimental sets (R and p for Pearson correlation, trend-line in red). Red dots mark tRNA coding for A/T ending codons; blue dots mark tRNA coding for G/C ending codons; black dots mark wobble/other tRNAs.



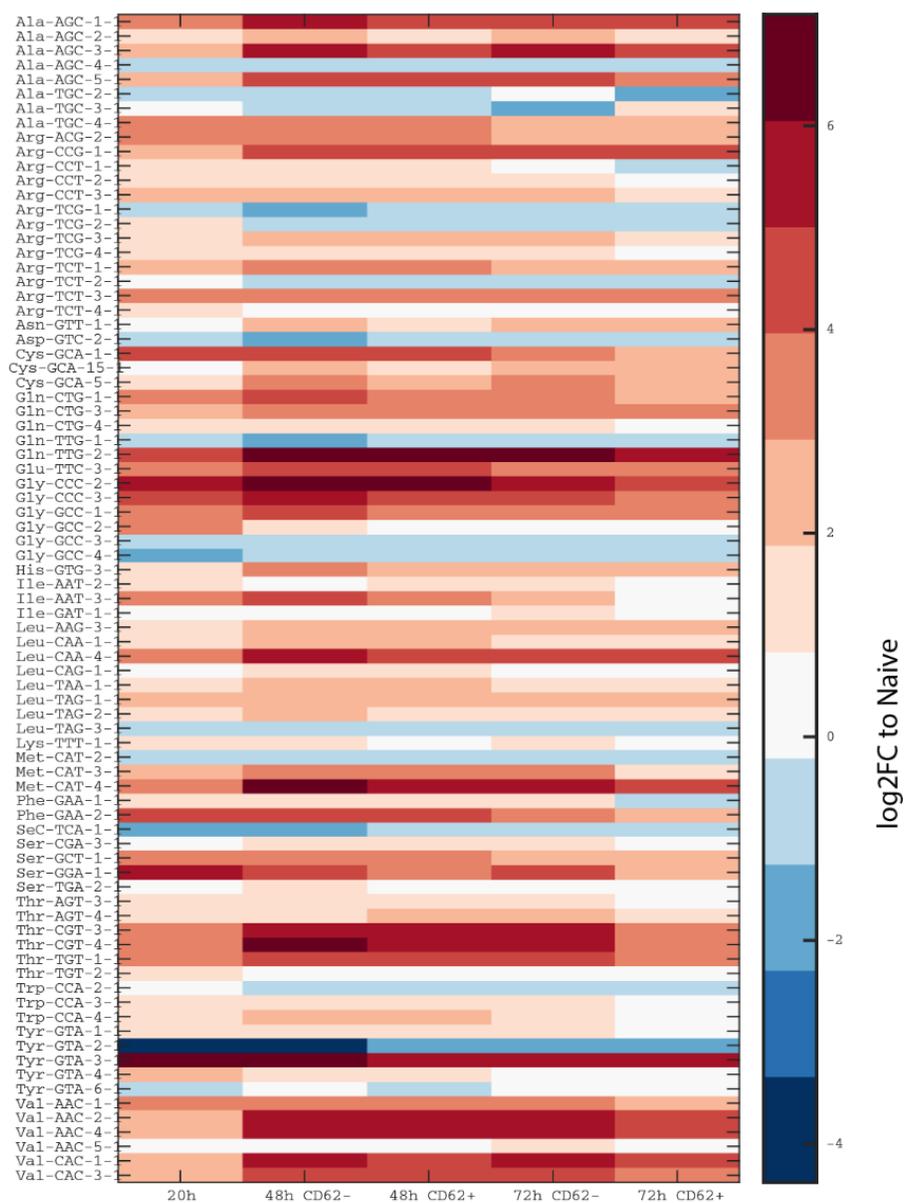
**Sup Figure 2-2- Dynamics of tRNA expression during T-cell activation.** A. Heat-map representation of changes in tRNA expression at each time point, normalized by expression in Naïve T-cells. The analysis is similar to shown in Figure 2B, performed on an independent biological experiment. Rows are sorted based on the tRNA expression fold-change between naïve and 48h after activation. Arrows indicate tRNA measured by qRT-PCR. B-D. qRT-PCR for the indicated tRNA, normalized by qRT-PCR of tRNA GTG-His. E. Coverage of reads aligned to tRNA-Ala-AGC at each position.



**Sup Figure 2-3- Expression of iMet (initiator-Methionine) and SeC (Seleno-Cysteine) tRNAs and translation initiation factors during T-cells activation process.** A. Expression of iMet-tRNA (fraction of cytosolic tRNA). B. expression of SeC tRNA (fraction of cytosolic tRNA). C. mRNA expression of translation initiation factors. Shown here mean of Z-transformed expression. Expression patterns are clustered into two main clusters- the lower panel (n=12 genes) shows genes with expression patterns correlative to iMet, while the upper panel has an opposing expression pattern (n=31 genes).

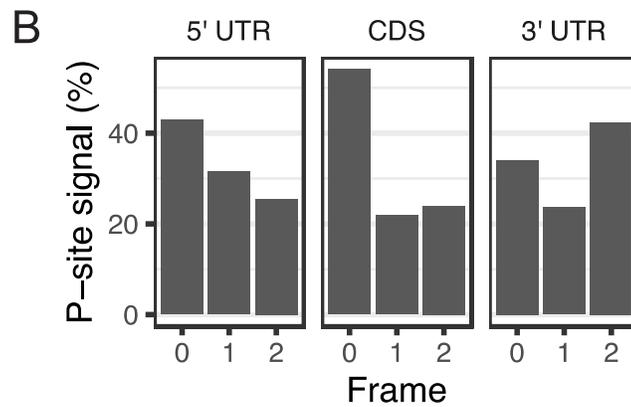
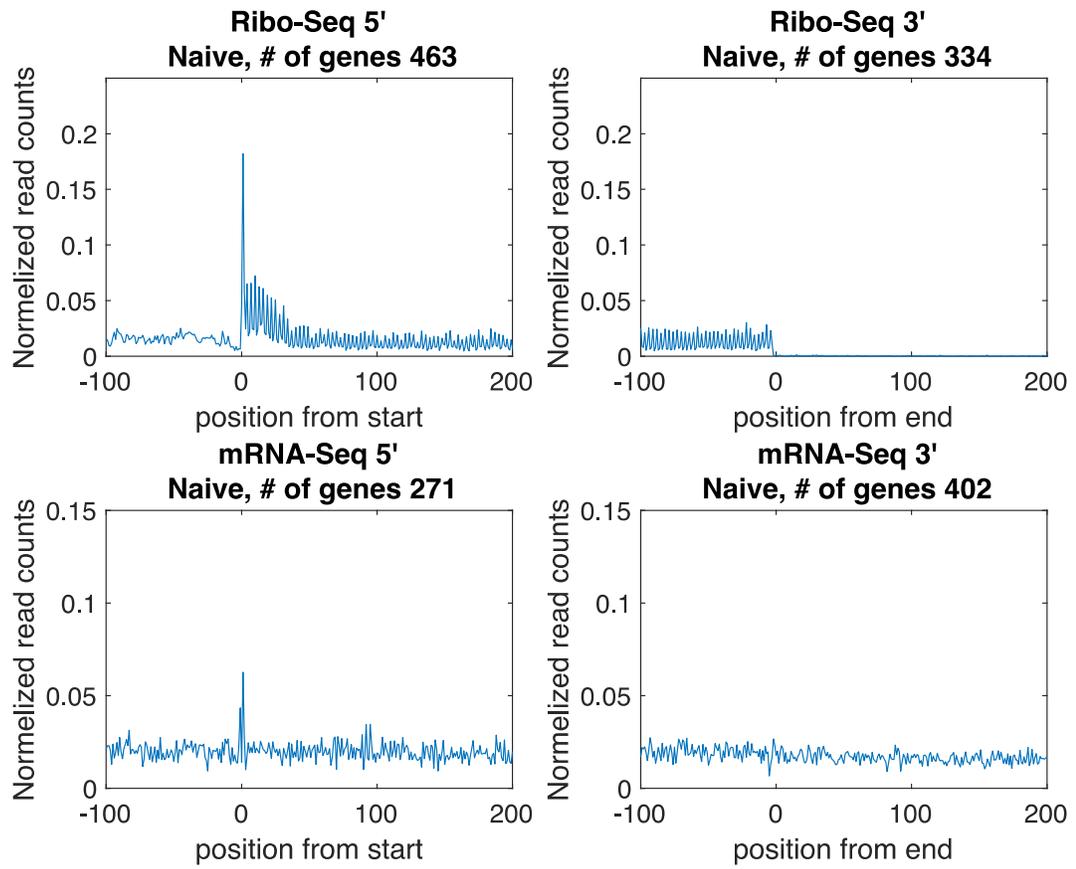


**Sup Figure 2-4- Expression of tRNA isodecoders during T-cells activation process.** Shown here are stacked bars representing the normalized read count of the indicated anti-codon, averaged over three repeats at the indicated time point. colors representing tRNA genes with unique sequences (iso-decoders). X labels are: N – Naïve, 20 – 20 hours, 48- -48 hours CD62L<sup>-</sup>, 48+ -48 hours CD62L<sup>+</sup>, 72- -72 hours CD62L<sup>-</sup>, 72+ -72 hours CD62L<sup>+</sup>.

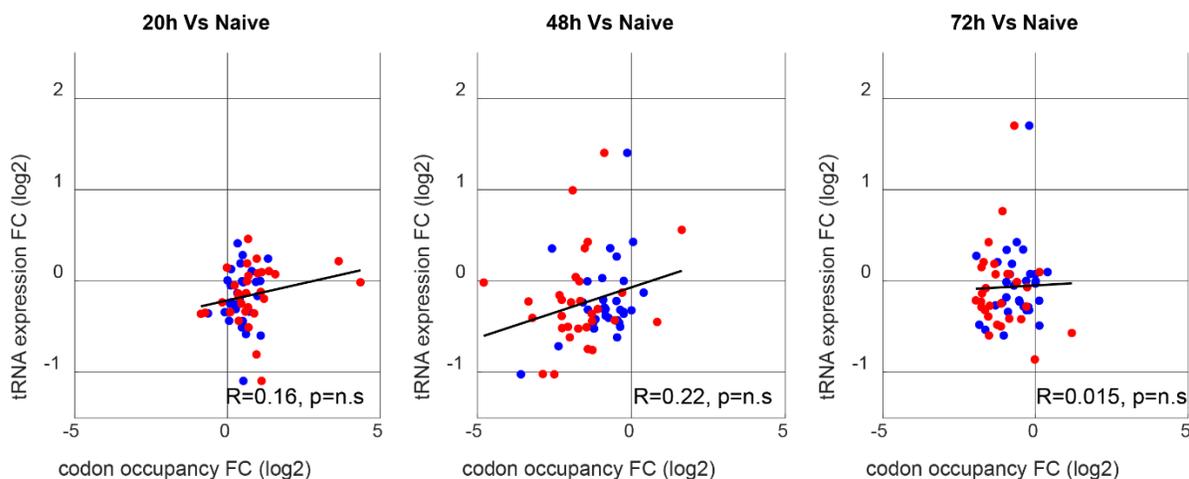


**Sup Figure 2-5- Expression of tRNA isodecoders during T-cells activation process.** A heat-map representing individual tRNA genes (grouped by tRNA copies identical in sequence) expression normalized to naïve. Shown are tRNA genes with expression changed 2-fold compared to naïve cells.

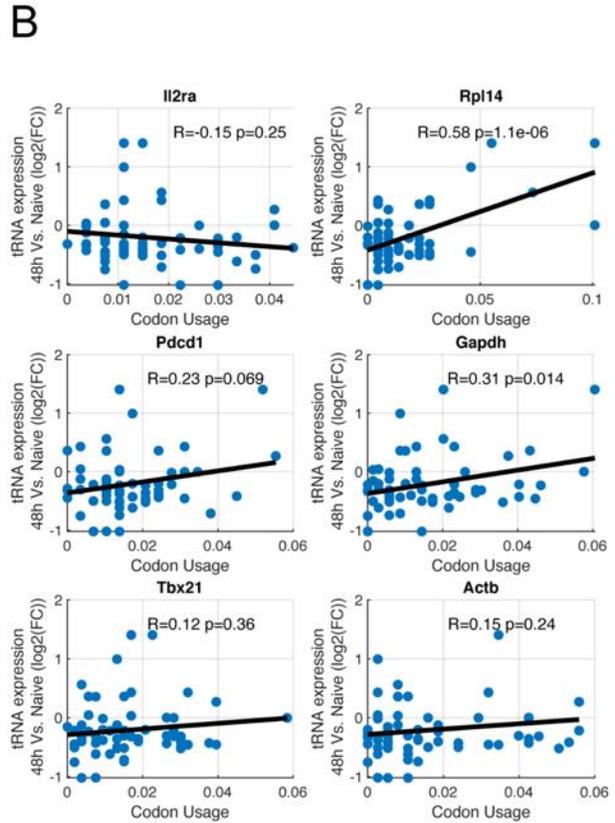
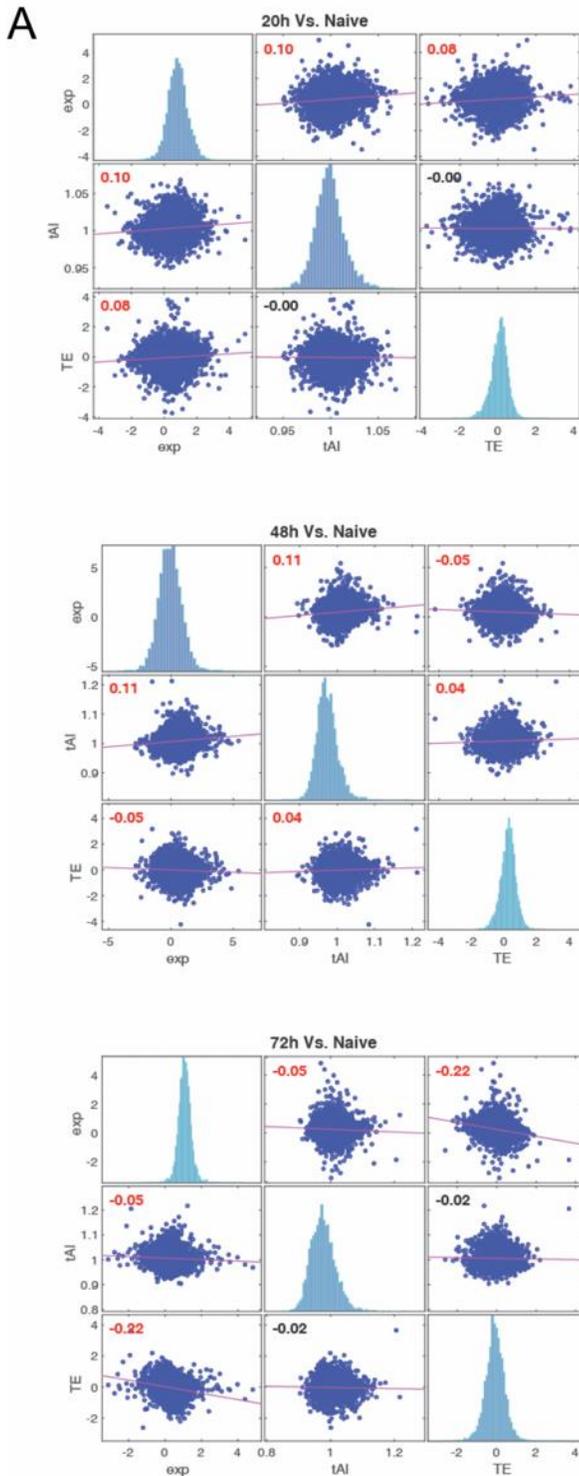
A



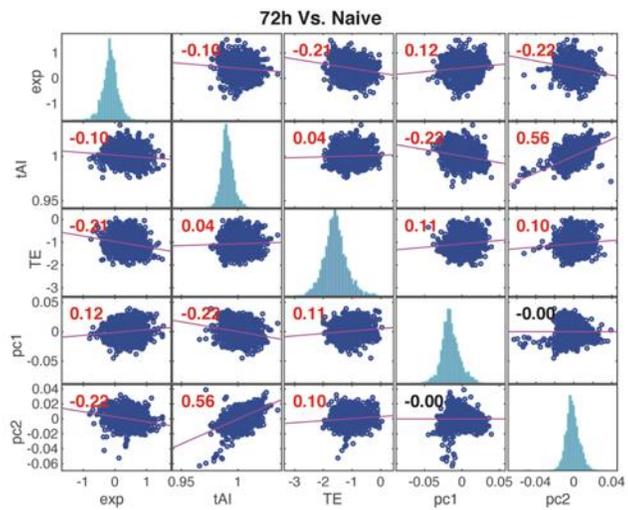
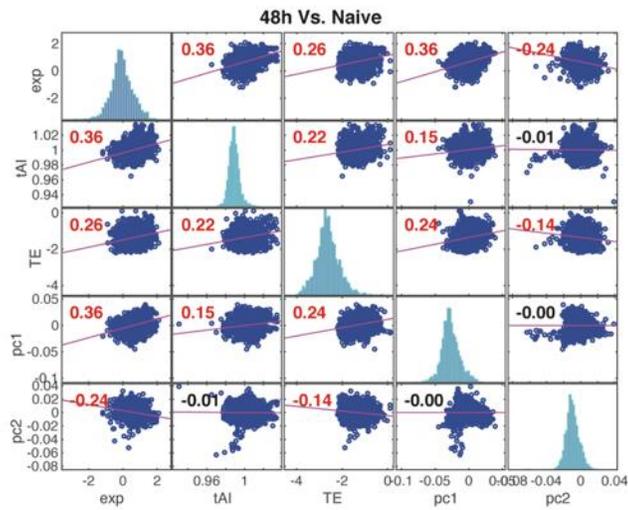
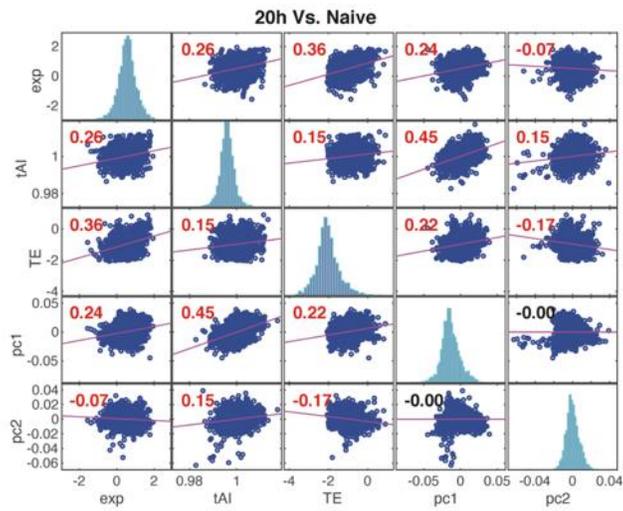
**Sup Figure 2-6- Ribo-seq quality assessments.** A. Plots showing reads density in ribo-seq (upper panel) and mRNA-seq (lower panel). The reads are aligned to the start codon (left) or the stop codon (right). The Ribo-seq reads are enriched at the start and stop codon and are higher in the coding region compared to the UTR. B. Ribosome reading frame in 5' UTR, coding sequence, and 3'UTR. The frame is calculated based on the P site position assigned for each read according to ribosome protected read length using RiboWaltz<sup>10</sup> (see M&M).



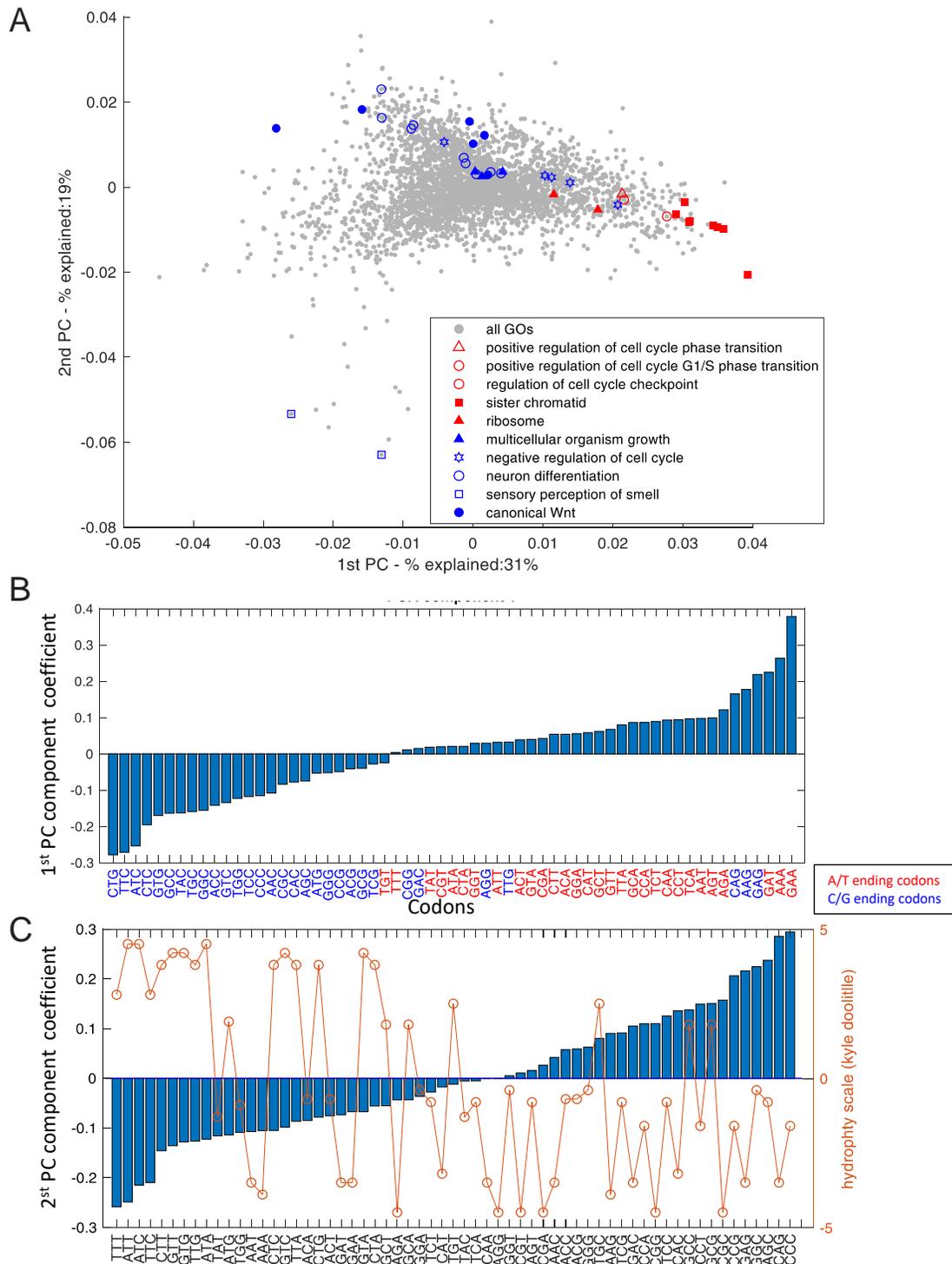
**Sup Figure 2-7- Changes in codon occupancy and tRNA expression.** Correlation of changes in A-site codon occupancy of the indicated samples compared to naïve (calculated using RiboWaltz<sup>10</sup>- Normalized A-site ribosome occupancy, log2FC), and the changes in tRNA expression (W-values, log2FC). A/T ending codons are shown as red dots, G/C ending codons are shown as blue dots. (R, p = Pearson correlation, trend-line in black). Red dots mark tRNA coding for A/T ending codons; blue dots mark tRNA coding for G/C ending codons.



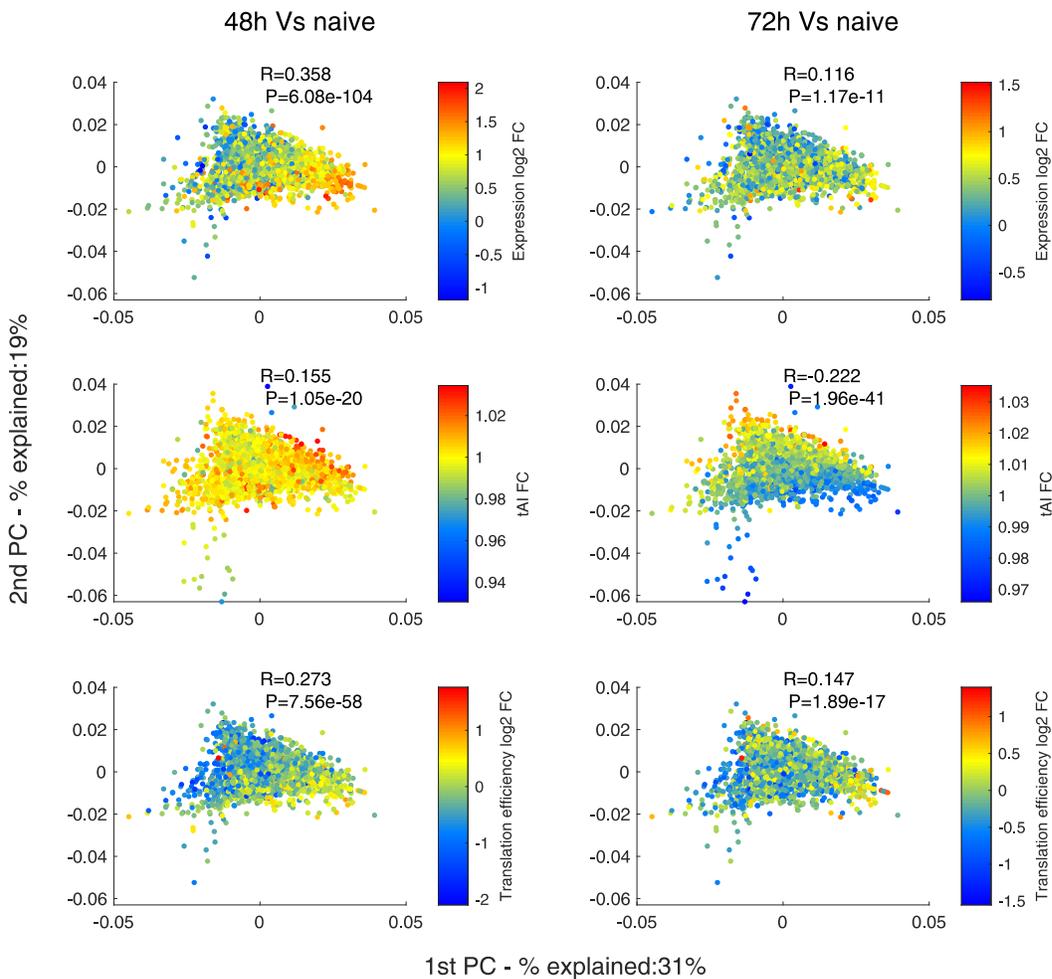
**Sup Figure 2-8- Gene level correlations- scatter plots of mRNA expression, tAI, and TE.** A. Each dot represents a gene; the axes represent mRNA expression (log2FC), tAI (fold change to Naive), and TE (log2FC). The different panels represent different time points during the T-cell activation process relative to Naive cells. Marked in red are significant correlations ( $p$ -value  $< 0.05$ ). B. Correlations between codon usage of specific genes (left- selected T-cells related genes, right- highly expressed genes) and fold-change in tRNA expression of 48h vs. naive T-cells (log2).



**Sup Figure 2-9- GO level correlations- scatter plots of mRNA expression, tAI, TE, and location on the PCA presented in figure 2D.** Each dot represents a gene set derived from a GO category. The axes represent the average value (Z-transformed) of expression, tAI and TE, and correlation to 1<sup>st</sup> and 2<sup>nd</sup> PC calculated based on codon usage of GO categories (as in Figure 2D). The different panels represent different time points during the T-cell activation process relative to Naïve cells. Marked in red are significant correlations (p-value <0.05).

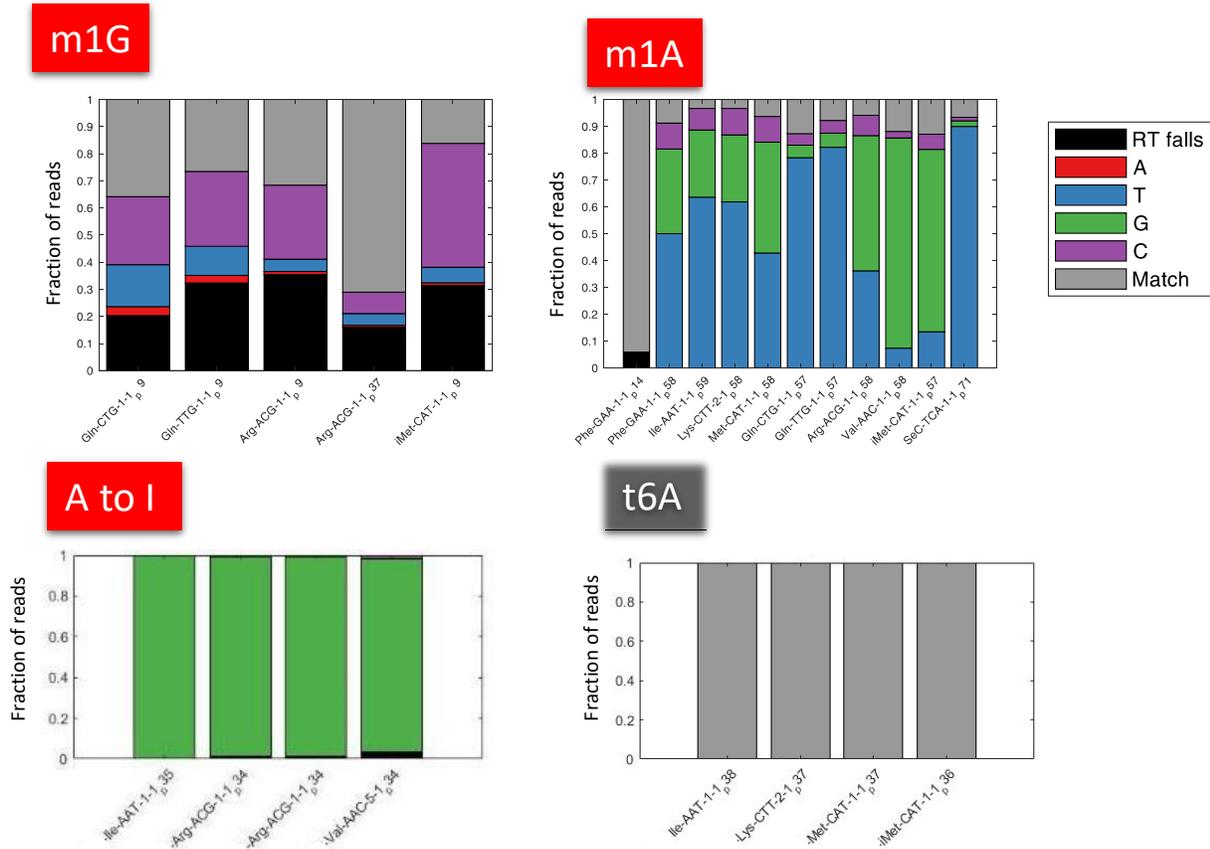


**Sup Figure 2-10- A PCA projection of the mouse codon usage of gene sets derived from GO categories.** A. The location of each gene set in this space is determined by the average codon usage of all the genes that belong to it. The % variance, out of the total original variance in the high-dimensional space, spanned by the first and second PCs is indicated on the x and y-axis, respectively. Marked in red are gene categories related to cell-autonomous functionality (proliferation-related). In blue are GO categories related to multi-cellularity functions (tissue specific) B. Codon coefficient contributes to the 1<sup>st</sup> PC. Marked in red are codons that end with A/T; marked in blue are codons with G/C. C. Codon coefficient contributes to the 2<sup>nd</sup> PC. In orange is the kyte-doolittle hydropathy score of the corresponding amino-acid<sup>24</sup>

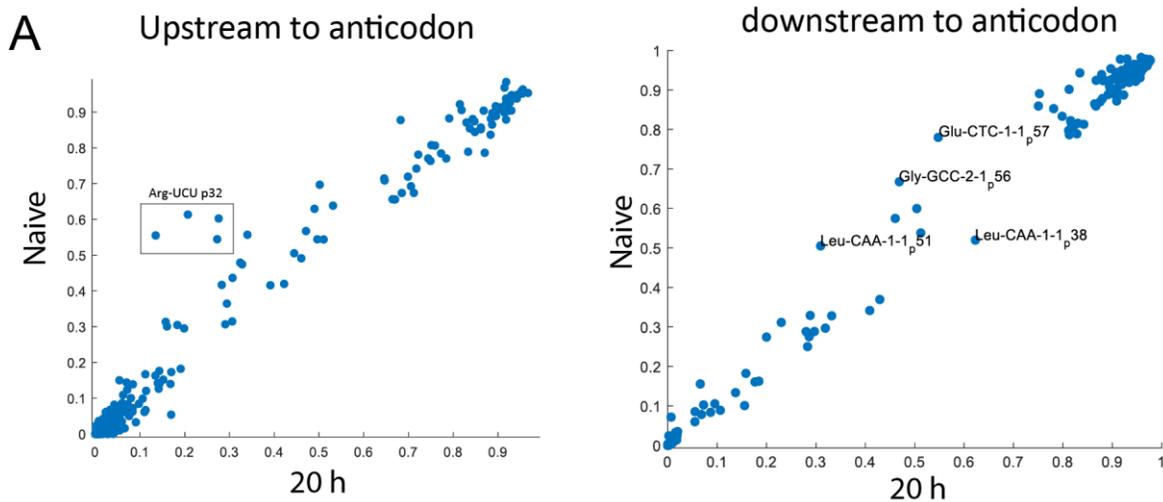


**Sup Figure 2-11- A PCA projection of the mouse codon usage of gene sets derived from GO categories.** Each point represents one gene set. Gene sets corresponding to tissue-specific GO terms are to the left side, and those corresponding to proliferation related GO terms are to the right. The color code in the upper panel represent changes at the mRNA level, averaged over all the genes in each gene set. In the middle panel, each gene category is color coded according to the relative change in availability of the tRNAs that correspond to the codon usage of its constituent genes, averaged over all genes in the gene set. A red color for a given gene set indicates that on average the genes in that set have codons that mainly correspond to the tRNAs that are induced in the condition, whereas a blue color indicates that the codon usage in the set is biased toward the tRNAs that were repressed in that given condition. In the lower panel the color code indicates translation efficiency (TE) calculated based on ribo-seq by ratio

of reads of the ribosome protected fragment (RPKM) divided by read count of mRNA-seq (RPKM). Right panels show 48h vs. Naïve, left panels show 72h vs. naïve.



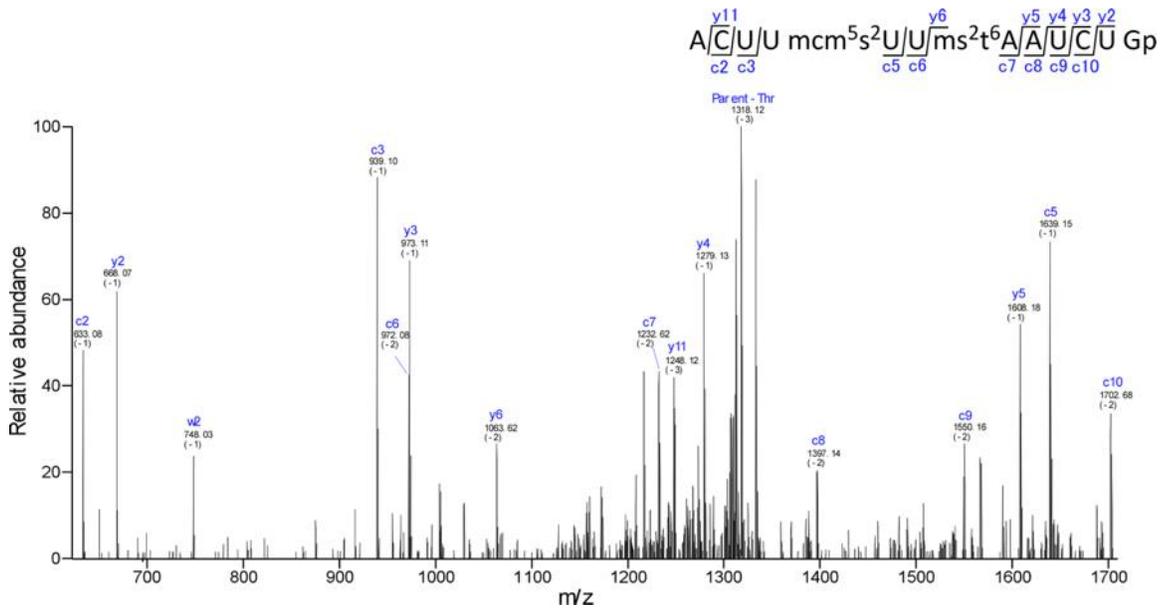
**Sup Figure 3-1- modification patterns identified by tRNA sequencing.** Each plot represents the indicated modification type, each bar is an annotated modification position based on MODOMICS database for mouse tRNA. Colors indicate a mismatch to the indicated base in the alignment, black indicate RT-abortion, and gray indicates a matched read.



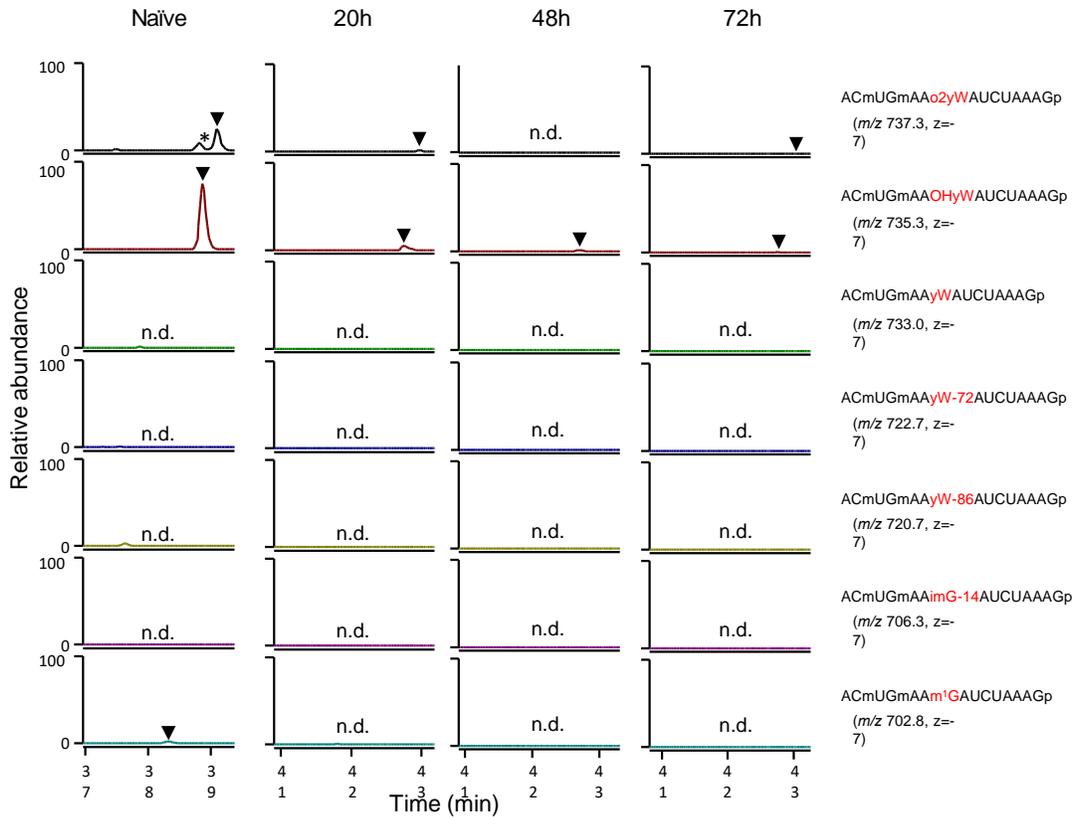
**Sup figure 3-2- modifications detection based on tRNA sequencing.** Average modification levels based on tRNA-seq from mice naïve and activated (20h) T-cells (n=3). The threshold for modification calling is minimum coverage of modified position = 20 reads, modification fraction above 10% in at least one sample. Left panel- upstream to anti-codon, right panel- downstream to anti-codon.



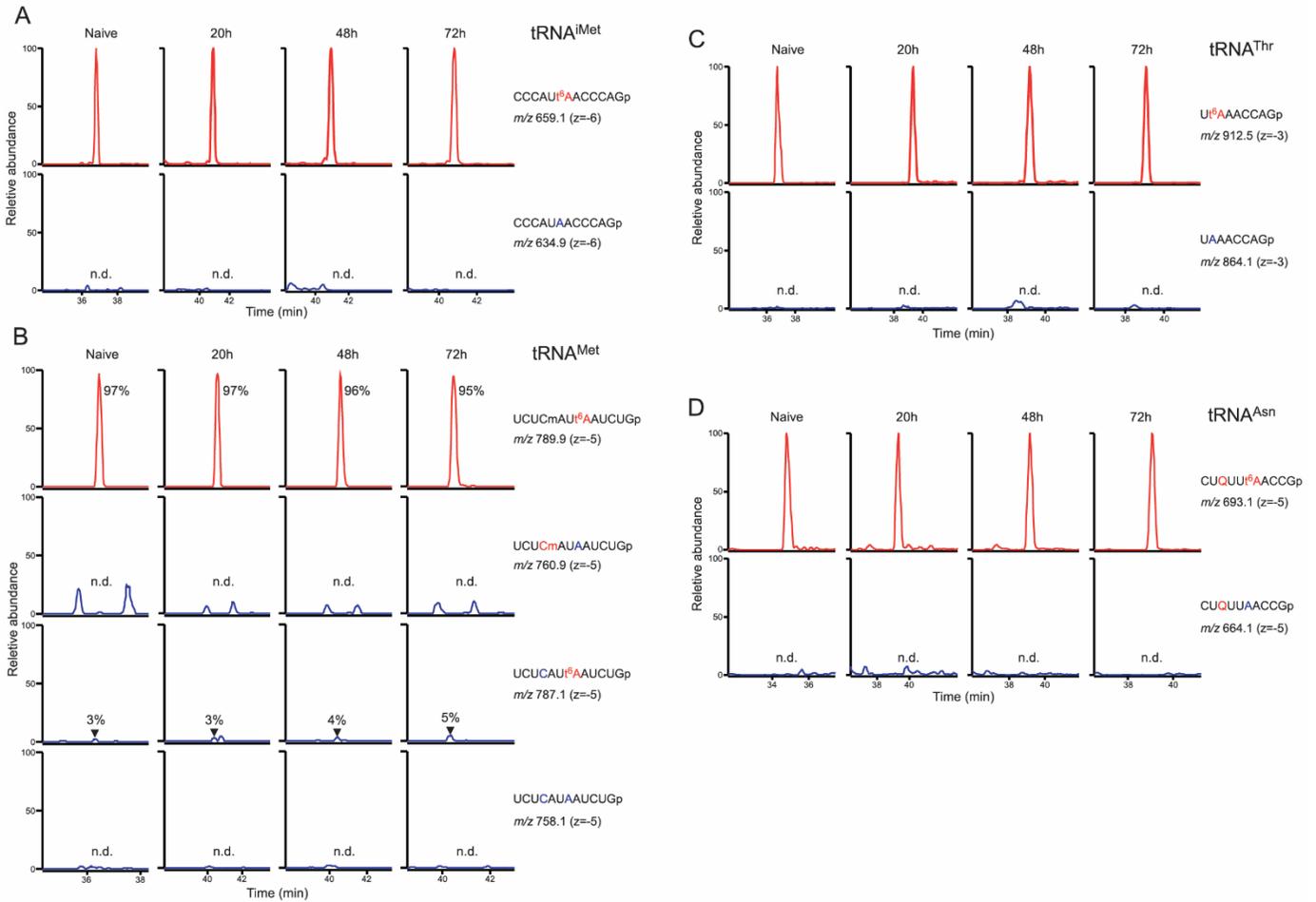
**Sup Figure 3-3- modification pattern detected by tRNA sequencing.** A. Bars show the mismatches and RT failures at position 37 in tRNA-Phe-GAA in all samples. Colors indicated mismatch to the alignment. Black indicates RT-abortion, and gray indicates a matched read. B. Bars show the mismatches and RT failures at position 37 in tRNA-Lys-TTT in all samples. Colors indicated mismatch in the alignment, black indicates RT-abortion, and gray indicated a matched read. C. m1A modification at position 58 remains constant throughout the activation process in tRNA-Lys-UUU (upper panel) and tRNA-Phe-GAA (lower panel).



**Sup Figure 3-4- Collision-induced dissociation (CID) spectrum of the anticodon-containing fragment of cytoplasmic tRNA<sup>Lys</sup> in mouse T-cells.** The triply-charged negative ion of the anticodon-containing fragment ( $m/z$  1358.15) bearing  $mcm^5s^2U34$  and  $ms^2t^6A37$  was used as a precursor ion for CID. The product ions were assigned as described <sup>25</sup>. Sequence of the parent fragment with assigned product ions is indicated in the inset.

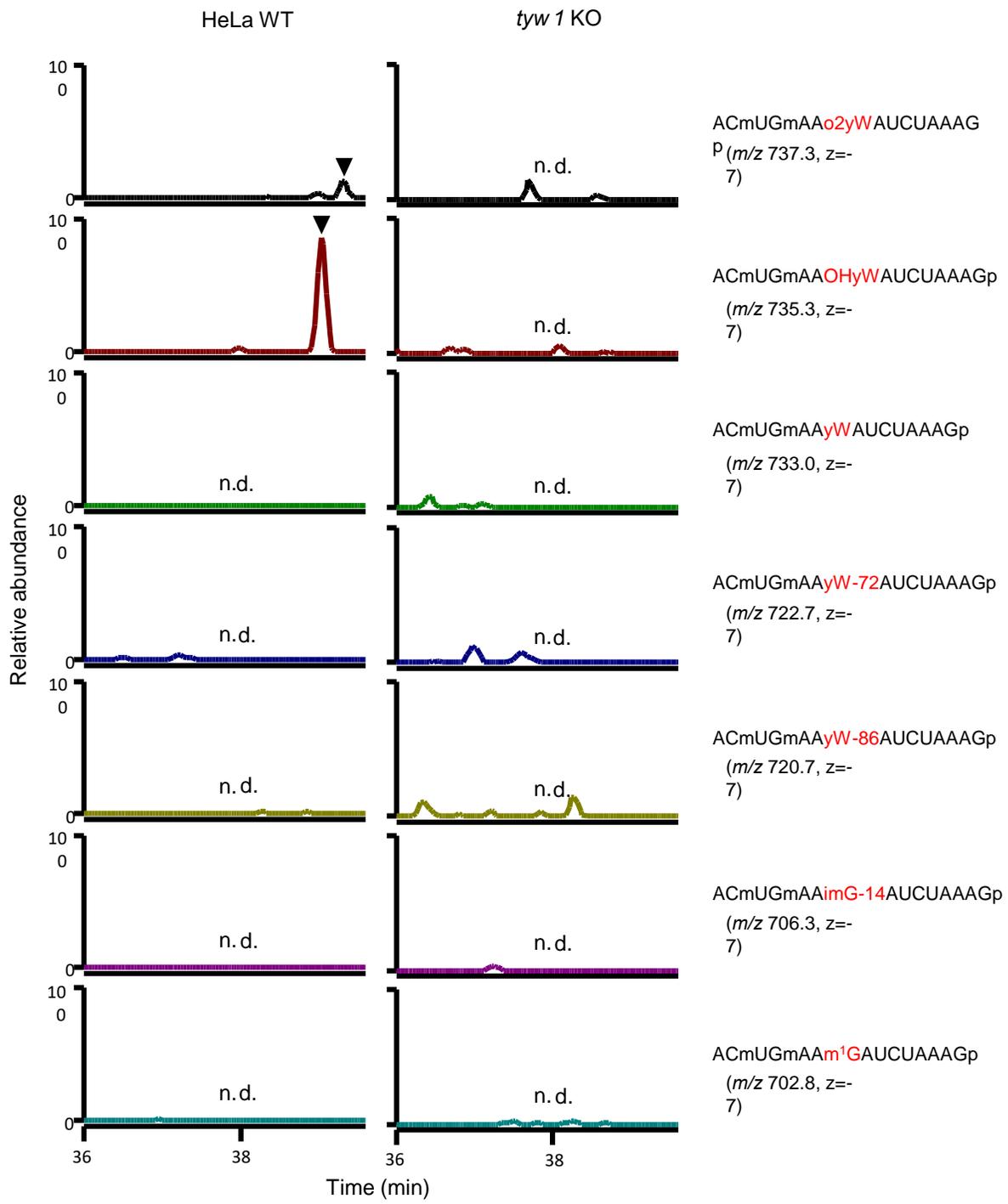


**Sup Figure 3-5- Mass spectrometric shotgun analysis of cytoplasmic tRNA modifications during T-cell activation.** Shotgun analysis of class I tRNA fraction from mouse CD4<sup>+</sup> T-cells collected at 0 (naive), 20, 48, and 72h after activation. The tRNA fraction was digested by RNase T1 and subjected to capillary LC/MS. Top to bottom panels represents XICs for negative ions of anticodon-containing fragments of cytoplasmic tRNA<sup>Phe</sup> bearing o2yW37, PHyW37, yW37, yW-72, yW-86, imG-14, and m1G, respectively. The peak heights in each condition are normalized by AAUAUAm<sup>2</sup>Gp (*m/z* 665.4, *z*=-4), an internal fragment of cytoplasmic tRNA<sup>Phe</sup>. n.d. stands for not detected. The actual peak is indicated by arrowhead. Non-specific peak is asterisked.

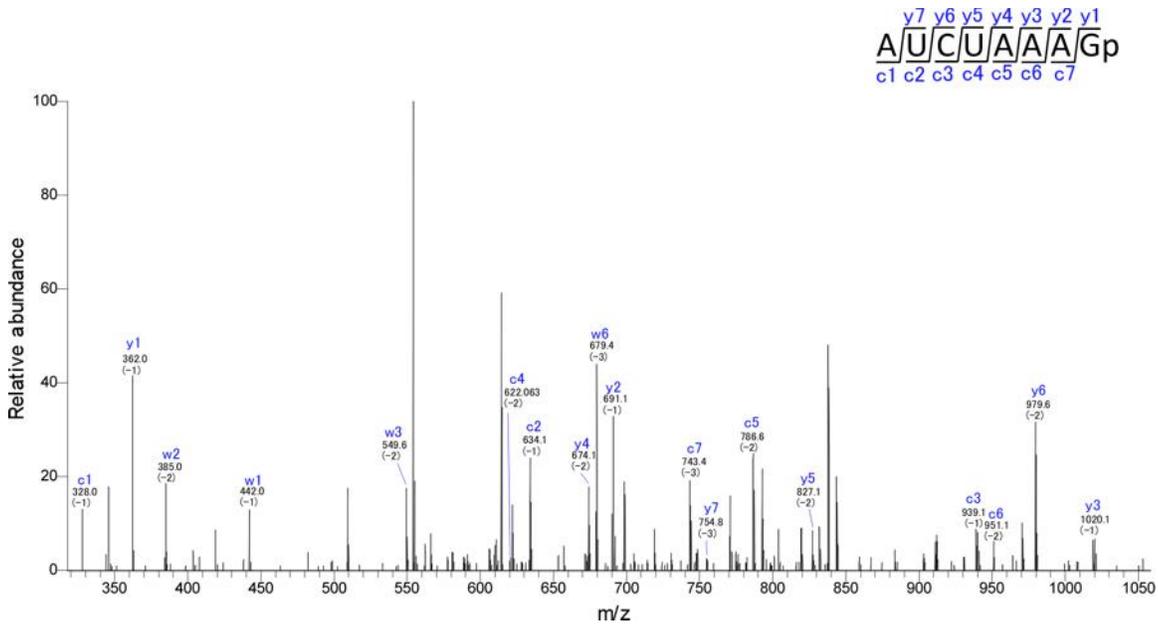


**Sup Figure 3-6. Mass spectrometric shotgun analysis of cytoplasmic tRNA modifications during T-cell activation.**

Shotgun analysis of class I tRNA fraction from mouse CD4<sup>+</sup> T-cells collected at 0 (naive), 20, 48, and 72h after activation. The tRNA fraction was digested by RNase T 1 and subjected to capillary LC/MS. Top to bottom panels represent XICs for negative ions of t 6 A37- containing fragments of cytoplasmic tRNAs for initiator Met (A), Met (B), Thr (C), and Asn (D). The percentage on each peak represents the relative abundance of each fragment.

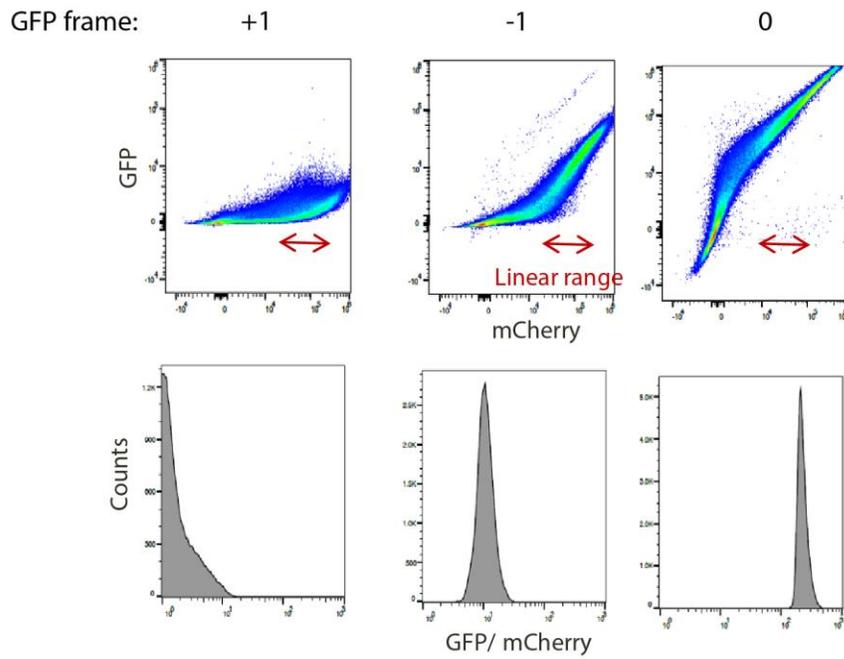


**Sup Figure 4-1- Mass spectrometric shotgun analysis of cytoplasmic tRNA<sup>Phe</sup>.** Shotgun analysis of class I tRNA fraction from HeLa WT (left panels) and *tyw1* KO (right panels) cells. The tRNA fraction was digested by RNase T1 and subjected to capillary LC/MS. Top to bottom panels represent XICs for negative ions of anticodon-containing fragments of cytoplasmic tRNA<sup>Phe</sup> bearing o2yW37, PHyW37, yW37, yW-72, yW-86, imG-14 and m1G, respectively. n.d. stands for not detected. Arrowheads indicate the actual peaks.

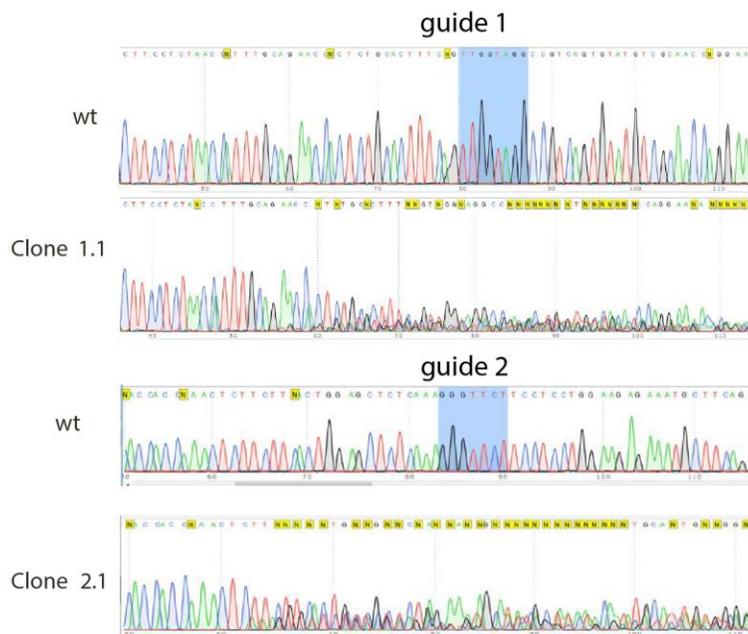


**Sup Figure 4-2- Collision-induced dissociation (CID) spectrum of the RNase T<sub>1</sub>-digested fragment of cytoplasmic tRNA<sup>Phe</sup> in HeLa *tyw1* KO cells.** The quadruply-charged negative ion of AUCUAAAGp (*m/z* 648.1, *z*=-4) was selected as a precursor ion for CID. The product ions were assigned as described<sup>25</sup>. Sequence of the parent fragment with assigned product ions are indicated in the inset.

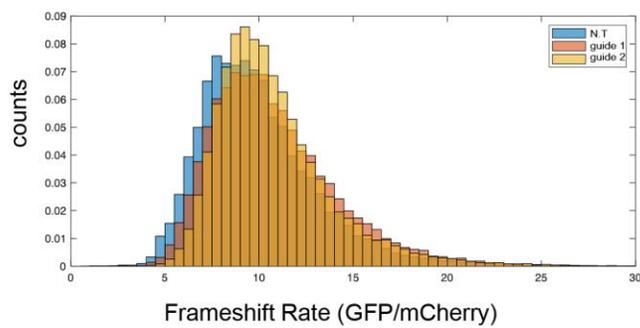
**A**



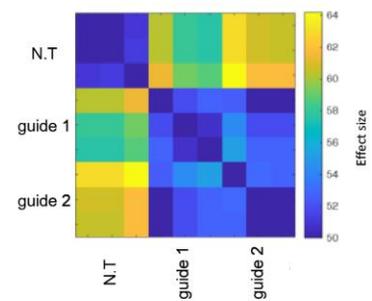
**B**



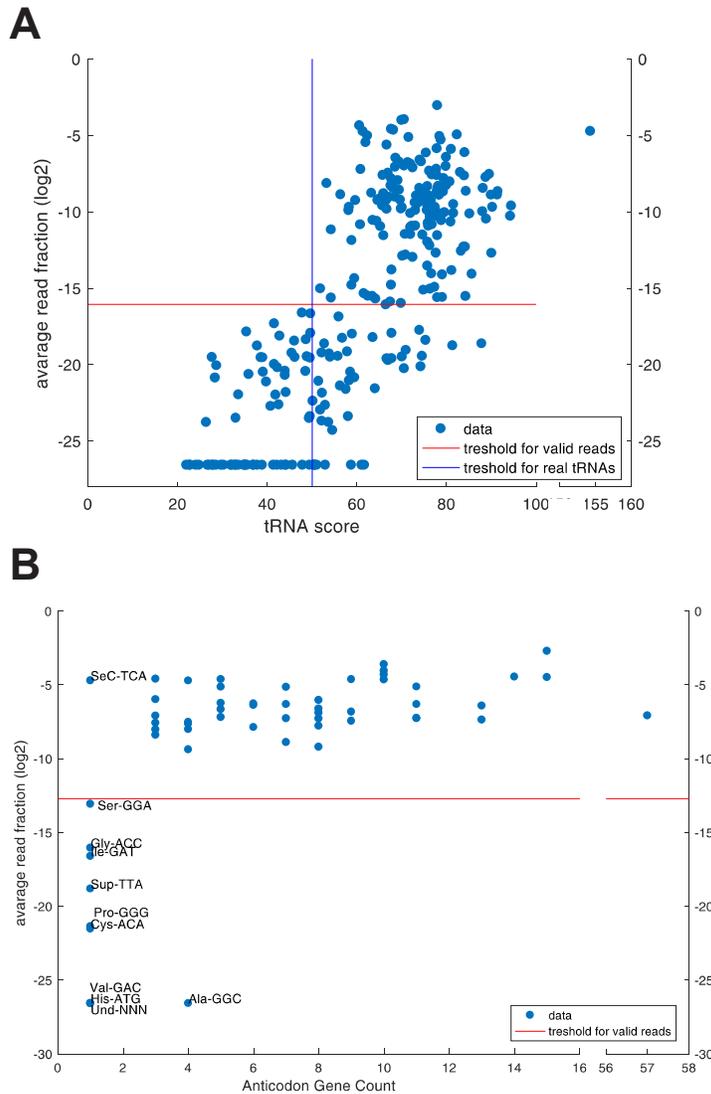
**C**



**D**



**Sup Figure 4-3- Frameshift detection in *tyw1*-KO cells.** A. Upper panel- scatter plots of GFP (y-axis) and mCherry (x-axis) intensity measured in HeLa cells expressing mCherry followed by HIV frameshift signal, followed by GFP in +1, -1 and 0 frame. Red arrows indicate the linear area we used to calculate the ratio between GFP and mCherry. Lower panel –histogram of GFP/mCherry ratio in the linear range as in upper panel. B. Validation of *TYW1*-KO in HeLa cells. Shown here are Sanger sequencing chromatograms of CRISPR edited cells using two different gRNA for *tyw1*, compared with non-targeting (N.T) gRNA. C. Histograms of GFP/mCherry ratio (as in A.) measured in *TYW1*-KO and N.T. HeLa cells. D. Shown are the calculated effect size for all comparisons. All comparisons of *TYW1* vs. N.T were higher than 56, indicating a medium effect size (rank-sum p-value  $<10^{-5}$ ).



**Sup figure 5-1- tRNA Deep sequencing in T-cells shows agreement with codon usage and tRNA score.** A. tRNA read count agrees with tRNA score. Shown here is a scatter plot of the predicted tRNA score as calculated by tRNAscan-SE<sup>16</sup> compare to the fraction of reads of each unique tRNA. tRNAs with a score below 50 consider to be pseudo-tRNAs, and indeed few reads were aligned to those tRNAs. B. Anti-codon read count. Shown here is a scatter plot of

the tRNA gene copy number, calculated based on the number of tRNA of the same anti- codon with score above 50, compared to the fraction of reads of each tRNA anti-codon group

## Supplementary Tables

| Detectable modification | Undetectable modification |
|-------------------------|---------------------------|
| I                       | Ar(p)                     |
| N                       | Cm                        |
| acp3U                   | D                         |
| cmnm5Um                 | Gm                        |
| k2C                     | QtRNA                     |
| m1A                     | Um                        |
| m1G                     | Y                         |
| m1I                     | Ym                        |
| m2,2G                   | ac4C                      |
| m3C                     | cmnm5s2U                  |
| ms2i6A                  | cmo5U                     |
| o2yW                    | f5C                       |
| yW                      | galQtRNA                  |
|                         | gluQtRNA                  |
|                         | i6A                       |
|                         | m1Y                       |
|                         | m2A                       |
|                         | m2G                       |
|                         | m5C                       |
|                         | m5U                       |
|                         | m5Um                      |
|                         | m6A                       |
|                         | m6t6A                     |
|                         | m7G                       |
|                         | m8A                       |
|                         | mnm5U                     |
|                         | mnm5s2U                   |
|                         | s2C                       |
|                         | s4U                       |
|                         | t6A                       |
|                         | xA                        |
|                         | xG                        |
|                         | xU                        |

**Table S1- Detectable and undetectable RNA modifications.** Using tRNA sequencing data from *E.coli*, human, and mouse, we compared known modification positions (based on MODOMICS annotations) with the signal of mismatches and RT failures. The modification types annotated as “detectable modification” showed a clear mismatches/RT abortion signal in at least one position, at higher than 10% of the reads. “Undetectable modifications” did not show a pattern of mismatches/RT failure at any annotated position

| U-C Codon box<br>(avoidance of U<br>ending) | Amino Acid  | Main Gos   |
|---|-------------|--|
| UUU/UUC                                     | Phe         | cell killing, antigen processing and presentation, canonical Wnt signaling pathway, T cell activation                                      |
| UAU/UAC                                     | Tyr         | G protein-coupled receptor signaling pathway, brown fat cell differentiation, regulation of excretion,                                     |
| UGU/UGC                                     | Cys         | regulation of renal system process, metal ion homeostasis, G protein-coupled receptor signaling pathway                                    |
| GAU/GAC                                     | Asp         | positive regulation of small molecule metabolic process, L-amino acid transport, regeneration, canonical Wnt signaling pathway             |
| CAU/CAC                                     | His         | oxidative DNA demethylation, negative regulation of cell fate commitment, heat generation  |
| AGU/AGC                                     | Ser (6-box) | regulation of carbohydrate catabolic process, acrosome reaction, DNA packaging, positive regulation of antigen processing and presentation |
| AAU/AAC                                     | Asn         | chromatin organization, fatty acid elongation, negative regulation of multicellular organismal process                                     |

**Table S2- Go enrichment of genes avoiding the use of XXU vs XXC codons.**

We sorted genes based on the use of the indicated codons, and performed GO enrichment analysis using Gorilla. The significant GO categories that were enriched in this analysis are presented here for each codon pair.

### Supplementary Dataset Legends

**Dataset S1- GO categories expression in correlation with samples' 1<sup>st</sup> PC.** Correlation of gene set expression belongs to the indicated GO categories with PC1, calculated based on active codon usage.

**Dataset S2- tRNA modification in activated T-cells. Summary of all positions with mismatches and RT-abortions in all samples.** The threshold for inclusion in the table is minimum coverage of modified position = 20 reads, Modification fraction above 10%.

**Dataset S3- mRNA differential expression in activated T-cells.** Differential expression based on mRNA sequencing, as calculated by DEseq2. Marked in yellow are wybutosine biosynthesis enzymes.

**Dataset S4- mRNA expression, translation efficiency, and tAI in activated T-cells.** Summarized are the mRNA expression (based on mRNA-Seq), translation efficiency (based on Ribo-Seq), and calculated tAI (based on tRNA-sequencing) for each gene, at all time points. Sheet #2 show the fold change to Naïve sample..

1. Mikl, M., Alon, A., Mordret, E., Pilpel, Y. & Segal, E. Extensive programmed ribosomal frameshifting in human as revealed by a massively parallel reporter assay. *bioRxiv* 469692 (2018) doi:10.1101/469692.
2. Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–9 (2014).
3. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
4. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
5. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
6. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
7. Koster, J. & Rahmann, S. Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
8. Finkel, Y. *et al.* Comprehensive Annotations of Human Herpesvirus 6A and 6B Genomes Reveal Novel and Conserved Genomic Features. *bioRxiv* (2019) doi:10.1101/730028.
9. Tirosh, O. *et al.* The Transcription and Translation Landscapes during Human Cytomegalovirus Infection Reveal Novel Host-Pathogen Interactions. *PLoS Pathog.* **11**, (2015).
10. Lauria, F. *et al.* riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput. Biol.* **14**, (2018).
11. Zheng, G. *et al.* Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* **12**, 835–7 (2015).
12. Duttke, S. H., Chang, M. W., Heinz, S. & Benner, C. Identification and dynamic quantification of regulatory elements using total RNA. *Genome Res.* **29**, 1836–1846 (2019).
13. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
14. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
15. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
16. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–64 (1997).
17. Torres, A. G., Reina, O., Stephan-Otto Attolini, C. & Ribas de Pouplana, L. Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 8451–8456 (2019).
18. Thornlow, B. P. *et al.* Predicting transfer RNA gene activity from sequence and genome context. *Genome Res.* **30**, 85–94 (2020).
19. dos Reis, M., Savva, R. & Wernisch, L. Solving the riddle of codon usage preferences: A test for translational selection. *Nucleic Acids Res.* **32**, 5036–5044 (2004).

20. Gingold, H. *et al.* A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell* **158**, 1281–1292 (2014).
21. Crick, F. H. C. Codon—anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* **19**, 548–555 (1966).
22. Sakai, Y., Miyauchi, K., Kimura, S. & Suzuki, T. Biogenesis and growth phase-dependent alteration of 5-methoxycarbonylmethoxyuridine in tRNA anticodons. *Nucleic Acids Res.* **44**, 509–523 (2016).
23. Miyauchi, K., Kimura, S. & Suzuki, T. A cyclic form of N 6-threonylcarbamoyladenosine as a widely distributed tRNA hypermodification. *Nat. Chem. Biol.* **9**, 105–111 (2013).
24. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
25. Mcluckey, S. A., Van Berkel, G. J. & Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **3**, 60–70 (1992).