Genome-wide analysis of differential translation and differential translational buffering using anota2seq

Christian Oertlin[1], Julie Lorent[1], Valentina Gandin[2] Carl Murie[1], Laia Masvidal[1], Marie Cargnello[2], Luc Furic[3,4], Ivan Topisirovic[2*] and Ola Larsson[1*]

1. Department of Oncology-Pathology, Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden.
2. Lady Davis Institute, SMBD Jewish General Hospital, Montreal, Canada; Departments of Oncology, Experimental Medicine, and Biochemistry McGill University, Montreal, Canada.
3. Cancer Program, Biomedicine Discovery Institute and Department of Anatomy & Developmental Biology, Monash University, VIC, Australia.
4. Prostate Cancer Translational Research Laboratory, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia.

* Correspondence: ivan.topisirovic@mcgill.ca and ola.larsson@ki.se

ABSTRACT

Genome-wide analysis of mRNA translation (translatomics) can improve understanding of biological processes and pathological conditions including cancer. Techniques used to identify the effects of various stimuli on the translatomes such as polysome- and ribosome-profiling necessitate adjustment for changes in total mRNA levels to capture *bona fide* alterations in translational efficiency. In addition to changes in translational efficiency, which affect protein abundance, translation can also act as a buffering mechanism whereby protein levels remain constant despite changes in mRNA levels (translational buffering). Herein, we present anota2seq, an algorithm for analysis of translatomes, which is applicable to DNA-microarray and RNA sequencing data. In contrast to anota2seq, current methods for analysis of translational efficiency using RNA sequencing data as input identify false positives at high rates and/or fail to distinguish changes in translation efficiency from translational buffering when two conditions (e.g. stimulated and non-stimulated cells) are compared. Moreover, we identify data-acquisition specific thresholds, which are necessary during anota2seq analysis. Finally, we employ anota2seq to show that insulin affects gene expression at multiple levels in a largely mTOR-dependent manner. Thus, the universal anota2seq algorithm allows efficient interrogation of the

translatomes, including distinction between the changes in translation efficiency and buffering.

INTRODUCTION

Regulation of gene expression is a multi-step process including transcription, mRNA-processing, -transport, -stability, and –translation[1]. Although the precise relative contribution of each of these processes in the regulation of gene expression remains controversial[2] and likely context dependent, several studies have implicated mRNA translation as a key step that determines the composition of the proteome[3,4]. Notably, adaptation to changes in cellular environment requires rapid adaptation of the proteome, which is in addition to protein degradation, largely accommodated by altering translational efficiency of mRNAs before affecting their de-novo synthesis[5]. Direct genome wide quantification of mRNA translation is therefore required to improve the understanding of how protein levels are regulated in response to a variety of stimuli and stressors as well as in normal vs. diseased cells. Polysome- and ribosome- profiling are techniques that are used to study translatomes (i.e. the genome-wide pool of translated mRNA)[6]. Polysome-profiling allows separation of mRNAs based on the number of ribosomes that they are associated with, which is directly proportional to translational efficiency. Efficiently translated mRNAs (typically defined as transcripts associated with >3 ribosomes) and cytosolic mRNA are then quantified using DNA-microarrays or RNA sequencing (RNAseq)[7]. Ribosome profiling involves isolation of ribosome-protected fragments (RPFs), generated by RNase treatment, irrespective of the number of ribosomes bound to each individual mRNA molecule. Ribosome-profiling thereby, in conjunction with RNAseq, allows mapping of the position of the ribosome on the mRNA with a single nucleotide resolution[8–10]. Recently, it has been reported that interpretation of ribosome profiling data may be challenging due to biases introduced by commonly used translation inhibitors including cycloheximide and/or lack of appropriate quality control [11–15]. Application of ribosome-profiling, also leads to greater biases as compared to polysome-profiling when employed to quantify changes in translational efficiencies, which is a consequence of strong dependence of ribosome-profiling on the magnitude of the shift of the

mRNA distribution across the polysomes, as well as mRNA abundance[16]. This dramatically favours identification of highly abundant mRNAs showing large shifts in translational efficiency such as terminal oligopyrimidine (TOP) mRNAs as being translationally regulated, over those of lower abundance that exhibit lesser shifts[16]. Thus, while ribosome-profiling provides accurate information on ribosome positioning on mRNA[6,16] and has thus understanding of multiple aspects of protein synthesis [9,10], polysome-profiling appears to be more reliable for analysis of translational efficiency on a genome- wide scale. We will therefore focus on analysis of translatomes using the polysome-profiling strategy, but notably the methodology that we describe herein can also be applied to assessing translational efficiencies using ribosome-profiling data.

In addition to a scenario whereby changes in translational efficiency impacts on protein levels, translation may be altered as a compensatory mechanism to maintain constant protein levels when total mRNA levels change. This phenomenon, referred to as translational buffering was initially described in yeast[17–19] but recent reports suggest that it may also be a prominent, yet unexplored, regulatory mode in mammals[20,21]. Statistical methods that distinguish between changes in translational efficiency that results in changes in protein levels (herein defined as differential translation) from those that do not (herein defined as differential translational buffering) are therefore warranted.

Alterations in translational efficiency are reflected by the changes in the number of ribosomes associated with mRNA between two or more different experimental conditions. Subclasses of mRNAs, however, exhibit dramatic differences in ribosome-association between conditions, which are also dependent on the type of the stimulus. For instance, TOP mRNAs that encode components of the translational machinery are amongst transcripts associated with the highest number of ribosomes under conditions wherein the mammalian/mechanistic target of rapamycin (mTOR) is active, while mTOR inhibition results in their near complete dissociation from ribosomes (on-off regulation). In stark contrast, under conditions when mTOR is activated,

mRNAs encoding proliferation-promoting (e.g. cyclins), survival (e.g. survivin) and mitochondria-related factors (e.g. ATP5O) are associated with prominently less ribosomes relative to TOP mRNAs, and upon mTOR inhibition the reduction in their ribosome association is less dramatic as compared to TOP mRNAs[16]. Identification of both types of changes in translation efficiency can be masked by the mechanisms that act upstream of translation (e.g. transcription, mRNA stability) that indirectly affect polysome-association of mRNAs by inducing alterations in corresponding mRNA levels (hereafter referred to as…total mRNA or cytosolic mRNA; herein we will refer to such mRNA as cytosolic mRNA). To identify differential translation, the Anota algorithm[22] applies per-gene analysis of partial variance (APV)[23] coupled with variance shrinkage. This approach is superior to methods interpreting differences between log ratios inasmuch as log ratio based approaches do not efficiently adjust changes in polysome-association or RPFs for changes in cytosolic mRNA levels due to a phenomenon referred to as spurious correlation[24]. Spurious correlations were initially described by K. Pearson and imply that a comparison of ratios, which in this case is the difference score of log ratios between polysome-associated mRNA (or RPF) to cytosolic mRNA, can systematically correlate with cytosolic mRNA levels[25]. A change in cytosolic mRNA level may therefore lead to false positive identification of differential translation. Indeed such spurious correlations are abundant in both polysome- and ribosome-profiling studies suggesting that log ratio based approaches should be avoided[24]. A recent study[26] evaluated anota in the context of RNAseq data. Anota[22], however, was developed for normalized DNA microarray data which have a continuous scale, and thus cannot be applied on non-transformed count data originating from RNAseq studies. This highlights an emerging issue in quantitative biology, whereby analytical methods are inappropriately used, emphasizing the importance of facilitating correct and efficient application of methodologies.

Several methods have been specifically developed for analysis of differential translation using RNAseq data as input, including babel[27] and Xtail[26]; or were developed for analysis of differential expression using RNAseq data and can be utilized for analysis of differential translation, including DESeq2[28]. Babel[27]

follows an errors-in-variables regression analysis comparing polysome-associated mRNA changes to a background model. DESeq2[28] models read counts using generalized linear models (details in materials and methods). Xtail[26] creates fold-change probability matrices for both polysome-associated and cytosolic mRNA and a joint probability matrix which is used to assess gene specific p-values. Because all these methods are based on interpretation of log-ratio difference scores and are therefore likely to entail spurious correlations, we aimed to develop an anota-like approach for count data (i.e. RNAseq). Initially we considered an approach using edgeR[29] or DESeq2[28]. However, because these methods cannot apply gene specific covariates, this was unachievable[28,29]. Instead, we identified suitable data transformations allowing application of the anota algorithm (i.e. passing quality control steps for efficient analysis within anota)[22] and among several newly added features updated the algorithm such that it can also be employed to identify translational buffering, which cannot be achieved by current methodology. Herein, we present the universal anota2seq algorithm that can be applied to analyse both differential translation and differential translational buffering using data on a continuous scale or count data. We show that anota2seq outperforms current methods for analysis of differential translation. Moreover, we generated polysome-profiling data sets using DNA-microarrays or RNAseq in parallel, wherein we studied the effects of insulin on gene expression in MCF7 cells. Using these data we established key quantification- and transformation-related aspects that allow broad application of anota2seq, and directly compared the performance of DNA-microarrays to RNAseq in polysome-profiling analysis. These experiments demonstrated that insulin induces changes in multiple gene expression programs which entail modulation of steady state mRNA levels, differential translation and differential translational buffering which are largely mediated by the mechanistic/mammalian target of rapamycin (mTOR).

MATERIALS AND METHODS

RNA isolation for polysome-profiling

MCF7 cells were obtained from American Type Culture Collection (ATCC) and maintained in RPMI-1640 supplemented with 10% fetal bovine serum, 1%

penicillin/streptomycin and 1% L-glutamine (all from Wisent Bio Products). Cell lines used were verified by profiling 17 short tandem repeat (STR) and 1 gender determining loci using ATCC Cell Line Authentication Service and the results were as following: MCF7 (100% match to ATCC MCF7 cells # HTB-22): D5S818 (11, 12 versus 11, 12); D13S317 (11 versus 11); D7S820 (8, 9 versus 8, 9); D16S539 (11, 12 versus 11, 12); vWA (14, 15 versus 14, 15); THO1 (6 versus 6); AMEL (X versus X); TPOX (9, 12 versus 9, 12), CSF1PO (10 versus 10). Cells were serum starved for 16h hours followed by 4 hour treatment with either DMSO (4.2nM), insulin (4.2nM) or insulin (4.2nM) + torin1 (250nM), washed with ice cold PBS containing 100µg/mL cycloheximide, collected and lysed in a hypotonic lysis buffer [5 mM Tris-HCl (pH 7.5), 2.5mM MgCl2, 1.5 mM KCl, 100 µg/L cycloheximide, 2mM DTT, 0.5% Triton X-100, and 0.5% sodium deoxycholate] (all chemicals used were obtained by Sigma unless otherwise stated). A lysate sample was collected and cytosolic RNA was isolated using TRIzol (Invitrogen). Lysates were loaded onto 10-50% (wt/vol) sucrose density gradients [20 mM Hepes-KOH (pH 7.6), 100 mM KCl, 5 mM MgCl2] and centrifuged at 36,000 rpm [SW 40 Ti rotor (Beckman Coulter, inc)] for 2h at 4 °C. Gradients were fractioned and the optical density at 254 nm was continuously recorded using an ISCO fractionator (Teledyne ISCO). RNA from each fraction was isolated using TRIzol (Invitrogen). Fractions with mRNAs associated with >3 ribosomes were pooled (polysome-associated mRNA). The experiment was performed in 4 biological replicates.

Protein isolation and Western Blotting

MCF7 were serum-starved overnight with RPMI supplemented with 0.1%FBS. Where indicated, cells were pre-treated with torin1 (150nM) for 15 min and stimulated for 4 hours with insulin (4.2nM). Cells were washed twice in ice-cold PBS and lysed on ice in RIPA buffer [20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% (v/v) NP-40, 1% (w/v) Na-deoxycholate, 0.1% (w/v) SDS, 2.5 mM Na-pyrophosphate, 50 mM NaF, 20 mM β-glycerophosphate, 1mM Na_3VO_4] supplemented with 1X complete protease inhibitors (Roche).

Protein concentrations in cell extracts were determined using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and lysates were prepared in 4X Laemmli buffer [250mM Tris/SDS (pH 6.8), 40% glycerol, 8% SDS, 400mM DTT and 0.01% bromophenol blue]. Proteins of interest were detected using the following primary antibodies: anti-β-actin (Clone AC-15) #A1978 from Sigma (Saint Louis, Missouri, USA) used at 1:5,000 dilution; and anti-4E-BP1(53H11) #9644, anti-p-4E-BP1 (S65) (174A9) #9456, anti-rpS6 #2217, anti-p-rpS6 (S240/244) #2215, anti-Cyclin D3 (DCS22) #2936, anti-S6K1/2 #9202, anti-p-S6K1/2 (T389) #9205 all used at 1:1,000 dilutions from Cell Signaling Technologies (Danvers, MA, USA). Secondary antibodies (Amersham) were used at 1:10,000, and signals were revealed by chemiluminescence (ECL, GE Healthcare).

Microarray data processing

DNA-microarray based expression analysis (polysome-associated and cytosolic RNA) was performed using the Human Gene ST 1.1 array plate (Gene Titan from Affymetrix) by the Bionformatics and Expression Analysis (BEA, Karolinska Institutet, Stockholm). Gene expression was summarized and normalised using Robust Multi-array Average (RMA)[30] using the custom probe set annotation (hugene11st_Hs_ENTREZG) as these provide improved precision and accuracy[31–33]. Microarray quality control was performed using standard methods[34] and all arrays were considered of high quality.

RNAseq data processing

RNAseq libraries were obtained according to the strand specific TruSeq protocol including ribosomal RNA depletion (Ribo Zero) and sequenced using Illumina HiSeq 2000 High Output platform obtaining single end reads for both polysome-associated and cytosolic mRNA. Library preparation and sequencing were performed at Scilifelab (Stockholm). All RNAseq reads were subjected to and passed quality control (Phred scores >30) and aligned to the hg19 reference genome using HISAT[35] with default settings. The average percentage of reads mapped to the genome was 83% (total number of reads

was 92 million). Reads mapped to multiple places in the genome were removed. Gene expression was quantified using the RPKMforgenes.py[36] script from which raw gene counts were obtained (version last modified 11.04.2013). Genes that overlap and/or had 0 counts in at least one sample were removed from the analysis. Genes were annotated using RefSeq[37] database. A total of 11,623 unique genes were represented in the RNAseq data. Raw counts were either rlog[28] or voom[38] transformed.

RNAseq data simulation using empirical data

Method performance comparison was done using simulated data. Data was simulated by sampling from a negative binomial (NB) distribution using estimated means and dispersions from empirical RNAseq data produced in this study (control and insulin condition only). We used a similar approach as described previously[39]. For polysome profiling data, in addition to non-regulated genes (i.e. evenly expressed), three regulatory modes were simulated: differential translation (a change in polysome-associated RNA independent of changes in cytosolic RNA), differential translational buffering (a change in cytosolic RNA that is not reflected by a similar change in polysome-associated RNA) and changes in mRNA abundance (i.e. a congruent change in polysome-associated and cytosolic RNA). Each regulatory mode is represented by 5% of all genes (total number of simulated genes was 10,748). Four replicates of control and treatment conditions were simulated using the following parameterization for the NB distribution:

$$Y_{gi} \sim NB(mean = \mu_{gi}, var = 1/\phi_{gi})$$

where *Y* is the simulated gene count for gene *g* and RNA source (i.e. polysome-associated or cytosolic mRNA) *i and* ϕ_{gi} is the dispersion.

The mean μ_{gi} and dispersion ϕ_{gi} were estimated from the empirical data using maximum likelihood estimates $\hat{\mu}_{gi}$ and $\hat{\phi}_{gi}$ with

$$\hat{\mu}_{gi} = \frac{1}{R * C} \sum_{r,c} K_{girc}$$

where R is the number of replicates, C the number of conditions and K_{girc} the read count for the empirical gene *g*, RNA source *i*, replicate *r* and condition *c*.

$\hat{\phi}_{gi}$ was obtained by maximizing the log-likelihood function described[39]. $\hat{\mu}_{gi}$ and $\hat{\phi}_{gi}$ were then directly used as parameters of a NB distribution to simulate 4 replicates control conditions. We used the rnbinom function of the stats R package with parameter size being $1/\hat{\phi}_{gi}$.

For non- regulated genes, the treatment condition was sampled from a distribution with equal $\hat{\mu}_{gi}$ and $\hat{\phi}_{gi}$ identical to the control condition for that gene. For regulated genes, these estimates were used as base parameters and modified according to the different regulatory modes. For differential translation (537 genes, similar number of up- and down-regulation), the base parameters were used for simulating cytosolic mRNA of both conditions and polysome-associated mRNA of the control condition. The mean and dispersion parameters used to simulate polysome-associated mRNA under treatment condition were modified as follows: an effect size parameter α_g for upregulation was sampled from a vector containing values from 1.5 to 3 with steps of 0.2. For down regulation the parameter α_g was modified to $1/\alpha_g$. The modified mean parameter was then $\alpha_g\hat{\mu}_{gi}$ (or $1/\alpha_g\hat{\mu}_{gi}$ for down regulation). In order to keep the mean-variance relationship as similar as possible to the empirical data, the modified dispersion was taken as the dispersion of the gene from the empirical data having the closest mean estimate to $\alpha_g\hat{\mu}_{gi}$ (or $1/\alpha_g\hat{\mu}_{gi}$ for down regulation). Similarly, for differential translational buffering (537 genes), base parameters were used for polysome-associated mRNA (both conditions) and cytosolic mRNA (control condition). A random effect size was introduced for simulating cytosolic mRNA under treatment condition and applied as described for differentially translated genes. Finally, for mRNA abundance base parameters were used for cytosolic and polysome-associated mRNA under control condition and the same effect size was introduced to modify the parameters of the distribution from which cytosolic and polysome-associated mRNA were sampled under the treatment condition; and applied as described for differentially translated genes.

Genes that had a simulated count of zero in any sample were removed before analysis. To assess the ability of methods to control for type I error/false discovery rate in the absence of any true regulation between control and treatment, we also simulated a null dataset using base parameters for all genes in both conditions. Furthermore for additional testing, we simulated datasets with additional 5%, 10% and 15% noise, where the noise is a percentage of each count, which is added or subtracted (same probability to add or substract); and datasets with different sequencing depths (5, 10, 15 and 25 million counts per sample). The average total count across all conditions and RNA types of the simulated data are 42 million counts.

Comparison of methods for analysis of differential translation using simulated data

We compared five tools for the analysis of the translatome: anota2seq, babel[27] (version 0.3.0), DESeq2[28] (version 1.10.1), translational efficiency score (TE score) and Xtail[26] (version 1.1.3). All analyses were performed using R (version 3.2.3). Simulated raw count data were used as input for babel[27], DESeq2[28] and Xtail[26]. For TE score analysis, counts were normalized using DESeq2 (normalization for library size using the median ratio method)[40] and log2 transformed. For anota2seq, counts were either rlog[28] or voom[38] transformed. Similar to anota[22], anota2seq combines analysis of partial variance (APV)[23] and the Random Variance Model (RVM)[41] and uses a two-step process that firstly assesses the model assumptions for 1) absence of highly influential data points that affect APV assumptions 2) common slopes of sample classes, 3) homoscedasticity of residuals and 4) normal distribution of per gene residuals. This is followed by analysis of differential translation or differential translational buffering using APV[23] and RVM[41]. Babel[27] uses an errors-in-variables regression consisting of two steps. Using a NB distribution in both steps, cytosolic mRNA levels are modelled as a first step followed by modelling of the RPFs as a second step. To estimate the mean in the second part of the model a trimmed least-squares approach is used. Following a parametric bootstrap, for every gene, a test is performed for the null hypothesis that observed RPF levels are as expected from cytosolic mRNA levels. Per gene per sample p-values are then combined into one single p-

value using the arithmetic mean of p-values. DESeq2[28] uses a generalised linear model approach to quantify translational changes. For all analyses done in this study we used the following implementation:

$$\sim 0 + RNA + Condition + RNA{:}Condition$$

Here RNA represents either polysome-associated or cytosolic mRNA; condition represent the control or treatment groups and RNA:Condition the interaction effect. In this model RNA:Condition is the measurement for differential translation. The TE score is the difference between conditions in log2 ratios (between polysome-associated and cytosolic mRNA). Statistics for changes in TE were calculated using Student's t-test. Xtail[26] is a method based on comparisons of fold- changes between conditions using data from polysome-associated mRNA, cytosolic mRNA or between polysome-associated and cytosolic mRNA. For each comparison, a fold change probability distribution is generated. Using these probability distributions a joint probability matrix is derived and p-values for each gene assessed. For all methods, we used either the reported FDRs for all genes when applying default parameters or the selection of filtered genes (filtered using default parameters for each method) under a specific FDR threshold. These tools were compared using receiver operator curves on simulated data.

Analysis of the empirical data set using anota2seq

DNA- microarray and RNAseq data were analysed for differential polysome-associated mRNA, differential cytosolic mRNA, differential translation and differential translational buffering using the anota2seq algorithm. Genes where considered differential translated or translationally buffered when passing default filtering criteria in anota2seq (-0.5 < slope < 1.5, FDR < 0.05 and deltaPT > 0.15). Filtering criteria were used for both DNA- microarray and RNAseq analysis. When comparing DNA- microarray and RNAseq data, genes measured on both platforms were used, leading to a total of 10679 genes.

R-packages and settings

Hierarchical clustering, scatter plots and PCA plots were generated using the hclust (default settings), smoothScatter and prcomp (center=TRUE) R-functions, respectively. Receiver operator curves (ROC), area under the curve (AUC), partial AUC (pAUC) and precision recall curves were generated using the ROCR R-package (version 1.0-7)[42]. P-values from each method were used in the ROC analysis. pAUC were obtained at 5 and 15% false positive rate by using the fpr.stop parameter from the ROCR package, followed by division of the so acquired AUC with the corresponding fpr cut-off rate. The precision is the positive predictive value and the recall is the true positive rate or sensitivity.

RESULTS

Babel and Xtail algorithms underperform under a NULL model

To evaluate the performance of current methods relative to anota2seq for analysis of the translatomes by RNA-seq, we employed a simulated data approach (see materials and methods for details). To identify appropriate data transformations for application of anota2seq we evaluated several normalization and transformation approaches including rlog[28] and voom[38]. We then compared the performance of anota2seq using rlog and voom data transformations to other methods including babel[27], DESeq2[28], translational efficiency (TE) score and Xtail[26]. An essential aspect in such analysis is the control of type I error/false discovery rates (FDR) under a NULL model wherein no differences are introduced, the distribution of the p-values is expected to be uniform resulting in FDRs for all genes to equal 1, after adjusting for multiple testing. To evaluate this we simulated a dataset with two conditions sampled from the same distribution i.e. where all the genes were set not to be regulated (i.e. no differential translation, changes in mRNA abundance or differential translational buffering; see material and methods; Fig. S1). Consistent with observations from empirical data sets[43], simulated data for cytosolic mRNA or polysome-associated mRNA differed when visualized using hierarchical clustering (Fig. 1A). Moreover, as expected under a NULL model, there was no separation of the treatment and control groups (Fig. 1A). The resulting density plots of p-values for differential translation revealed uniform distributions for all methods except Xtail, which has a strong overrepresentation of low p-values, and babel, which shows an overrepresentation of high p-values and a local enrichment of low p-values under the NULL model (Fig. 1B). Accordingly both Xtail and babel assign low FDRs to a subset of genes even when analysing data sets which were a priori set to be devoid of any changes in translational efficiency (Fig 1B). This indicates that babel and Xtail are unable to control FDRs.

Anota2seq outperforms current methods for identification of differential translation

To evaluate the performance of the various methods in identification of differential translation, we next simulated a data set with two conditions and three regulatory patterns (537 genes per regulatory pattern along with 9137 evenly expressed genes): differential translation ("dTE"), differential translational buffering ("TB") and changes in total mRNA abundance ("Tot") (Fig. 2A). Hierarchical clustering showed the expected pattern whereby polysome-associated mRNA and cytosolic mRNA cluster separately and treatment and control groups form separated clusters within each of these clusters - thus capturing the complex structure of polysome-profiling data (which is similar to ribosome-profiling data[43]; Fig. 2B). We next set to determine the performance of each method in detecting differential translation. Therefore, for each method identification of a gene belonging to the differential translation pattern ("dTE") was considered as a true positive event, whereas identification of genes whose translation remains unchanged or those belonging to "TB" and "Tot" regulatory patterns was considered as a false positive event (Fig. 2A). To assess the performance of the methods, we used area under the curve (AUC) and partial AUC (pAUC) of receiver operator curves (ROC)[44]. The methods were applied using default settings on 5 replicate data sets simulated as in figure 2A (to assess the variability of the simulation) followed by assessment of their outputs prior to any filtering (see material and methods). ROC curves show that anota2seq analysis on rlog or voom transformed data performs equally well and outperforms all other methods as judged by AUC and pAUCs (Fig. 2C, table 1). In addition, precision recall curves reveal high initial precision values for anota2seq compared to other methods (Fig. 2C). This can be explained by the analysis principle of the other methods. TE-score, babel, DESeq2 and Xtail are indifferent to whether non-congruent changes in cytosolic mRNA levels or polysome-association are due to differential translation (i.e. changes in polysome-associated mRNAs but not in cytosolic mRNA levels) or differential translational buffering (i.e. no difference in polysome-associated mRNA despite a change in cytosolic mRNA levels) whereas anota2seq can make this distinction. Indeed, this property likely contributes to the reported superior performance of anota as compared to TE-score in reflecting changes in the proteome[45].

To further characterize the performance of the methods at common FDR thresholds, we counted the number of identified genes from each regulatory pattern (differential translation, differential translational buffering and differential mRNA abundance) and genes whose expression did not change identified at a 5%, 10% or 15% FDR threshold (Fig. 2D) using the default settings of each method. Babel and TE score identified fewer true positive events of differential translation than the other methods. Anota2seq identified approximately the same number of true positives at the 15% FDR threshold as DESeq2 while Xtail identified slightly more true positive events. The number of genes identified from the set of buffered genes by each method reveals that only anota2seq can efficiently distinguish between differential translation and differential translational buffering. Notwithstanding that all methods perform similarly in terms of rejecting genes that show changes in mRNA abundance (i.e. congruent changes in cytosolic and polysome-associated mRNA), there were dramatic differences in terms of identification of equally expressed genes (i.e. false positives). Anota2seq and TE-score identified few such genes while babel and DEseq2 identified a sizeable number. Strikingly, Xtail identified >600 such genes at FDR<15%. The latter finding is consistent with the poor performance of Xtail[26] under the NULL model (Fig. 1B). To assess differences between applying a 5% and 15% FDR threshold, we visualized the disparity between the numbers of identified genes from each regulatory pattern (Fig. 2E). For anota2seq, there was a gain in true positives at the cost of a smaller increase in false positives, whereas for other methods, especially Xtail, increasing the FDR threshold introduces more false positives than additional true positives.

The above simulation included equal numbers of genes regulated via differential translation, differential translational buffering and differential mRNA abundance. Under certain conditions, however, differential translation or differential translational buffering may be the predominate mechanism affecting changes in translational efficiency and hence it was of interest to compare the methods using a simulated data set containing differential translation (537 genes) and differential mRNA abundance (537 genes) along

with unchanged genes (9674 genes), but without differential translational buffering (Fig. S2A-B). As expected from their inability to separate differential translation from differential translational buffering (Fig. 2D), under these conditions babel[27], DESeq2[28], TE score and Xtail[26] showed improved performance as judged by pAUC and AUC, which was comparable to anota2seq (e.g. at pAUC 15% FPR and AUC; table S1). Moreover, removal of differential translational buffering from simulated data resulted in an increase in precision of the aforementioned methods (compare Fig. S2C to Fig 2C). As expected, the performance under different FDR thresholds was similar to the full simulation (compare Fig. S2D-E to Fig. 2D-E). This included similar performance of anota2seq and DEseq2[28] for identification of differentially translated genes and elevated identification of non-regulated genes for babel[27], DESeq2[28] and Xtail[26]. Similar to the full simulation, TE-score identified fewer genes as differentially translated but did not assign evenly expressed genes low FDRs. These results demonstrate that while anota2seq shows superior performance under conditions wherein gene expression is regulated via both differential translation and differential translational buffering, it also outperforms other methods even in the absence of differential translational buffering. Therefore, anota2seq can be applied to efficiently identify changes in differential translation, which correspond to alterations in protein levels, when analyzing data sets quantified by RNAseq independent of what the underlying regulatory patterns are.

These results appear to contradict the recently report which described development of Xtail[26]. In the Xtail development study[26] it is clearly stated to assess performance to correctly identify differential translation and reports good ROC and precision/recall performance for babel, TE score and Xtail and poor performance for anota. The reported poor performance of anota [26] is caused by inappropriate application of anota on non-transformed counts. Moreover, the assessment of performance of other methods in Xtail study does also not completely agree with this study. Our analysis reveals poor precision recall performance for babel, TE score and Xtail for determination of differential translation. We therefore examined the simulated data set used during development of Xtail in detail (Fig. S3A). This revealed that genes

selected as true positives for differential translation belonged to any regulatory pattern (differential translation, differential translational buffering or differential mRNA abundance; and also included seemingly genes with unchanged expression) (Fig. S3B top left). Moreover, there were genes which showed increased polysome-association but decreased mRNA levels (or vice versa), which represent unlikely biological events that were not observed in any of the empirical data sets examined[26,46] (Fig. S3B top right and lower panel; notably this is distinct from translational buffering wherein levels of polysome-associated mRNA remain unchanged despite alterations in mRNA levels). We therefore reclassified genes (Fig. S3C) and only considered those showing differential translation (i.e. those where the change in translational efficiency is expected to be reflected by changes in protein levels) as truly differentially translated. We then examined the population of genes that were identified at different FDR thresholds with largely similar results as observed using our simulated datasets, except for babel, which identified very few regulated events (Fig. S3D). Thus the optimal performance of anota2seq is not restricted to the single simulated dataset.

Identification of differential translational buffering using anota2seq

In addition to differential translation, differential translational buffering also holds potentially important information about translational control as it is thought to function as a homeostatic mechanism that maintains protein levels despite changes in mRNA expression. Methods to capture translational buffering however have not been developed to date. We, therefore implemented analysis of differential translational buffering as part of the anota2seq software. Analysis of differential translational buffering in anota2seq is based on the same principle as analysis of differential translation (i.e. analysis of partial variance coupled with variance shrinkage) except that it captures changes in cytosolic mRNA levels, which do not lead to changes in levels of polysome- associated mRNAs. To assess the performance of anota2seq for analysis of translational buffering, we used the same data set as in Fig. 2A-B (i.e. with equal numbers of genes for each pattern [differential translation, differential translational buffering, or differential mRNA abundance]). In contrast to the analysis in figure 2A-B, for each method

identification of genes from the differential translational buffering group was considered as a true positive event, whereas identification of evenly expressed genes or genes belonging to other regulatory modes was considered as a false positive event. As expected, pAUC and AUC for differential translational buffering were similar to the performance of anota2seq for analysis of differential translation efficiency (Fig. 3A and table 2). Critically, very few genes whose translational efficiency changed were identified among those showing differential translational buffering under the FDRs assessed (Fig. 3B). Increasing the FDR threshold primarily lead to additional identification of differentially translationally buffered genes (Fig. 3C). Thus in addition to analysis of changes in translation efficiency anota2seq can be efficiently applied to identify differential translational buffering.

Assessing robustness against noise and varying sequencing depth

Experimental and technical challenges and/or study setups may give rise to datasets exhibiting dramatically different characteristics. This includes data sets with higher noise level and varying sequencing depth [47,48]. To test the effects of variations in noise or sequencing depth on performance of translatome analyses, we simulated datasets with increasing noise levels (5, 10 or 15%; Fig. S4A) or increasing number of RNAseq reads (5, 10, 15 or 25 million reads mapping to mRNAs per sample). Increased noise leads to a decrease in the number of true positives at an FDR of 15% especially for TE score (Fig. S4B-C). This is associated with a decrease in performance as assessed by ROC and precision recall curves (Fig. S5A-F). Judged by pAUC and AUC, anota2seq outperforms other methods at all noise levels in analysis of differential translation using a simulated data set including all regulatory patterns (Fig. S4; table S2). Reduced sequencing depth had no clear effect on ROC analysis within methods (Fig. S6). Nevertheless, the number of identified genes decreased ~10-50 genes with reduced sequencing depth depending on method (Fig. S7A-B). In our data set about 49% of the sequenced reads mapped uniquely to mRNA (Fig. S8A) indicating that an appropriate sequencing depth is ~10 million to obtain 5 million reads mapping

to mRNA although this is likely to depend on sample and protocol used for preparation of RNAseq libraries.

Incorporating batch adjustment in anota2seq

Hybridization based DNA-microarrays and count based RNAseq are commonly used to estimate cytosolic and polysome-associated mRNA levels. Both platforms have their intrinsic strengths and weaknesses whereby DNA microarrays have higher background signals but are less susceptible to issues associated with very high expression of only a limited number of genes that may dominate the output, as is seen in RNAseq[49]. Nevertheless, it remains to be determined how these technical traits affect genome- wide analysis of the translatome.

To directly test the performance of anota2seq on datasets obtained via DNA-microarrays vs. RNAseq, we performed a polysome-profiling study to assess the effect of insulin on steady-state mRNA levels (reflective of translation and mRNA degradation) and translation and investigated the role of mTOR complex 1 (mTORC1) in mediating these effects. mTORC1 is activated by insulin via the PI3K/AKT pathway and acts as a major anabolic pathway in the cell that bolsters lipid, protein and nucleotide synthesis[50–52]. .To this end MCF7 cells were serum starved for 16 hours and treated with DMSO (control), insulin or insulin in combination with the active-site mTOR inhibitor torin1 for 4 hours (n=4; Fig. 4A). Despite harbouring activating PIK3CA, mTORC1 signalling in MCF7 cells is highly responsive to insulin as illustrated by a dramatic induction in phosphorylation of mTORC1 downstream targets eIF4E-binding protein 1 (4E-BP1) and ribosomal protein S6 kinases (S6K), and its downstream target ribosomal protein S6 (rpS6) which was abolished by torin1 (Fig S8A). Absorbance profiles revealed that as expected, insulin induced global mRNA translation, as evidenced by decrease in monosome peak and concomitant increase in polysome area as compared to control, which was also reversed by torin1 (Fig S8B). Total and heavy polysome-associated mRNA (i.e. associated with more than 3 ribosomes) was isolated and subjected in parallel to DNA-microarrays and RNAseq to generate two data

sets from the exact same samples. The resulting data were analysed using anota2seq.

Principal component analysis of the DNA-microarray data set revealed the expected distinction between RNA source (polysome-associated or cytosolic) and treatment (Fig 4B), as previously reported in[7]. The same analysis using RNAseq data revealed a similar separation based on RNA source but also a batch effect illustrated by a consistent pattern between the treatment groups within each replicate that was shifted across replicates (Fig. 4B). Such batch effects must be considered for efficient analysis[54]. We therefore implemented a batch effect parameter in anota2seq, which is applied during APV and affects variance shrinkage[23]. To assess the impact of the batch effect correction on analysis of translational efficiencies, we analysed the RNAseq data with and without batch effect correction. We then selected genes with an FDR < 0.05 (and passing filtering criteria, see materials and methods) in analysis of changes in polysome-associated mRNA, cytosolic mRNA, differential translation and differential translational buffering (we added a feature in anota2seq which allows for differential expression analysis of polysome-associated or cytosolic mRNA only using the same model as applied during analysis of translational efficiencies). When correcting for batch effects during anota2seq analysis, there is an increase in the number of genes with an FDR < 0.05 (Fig. S9B). Thus it is critical to enable correction for batch effects during analysis of translational efficiencies[54]. Notably, among all of the assessed methods for translatome analysis only Xtail is not equivalent to and does not have the option to perform an analysis incorporating batch effects.

RNAseq allows identification of differential translation across a broader dynamic range as compared to DNA-microarrays

We then characterized similarities and differences between anota2seq analysis of the DNA-microarray and RNAseq data. We first observed that RNAseq quantification (rlog or voom transformed) leads to identification of more mRNAs showing differential polysome-association or cytosolic levels between control and insulin stimulated cells as compared to DNA-

microarrays. Consistently more genes were identified as differentially translated or differentially translationally buffered when using RNAseq data as input relative to DNA-microarray data, which was independent of transformation (Fig. 4C). We next identified differentially translated genes from DNA-microarray data or rlog or voom transformed RNAseq data (FDR < 0.05 and passed filtering as indicated in materials and methods) using anota2seq. This was followed by comparison of expression ranges for genes identified as differentially translated. RNAseq uniquely identifies a set of genes with low signals on the DNA-microarray as differentially translated. When comparing rlog vs voom transformed RNAseq data, rlog identified a set of genes with lower expression levels as differentially translated as compared to voom transformation (Fig. 4D).

Commonly, differential translated genes are defined using both FDR and fold change (FC) thresholds[55,56]. We therefore assessed the impact of such thresholds on anota2seq analyses of RNAseq or DNA-microarray data by comparing commonly or uniquely identified genes across platforms and data transformations using multiple thresholds (FDR < 0.05; FC > 1.5; or FDR < 0.05 and FC > 1.5). Voom transformation showed a larger overlap to DNA microarray as compared to rlog transformation at FDR<0.05 although the majority of genes were distinct between voom-transformed RNAseq and DNA-microarrays (Fig 5A). The FC thresholds dramatically affect the overlap (Fig. 5A). To further illustrate this, volcano plots were generated showing genes identified by anota2seq as being differentially translated (FDR < 0.05, filtering) (Fig. 5B). Thus although genes selected solely by FDR in the voom based and DNA-microarray anota2seq analysis show substantial overlap, fold-changes are not comparable across platforms and transformations.

Assessing the need for replication for efficient analysis of differential translation using anota2seq

As noted above, data on translatomes can include substantial variance, thereby highlighting that sufficient replication is essential for efficient analysis. Because this study includes four biological replicates and anota2seq requires two replicates when there are three treatment groups (3 replicates in the case

of just two treatment groups), we determined the effect of reducing the number of replicates on anota2seq performance. Data sets were generated with all combinations of 2 (n=6) and 3 (n=4) replicates and analyzed differential translation using anota2seq to monitor genes that passed the filtering and FDR < 0.15. When using 2 replicates per sample, we observed a drastic decrease in the number of identified genes compared to scenarios where 3 or 4 replicates were used (Fig. S9C). We next monitored the FDRs of all genes with a FDR < 0.15 (all replicates) using 2 or 3 replicates. This illustrated that with 3 replicates per condition, FDRs are below 0.15 for the majority of such genes (Fig. S9D). Consequently, for the efficient analysis using anota2seq, a minimum of 3 replicates should be applied - which is consistent with what has been proposed previously for analysis using babel[27]. It should be noticed, however, that we used batch adjustment in the analysis, which will reduce the degrees of freedom for the residual error in the APV[23] model, and fewer replicates may therefore be sufficient when analysing data sets without batch effects and/or with >3 conditions.

Insulin regulates gene expression by modulating multiple steps in the gene expression pathway in a largely mTOR dependent manner

Next we determined how 4 hour insulin treatment of MCF7 cells affects gene expression at multiple levels including differential translation, differential translational buffering and differential mRNA abundance. This revealed that the major effect of 4 hour insulin treatment on gene expression occurs at the level of mRNA abundance followed by changes in translational efficiency and, to a lesser extent, translational buffering (FIG 6A). Moreover, a comparison to the number of differentially regulated genes in insulin vs. insulin and torin1 treated cells as compared to the control, revealed that the vast majority of the effects of insulin on gene expression are reversed by torin1 (Fig. 6A). To further illustrate this, expression levels for polysome-associated and cytosolic mRNA under both insulin and insulin+torin1 conditions for genes with an FDR < 0.15 (insulin vs control for polysome-associated mRNA) were visualized using hierarchical clustering (while monitoring which genes were regulated via translation; Fig. 6B). Indeed this revealed an almost complete reversal of the changes in gene expression induced by insulin in the presence of torin1. Thus

insulin largely regulates changes in mRNA abundance, translational efficiency and buffering via mTOR.

DISCUSSION

The importance of mRNA translation in regulation of gene expression has been exemplified in numerous biological and pathological process ranging from stress response, cancer[57] to learning and memory[58]. Identification of specific genes that are under translational control requires interrogations of translatomes. Yet, translatomes are vastly under-studied in comparison to transcriptomes (which reflect mRNA abundance determined at the level of transcription or mRNA stability)[6]. Further research is thus required to understand the full complexity of translatomes and their impact on cell physiology and pathology. Such understanding largely relies on stringent and efficient application of genome wide methods to measure changes in translational efficiencies. Moreover, downstream of methods applied to generate such data, sufficient replication and efficient data analysis is crucial for deriving valid conclusions. Notwithstanding that the noisy nature of polysome- and ribosome-profiling data is widely acknowledged, it is often paradoxically suggested that methods for identification of differential translation that do not require replication are warranted[26]. This does not seem to be consistent with concerns about reproducibility in quantitative biology, which instead suggests that sufficient replication is essential to derive meaningful conclusions.

Currently, polysome- and ribosome-profiling are most commonly used to interrogate translatomes on a genome-wide scale whereby polysome-profiling is more efficient for identification of translational efficiencies[6,16], while ribosome-profiling allows genome-wide information about ribosome positioning[9,10]. A powerful unique property of polysome-profiling is that it allows examination of 5' and/or 3'UTRs of mRNAs that are translated thereby facilitating identification of regulatory elements as well as potential differences in translation of mRNA isoforms that are co-expressed in the cells but differ in their 5' or 3'UTRs[16]. Thus, ribosome- and polysome-profiling represent compatible methodologies that provide ample opportunity to study

translatomes. There is therefore a heightened interest to develop efficient methodologies to analyse polysome- and ribosome- profiling data. Such analysis has to be adopted to advances in technology that bear distinct characteristics, such as the count nature of RNAseq data, but also to our emerging understanding of mechanisms regulating mRNA translation. Translational buffering represents one such emerging mechanism of translation control wherein alterations in mRNA levels are buffered at the level of translation such that levels of polysome-associated mRNAs are not affected by the changes in mRNA abundance. While the nature of such regulation suggests a role of mRNA translation in homeostatic regulation of protein levels, the mechanisms, extent and importance of this phenomenon are yet to be established. Nevertheless, analysis of translatomes requires efficient distinction between differential translation and differential translational buffering because the former is expected to result in changes in protein levels while the latter is expected to maintain protein levels constant in spite of alteration in cytosolic RNA levels.

To meet these requirements, we developed the anota2seq algorithm, capable of analysis of DNA- microarray and RNAseq data that efficiently identifies and separates differential translation from differential translational buffering. Evaluation of anota2seq compared to other methods for translatome analyses indicated superior performance of anota2seq in detecting differential translation with low type- 1-error rates and robustness against noise and varying sequencing depths. One unexpected finding from these studies was the poor performance of Xtail under the NULL condition inasmuch as a large number of genes were identified as differentially translated despite no true changes in their translational efficiency (Fig. 1B and 2D). This possibly stems from the usage of the same data multiple times during Xtail analysis (e.g. both polysome-associated mRNA and the ratio between polysome-associated RNA and cytosolic mRNA) thereby not meeting the independence criterion. This furter underlines the importance of assessing the performance under the NULL condition during method development. Moreover, anota2seq has several distinct features as compared to other methods: i) it is not based on interpretation of difference scores between log ratios and hence will not be

affected by spurious correlations; ii) it distinguishes differential translation from differential translational buffering; iii) it allows for batch correction; and iv) it permits analysis of polysome- associated and cytosolic mRNA changes using the same analytical methods thereby allowing for simple and comparable identification of changes in polysome-associated mRNA, cytosolic mRNA, differential translation and differential translational buffering. Our comparisons between data obtained from DNA-microarrays and RNAseq transformed by rlog or voom suggests that voom is the preferred transformation based on its closer resemblance to DNA-microarrays. Nevertheless, rlog and voom are themselves comparable and a major impact on the overlaps is the estimated fold-changes where rlog is associated with smaller fold-changes as compared to DNA-microarrays and voom. Differences between the platforms seem to at least in part originate from the broader dynamic range of RNAseq compared to DNA- microarray data. Thus, due to technological biases and differences in data transformation methods, prudence is advised when selecting transformations and filtering thresholds for differential expression analysis. Anota2seq therefore incorporates both voom and rlog but also allows the user to supply custom transformed and normalized data. Although we assessed the performance of anota2seq for analysis of polysome-profiling data, it is equally applicable data from ribosome-profiling.

We[59] and others [43,60] have recently reported that mTOR primarily regulates gene expression by modulating mRNA translation. Yet, our current study suggests that while insulin primarily regulates gene expression depending on mTOR it does to a larger extent affect cytosolic mRNA levels as compared to mRNA translation (although mRNA translation is modulated; Fig. 6). This discrepancy may be related to the experimental set-ups applied (serum vs. insulin) and different time points evaluated. Nevertheless, these findings indicate that insulin regulates gene expression at multiple levels and understanding how regulation of cytosolic mRNA levels is coordinated with changes in mRNA translation during the course of the response to insulin may uncover principles for how these processes are orchestrated. In summary, we designed anota2seq for analysis of differential translation and differential

translational buffering which can be applied independent of platform used for quantification. Application of such statistically stringent analyses holds a substantial promise to facilitate our understanding of dynamic regulation of translatomes in health and disease.

References

1. Komili, S. Coupling and coordination in gene expression processes: a systems biology view. *Nat. Rev. Genet.* **9,** (2008).
2. Li, J. J., Bickel, P. J. & Biggin, M. D. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ* **2,** e270 (2014).
3. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–42 (2011).
4. Kristensen, A. R., Gsponer, J. & Foster, L. J. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol. Syst. Biol.* **9,** 689 (2013).
5. Piccirillo, C. A., Bjur, E., Topisirovic, I., Sonenberg, N. & Larsson, O. Translational control of immune responses: from transcripts to translatomes. *Nat. Immunol.* **15,** 503–511 (2014).
6. Piccirillo, C. A., Bjur, E., Topisirovic, I., Sonenberg, N. & Larsson, O. Translational control of immune responses: from transcripts to translatomes. *Nat. Immunol.* **15,** 503–511 (2014).
7. Gandin, V. *et al.* Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale. *J. Vis. Exp.* (2014). doi:10.3791/51455
8. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324,** 218–23 (2009).
9. Andreev, D. E. *et al.* Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. *Nucleic Acids Res.* gkw1190 (2016). doi:10.1093/nar/gkw1190

10. Ingolia, N. T. Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* **165,** 22–33 (2016).
11. O'Connor, P. B. F. *et al.* Comparative survey of the relative impact of mRNA features on local ribosome profiling read density. *Nat. Commun.* **7,** 12915 (2016).
12. Mohammad, F., Woolstenhulme, C. J., Green, R. & Buskirk, A. R. Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep.* **14,** 686–94 (2016).
13. Pelechano, V., Wei, W. & Steinmetz, L. M. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* **161,** 1400–12 (2015).
14. Gerashchenko, M. V & Gladyshev, V. N. Ribonuclease selection for ribosome profiling. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw822
15. Gerashchenko, M. V & Gladyshev, V. N. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res.* **42,** e134 (2014).
16. Gandin, V. *et al.* nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs. *Genome Res.* **26,** 636–48 (2016).
17. Wang, Z. *et al.* Evolution of gene regulation during transcription and translation. *Genome Biol. Evol.* **7,** 1155–67 (2015).
18. McManus, C. J., May, G. E., Spealman, P. & Shteyman, A. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res.* **24,** 422–30 (2014).
19. Bader, D. M. *et al.* Negative feedback buffers effects of regulatory variants. *Mol. Syst. Biol.* **11,** 785 (2015).
20. Cenik, C. *et al.* Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res.* **25,** 1610–21 (2015).
21. Jovanovic, M. *et al.* Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347,** 1259038 (2015).
22. Larsson, O., Sonenberg, N. & Nadon, R. anota: Analysis of differential translation in genome-wide studies. *Bioinformatics* **27,** 1440–1 (2011).
23. Schleifer, S. J., Eckholdt, H. M., Cohen, J. & Keller, S. E. Analysis of partial variance (APV) as a statistical approach to control day to day

variation in immune assays. *Brain. Behav. Immun.* **7,** 243–52 (1993).

24. Larsson, O., Sonenberg, N. & Nadon, R. Identification of differential translation in genome wide studies. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 21487–92 (2010).
25. Pearson, K. Mathematical Contributions to the Theory of Evolution.--On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. *Proc. R. Soc. London* **60,** 489–498 (1896).
26. Xiao, Z., Zou, Q., Liu, Y. & Yang, X. Genome-wide assessment of differential translations with ribosome profiling data. *Nat. Commun.* **7,** 11194 (2016).
27. Olshen, A. B. *et al.* Assessing gene-level translational control from ribosome profiling. *Bioinformatics* **29,** 2995–3002 (2013).
28. Love, M. I. *et al.* Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
29. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–40 (2010).
30. Lim, W. K., Wang, K., Lefebvre, C. & Califano, A. Comparative analysis of microarray normalization procedures: Effects on reverse engineering gene networks. *Bioinformatics* **23,** 282–288 (2007).
31. Ferrari, F. *et al.* Novel definition files for human GeneChips based on GeneAnnot. *BMC Bioinformatics* **8,** 446 (2007).
32. Dai, M. *et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **33,** e175 (2005).
33. Sandberg, R. & Larsson, O. Improved precision and accuracy for microarrays using updated probe set definitions. *BMC Bioinformatics* **8,** 48 (2007).
34. Eijssen, L. M. T. *et al.* User-friendly solutions for microarray quality control and pre-processing on ArrayAnalysis.org. *Nucleic Acids Res.* **41,** W71-6 (2013).
35. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12,** 357–360 (2015).
36. Ramsköld, D., Wang, E. T., Burge, C. B. & Sandberg, R. An Abundance

of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data. *PLoS Comput. Biol.* **5,** e1000598 (2009).

37. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33,** D501-4 (2005).
38. Law, C. W. *et al.* voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15,** R29 (2014).
39. Robles, J. A. *et al.* Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing. *BMC Genomics* **13,** 484 (2012).
40. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).
41. Wright, G. W. & Simon, R. M. A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* **19,** 2448–55 (2003).
42. Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. ROCR: visualizing classifier performance in R. *Bioinformatics* **21,** 3940–1 (2005).
43. Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485,** 55–61 (2012).
44. Murie, C., Woody, O., Lee, A. Y. & Nadon, R. Comparison of small n statistical tests of differential expression applied to microarrays. *BMC Bioinformatics* **10,** 45 (2009).
45. Colman, H. *et al.* Genome-wide analysis of host mRNA translation during hepatitis C virus infection. *J. Virol.* **87,** 6668–77 (2013).
46. Tebaldi, T., Dassi, E., Kostoska, G., Viero, G. & Quattrone, A. tRanslatome: an R/Bioconductor package to portray translational control. *Bioinformatics* **30,** 289–91 (2014).
47. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10,** 1093–1095 (2013).
48. Sims, D., Sudbery, I., Ilott, N. E., Heger, A. & Ponting, C. P. Sequencing depth and coverage: key considerations in genomic analyses. *Nat. Rev. Genet.* **15,** 121–32 (2014).
49. Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and

opportunities. *Nat. Rev. Genet.* **12,** 87–98 (2011).

50. Wang, X. & Proud, C. G. The mTOR Pathway in the Control of Protein Synthesis. *Physiology* **21,** (2006).
51. Laplante, M. & Sabatini, D. M. An Emerging Role of mTOR in Lipid Biosynthesis. *Curr. Biol.* **19,** R1046–R1052 (2009).
52. Ben-Sahra, I., Ricoult, S., Howell, J., Asara, J. & Manning, B. mTORC1 stimulates nucleotide synthesis through both transcriptional and post-translational mechanisms. *Cancer Metab.* **2,** P6 (2014).
53. Proud, C. G. Regulation of protein synthesis by insulin. *Biochem. Soc. Trans.* **34,** (2006).
54. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* **11,** 733–739 (2010).
55. Mao, Y. *et al.* IL-15 activates mTOR and primes stress-activated gene expression leading to prolonged antitumor capacity of NK cells. *Blood* **128,** (2016).
56. Livingstone, M. *et al.* Assessment of mTOR-Dependent Translational Regulation of Interferon Stimulated Genes. *PLoS One* **10,** e0133482 (2015).
57. Topisirovic, I. & Sonenberg, N. mRNA translation and energy metabolism in cancer: the role of the MAPK and mTORC1 pathways. *Cold Spring Harb. Symp. Quant. Biol.* **76,** 355–67 (2011).
58. Bramham, C. R. & Wells, D. G. Dendritic mRNA: transport, translation and function. *Nat. Rev. Neurosci.* **8,** 776–789 (2007).
59. Larsson, O. *et al.* Distinct perturbation of the translatome by the antidiabetic drug metformin. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 8977–82 (2012).
60. Thoreen, C. C. *et al.* A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485,** 109–113 (2012).
61. Parent, R., Kolippakkam, D., Booth, G. & Beretta, L. Mammalian Target of Rapamycin Activation Impairs Hepatocytic Differentiation and Targets Genes Moderating Lipid Homeostasis and Hepatocellular Growth. *Cancer Res.* **67,** (2007).

Figure descriptions

Figure 1. Analysis of differential translation with xtail or babel is associated with increased false positive findings. (A) Hierarchical clustering of a simulated treatment and control dataset including polysome-associated mRNA and cytosolic mRNA levels where no genes are regulated (i.e. a NULL model). (B) P-value and FDR density plots for analysis of differential translation comparing treatment to control for each method using the data from (A).

Figure 2. Anota2seq outperforms other methods for analysis of differential translation using RNAseq data as input. (A) Scatterplot of polysome-associated and cytosolic mRNA log2 fold changes between treatment and control groups for a simulated data set. Simulated genes regulated by changes in translational efficiency (i.e. translation and buffering) and mRNA abundance are indicated. (B) Hierarchical clustering of the dataset from (A). (C) Receiver operator curves for analysis of differential translation in a simulated dataset (i.e. from A-B; top). Precision recall curves for analysis for differential translation in the simulated dataset (bottom). For both analyses, identification of a gene simulated as differentially translated was considered as a true positive event. Vertical lines indicate 5% and 15% false positive rates. (D) Number of genes identified as differentially translated that were differentially translated (true positives), differentially buffered (false positives (FP)), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at several FDR thresholds (mean and standard deviations from 5 simulated data sets are indicated) (E) Differences in the number of genes identified as differentially translated belonging to the 4

categories in (D) when changing the FDR threshold from 5% to 15% (mean from 5 simulated data sets).

Figure 3. Identification of translational buffering using anota2seq. (A) Receiver operator curve (left) and precision recall curve (right) for analysis of translational buffering using anota2seq on a simulated dataset. The dataset is the same as in 2A-B but identification of a genes simulated as differentially translationally buffered was considered as a true positive event. Vertical lines indicate 5% and 15% false positive rates. (B) Number of genes identified as differentially buffered that were differentially translated (FP), differentially buffered (TP), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at several FDR thresholds (mean and standard deviations from 5 simulated data sets are indicated) (C) Difference in the number of genes identified as differentially translationally buffered belonging to the 4 categories in (B) when changing the FDR threshold from 5% to 15% (mean from 5 simulated data sets).

Figure 4. Platform and data transformation characteristics during analysis of differential translation. (A) Experimental setup for study on insulin's effect on gene expression patterns. Serum starved MCF7 cells were treated with DMSO (ctrl), insulin or insulin and torin1 (insulin_T1) followed by isolation of polysome-associated and cytosolic mRNA and their quantification using DNA-microarrays or RNAseq. (B) Principal component analysis using normalized DNA-microarray data as input (circles indicate clusters corresponding to RNA source and treatment [left]); and voom transformed RNAseq data (lines connect samples for each RNA source from the same replicate, i.e. illustrates a batch effect [right]). (C) FDR density plots from anota2seq analysis of changes in polysome- associated mRNA, cytosolic mRNA, differential translation and differential translational buffering using data from DNA-microarray or RNAseq data transformed with rlog or voom. (D) Density plots of mean expression levels for mRNAs identified as differentially translated using DNA-microarray and RNAseq data transformed with rlog or voom. Genes with FDR<0.05 that passed anota2seq filtering (material and methods) were included. The number of genes passing these criteria and

where there is agreement between platforms or data transformation (shared) and platform or data transformation unique genes are indicated.

Figure 5. Fold-change filtering is a major determinant affecting overlaps of identified differentially translated genes between platforms and transformations. (A) Venn diagrams of differentially translated genes identified by anota2seq using either FDR<0.05, FC>1.5 or FDR<0.05 & FC>1.5 thresholds and passing anota2seq filtering (see materials and methods). (B) Volcano plots of the anota2seq output for analysis of differential translation using DNA-microarray data or RNAseq data transformed by rlog or voom. Genes with FDR<0.05 that passed anota2seq filtering were included. Horizontal and vertical lines indicate 5% FDR or positive and negative log2 FC of 1.5 respectively. Unique and shared genes are indicated as in figure 4D.

Figure 6. Insulin regulates gene expression at multiple levels in an mTOR dependent manner. (A) Number of genes identified as showing differential translation differential translational buffering, or mRNA abundance following insulin treatment as analysed by anota2seq. The number in each category that depended on mTOR (i.e. reversed by torin1) are also indicated. (B) Genes changing their polysome-associated mRNA level following insulin treatment are included and the fold-changes for insulin vs control and insulin plus torin1 vs control are shown for both polysome-associated and cytosolic mRNA using hierarchical clustering. Genes identified as differentially translated according to anota2seq are indicated with black bars on the left. Insulin (ins); insulin plus torin1 (ins+T1); polysome-associated mRNA (P); cytosolic mRNA (C).

Supplement figure 1. Comparison of empirical to simulated data. (A) Scatterplots of the mean log2 counts and variance of the log2 counts for empirical RNAseq data and simulated null distributed data. (B) Density plot of the mean log2 counts for empirical data and simulated null distributed datasets (lines completely overlap).

Supplement figure 2. Anota2seq outperforms other methods for analysis of differential translation using RNAseq data as input also in the absence of differential buffering. (A) Scatterplot of polysome-associated and cytosolic mRNA log2 fold changes between treatment and control groups for a simulated data set. Simulated genes regulated by changes in translational efficiency (i.e. translation and buffering) and mRNA abundance are indicated. (B) Hierarchical clustering of the dataset from (A). (C) Receiver operator curves for analysis of differential translation in the simulated dataset (i.e. from A-B; top). Precision recall curves for analysis of differential translation in the simulated data set (bottom). For both analyses identification of a gene simulated as differentially translated was considered as a true positive event. Vertical lines indicate 5% and 15% false positive rates. (D) Number of genes identified as differentially translated that were differentially translated (TP), differentially buffered (FP), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at several FDR thresholds (mean and standard deviations from 5 simulated data sets are indicated) (E) Differences in the number of genes identified as differentially translated belonging to the 4 categories in (D) when changing the FDR threshold from 5% to 15% (mean from 5 simulated data sets).

Supplement Figure 3. Evaluation and analysis of simulated data used in Xiao et al.[26].(A) Hierarchical clustering of simulated dataset used in Xiao et al.[26] (B) Scatterplots of polysome- associated mRNA and cytosolic mRNA log2 fold-changes comparing treatment to control from (A) (top left) and three empirical data sets (polysome-profiling RNAseq data comparing insulin to control [current study; top right], polysome-profiling DNA-microarray data comparing differentiated and non differentiated human hepatocytes[61] [bottom left] and ribosome profiling data comparing mTOR inhibition to control data used in A. Hsieh et al.[43] [bottom right]). Orange points indicate the genes considered differentially translated in prior analysis[26]. (C) Scatterplot of polysome-associated mRNA and cytosolic mRNA log2 fold-changes from the dataset in (A) reclassified into regulatory modes corresponding to figure 2A. (D) Number of genes identified as differentially translated that were

differentially translated (TP), differentially buffered (FP), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at several FDR thresholds. (E) Differences in the number of identified genes belonging to the 4 categories in (D) when changing FDR threshold from 5% to 15% using the simulated data in Xiao et al.[26].

Supplement figure 4. Evaluation of noise sensitivity for identification of differential translation. (A) Heatmap of correlations within and between simulated datasets with 5%, 10% and 15% additional noise (Note that the set off differentially expressed will change between simulations). (B) Number of genes identified as differentially translated that were differentially translated (TP), differentially buffered (FP), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at a 15% FDR threshold (mean and standard deviations from 5 simulated data sets are indicated) (C) Difference in the number of genes identified as differentially translated (FDR<0.15) belonging to the 4 categories in (B) when changing the noise from 15% additional noise to no additional noise (mean from 5 simulated data sets).

Supplement figure 5. Assessing noise sensitivity for identification of differential translation. (A-F) Receiver operator curves and precision recall curves of analysis of differential translation for all methods using simulated dataset with 5%, 10% and 15% additional noise. Identification of a gene simulated as differentially translated was considered as a true positive event. P-values were used as input. Vertical lines indicate 5% and 15% false positive rates.

Supplement figure 6. Assessing sensitivity to decreasing RNAseq depth for identification of differential translation - (A-F) Receiver operator curves and precision recall curves of analysis of differential translation for all methods using simulated datasets with increasing sequencing depth (number of reads uniquely mapped to mRNA as defined by refseq). Identification of a genes simulated as differentially translated was considered as a true positive event.

P-values were used as input. Vertical lines indicate 5% and 15% false positive rates.

Supplement figure 7. Assessing sensitivity to decreasing RNAseq depth for identification of differential translation. (A) Number of genes identified as differentially translated that were differentially translated (TP), differentially buffered (FP), had differential mRNA abundance (FP) or were evenly expressed (FP) are indicated for each method at a 15% FDR threshold (mean and standard deviations from 5 simulated data sets are indicated). (B) Differences in the number of genes identified as differentially translated (FDR<15%) belonging to the 4 categories in (A) when changing the sequencing depth from 25 million reads to 5 million reads.

Supplement figure 8. The effects of insulin and torin1 on mTOR signalling and polysome assembly.

(A) MCF7 were serum-starved for 16 hours, pre-treated with torin1 (250nM) for 15 min and stimulated for 4 hours with insulin (4.2nM) in the presence or absence of torin 1 as shown in the figure. Expression and phosphorylation status of indicated proteins were assessed by Western blotting using indicated antibodies (see material and methods). β-actin served as a loading control. (B) Polysome profiling tracings of cells treated as described in (A) were obtained by 5-50% sucrose gradient centrifugation followed by fractionation and continuous monitoring of UV absorbance (254nm). Positions of 40S and 60S ribosomal subunits, monosome (80S) and polysomes are indicated. 0.1% FBS=starved cells (red); insulin=insulin treated cells (green); insulin + torin1=insulin and torin1 treated cells (black).

Supplement figure 9. Evaluation of RNA sequencing characteristics; and effects from adjusting for batch effects or increasing number of biological replicates in the study of insulin's effects on gene expression. (A) Mean number of total reads, reads mapped to the genome and reads uniquely mapped to refSeq[37] mRNA. (B) Number of genes with an FDR<0.05

for a difference in polysome-associated mRNA, cytosolic mRNA, differential translation and differential translational buffering as identified by anota2seq with or without adjustment for batch effects. (C) Number of genes identified as differentially translated (insulin vs control) by anota2seq using data from DNA-microarrays or RNAseq counts transformed using rlog or voom. Anota2seq analysis was performed on all possible combinations (n=2 and n=3) from the 4 biological replicates. Mean number of genes with FDR<0.15 across all combinations are indicated. (D) Boxplots of mean FDRs for differential translation as analysed by anota2seq across all possible combinations (as in C) for genes with an FDR<0.15 when analysing the complete experiment (4 replicates).

Table Descriptions

Table 1. Mean and standard deviation (sd) of pAUCs and AUCs for an ROC analysis assessing performance for identification of differential translation in simulated datasets (n=5) including evenly expressed genes and genes with changes in translation, translational buffering and mRNA abundance.

Table 2. Mean and standard deviation (sd) of AUCs and pAUCs for an ROC analysis assessing performance for identification of differential buffering in simulated datasets (n=5) including evenly expressed genes and genes with changes in translation, translational buffering and mRNA abundance.

Supplement table 1. Mean and standard deviation (sd) of AUCs and pAUCs for an ROC analysis assessing performance for identification of differential translation in simulated datasets (n=5) including evenly expressed genes and genes with changes in translation and mRNA abundance.

Supplement table 2. Mean and standard deviation (sd) of AUCs and pAUCs for an ROC analysis assessing performance for identification of differential translation in simulated datasets with 5%, 10% and 15% additional noise

(n=5) including evenly expressed genes and genes with changes in translation, translational buffering and mRNA abundance.

Figure 1

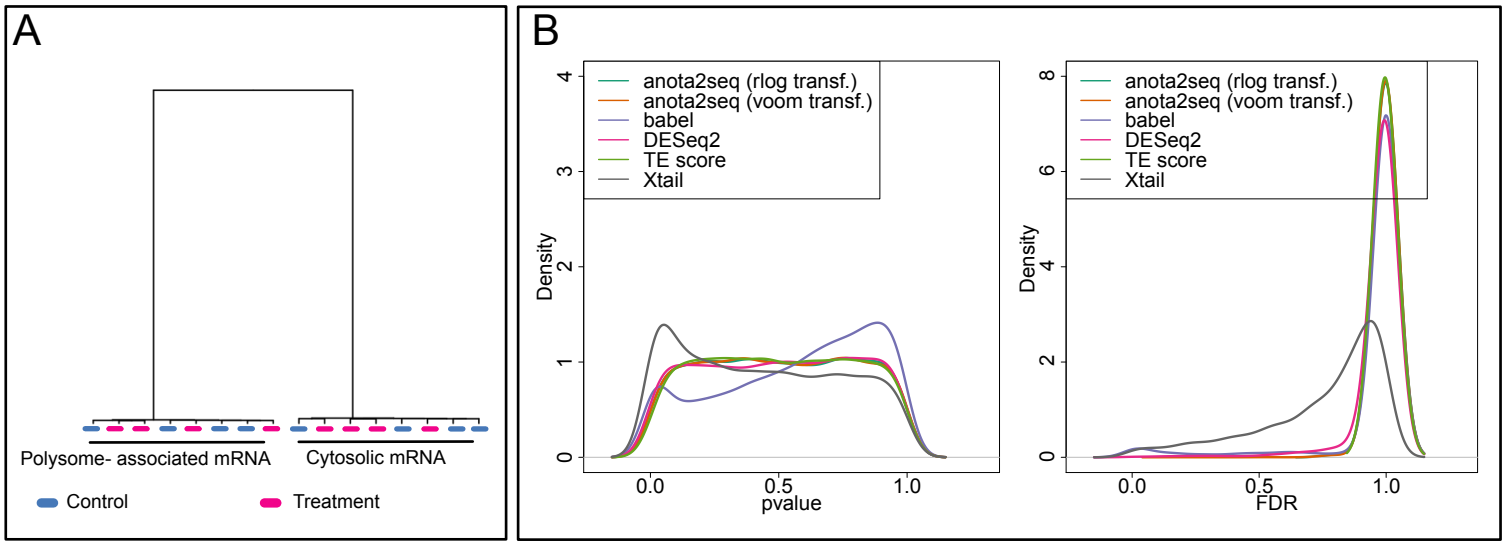

Figure 2

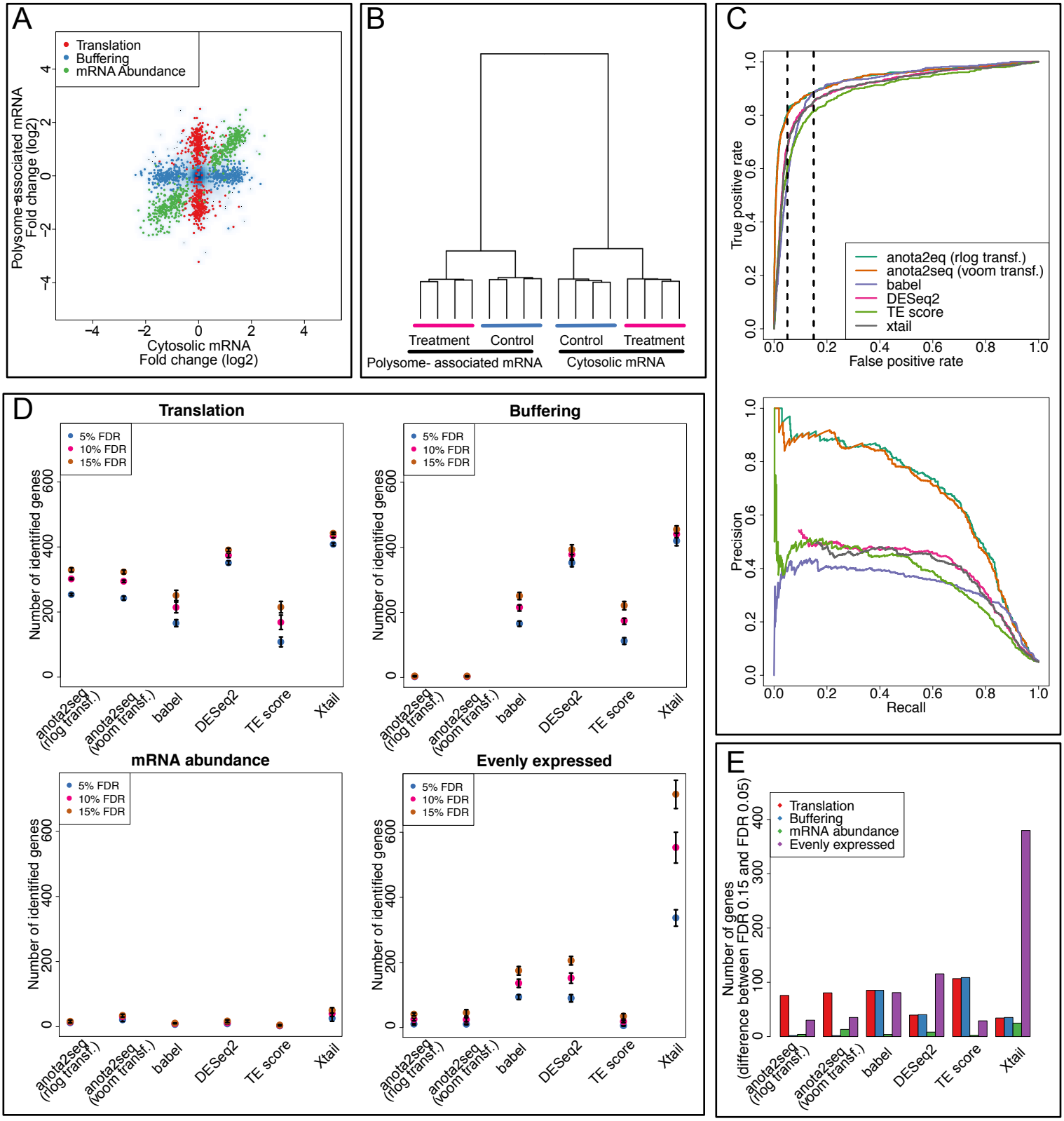

Figure 3

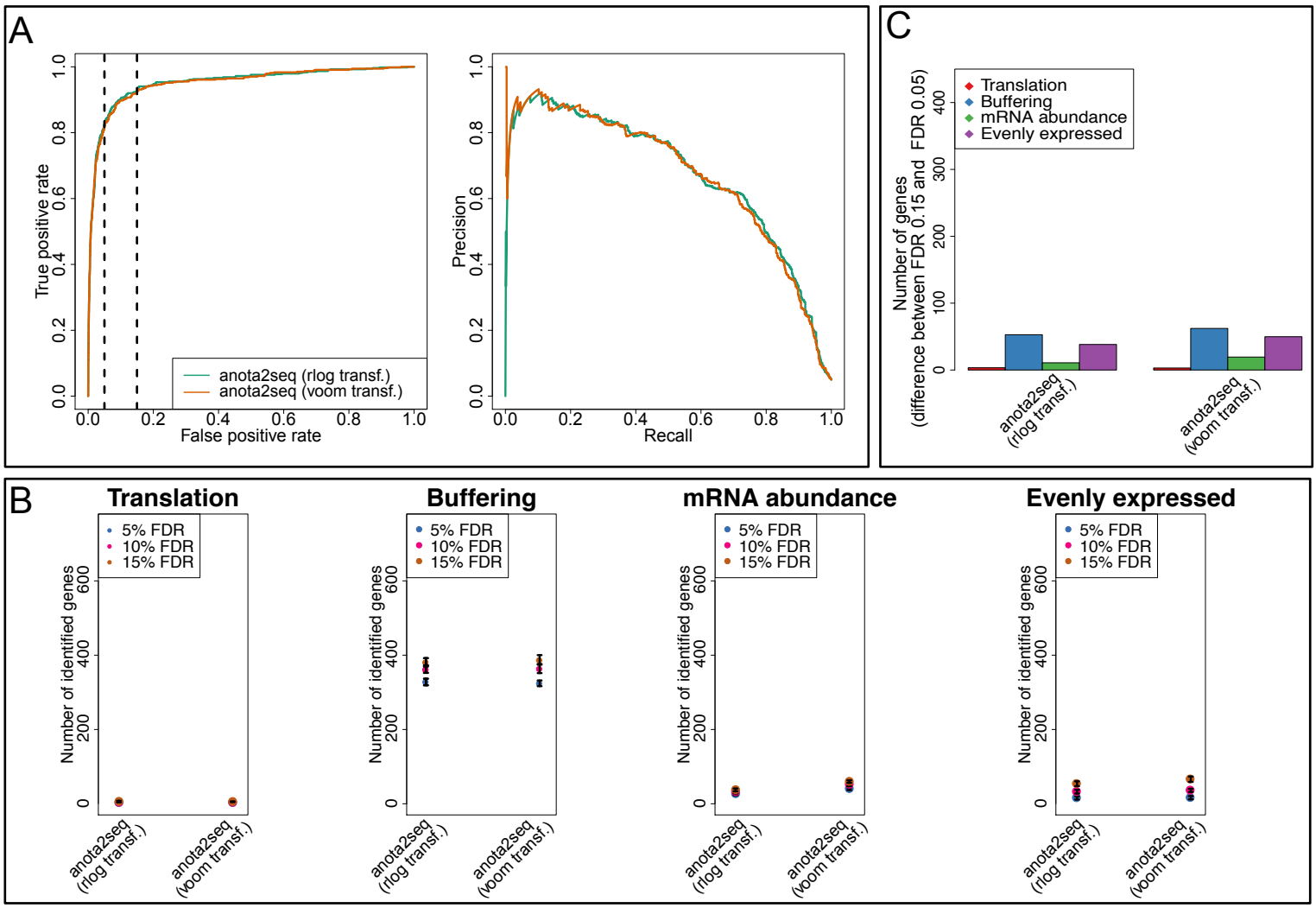

Figure 4

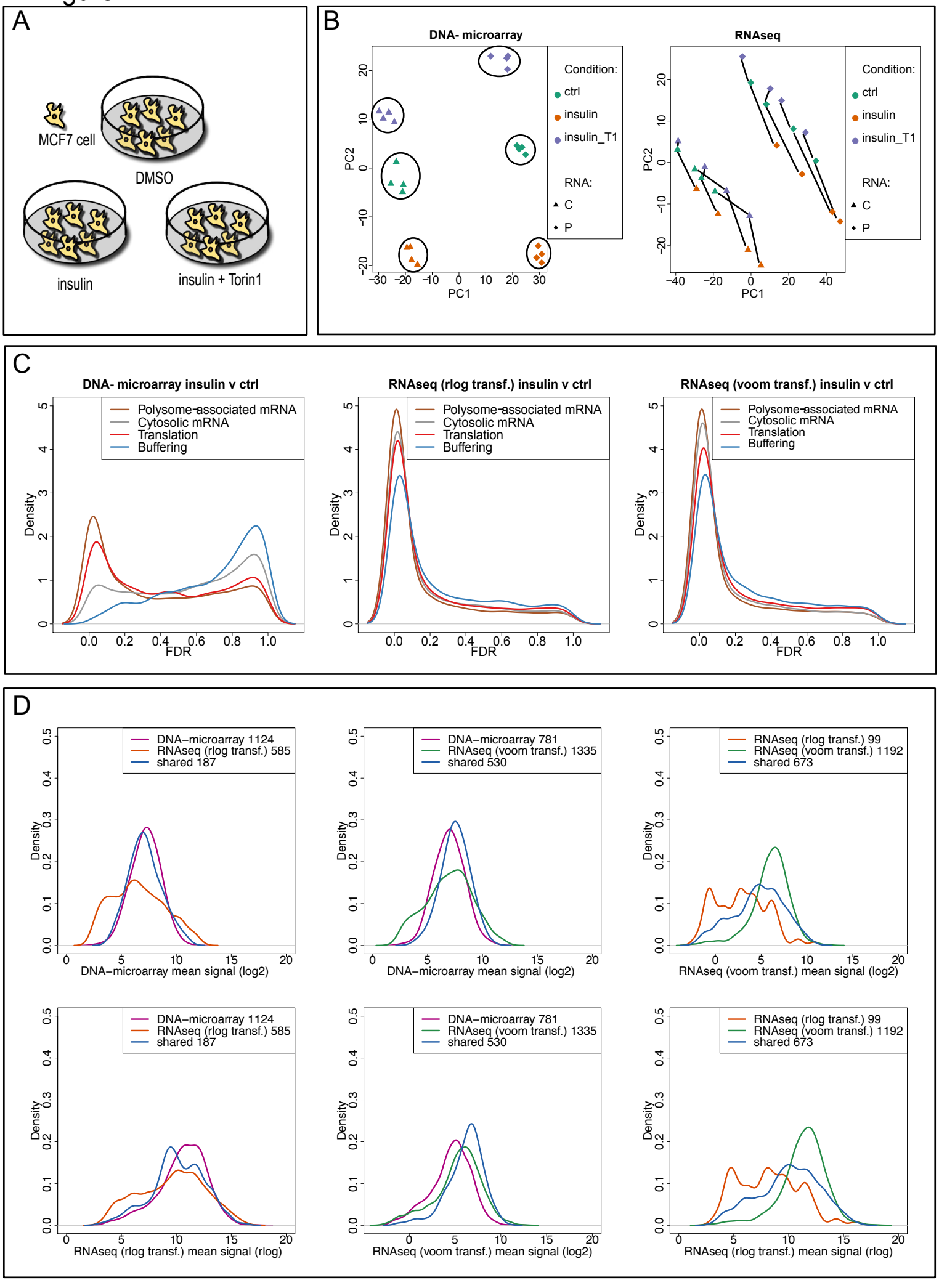

Figure 5

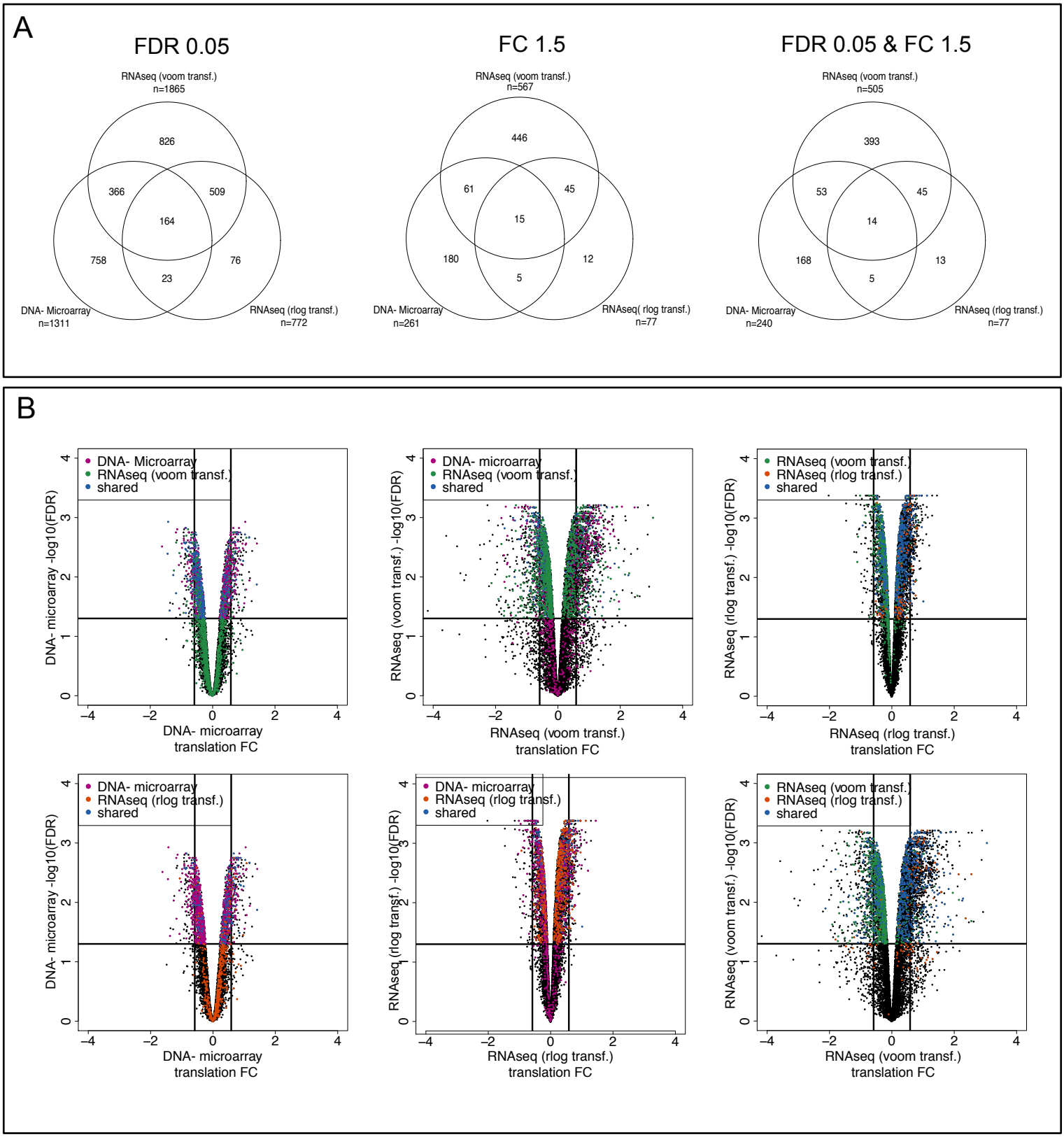

Figure 6

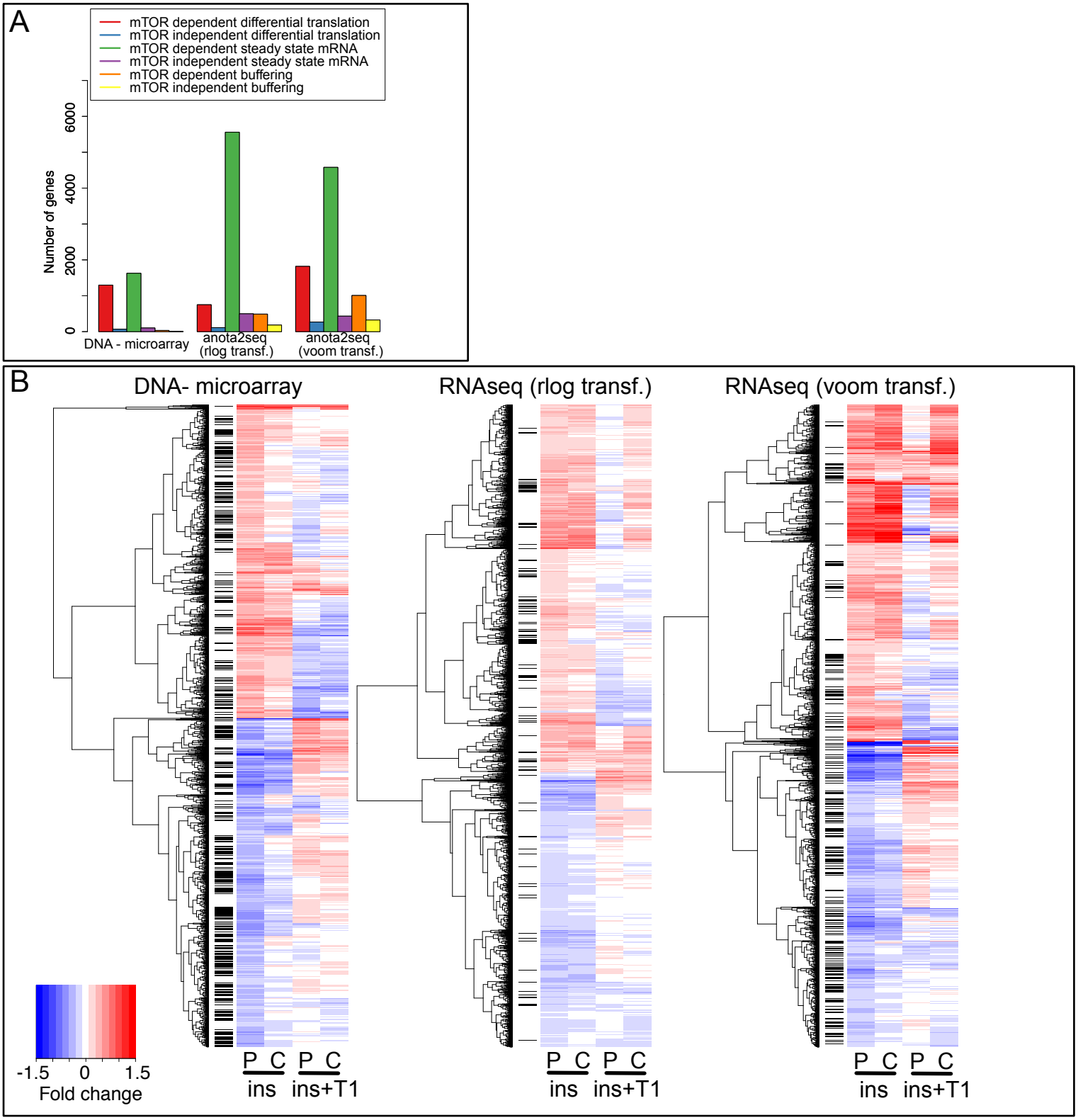

Table 1

		pAUC FPR 5%		pAUC FPR 15%		AUC	
Method	n	mean	sd	mean	sd	mean	sd
anota2seq (rlog transf.)	5	0.6700	0.0159	0.8073	0.0140	0.9486	0.0069
anota2seq (voom transf.)	5	0.6601	0.0131	0.8014	0.0131	0.9462	0.0067
babel	5	0.2860	0.0144	0.5930	0.0147	0.9124	0.0044
DESeq2	5	0.4080	0.0106	0.6679	0.0052	0.9145	0.0053
TE score	5	0.3602	0.0201	0.6072	0.0105	0.8945	0.0050
Xtail	5	0.3893	0.0117	0.6563	0.0056	0.9136	0.0059

Table 2

		pAUC FPR 5%		pAUC FPR 15%		AUC	
Method	n	mean	sd	mean	sd	mean	sd
anota (rlog transf.)	5	0.651	0.021	0.805	0.021	0.952	0.010
anota (voom transf.)	5	0.646	0.021	0.800	0.021	0.949	0.011