Limited but gene-class specific correlation of mRNA and protein expression in human CD8$^+$ T cells

Benoît P. Nicolet[1] and Monika C. Wolkers[1,*]

[1] Department of Hematopoiesis, Sanquin Research, Landsteiner Laboratory, Amsterdam UMC, University of Amsterdam, and Oncode Institute, Amsterdam, The Netherlands.

***** Corresponding author: M. Wolkers, Sanquin Research, Department of Hematopoiesis, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands. Tel: +31-20-5127003; Fax: +31-20-5123474; Email: m.wolkers@sanquin.nl

SUMMARY

T cell differentiation and activation induces substantial alterations in gene expression. While RNA sequencing and single cell RNA sequencing analysis provided important insights in the gene expression dynamics of T cells, it is not well understood how the mRNA expression translates into the protein landscape. By combining paired RNA-sequencing and mass spectrometry data of primary human $CD8^+$ T cells, we found that mRNA expression is a poor proxy for the overall protein output. Irrespective of the differentiation or activation status, the correlation coefficient of human $CD8^+$ T cells reached a mere 0.41-0.43. Only gene classes that mediate conserved cellular processes such as protein translation or cellular metabolism showed a strong correlation of mRNA with protein expression. In contrast, the mRNA expression and protein output of transcription factors, cell surface molecules, and secreted proteins - including cytokines - only mildly correlated. Conversely, highly conserved genes correlated well with the protein output. This was also true for the presence of AU-rich elements in the 3'untranslated region, in particular for mRNAs that encode secreted proteins. In conclusion, the in-depth characterization of the transcriptome and proteome in human $CD8^+$ T cells emphasizes the need of combined mRNA and protein analysis for our understanding of T cell biology and function.

KEYWORDS: mRNA-protein correlation; gene-class; AU-rich elements; CD8+ T cells; activation; differentiation; secreted proteins

INTRODUCTION

$CD8^+$ T cell responses are critical to clear and control viral infections and tumor cells. To exert their effector function, naïve T cells become activated in the draining lymph nodes by dendritic cells to differentiate into effector T cells (reviewed in (Harty and Badovinac, 2008)). Activated T cells then travel to the infected tissue where they clear antigen-presenting target cells by producing ample amounts of effector molecules. T cells also mature into different subsets of memory T cells, which protect us from recurring infections (Amsen et al., 2018; Kaech and Cui, 2012). Protection is achieved by their capacity to become rapidly reactivated upon antigen recognition, which results in the rapid production of effector molecules, such as the key proinflammatory cytokines IFN-γ, TNF-α and IL-2 (Han et al., 2012; Harty and Badovinac, 2008; Masopust, 2001; Reinhardt et al., 2001)

T cell differentiation and activation depends on substantial alterations of the transcriptional landscape, and on metabolic reprogramming (Best et al., 2013; van der Windt et al., 2012). To study these alterations in T cell differentiation and activation, analysis of the transcriptome, DNA methylation, histone modification, and the proteome have been instrumental (van Aalderen et al., 2017; Best et al., 2013; Moskowitz et al., 2017; Ross et al., 2016; Rudra et al., 2012; Stubbington et al., 2015; Vahedi et al., 2012). Dynamic genome-wide alterations were observed in murine and human $CD8^+$ T cells when responding to infection (Best et al., 2013; Roychoudhuri et al., 2016). While these findings were insightful, data integration of the transcriptome and proteome revealed a significant discrepancy between transcript and protein levels. In fact, model systems including NIH3T3 mouse fibroblasts, HeLa cells, or human and mouse T cells all showed a correlation coefficient of a mere ~ 0.4 (Cuadrado et al., 2018; Hukelmann et al., 2016; Nagaraj et al., 2011; Schwanhäusser et al., 2011). Whether this limited correlation coefficient is a feature of these reported steady state conditions, or whether it also holds true during T cell differentiation and upon T cell activation, is to date not known.

Evidence is accumulating that in addition to transcriptional and metabolic alterations, post-transcriptional and post-translational mechanisms are critical determinants for T cell effector function (Salerno et al., 2020). In activated Jurkat T cells, for instance, as much as 50% of the transcriptome changes was attributed to alterations in mRNA stability (Cheadle et al., 2005). Furthermore, the production of IFN-γ, TNF-α and IL-2 in primary $CD8^+$ T cells is subject to

the type of signal and the signal strength that T cells receive (Nicolet et al., 2017; Salerno et al., 2016, 2017). These signals define not only transcriptional regulation but also post-transcriptional events and alter the stability and translation efficiency of cytokine mRNA, which can result in discrepancies in mRNA and protein levels for these cytokines (Salerno et al., 2017, 2019b, 2019a). Untranslated regions (UTRs) are key regulatory elements in this process, as they are required for the modulation of mRNA transcript stability, translation initiation, and translation efficiency (Mayr, 2017). Several motifs within the 3'UTR contribute to this regulation. For example, AU-rich elements (AREs) are hubs for RNA binding proteins which in turn determine the fate of mRNA (Caput et al., 1986; Kontoyiannis et al., 1999; Stoecklin et al., 2003). Whether the observed discrepancy of mRNA and protein levels for IFN-γ, TNF-α and IL-2 also holds true for other gene products and whether gene regulation is defined by specific gene classes, or whether specific features in mRNA influence the protein output in T cells is to date not well understood.

Here, we compared the transcriptome and proteome expression of human $CD8^+$ T cells during differentiation, and upon activation. Independent of the differentiation status of the T cells, and irrespective of increased protein levels in activated T cells, the overall correlation coefficient of mRNA and protein remained unchanged. Nevertheless, we found that the relation of mRNA expression with protein expression was gene-class specific and depended on specific sequence characteristics. In conclusion, even though the overall correlation of mRNA and protein is poor, specific features can more accurately predict the mRNA-protein ratio during T cell differentiation and upon T cell activation.

RESULTS

Limited correlation of mRNA and protein abundance during $CD8^+$ T cell differentiation and upon T cell activation

We first asked how the mRNA and protein expression related during T cell differentiation. To this end, we used previously published RNA-sequencing (RNA-seq) and mass spectrometry (MS) data of peripheral blood-derived human naïve (T_N; $CD45RA^+$ $CD62L^{high}$), central memory (T_{CM}; $CD45RO^+$ $CD62L^{high}$ $CX3CR1^-$) and effector-memory (T_{EM}; $CD45RO^+$, $CD62L^{low}$, $CX3CR1^{high}$) $CD8^+$ T cells (Böttcher et al., 2015). After quasi-mapping ~ $3.5x10^6$ RNA-seq reads on average per sample onto the transcriptome, normalized counts (TPM) were extracted. TPM corrects for transcript length and sequencing depth, a critical feature for accurate quantification of RNA abundance. We detected 14,491, 14,705 and 14,532 genes expressed in T_N, T_{CM} and T_{EM} populations, respectively, representing ~70% of all protein-coding genes. Mass spectrometry measurements identified 6,069 proteins across all populations. To estimate the protein abundance, we computed the protein copy number (CN) from the MS data using the proteomic ruler methodology (Wiśniewski et al., 2014). mRNA and protein abundance spanned over 7.2 and 6.2 orders of magnitude, respectively (Figure S1A, B; Table S2).

CN and TPM units enable the integration of mRNA expression levels with protein measurements. 6,013 gene products were detected at mRNA and protein level, representing ~99,1% of detected protein and ~42% of all genes detected at mRNA level. Integration of mRNA and protein abundance revealed that irrespective of the differentiation status of the T cells, the mRNA-protein correlation was almost identical (Pearson's coefficient = 0.41-0.42; Spearman's coefficient: 0.36-0.37; Figure 1A). We also calculated the overall mRNA-protein slope, which reflects how the mRNA abundance translates into protein abundance (Figure S1C). The mRNA-protein slope, which is visualized by the regression line, was 0.62-0.63 in all three T cell subsets. This finding shows that the overall mRNA expression levels explains only ~62% of the protein levels (Figure 1A). Thus, the correlation coefficient and the slope of mRNA and protein abundance show some, but not a full overlap of mRNA and protein, a feature that does not alter during T cell differentiation.

We next investigated how T cell activation influenced the overall mRNA and protein abundance. This question is of particular interest because effector and memory T cells can swiftly adjust their transcriptional profile upon T cell activation and produce ample amounts

of effector molecules (Best et al., 2013; Han et al., 2012; Nicolet et al., 2017). In addition, the total protein amount nearly doubled in *in vitro* activated $CD8^+$ T cells from 47.35 pg/cell (n=16) of peripheral blood-derived $CD45RO^+$ $CD8^+$ T cells to 93.75 pg/cell (n=12; Figure 1B), a feature that is also found in activated murine T cells (Marchingo et al., 2020). How this increase in protein mass is reflected in the relation of mRNA and protein abundance, however, is not known. To address this question, we used our recently published RNA-seq and MS dataset on *in vitro* cultured $CD8^+$ T cells that were stimulated for 4h with PMA-Ionomycin ((Nicolet et al., 2020), *see methods*). RNA-seq libraries were prepared with ribosomal RNA depletion, therefore non-poly adenylated transcripts such as histones could also be quantified (Marzluff et al., 2008). Furthermore, protein secretion was blocked by Monensin during PMA-Ionomycin activation, thus making accurate measurements of most secreted proteins possible.

After quasi-mapping ~ $6.4x10^6$ RNA-seq reads per sample onto the transcriptome, TPM normalized counts per gene were extracted. Similar to T_N, T_{CM} and T_{EM} cells, we detected 13,055 protein-coding genes in activated $CD8^+$ T cells (data from (Nicolet et al., 2020)), which represents ~64% of all protein-coding genes (Figure S1D; Table S1). MS analysis detected a total of 4,626 proteins expressed in at least 2 samples (Figure S1E; Table S1; (Nicolet et al., 2020)), of which 4,583 proteins were co-detected at mRNA level (99.07%; Figure S1F). Thus, of 35.1% of protein-coding genes, protein expression was detected (Figure S1F). Notably, even though the mass per cell nearly doubled in activated T cells (Figure 1B), the correlation between mRNA and protein remained unchanged compared to resting *ex vivo* $CD8^+$ T cells (Figure 1C; Pearson's coefficient = 0.44, Spearman's coefficient = 0.39), just like the overall mRNA-protein slope (0.61; Figure 1C). In conclusion, the overall abundance of mRNA in human $CD8^+$ T cells only moderately correlates with protein expression, a feature that does not alter during T cell differentiation, or upon T cell activation.

Transcriptome and proteome of $CD8^+$ T cells are dominated by a few gene products

We then questioned how the overall mRNA and protein expression was distributed. In PMA-Ionomycin activated $CD8^+$ T cells, we observed a broad range of expression levels. The overall RNA abundance from the least to the most abundant mRNA spanned over 6 orders of magnitude (6.58 log10 span; Figure S1D). The mRNA landscape was dominated by a few gene products. Only 105 mRNAs (0.8% of 13,055) constituted 25% of the transcriptome, and

585 mRNAs (4.5%) represented 50% of the transcriptome (Figure S1G; Table S1). Histone mRNA belonged to the top 5 most abundant mRNAs, together with Ankyrin Repeat Domain 30B Like (ANKRD30BL), GM-CSF (CSF2), WD Repeat Domain 74 (WDR74), Mitochondrially Encoded Cytochrome C Oxidase II (MT-CO2), and X-C Motif Chemokine Ligand 2 (XCL2). Also the protein abundance spanned nearly 6 orders of magnitudes (Figure S1E; 5.62 log10 span). Nevertheless, a mere 6 proteins (0.13% of the 4,626 detected proteins), including histones and β-Actin (ACTB), accounted for 25% of the total proteome and 75 proteins (1.6%) constituted 50% of the proteome (Figure S1G; Table S1). Because the RNA-seq library preparation set of T_N, T_{CM} and T_{EM} cells was prepared with poly-A selection (Böttcher et al., 2015), the detection of histone mRNA was excluded from the analysis. Nevertheless, the remaining proteins showed a similar broad distribution of gene products (Table S2). Thus, a small fraction of gene products constitute the majority of mRNA and proteins in human $CD8^+$ T cells.

Distinct gene functions show differential correlation of mRNA with protein expression

We next asked whether the limited correlation of mRNA expression with protein output was equal for all genes in $CD8^+$ T cells, or whether different gene classes displayed specific features. We therefore compared the mRNA abundance with protein abundance for secreted proteins, CD molecules, transcription factors, RNA-binding proteins (RBPs), ribosomal proteins and Tricarboxylic (citric) acid cycle gene products (TCA), (Table S1, S2).

For this analysis, we pooled all $CD45RO^+$ human $CD8^+$ T cells (T_{MEM}) from ((Böttcher et al., 2015); see methods). T_{MEM} showed a Pearson's correlation coefficient of 0.54 and 0.57 for the 29 secreted proteins and the 105 CD molecules, respectively, which was higher than that of all genes combined (~0.41; Figure 1D, 2A). Conversely, the 278 transcription factors and 1,320 RNA-binding proteins followed the overall correlation coefficient with 0.32 and 0.39, respectively (Figure 2A). Interestingly, whereas mitochondrial ribosomal proteins showed limited correlation of mRNA and protein (coefficient=-0.34), the highly abundant mammalian ribosomal proteins displayed a high correlation coefficient (0.69). This high correlation coefficient was also observed for mRNA and protein expression of TCA cycle genes (0.6). Of note, the correlation coefficient of these 6 gene classes was comparable to those in PMA-Ionomycin activated $CD8^+$ T cells ($T_{Activated}$; Figure 2B). Thus, mRNA and protein relation in T_{MEM} and in activated $CD8^+$ T cells differs per gene class.

We next determined the ratio of mRNA and protein expression for each gene class in T_{MEM} $CD8^+$ T cells. Interestingly, not all gene classes had the same protein output for similar mRNA levels (Figure 2C, top panel). With the exception of mammalian ribosomal proteins, all gene classes had an average of ~30 TPM of mRNA expression. Despite the comparable average mRNA levels across all gene classes, TCA genes had with ~140,000 copy numbers a much higher protein output than all other gene classes (~3,000-30,000 copy numbers; Figure 2C, top panel)). Similar patterns of mRNA/protein expression were observed in activated T cells, with the exception of secreted molecules that showed increased mRNA expression (Figure 2C, bottom panel). Nevertheless, we did not detect alterations in average protein levels for this gene class (Figure 2C, bottom panel). We then determined the mRNA-protein slope per gene class in T_{MEM} $CD8^+$ T cells. Mitochondrial ribosomal proteins had with -0.36 the lowest slope of all gene classes (Figure 2D), and mRNA expression levels did not explain protein abundance at all. For transcription factors and secreted molecules, mRNA levels explained ~40% of protein levels (slope: 0.4-0.41). CD molecules, mammalian ribosomal proteins and RBPs mRNA levels explained ~60% of protein levels (slopes: 0.58-0.68; Figure 2D). Again, TCA cycle genes stood out by a very strong relation between mRNA and protein levels, with a slope of 1.22. This mRNA-protein slope did not substantially alter in *in vitro* activated $CD8^+$ T cells (Figure 2D; Table S1). Only the slope of secreted proteins increased from 0.4 in T_{MEM} $CD8^+$ T cells to 0.61 in activated $CD8^+$ T cells, while the slope of CD molecules decreased (0.58 in T_{MEM} to 0.41 in activated T cells; Figure 2D). In conclusion, each gene class shows distinct correlation coefficients, different mRNA-protein ratios and slopes, a feature that is maintained for most gene classes in T_{MEM} and in activated $CD8^+$ T cells, except for secreted proteins and CD molecules.

Discrepancies of mRNA and protein abundance within gene classes

Importantly, within gene classes, gene-specific discrepancies in mRNA abundance and protein expression were also apparent. For instance, the mRNA levels of the secreted proteins GM-CSF (CSF2) and IL-8 (CXCL8) differed by 37-fold, yet their protein levels only differed by a 3-fold (Figure 3A). Similarly, Granzyme A and B (GZMA & GZMB) mRNA levels differed by 38-fold but protein levels remained unchanged (Figure 3B). Of note, monensin blocks de novo protein production by inhibiting golgi transport, and may not block the secretion of preformed granzymes. For surface molecules, similar mRNA expression of *CD8A* and *IL2RB* (CD122) also did not translate in equal protein levels. The protein

abundance of CD8α was 27-fold higher than that of IL-2Rβ (Figure 3C). The mRNA and protein abundance of the IL2-receptor family members α:β:γ correlated well and showed a similar protein stoichiometry with 3.7:1:6.5. This was in contrast to the reported differential stoichiometry 92:1:2 in murine $CD8^+$ T cells (Figure 3D; (Hukelmann et al., 2016)), suggesting a possible differential sensitivity to IL-2 between human and mouse. The mRNA levels of two key T cell transcription factors *STAT1* and *TBX21* (T-bet) was equally abundant, but the protein expression of STAT1 was 153 times higher than that of T-bet (Figure 3E). This gene-specific discrepancy in mRNA and protein levels was also apparent for the RBPs ILF3 (NF90), ZFP36L1 (Tis11b) and ELAVL1 (HuR) (Figure 3F). Even through *ZFP36L1* and *ILF3* mRNA levels were 8-fold higher than that of *ELAVL1* mRNA, ZFP36L1 protein levels were 22-fold lower than those of ILF3, and ELAVL1 (Figure 3F). Thus, mRNA and protein levels within gene classes can substantially vary.

Gene and transcript characteristics influence mRNA and protein abundance

We next sought to identify which features in the mRNA could explain the discrepancy of mRNA and protein expression. High sequence conservation has been associated with high gene expression levels in mammals (Villanueva-Cañas et al., 2017). We therefore questioned whether this was also true for both mRNA and protein abundance in PMA-Ionomycin activated $CD8^+$ T cells. Of the 4,583 gene products that were co-detected at mRNA and protein level in PMA-Ionomycin activated $CD8^+$ T cells, 4,087 (89.2%) gene products were present in both human and zebra fish (*Danio rerio*). We subdivided these 4,087 gene products into genes with high (top 25% of genes: >76.9% conservation), intermediate (25-75% of genes: 50%-76.9% conservation), and low gene (pre-mRNA) sequence homology (bottom 25% of genes: <50% conservation; Figure 4A, S2A; Table S1). Genes that were highly conserved also had the highest mRNA abundance and protein abundance in activated T cells (Figure 4A). This translated in a high mRNA-protein slope with 0.79 for these well-conserved transcripts, while it only reached 0.47 and 0.42 for genes with intermediate and low conservation, respectively (Figure 4B). Evolutionary sequence conservation can also result in a bias of codon usage and altered GC-content of the mRNA coding region, translated into higher mRNA levels and translation efficiency (Kudla et al., 2006; Rudolph et al., 2016). Indeed, we observed that genes with a high GC content (top 25% of genes: GC>50.9%) had a significantly higher mRNA and protein abundance compared to genes with low or intermediate GC content (bottom 25% of genes: GC<39.7%, and intermediate 50% of genes:

39.7%<GC<50.9%, respectively; Figure S2A, B; Table S1). Also the mRNA-protein slope increased with high GC content, with 0.66 and 0.63 for high and intermediate GC content, and 0.43 for genes with a low GC content (Figure S2C; Table S1).

Another potent regulator of mRNA expression and protein production are the untranslated regions (UTR) of an mRNA (Mayr, 2017). To define whether the 5'UTR length influenced the mRNA and protein abundance, we first stratified the genes into short (shortest 25%: < ~50nt), intermediate (25-75%), and long 5'UTRs (longest 25%: > ~200nt; Figure 4C, D, S2A; Table S1). mRNA abundance was not affected by the length of the 5'UTR (Figure 4C). In line with previous studies in the yeast strain *Saccharomyces cerevisiae* (Dvir et al., 2013), slight but significantly higher protein levels were measured for genes with a short 5'UTR (Figure 4C). Furthermore, genes encoding short 5'UTRs showed a slightly higher mRNA-protein slope than genes with longer 5'UTRs (short: 0.69, intermediate: 0.61, long: 0.53; Figure 4D). For the analysis of the 3'UTR, we divided genes based on short (bottom 25%, < ~100nt), intermediate (25-75%), and long (top 25%, > ~1,000nt) 3'UTRs (Figure 4E, F, S2A; Table S1). Long 3'UTRs slightly suppressed the mRNA abundance (Figure 4E, F). In concordance with previous work in murine T cells (Sandberg et al., 2008), this also translated into lower protein abundance and a lower slope for mRNAs with long 3'UTRs compared to short 3'UTRs (long: 0.44, intermediate: 0.63, short: 0.72; Figure 4E, F). Thus, the length of the UTRs - and in particular of the 3'UTR - influences the mRNA and protein level in activated T cells.

One of the key regulatory elements in 3'UTRs are A-U rich elements (AREs) (Caput et al., 1986; Chen and Shyu, 1995; Mayr, 2017). To define if AREs modify mRNA and protein abundance in activated $CD8^+$ T cells, we applied ARED-plus AU-rich element annotations from (Bakheet et al., 2018). ARED-plus annotations define 1 cluster as WWWW(AUUUA)WWWW (where W is A or U), 2 clusters as AUUUAUUUAU or UAUUUAUUUA flanked by W or WW, and 3 and over clusters are defined as AUUUAUUUAUUUA and additional repeats of AUUUA (Bakheet et al., 2018). The mRNA abundance was not different in the presence or absence of ARE clusters (Figure 4G; Table S1). In contrast, independent of the number of AREs present, ARE clusters within the 3'UTR significantly affected the protein abundance (Figure 4G). Consequently, ARE cluster-containing transcripts displayed a lower mRNA-protein slope with 0.34-0.44 than transcripts without ARE clusters (0.68; Figure 4H). In conclusion, sequence conservation is the best

predictor of mRNA and protein abundance and is influenced by the 3'UTR length and the presence of ARE clusters within the 3'UTR.

T cell activation alters the effect of ARE clusters on mRNA and protein abundance

To determine whether ARE-mediated regulation of mRNA and protein abundance depended on the activation status of T cells, we compared the mRNA expression of all genes detected in T_{MEM} $CD8^+$ T cells (n= 15,038 genes) and in PMA-Ionomycin activated T cells (n= 13,055 genes), irrespective of whether protein expression was measured or not (Figure 5A, B, ; Table S1, S2). mRNAs containing one or more ARE clusters in the 3'UTR expressed higher mRNA levels in T_{MEM} cells than mRNA containing no ARE (Figure 5A). However, activated T cells showed no differences in mRNA expression whether ARE clusters were present or not (Figure 5B). In contrast, protein abundance of the n=6,069 and n=4,626 detected proteins in T_{MEM} and activated T cells, respectively, was substantially affected by the presence of ARE clusters irrespective of the T cell activation status (Figure 5A, B), indicating that AREs consistently modulate the protein output in resting and in activated T cells.

We also investigated whether the effect of ARE clusters on mRNA expression and protein output is gene-class specific. For CD molecules and transcription factors, we did not detect any influence of ARE clusters on the mRNA or protein levels in T_{MEM}, or in PMA-Ionomycin activated T cells (Figure S3 A-D). Furthermore, the slope of mRNA-protein of CD molecules and of transcription factors only showed marginal, if any, changes when comparing ARE-containing gene products in T_{MEM} with activated T cells (Figure S3 E- H). Only genes not containing ARE clusters were subject to different mRNA-protein slopes between T_{MEM} and activated T cells, pointing to ARE-independent effects for these genes (Figure S3G, H).

When we measured the effect of ARE-clusters on secreted proteins, no significant effect was found the mRNA abundance in T_{MEM} $CD8^+$ T cells (Figure 5C). Yet, the average mRNA levels went down with increasing numbers of ARE clusters (means: 0 ARE: 1.2 TPM; 1 ARE: 0.38 TPM; >1 ARE: 0.14 TPM; Figure 5C). Interestingly, even though T cell activation resulted in overall increased mRNA levels, this was most apparent for mRNAs with >1 ARE cluster (>1ARE: mean 30.4 TPM; Figure 5D). None of the 23 mRNAs containing >1 ARE clusters found in T_{MEM} $CD8^+$ T cells had detectable protein levels (Figure 5C). This changed however upon T cell activation. 6 out of the 17 mRNAs with >1 ARE

clusters detected in activated T cells, i.e. CCL3, CSF2, CXCL8, IL2, PTGS2, TNF, now showed high protein abundance (Figure 5D). Thus, ARE clusters suppress protein translation of several transcripts present in T_{MEM} cells, and T cell activation increases mRNA and protein expression of the ARE cluster containing transcripts.

We also calculated the slope of mRNA and protein abundance for secreted proteins. In T_{MEM} cells, mRNA with >1 ARE clusters did not express protein (Figure 5A), which impeded slope calculation. The slope of transcripts in T_{MEM} with 0 or 1 ARE cluster was identical with 0.39 (Figure 5C). However, in activated T cells, the slope of mRNA and protein abundance increased irrespective of the ARE cluster content (Figure 5D). Even though only 7 and 6 mRNA/proteins were co-detected 1 and >1 ARE clusters, respectively, the slope differed between these two groups (1 ARE: 1.38; >1 ARE: 0.35; Figure 5F). This finding suggests that in particular ARE cluster -containing transcripts are subject to post-transcriptional regulation.

DISCUSSION

With an integrated analysis of transcriptome and proteome data from primary human CD8+ T cells, we show here that the mRNA expression only moderately correlates with protein expression in primary human $CD8^+$ T cells. This feature is independent of the differentiation status of T cells. Furthermore, even though T cell activation doubles the protein content in T cells, this is not affecting the overall correlation coefficient of mRNA and protein abundance. Of note, we detected peptides for about ~35% of the protein-encoding mRNAs by MS analysis, which allowed for the correlation of mRNA and protein abundance of only this subset of gene products. Thus, even though our analysis is restricted to this subset of gene products, the observed limited correlation at least holds true for proteins that are highly expressed and/or that contain peptides suitable for MS detection.

The limited correlation in mRNA and protein abundance was previously reported in other model systems (Cuadrado et al., 2018; Hukelmann et al., 2016; Nagaraj et al., 2011; Schwanhäusser et al., 2011), and has been attributed to differences in transcriptional rates, mRNA stability for mRNA abundance, and to protein synthesis and degradation rates for protein abundance (Vogel and Marcotte, 2012). Which of these parameters is the prime driver of the mRNA expression levels and protein output could be gene specific and may depend on the differentiation and activation status of a cell. Indeed, whereas the overall correlation between mRNA and protein was limited, we found gene-class specific features of mRNA and protein abundance. Mitochondrial ribosomal proteins are only stable when bound to the mitochondrial ribosome (Bogenhagen et al., 2018). Degradation of unbound mitochondrial ribosomal proteins could thus lead to the observed discrepancy in mRNA and protein abundance. In contrast, the mRNA of mammalian ribosomal proteins is considered stable (Perry, 2007), and mammalian ribosomal proteins are known to be expressed in unbound form as well (Simsek et al., 2017). Similarly, TCA genes not only expressed high levels of mRNA and protein but also showed a strong correlation of mRNA abundance with protein abundance. It is conceivable that conservation of these so-called 'housekeeping' genes resulted in high transcription and translation efficiency (Berthelot et al., 2018).

Protein stability is another important determinant of the overall protein levels (Kristensen et al., 2013; Schwanhäusser et al., 2011). Mammalian ribosomal proteins and nucleosomes are more stable than secreted proteins, or zinc-finger containing proteins (including RNA-

binding proteins and transcription factors (Cassandri et al., 2017; Kristensen et al., 2013)). Although we did not directly address protein stability in differentiating and activated human $CD8^+$ T cells, the high correlation for mRNA and protein expression for TCA proteins and ribosomal protein match with these studies and reiterate the predictive power of gene sequence conservation on mRNA and protein levels.

We also observed that a high GC content is a good predictor of higher mRNA and of high protein levels. In line with this finding, GC content manipulation of the IL6 and IL2 gene resulted in higher mRNA stability, ribosome occupancy and protein output (Hia et al., 2019; Kudla et al., 2006). GC content also influences codon usage and secondary structures which in turn increase translation efficiency and protein output (Chen et al., 2004; Fang et al., 2018; De Smit and Van Duin, 1990; Vogel et al., 2010).

Not only sequence conservation but also the length of the 5' and 3'UTR influences the mRNA and/or protein abundance. The length of the UTR alone however is unlikely to be the sole factor that influences mRNA and protein abundance. The likelihood of containing regulatory motifs or structures also increases with length, as we observed for the increased presence of ARE clusters in longer 3'UTRs (*data not shown*). It is likely that this also holds true for other regulatory motifs could also be associated with longer 3'UTRs. It is important to note that 48h activation of murine T cells with aCD3/CD28 beads shortens the length of 3'UTRs (Sandberg et al., 2008). Whether this also occurs after 4h PMA-Ionomycin activation is yet to be determined. It is however conceivable that alterations in 3'UTR length could also contribute in defining mRNA expression in activated T cells, at least in part by altering the content of miRNA seeds and thus mRNA stability (Mayr and Bartel, 2009).

Also the presence of AU-rich elements was a strong predictor of mRNA and protein levels (Kontoyiannis et al., 1999; Salerno et al., 2018, 2019a). Interestingly, ARE-mediated regulation in CD molecules and transcription factors did not alter mRNA and protein expression, irrespective of the number of ARE in cluster or the activation status. Only ARE-containing transcripts encoding secreted proteins had high mRNA levels while having low to undetectable protein levels in memory T cells. Upon activation, however these transcripts showed high mRNA and high protein levels, reflecting the discrepancy of mRNA and protein levels of IFN-γ in murine memory and tumor-infiltrating T cells due to a ARE-dependent translational block (Salerno et al., 2018, 2019a). What makes transcripts susceptible to this

regulation is yet to be determined, but it could possibly stem from subtle differences in position and spacing between AREs, the surrounding RNA structures, or the interplay with other regulatory mechanisms on the same transcript.

In summary, we here provide the methodology to study the transcriptome and proteome regulation translated to primary human $CD8^+$ T cells. Our findings demonstrate that the limited relation between the mRNA and protein levels is gene class specific and determined by sequence conservation and by sequences and structures of untranslated regions, and at least for ARE cluster containing gene products by T cell activation. Our study will therefore serve as valuable resource for further studies of $CD8^+$ T cell biology and function.

ACKNOWLEDGMENTS

We would like to thank B.P. Popovic and N.D. Zandhuis for critical reading of this manuscript.

AUTHOR CONTRIBUTION

B.P.N. and M.C.W. designed experiments; B.P.N. performed experiments and analyzed data. M.C.W. directed the study; B.P.N. and M.C.W wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FUNDING

This research was supported by the Oncode Institute, and the European research council (ERC-Printers 817533).

FIGURE LEGENDS

Figure 1: Transcriptome and proteome integration in human $CD8^+$ T cells

(**A**) mRNA abundance in transcript per kilobase per million (TPM) and protein abundance data in protein copy number (CN) from (Böttcher et al., 2015) of blood-derived $CD8^+$ T cells. Compiled data of naïve (T_N; $CD45RA^+$ $CD62L^{high}$; n=3), central memory (T_{CM}; $CD45RO^+$ $CD62L^{high}$ $CX3CR1^-$; n=3) and effector-memory (T_{EM}; $CD45RO^+$, $CD62L^{low}$, $CX3CR1^{high}$; n=4) T cell subsets are depicted. (**B**) Estimated protein mass per cells using the proteomic ruler methodology (Wiśniewski et al., 2014) for all $CD45RO^+$ memory T cell subsets combined (T_{MEM}; n=16; from (Böttcher et al., 2015)) and for $CD8^+$ T cells that were activated for 2 days with αCD3/αCD28, cultured for 4 days, and were then re-activated for 4h with PMA-Ionomycin ($T_{Activated}$; n=12; from (Nicolet et al., 2020)). (**C**, **D**) mRNA abundance and protein abundance of (**C**) $T_{Activated}$ (n=12), and (**D**) T_{MEM} (n=16). The Pearson's correlation coefficient, the linear regression (grey line), and its slope (Figure S1C) are indicated in each graph. The p-value resulting from two-tailed t-test is indicated in (B).

Figure 2: Gene class-intrinsic relation of mRNA and protein abundance

(**A-B**) Compiled mRNA (TPM) and protein abundance (CN) of secreted proteins, CD molecules, transcription factors, RNA-binding proteins (RBPs), mammalian (light green) and mitochondrial (dark green) ribosomal proteins, and TCA (citric acid) cycle proteins in (**A**) of $CD45RO^+$ $CD8^+$ T_{MEM} cells, and in (**B**) of PMA-Ionomycin activated $CD8^+$ T cells ($T_{Activated}$). The number of genes per class, the Pearson's correlation coefficient, and regression line (linear model) were calculated for each class, and are indicated in the panels. (**C**) Average mRNA and protein abundance of gene classes indicated in (A, B) of T_{MEM} (top panel) and $T_{Activated}$ (bottom panel). (**D**) The linear regression slope was calculated, clustered and plotted for each gene class.

Figure 3: Heterogenous mRNA and protein abundance within gene classes

(**A-F**) mRNA (TPM) and protein (CN) abundance in PMA-Ionomycin activated $CD8^+$ T cells of indicated genes. Genes belong to secreted proteins (A, B), surface molecules (C, D), RNA-binding proteins (E), and transcription factors (F). Numbers in graph depict the fold difference in expression levels. Each dot indicates the value measured in one individual sample (n=12).

Figure 4: Altered mRNA and protein abundance is associated with gene characteristics

(**A**, **B**) The effect of sequence conservation on mRNA (A, left panel) and protein abundance (A, right panel) and the relation thereof (B) in PMA-Ionomycin activated $CD8^+$ T cells. Genes products were stratified according to Figure S2A, in: (**A, B**) high (>76.9%), intermediate (between 50-76.9%) and low (<50%) orthologous sequence conservation between Human and Zebra-fish (*Danio rerio*); (**C, D**) short (<~50nt), intermediate (~50nt-~200nt), and long 5'UTRs (>~200nt); (**E, F**) short (<~100nt), intermediate (~100- ~1,000nt), and long (>~1,000nt) 3'UTRs; (**G, H**) number of ARE in clusters according to ARED-plus annotation (see methods). The slope of the regression line of each group is color-coded in B, D, F, H. Differences between groups were assessed with a two-tailed t-test followed by adjustment of p-value using the Benjamini-Hochberg procedure. Adjusted p-value are indicated. Only gene products that are detected at mRNA and at protein levels are shown. TPM: Transcripts per kilo-base per million; CN: Protein copy number.

Figure 5: ARE clusters differentially modulate mRNA and protein levels in resting and activated $CD8^+$ T cells

(**A-D**) mRNA (TPM) and protein (CN) abundance of all genes grouped according to the number of AU-rich element in cluster within the transcript's 3'UTR, for (**A**) T_{MEM} or (**B**) $T_{Activated}$ $CD8^+$ T cells. (**C**, **D**) similar to (A, B) for all secreted proteins. (**E**, **F**) Integration of mRNA and protein abundance from (C, D), for genes detected both at mRNA and protein levels. The linear regression and its slope are color-coded for each group. Difference between groups were assessed with a two-tailed t-test followed by adjustment of p-value using the Benjamini-Hochberg procedure. Adjusted p-value are indicated. TPM: Transcripts per kilo-base per million; CN: Protein copy number.

MATERIALS AND METHODS

Data acquisition

For analysis of mRNA and protein abundance during T cell differentiation, we used MS data and paired RNA-seq data retrieved, respectively, from *Böttcher et al.* (Supplemental table 3 of: (Böttcher et al., 2015)), and on the National Center for Biotechnology Information (NCBI) gene expression omnibus (GEO) repository under the accession no: GSE63144. Data of the biological replicate of human naïve (T_N; $CD45RA^+$ $CD62L^{high}$), central memory (T_{CM}; $CD45RO^+$ $CD62L^{high}$ $CX3CR1^-$) and effector-memory (T_{EM}; $CD45RO^+$, $CD62L^{low}$, $CX3CR1^{high}$) $CD8^+$ T cells were averaged per group for analysis. For all other analyses, $CD45RO^+$ $CD8^+$ T cells (T_{MEM}) samples including all different memory subsets were averaged for analysis. Sequencing runs of the same samples were concatenated. For analysis on activated $CD8^+$ T cells, we used our previously published RNA-seq and MS dataset (Nicolet et al., 2020), available on the NCBI GEO repository under the accession no: GSE125497, and on the Proteome Xchange Consortium via the PRoteomics IDEntifications Database (PRIDE) repository with the dataset identifier no: PXD012874. RNA-seq and MS data from each biological replicate containing IFN-γ producers, IL-2 producers, IFN-γ/IL-2 double producers, and double negative producers were pooled, and the average expression was used.

RNA-seq analysis

An average of 10.5 million (*Böttcher et al.*) and 21.4 million (*Nicolet et al.*) reads per sample were used for quasi-mapping using Salmon (version 0.13) on the human coding transcriptome hg38-release 92 from ENSEMBL (Zerbino et al., 2018). The normalized counts in 'Transcripts per kilobase per million' (TPM; which corrects for transcript length and library size) was calculated by Salmon, and ENSEMBL annotations were used to filter for protein-coding transcripts, and to sum up the TPMs per genes. Genes with TPM>0.1 (*Böttcher et al*) and TPM>0.01 (*Nicolet et al*) were considered as expressed. Due to lower sequencing depth in samples from *Böttcher et al*, a higher cut-off was used. For each study separately, TPMs per gene were log10-tranformed and averaged across samples as describe above.

MS analysis

For samples from *Nicolet et al* (Nicolet et al., 2020), the RAW mass spectrometry files were processed as described before (Nicolet et al., 2020). The Perseus (version 1.6.0.7) plugin for Protein ruler methodology (Wiśniewski et al., 2014) was used to calculate protein copy number (CN) and protein content per cell (in pg per cell). The CN values of all 12 biological replicates were averaged, and used to integrate with mRNA data. For samples from (Böttcher et al., 2015), the raw data were not available. LFQ values provided by the authors were used to compute the CN values in Perseus. Values from all samples were averaged to integrate with mRNA data.

RNA-protein relation in different gene classes

Genes encoding secreted proteins were extracted from https://www.proteinatlas.org/humanproteome/tissue/secretome (downloaded Jan 2018), and manually curated the list for obvious mis-annotations (histones, membrane, TF, RNA-binding, CD molecules, protein without protein evidence, collagen, nuclear, nucleus, ribosomal, mitochondrial). Similarly, genes and proteins encoding CD molecules were extracted with the corresponding gene and protein names (www.uniprot.org/docs/cdlist.txt; downloaded July 2017). Genes of experimentally validated RNA-binding proteins were obtained from (Castello et al., 2016; Gerstberger et al., 2014; Perez-Perri et al., 2018). Genes encoding for transcription factors, ribosomal proteins, and TCA-cycle were obtained from https://www.proteinatlas.org/humanproteome/proteinclasses (downloaded Jan. 2019). A-U rich element (ARE) cluster content of 3'UTRs was obtained from ARED-plus (Bakheet et al., 2018). Genes with more than one ARE cluster were annotated as ">1 ARE". GC content of genes (pre-mRNA sequence), orthologous (pre-mRNA) sequence conservation between Zebra fish (*Danio rerio*) and human, 5' and 3'UTR length per gene, were obtained from Ensembl BioMART (ENSEMBL; (Kinsella et al., 2011)). For each studied parameter (Figure 4 and S2), data were subdivided into 3 groups based on the distribution of the parameter: bottom 25%, intermediate 50%, and top 25%.

Statistical analysis and Data visualisation

The mRNA-protein slope is the slope coefficient defined by the linear model (*lm()* function of R 'stats' package) using the formula '*lm(protein ~ mRNA)*' where protein is the log10(*copy number*) and mRNA is the log10(*TPM*). The mRNA protein ratio is calculated using the average mRNA in log10(*TPM*) plotted against the average protein in log10(copy number), for each gene class. Statistical analysis was performed in R using a two-tailed t-test

and the resulting p-values were adjusted using Benjamini-Hochberg procedure. Differences were considered significant if adjusted p-value <0.05, and depicted in each graph. Plots were generated with ggplot2 version 3.0, and with GraphPad PRISM version 8. Heatmaps were generated with pHeatmap version 1.0.10.

SUPPLEMENTAL INFORMATION:

Figure S1: mRNA and protein abundance in *in vitro* activated CD8$^+$ T cells.

(**A**, **B**) Distribution of (**A**) mRNA (TPM) and (**B**) protein (CN) abundance in CD8$^+$ T_{MEM}. (**C**) Graph representing mRNA and protein correlation (such as Pearson's correlation) and slope of a linear regression. Of note, the correlation coefficient represents a numerical integration of the dispersion of the data points, and is not a measure of the angle of the regression line. (**D**, **E**) Distribution of (**D**) mRNA (TPM) and (**E**) protein (CN) abundance in CD8$^+$ $T_{Activated}$. (**F**) Venn diagram depicting the overlap in gene (>0.01 TPM, n= 13,055) and protein (n=4,626) expressed in $T_{Activated}$. (**G**) Cumulative contribution of individual mRNA to the transcriptome (in blue) and of individual protein to the proteome (in red). mRNA and protein are ranked from high to low expression. The names of the top 5 most abundant mRNA and protein are indicated. TPM: Transcripts per kilo-base per million; CN: protein copy number.

Figure S2: Altered mRNA and/or protein abundance is associated with gene characteristics

(**A**) Distribution and color-coded cut-off (top 25%; mid. 50%; bottom 25% of all genes) for the stratification of orthologous gene sequence conservation between Human and Zebrafish in %, GC content in %, 5' and 3'UTR length. (**B**) Separate, or (**C**) integration of mRNA (TPM) and protein (CN) abundance in CD8$^+$ $T_{Activated}$. Only gene products that are detected both at mRNA and protein levels are shown. Genes products were stratified in: high (>50.9%), intermediate (39.7%<GC<50.9%) and low (<39.7%) GC content per gene. The linear regression line and its slope is color-coded for each group. Difference between groups were assessed with a two-tailed t-test followed by adjustment of p-value using the Benjamini-Hochberg procedure. Adjusted p-values are indicated.

Figure S3: ARE-mediated regulation of transcripts encoding CD molecules and transcription factors transcript does not alter upon T cell activation

(**A-H**) mRNA (TPM) and protein (CN) abundance (**A-D**) and integration (**E-H**) from T_{MEM}, or $T_{Activated}$ CD8$^+$ T cells, for CD molecules (A, B, E, F) and transcription factors (C, D, G, H) gene classes. Genes products are color-coded according to the number of AU-rich element in cluster within the transcript's 3'UTR. Difference between groups were assessed with a two-tailed t-test followed by adjustment of p-value using the Benjamini-Hochberg procedure. TPM: Transcripts per kilo-base per million; CN: protein copy number.

Table S1: Transcriptome and proteome data in $T_{Activated}$ from (Nicolet et al., 2020).

Table S2: Transcriptome and proteome data of T_N, T_{CM}, T_{EM} and T_{MEM} from (Böttcher et al., 2015)

REFERENCES

van Aalderen, M.C., van den Biggelaar, M., Remmerswaal, E.B.M., van Alphen, F.P.J., Meijer, A.B., ten Berge, I.J.M., and van Lier, R.A.W. (2017). Label-free Analysis of CD8+T Cell Subset Proteomes Supports a Progressive Differentiation Model of Human-Virus-Specific T Cells. Cell Rep. *19*, 1068–1079.

Acosta-Rodriguez, E. V., Rivino, L., Geginat, J., Jarrossay, D., Gattorno, M., Lanzavecchia, A., Sallusto, F., and Napolitani, G. (2007). Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. Nat. Immunol. *8*, 639–646.

Amsen, D., Van Gisbergen, K.P.J.M., Hombrink, P., and Van Lier, R.A.W. (2018). Tissue-resident memory T cells at the center of immunity to solid tumors. Nat. Immunol. *19*, 538–546.

Bakheet, T., Hitti, E., and Khabar, K.S.A. (2018). ARED-Plus: An updated and expanded database of AU-rich element-containing mRNAs and pre-mRNAs. Nucleic Acids Res. *46*, D218–D220.

Berthelot, C., Villar, D., Horvath, J.E., Odom, D.T., and Flicek, P. (2018). Complexity and conservation of regulatory landscapes underlie evolutionary resilience of mammalian gene expression. Nat. Ecol. Evol. *2*, 152–163.

Best, J.A., Blair, D.A., Knell, J., Yang, E., Mayya, V., Doedens, A., Dustin, M.L., Goldrath, A.W., Monach, P., Shinton, S.A., et al. (2013). Transcriptional insights into the CD8 + T cell response to infection and memory T cell formation. Nat. Immunol. *14*, 404–412.

Bogenhagen, D.F., Ostermeyer-Fay, A.G., Haley, J.D., and Garcia-Diaz, M. (2018). Kinetics and Mechanism of Mammalian Mitochondrial Ribosome Assembly. Cell Rep. *22*, 1935–1944.

Bonecchi, R., Bianchi, G., Bordignon, P.P., D'Ambrosio, D., Lang, R., Borsatti, A., Sozzani, S., Allavena, P., Gray, P.A., Mantovani, A., et al. (1998). Differential expression of chemokine receptors and chemotactic responsiveness of type 1 T helper cells (Th1s) and Th2s. J. Exp. Med. *187*, 129–134.

Böttcher, J.P., Beyer, M., Meissner, F., Abdullah, Z., Sander, J., Höchst, B., Eickhoff, S., Rieckmann, J.C., Russo, C., Bauer, T., et al. (2015). Functional classification of memory CD8 + T cells by CX 3 CR1 expression. Nat. Commun. *6*, 8306.

Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S., and Cerami, A. (1986). Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. Proc. Natl. Acad. Sci. U. S. A. *83*, 1670–1674.

Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., and Raschellà, G. (2017). Zinc-finger proteins in health and disease. Cell Death Discov. *3,17071*.

Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.-M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2016). Comprehensive Identification of RNA-Binding Domains in Human Cells. Mol. Cell *63*, 696–710.

Cheadle, C., Fan, J., Cho-Chung, Y.S., Werner, T., Ray, J., Do, L., Gorospe, M., and Becker, K.G. (2005). Control of gene expression during T cell activation: Alternate regulation of mRNA transcription and mRNA stability. BMC Genomics *6*, 1–16.

Chen, C.Y.A., and Shyu, A. Bin (1995). AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. *20*, 465–470.

Chen, S.L., Lee, W., Hottes, A.K., Shapiro, L., and McAdams, H.H. (2004). Codon usage between genomes is constrained genome-wide mutational processes. Proc. Natl. Acad. Sci. U. S. A. *101*, 3480–3485.

Cuadrado, E., van den Biggelaar, M., de Kivit, S., Chen, Y. yen, Slot, M., Doubal, I., Meijer, A., van Lier, R.A.W., Borst, J., and Amsen, D. (2018). Proteomic Analyses of Human Regulatory T Cells Reveal Adaptations in Signaling Pathways that Protect Cellular Identity. Immunity *48*, 1046–1059.e6.

Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L.B., Weinberger, A., and Segal, E. (2013). Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. Proc. Natl. Acad. Sci. *110*, E2792–E2801.

Fang, H., Huang, Y.F., Radhakrishnan, A., Siepel, A., Lyon, G.J., and Schatz, M.C. (2018). Scikit-ribo Enables Accurate Estimation and Robust Modeling of Translation Dynamics at Codon Resolution. Cell Syst. *6*, 180–191.e4.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Han, Q., Bagheri, N., Bradshaw, E.M., Hafler, D.A., Lauffenburger, D.A., and Love, J.C. (2012). Polyfunctional responses by human T cells result from sequential release of cytokines. Proc. Natl. Acad. Sci. *109*, 1607–1612.

Harty, J.T., and Badovinac, V.P. (2008). Shaping and reshaping CD8+ T-cell memory. Nat. Rev. Immunol. *8*, 107–119.

Hia, F., Yang, S.F., Shichino, Y., Yoshinaga, M., Murakawa, Y., Vandenbon, A., Fukao, A., Fujiwara, T., Landthaler, M., Natsume, T., et al. (2019). Codon bias confers stability to human mRNA s . EMBO Rep. *20*, 1–19.

Hukelmann, J.L., Anderson, K.E., Sinclair, L. V., Grzes, K.M., Murillo, A.B., Hawkins, P.T., Stephens, L.R., Lamond, A.I., and Cantrell, D.A. (2016). The cytotoxic T cell proteome and its shaping by the kinase mTOR. Nat. Immunol. *17*, 104–112.

Kaech, S.M., and Cui, W. (2012). Transcriptional control of effector and memory CD8+ T cell differentiation. Nat. Rev. Immunol. *12*, 749–

761.
Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., et al. (2011). Ensembl BioMarts: A hub for data retrieval across taxonomic space. Database *2011*, 1–9.
Kontoyiannis, D., Pasparakis, M., Pizarro, T.T., Cominelli, F., and Kollias, G. (1999). Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU- rich elements: Implications for joint and gut-associated immunopathologies. Immunity *10*, 387–398.
Kristensen, A.R., Gsponer, J., and Foster, L.J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. Mol. Syst. Biol. *9:689*.
Kudla, G., Lipinski, L., Caffin, F., Helwak, A., and Zylicz, M. (2006). High guanine and cytosine content increases mRNA levels in mammalian cells. PLoS Biol. *4*, 0933–0942.
Marchingo, J.M., Sinclair, L. V, Howden, A.J., and Cantrell, D.A. (2020). Quantitative analysis of how Myc controls T cell proteomes and metabolic pathways during T cell activation. Elife *9*, e53725.
Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: Life without a poly(A) tail. Nat. Rev. Genet. *9*, 843–854.
Masopust, D. (2001). Preferential Localization of Effector Memory Cells in Nonlymphoid Tissue. Science (80-.). *291*, 2413–2417.
Mathieson, T., Franken, H., Kosinski, J., Kurzawa, N., Zinn, N., Sweetman, G., Poeckel, D., Ratnu, V.S., Schramm, M., Becher, I., et al. (2018). Systematic analysis of protein turnover in primary cells. Nat. Commun. *9*, 1–10.
Mayr, C. (2017). Regulation by 3'-Untranslated Regions. Annu. Rev. Genet. *51*, 171–194.
Mayr, C., and Bartel, D.P. (2009). Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. Cell *138*, 673–684.
Moskowitz, D.M., Zhang, D.W., Hu, B., Le Saux, S., Yanes, R.E., Ye, Z., Buenrostro, J.D., Weyand, C.M., Greenleaf, W.J., and Goronzy, J.J. (2017). Epigenomics of human CD8 T cell differentiation and aging. Sci. Immunol. *2*, 1–14.
Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. *7*, 1–8.
Nicolet, B.P., Guislain, A., and Wolkers, M.C. (2017). Combined Single-Cell Measurement of Cytokine mRNA and Protein Identifies T Cells with Persistent Effector Function. J. Immunol. *198*, 962–970.
Nicolet, B.P., Guislain, A., van Alphen, F.P.J., Gomez-Eerland, R., Schumacher, T.N.M., van den Biggelaar, M., and Wolkers, M.C. (2020). CD29 identifies IFN-γ–producing human CD8 + T cells with an increased cytotoxic potential. Proc. Natl. Acad. Sci. *117*, 6686–6696.
Perez-Perri, J.I., Rogell, B., Schwarzl, T., Stein, F., Zhou, Y., Rettel, M., Brosig, A., and Hentze, M.W. (2018). Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nat. Commun. *9*, 4408.
Perry, R.P. (2007). Balanced production of ribosomal proteins. Gene *401*, 1–3.
Reinhardt, R.L., Khoruts, A., Merica, R., Zell, T., and Jenkins, M.K. (2001). Visualizing the generation of memory CD4 T cells in the whole body. Nature *410*, 101–105.
Ross, S.H., Rollings, C., Anderson, K.E., Hawkins, P.T., Stephens, L.R., and Cantrell, D.A. (2016). Phosphoproteomic Analyses of Interleukin 2 Signaling Reveal Integrated JAK Kinase-Dependent and -Independent Networks in CD8+ T Cells. Immunity *45*, 685–700.
Roychoudhuri, R., Clever, D., Li, P., Wakabayashi, Y., Quinn, K.M., Klebanoff, C.A., Ji, Y., Sukumar, M., Eil, R.L., Yu, Z., et al. (2016). BACH2 regulates CD8+ T cell differentiation by controlling access of AP-1 factors to enhancers. Nat. Immunol. *17*, 851–860.
Rudolph, K.L.M., Schmitt, B.M., Villar, D., White, R.J., Marioni, J.C., Kutter, C., and Odom, D.T. (2016). Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. PLoS Genet. *12*, 1–23.
Rudra, D., DeRoos, P., Chaudhry, A., Niec, R.E., Arvey, A., Samstein, R.M., Leslie, C., Shaffer, S. a, Goodlett, D.R., and Rudensky, A.Y. (2012). Transcription factor Foxp3 and its protein partners form a complex regulatory network. Nat. Immunol. *13*, 1010–1019.
Sakaguchi, S., Yamaguchi, T., Nomura, T., and Ono, M. (2008). Regulatory T Cells and Immune Tolerance. Cell *133*, 775–787.
Salerno, F., Guislain, A., Cansever, D., and Wolkers, M. (2016). TLR-Mediated Innate Production of IFN-γ by CD8+ T Cells Is Independent of Glycolysis. J. Immunol. *196*, 3695–3705.
Salerno, F., Paolini, N.A., Stark, R., von Lindern, M., and Wolkers, M.C. (2017). Distinct PKC-mediated posttranscriptional events set cytokine production kinetics in CD8 + T cells . Proc. Natl. Acad. Sci. *114*, 9677–9682.
Salerno, F., Engels, S., van den Biggelaar, M., van Alphen, F.P.J., Guislain, A., Zhao, W., Hodge, D.L., Bell, S.E., Medema, J.P., von Lindern, M., et al. (2018). Translational repression of pre-formed cytokine-encoding mRNA prevents chronic activation of memory T cells. Nat. Immunol. *19*, 828–837.
Salerno, F., Freen-van Heeren, J.J., Guislain, A., Nicolet, B.P., and Wolkers, M.C. (2019b). Costimulation through TLR2 Drives Polyfunctional CD8 + T Cell Responses . J. Immunol. *202*, 714–723.

Salerno, F., Guislain, A., Freen-Van Heeren, J.J., Nicolet, B.P., Young, H.A., and Wolkers, M.C. (2019a). Critical role of post-transcriptional regulation for IFN-γ in tumor-infiltrating T cells. Oncoimmunology *8*, 1–12.

Salerno, F., Turner, M., and Wolkers, M.C. (2020). Dynamic Post-Transcriptional Events Governing CD8+ T Cell Homeostasis and Effector Function. Trends Immunol. *41*, 240–254.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science (80-.). *320*, 1643–1647.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Simsek, D., Tiu, G.C., Flynn, R.A., Byeon, G.W., Leppek, K., Xu, A.F., Chang, H.Y., and Barna, M. (2017). The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. Cell *169*, 1051–1065.e18.

De Smit, M.H., and Van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. Proc. Natl. Acad. Sci. U. S. A. *87*, 7668–7672.

Stoecklin, G., Lu, M., Rattenbacher, B., and Moroni, C. (2003). A Constitutive Decay Element Promotes Tumor Necrosis Factor Alpha mRNA Degradation via an AU-Rich Element-Independent Pathway. Mol. Cell. Biol. *23*, 3506–3515.

Stubbington, M.J.T., Mahata, B., Svensson, V., Deonarine, A., Nissen, J.K., Betz, A.G., and Teichmann, S.A. (2015). An atlas of mouse CD4+ T cell transcriptomes. Biol. Direct *10*, 1–25.

Vahedi, G., Takahashi, H., Nakayamada, S., Sun, H., Sartorelli, V., Kanno, Y., and O'Shea, J.J. (2012). STATs Shape the Active Enhancer Landscape of T Cell Populations. Cell *151*, 981–993.

Villanueva-Cañas, J.L., Ruiz-Orera, J., Agea, M.I., Gallo, M., Andreu, D., and Albà, M.M. (2017). New genes and functional innovation in mammals. Genome Biol. Evol. *9*, 1886–1900.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. *13*, 227–232.

Vogel, C., De Sousa Abreu, R., Ko, D., Le, S.Y., Shapiro, B.A., Burns, S.C., Sandhu, D., Boutz, D.R., Marcotte, E.M., and Penalva, L.O. (2010). Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol. Syst. Biol. *6*, 1–9.

van der Windt, G.J.W., Everts, B., Chang, C.H., Curtis, J.D., Freitas, T.C., Amiel, E., Pearce, E.J., and Pearce, E.L. (2012). Mitochondrial Respiratory Capacity Is a Critical Regulator of CD8 + T Cell Memory Development. Immunity *36*, 68–78.

Wiśniewski, J.R., Hein, M.Y., Cox, J., and Mann, M. (2014). A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation without Spike-in Standards. Mol. Cell. Proteomics *13*, 3497–3506.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., et al. (2018). Ensembl 2018. Nucleic Acids Res. *46*, D754–D761.

Figure 1 Nicolet et al.

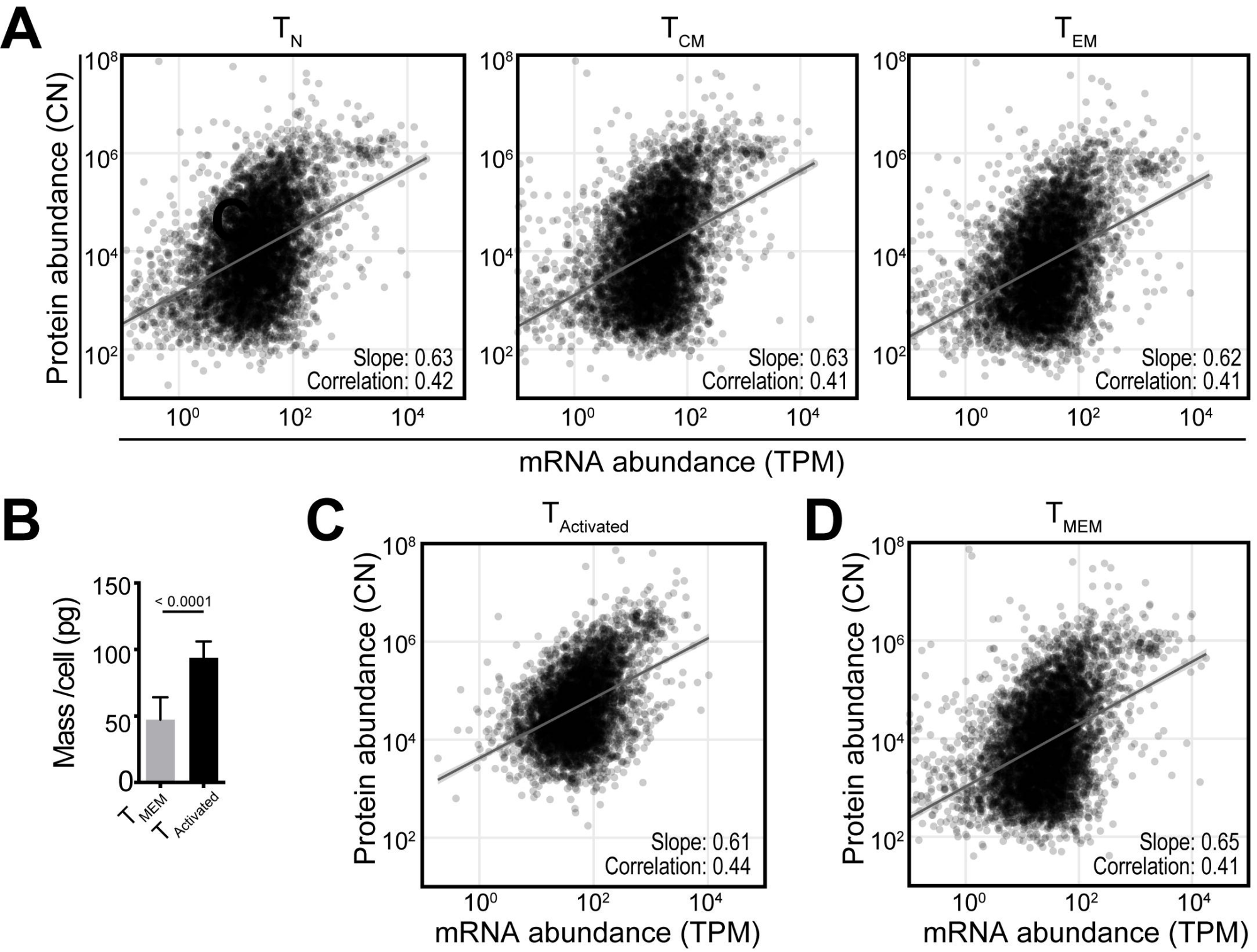

Figure 2 Nicolet et al.

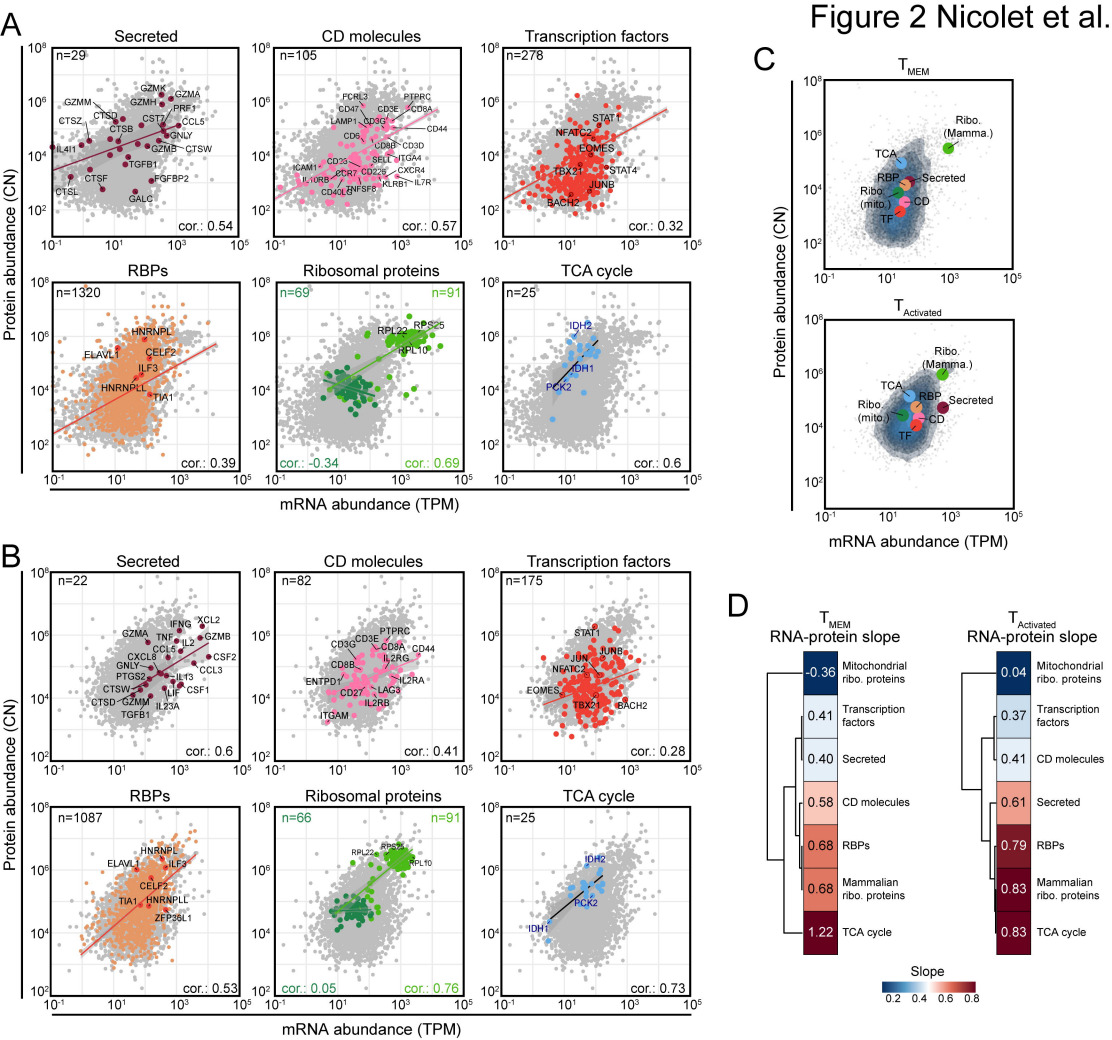

Figure 3 Nicolet et al.

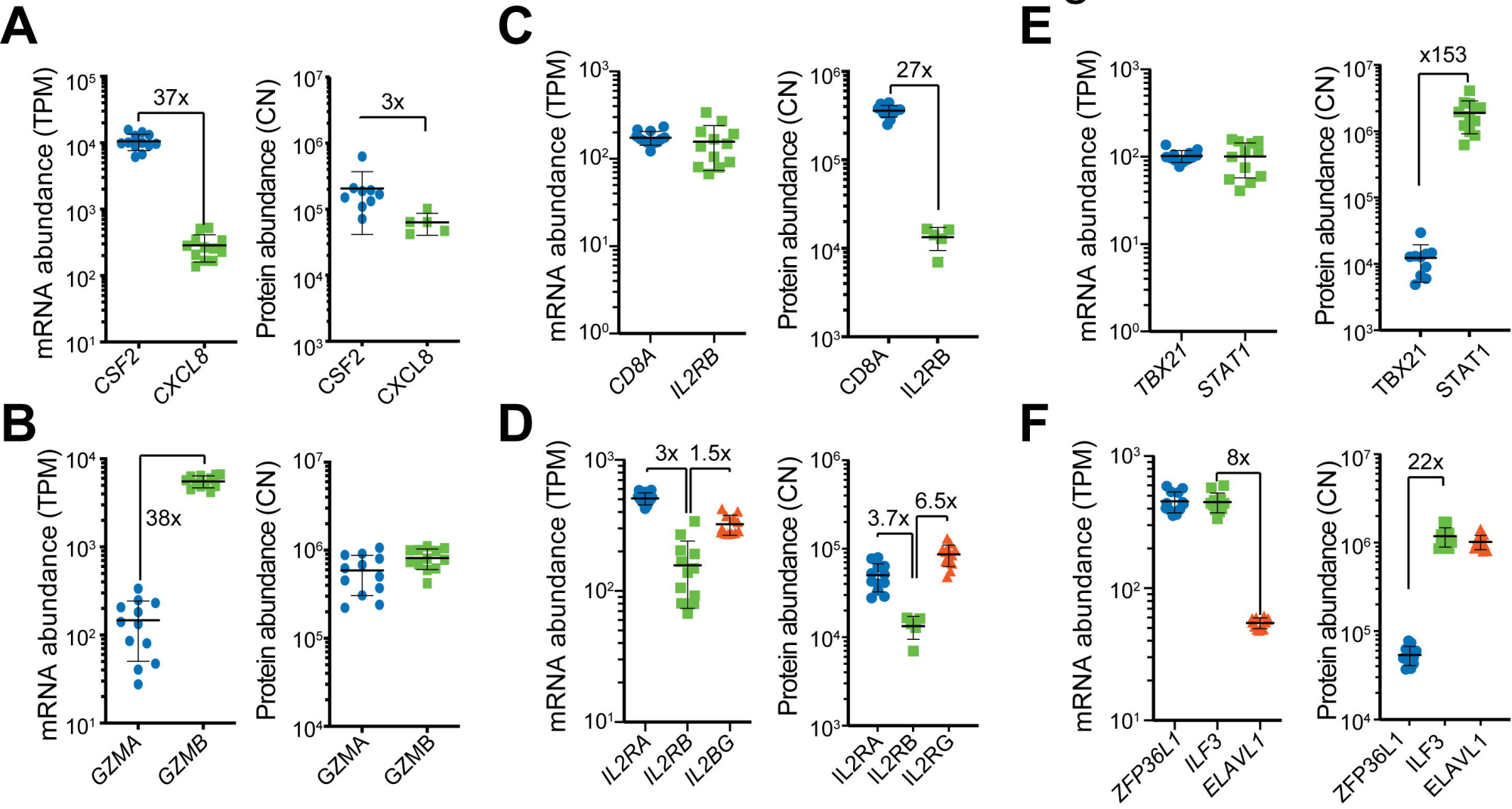

Figure 4 Nicolet et al.

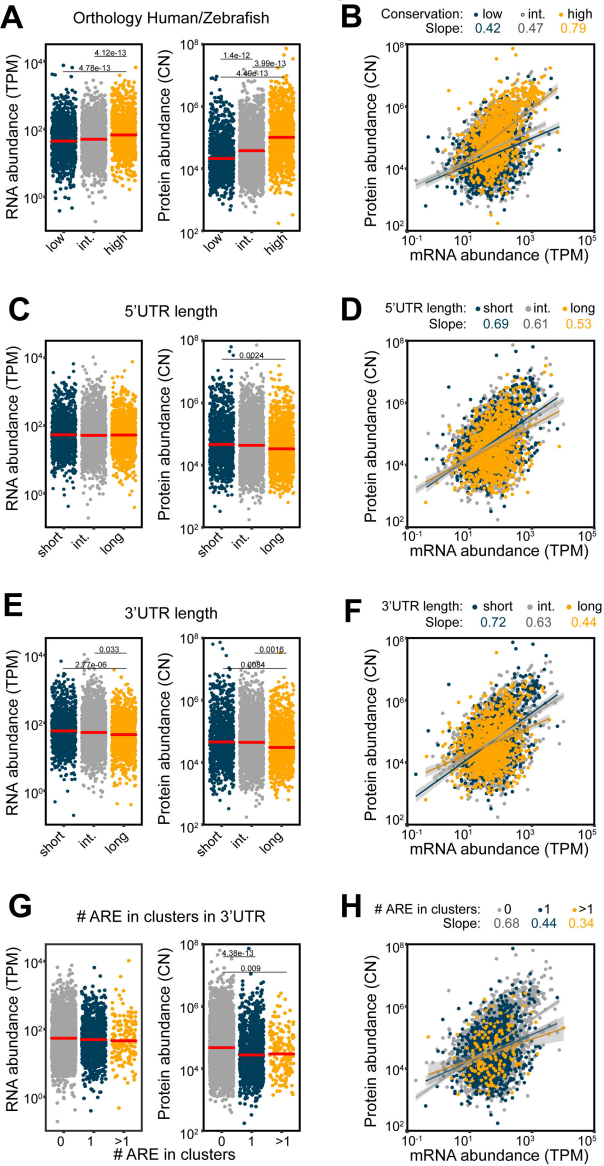
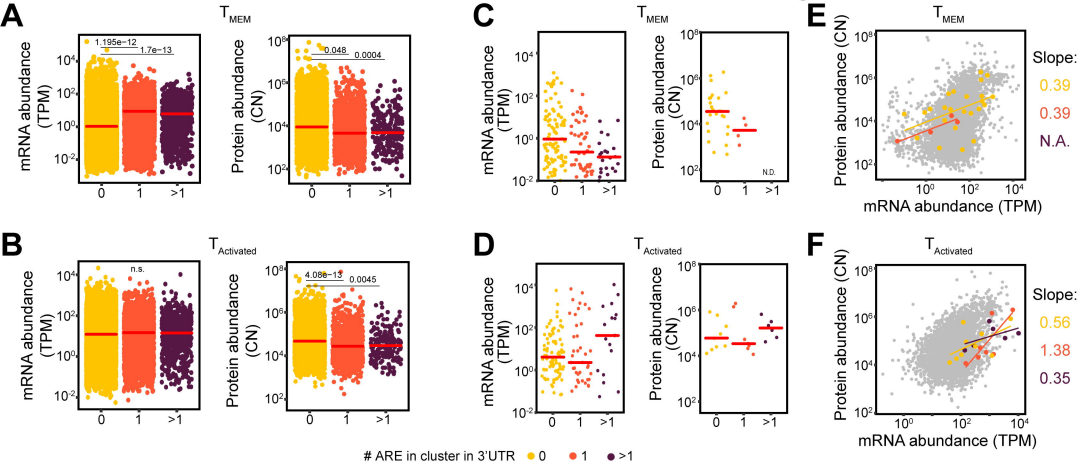
Figure 5 Nicolet et al.