

Identification of differential translation in genome wide studies

Ola Larsson^{a,1,2}, Nahum Sonenberg^a, and Robert Nadon^{b,c,2}

^aDepartment of Biochemistry, McGill University, Montreal, Quebec H3A 1A3, Canada; ^bDepartment of Human Genetics, McGill University, Montreal, Quebec H3A 1B1, Canada; and ^cMcGill University and Genome Quebec Innovation Centre, Montreal, Quebec H3A 1A4, Canada

Edited by Peter J. Bickel, University of California, Berkeley, CA, and approved October 26, 2010 (received for review May 20, 2010)

Regulation of gene expression through translational control is a fundamental mechanism implicated in many biological processes ranging from memory formation to innate immunity and whose dysregulation contributes to human diseases. Genome wide analyses of translational control strive to identify differential translation independent of cytosolic mRNA levels. For this reason, most studies measure genes' translation levels as log ratios (translation levels divided by corresponding cytosolic mRNA levels obtained in parallel). Counterintuitively, arising from a mathematical necessity, these log ratios tend to be highly correlated with the cytosolic mRNA levels. Accordingly, they do not effectively correct for cytosolic mRNA level and generate substantial numbers of biological false positives and false negatives. We show that analysis of partial variance, which produces estimates of translational activity that are independent of cytosolic mRNA levels, is a superior alternative. When combined with a variance shrinkage method for estimating error variance, analysis of partial variance has the additional benefit of having greater statistical power and identifying fewer genes as translationally regulated resulting merely from unrealistically low variance estimates rather than from large changes in translational activity. In contrast to log ratios, this formal analytical approach estimates translation effects in a statistically rigorous manner, eliminates the need for inefficient and error-prone heuristics, and produces results that agree with biological function. The method is applicable to datasets obtained from both the commonly used polysome microarray method and the sequencing-based ribosome profiling method.

differential expression | RIP-CHIP | random variance model | translomics

Regulation of gene expression is a multistep process that includes transcription, splicing, mRNA-export, -localization, -stability, and -translation. Translational control of gene expression can be achieved by modulating translation elongation, termination, or initiation, although the latter seems to be most frequent (1). Thus, differential translation usually involves a change in the number of ribosomes bound to each mRNA, leading to a change in the amount of synthesized protein per mRNA molecule and time unit. This regulation can be accomplished by a specific process, which targets individual or sets of mRNAs for regulation, or by a general process, which affects most mRNAs equally (2). Specific translational control is important in many biological processes, including cellular senescence (3) and diseases (e.g., cancer) (4).

The polysome microarray approach is the most commonly used method for studying genome wide translational control. In this approach, mRNAs associated with several ribosomes (usually >3) are separated from mRNAs associated with fewer ribosomes and probed with microarrays. A more recent method, which circumvents some of the polysome microarray approach's limitations (5), involves isolation and sequencing of RNA pieces that are physically protected by ribosomes (6). A critical interpretative difficulty with both methods, however, is that observed differential levels of actively translated mRNAs or protected mRNA pieces may be due to differential cytosolic mRNA levels. Correction for cytosolic mRNA level has historically been

achieved by dividing actively translated mRNA levels by cytosolic mRNA levels obtained in parallel and logging the ratios. We show that these log ratios do not actually correct for cytosolic mRNA levels and that they consequently generate substantial numbers of biological false positives and false negatives. Here we propose a more sensitive and specific method for analysis of translational activity that includes analysis of partial variance (APV) linear regression to control for cytosolic mRNA levels and a variance shrinkage method for improving statistical inference.

Results and Discussion

Common Issues in Genome Wide Analysis of Translational Control.

Two types of data are produced from each sample when studying translational activity: actively translated mRNAs ("translational activity data") and cytosolic mRNAs ("cytosolic mRNA data") obtained in parallel. Typically, log (translational activity data/cytosolic mRNA data) ratios are calculated (i.e., logged cytosolic mRNA data is subtracted from its associated logged translational activity data) with the idea of obtaining a cytosolic mRNA data-corrected estimate of translation activity and compared between classes (Fig. 1A). Because log ratios are differences between logged values, class comparison effects are estimated from difference scores. Log difference scores can be correlated with cytosolic mRNA data, however, leading to incorrect biological conclusions. Such is the case when translational activity data and cytosolic mRNA data are uncorrelated for technical or biological reasons (Fig. 1B and C). For example, false positives can arise from log ratios when mRNAs fail to reach the predetermined threshold for the number of ribosomes necessary to join the pool of actively translated mRNAs or when short poorly translated mRNAs, but not their paired cytosolic mRNAs, fail to achieve counts for protected mRNA pieces above the noise level. Log ratios for such mRNAs tend to produce false positives as they appear to be under translational control (Fig. 1B). False negatives can be produced by log ratios when translational activity is regulated independently of the cytosolic mRNA level (Fig. 1C). The phenomenon of a difference score ($Y - Z$) correlating with each of its terms (Y and Z) was first described by Pearson in 1897 (7) who labeled it "spurious correlation" because of the frequent practice of interpreting such correlations as substantive rather than artifactual. We will also refer to these correlations as spurious, although our focus will be on the inadequacy of difference scores to control for cytosolic mRNA data and as a consequence fail to correctly estimate class effects (e.g., genotype, treatment, disease). The following equation illustrates the mathematical

Author contributions: O.L., N.S., and R.N. designed research; O.L. performed research; O.L. analyzed data; and O.L., N.S., and R.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Department of Oncology-Pathology, Cancer Center Karolinska, Karolinska Institute, R8:01, 171 76 Stockholm, Sweden.

²To whom correspondence may be addressed. E-mail: ola.larsson@ki.se or robert.nadon@mcgill.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1006821107/-DCSupplemental.

necessity underlying the problem (8):

$$r_{(Y-Z)Z} = \frac{r_{YZ}s_Ys_Z - s_Z^2}{s_Z\sqrt{s_Y^2 + s_Z^2 - 2r_{YZ}s_Ys_Z}} \quad [1]$$

where Y is a vector of translational activity data for a specific mRNA, Z is a vector of paired cytosolic mRNA data for the same mRNA, r is the Pearson correlation coefficient, and s is the sample standard deviation. When translational and cytosolic levels are uncorrelated, Eq. 1 simplifies to:

$$r_{(Y-Z)Z} = \frac{-s_Z}{\sqrt{s_Y^2 + s_Z^2}}. \quad [2]$$

Eq. 2 makes the origin of these false negatives and false positives clear; under this circumstance the correlation between the log ratios and their corresponding log cytosolic mRNA data is a function of the standard deviations of translational activity data and cytosolic mRNA data replicates. When the standard deviations are also equal, Eq. 2 yields a correlation of -0.71 ($-1/\sqrt{2}$). That is, half (-0.71^2) of the variance associated with the log ratio would be due to cytosolic mRNA levels. A similar difficulty with log ratios arises in the more typical situation in which the correlation between translational activity data and cytosolic mRNA is nonzero. When standard deviations are also equal, Eq. 1 then simplifies to:

$$r_{(Y-Z)Z} = \frac{1}{\sqrt{2}} \frac{(r_{YZ} - 1)}{\sqrt{1 - r_{YZ}}}. \quad [3]$$

When the correlation between translational activity and paired cytosolic mRNA is 0.60, Eq. 3 yields a correlation of -0.45 . Thus, under various realistic scenarios, the correlation between log ratios and cytosolic mRNA is nontrivial and gives rise to biological false positives and negatives.

To examine the practical implications of problems with log ratios, we calculated the per gene correlations between logged translational activity data/cytosolic mRNA data ratios and log cytosolic mRNA data per gene across all samples in a set of published studies (9–28). The median of the per dataset medians for the spurious correlations was -0.61 with values covering almost the entire range between $+1$ and -1 (Fig. 1E). Although these results do not imply low data quality, they demonstrate that log ratios are inappropriate for controlling the confounding effects of cytosolic mRNA levels. It is noteworthy that using the union from the translational activity data only analysis and the log ratio approach resolves the false negative example (Fig. 1C), but leads to increased false positives (Fig. 1A and B). Using the intersection of the translational activity data only analysis and the log ratio approach limits identification of differential translation to situations that are shown in Fig. 1D and may fail to identify instances of translational regulation that are independent of cytosolic mRNA levels (Fig. 1C). One might argue that a set of heuristic thresholds for differences at the translational activity data and cytosolic mRNA data could be used either alone or in combination with the translational activity data only or the log ratio approach. Such heuristics suffer from numerous failings, however, including the absence of objective statistical thresholds and that any choice of subjective threshold generates unknown rates of false positives and false negatives. Thus, from both theoretical and practical standpoints, current approaches are inadequate for analysis of differential translation.

Anota Outperforms Current Approaches for Analysis of Differential Translation. APV (29) provides a more suitable analysis of translational activity data because APV-corrected translation mea-

surements do not, by definition, show spurious correlations. We have implemented this approach in the anota (analysis of translational activity) R-package. In anota, a common slope for all sample categories is identified for each gene from the least squares linear regression of translational activity data on cytosolic mRNA data. Class comparison effects are estimated by calculating differences between sample category intercepts; the sum of squares error for these comparisons is reduced by the sum of squares associated with the covariance between the translational activity data and the cytosolic mRNA data. We found that anota potentially correctly analyzes the different scenarios outlined in Fig. 1A–D. Anota corrects for confounding cytosolic mRNA levels (Fig. 1A), handles false positives arising from technical issues (Fig. 1B), identifies translational regulation that is independent of cytosolic mRNA levels (Fig. 1C), and identifies translational control when there is some confounding cytosolic mRNA difference (Fig. 1D). In summary, anota handles different problematic scenarios by automatically adapting the analysis to the situation at hand.

To further assess anota's performance, we analyzed genome wide data from both the polysome microarray approach and the sequencing-based approach in three studies [Ingolia et al. (24), Kitamura et al. (21), and Otulakowski et al. (9); see Fig. 1E]. Two sample categories were identified in each study. We compared results obtained from anota with those obtained from a t -test between the log ratios of the two sample classes. Scatter plots of effects and of p -values reveal large discrepancies between the two approaches for many genes (Fig. S1). We investigated this further by classifying mRNAs as either translationally activated or inactivated and examining the overlap among the top 5% of differentially translated mRNAs (ranked by nominal p -values) between the two analytic approaches. As shown in Fig. 1F, the number of genes that were identified as translationally regulated by one method but not the other was substantially higher than the number of shared genes. Genes were then classified into one of three sets: identified by anota only, log ratio only, or by both log ratio and anota. The top two genes (ranked by p -value) from each set from the Ingolia et al. study (24) illustrate the differences between the methods (Fig. 2). Of the top two genes identified by the log ratio approach, one was of the type exemplified in Fig. 1B with cytosolic mRNA data differences combined with smaller translational activity data differences (Fig. 2, rank 2) whereas the other showed a larger difference at the cytosolic mRNA level compared to the translational activity level, leading to the conclusion that a gene with lower translational activity data is more translationally active. This could be a case of “over-correction” and hence represents a false positive or possibly a biologically interesting finding. However, it is commonly assumed that a gene that is translationally activated would also show higher protein levels and this gene would therefore not be a primary target for follow-up studies. The top two genes that were identified by anota only showed translational regulation that was independent of cytosolic mRNA levels and accordingly belonged to the type exemplified in Fig. 1C. The genes that were identified by both anota and the log ratio approach showed translational regulation independent of cytosolic mRNA level (as exemplified in Fig. 1C) or a small difference in cytosolic mRNA level associated with a larger difference in translational activity data (as exemplified in Fig. 1D). The two other studies analyzed in Fig. 1F showed similar results (Figs. S2 and S3).

To extend the observations from Fig. 2, we compared all genes from the three sets (identified by anota only, log ratios only, or both log ratios and anota). Within each set, genes were classified further into one of three modes of regulation based on between-group changes in translational activity data and cytosolic mRNA data (shown as delta translation and delta transcription respectively, in Fig. 2). The first mode, labeled “translation > transcription,” identifies those genes for which the translational

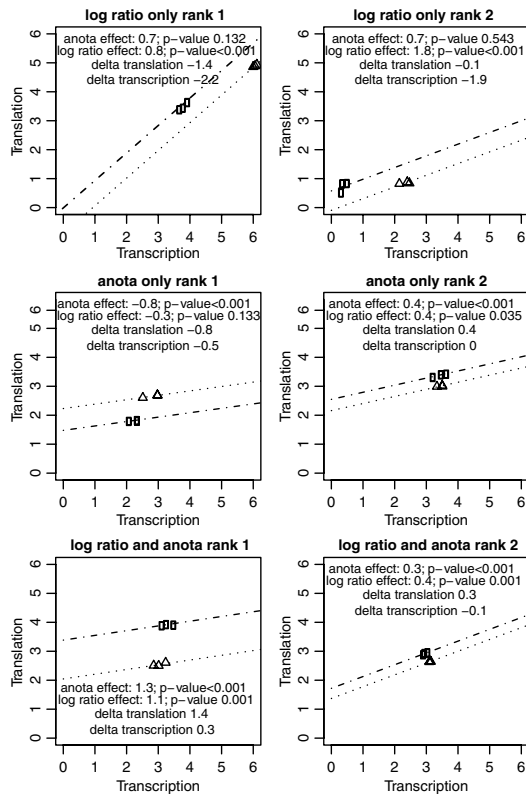


Fig. 2. A comparison of the top two mRNAs identified as differentially translated by the log-ratio approach and by anota. Translational activity data (indicated as translation on each y-axis) and cytosolic mRNA data (indicated as transcription on each x-axis) for the top two ranked genes (by *p*-value) for each set (log-ratio only; anota only; and both) are plotted using a common scale for all graphs. The effects obtained from anota and from the log-ratio approach are indicated together with their corresponding *p*-values. The differences in means between the sample classes for the translational activity data only and the cytosolic mRNA data only are indicated as delta translation and delta transcription, respectively. Lines represent the regression lines estimated by anota for the two sample categories.

activity data difference was larger than the opposing cytosolic mRNA difference; i.e., for a gene that is translationally activated, the activation originated primarily from an activation at the translational activity level and not from a decrease at the cytosolic mRNA level (all genes identified by anota only or by both anota and log ratios in Fig. 2 fit into this group). The second mode, labeled “no translation,” identifies those genes for which there was no difference in the translational activity data; i.e., for a gene defined as activated, no increase in the translational activity data (delta translation ≤ 0) was associated with a decrease in cytosolic mRNA data (the log ratio only genes from Fig. 2 would belong to this mode). The third mode, labeled “transcription > translation,” identifies those genes that show a translational activity data difference but for which the cytosolic mRNA data difference in the opposite direction is of larger magnitude. Of these three modes, the translation > transcription mode corresponds to the examples shown in Fig. 1 C and D and represents cases when translational regulation is plausible whereas the no translation and the transcription > translation modes are variants of the example in Fig. 1B when it is unclear if there is any translational regulation. Anota primarily identified genes belonging to the preferred translation > transcription mode whereas most of the genes identified by the log ratio approach belong to the undesired no translation or transcription > translation modes (Fig. 3). Interestingly, those genes that were identified by both anota and the log ratio approach showed a distribution among the three modes similar to that of anota and thus represent a sub-

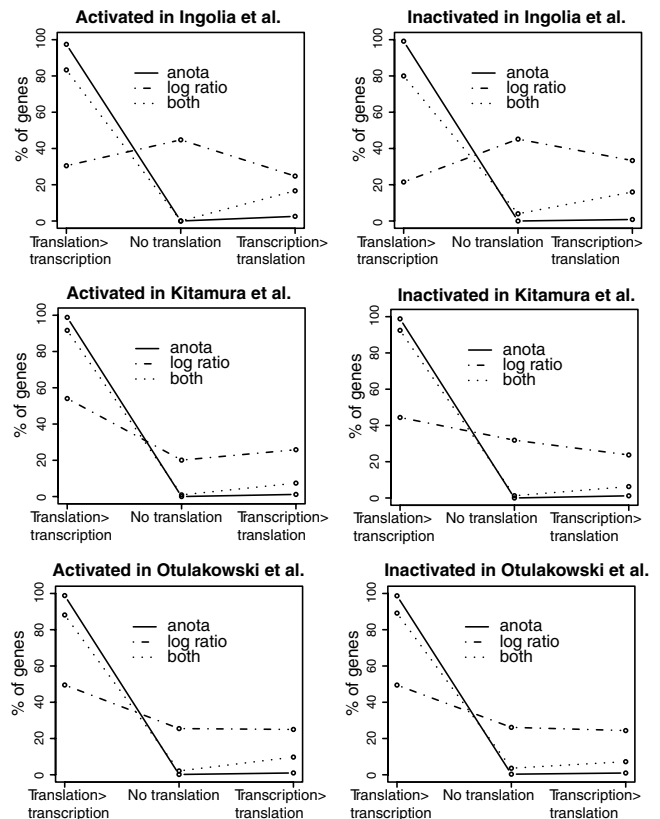


Fig. 3. A comparison of the mode of translational regulation of genes identified as differentially translated by the log-ratio approach and by anota. The top 5% of genes ranked by *p*-value from the three examined datasets (from Fig. 1F) were separated into translationally inactivated and translationally activated (i.e., negative and positive effects) and further divided into three sets each: those identified by log ratio only, those identified by anota only, and those identified by both log ratio and anota. The genes within each of these six sets were categorized based on their mode of regulation (the modes are further discussed in the main text): Translation > transcription (similar to Fig. 1 C and D); no translation (similar to Fig. 1B); Transcription > translation (similar to Fig. 1B). Shown is the percent of genes among the translationally activated or translationally inactivated belonging to each mode.

set of genes for which the log ratio approach obtained a result similar to anota. In summary, anota outperforms the log ratio approach and can be used both for microarray-based as well as sequencing-based datasets.

Improving the Performance of Anota Using Variance Shrinkage. In genome wide studies, gene-specific variances of many genes will be greatly under or overestimated due to chance factors. It is therefore possible that some of the genes shown in Fig. 2 have been identified as a result of unrealistically low variance estimates. This characteristic of high dimensional data has produced a consensus among methodologists to favor variance shrinkage methods that adjust error variances for statistical tests by weighting gene-specific variance estimates and a variance estimate obtained from all genes in the dataset (30). We have generalized one such method, the random variance model (RVM) (31), for APV within anota (referred to as “anota RVM”). We performed an analysis identical to that shown in Fig. 2 but with the RVM-adjusted error term. The top three genes that were selected by anota RVM but not by the log ratio approach in the Ingolia et al. study (24) are shown in Fig. 4A. These genes show larger effects than those genes identified without RVM (compare Fig. 4A to Fig. 2). To illustrate the effect across all genes, we generated “volcano” plots comparing anota and anota RVM. The volcano plots from the Ingolia et al. dataset show that smaller effects are less likely to be asso-

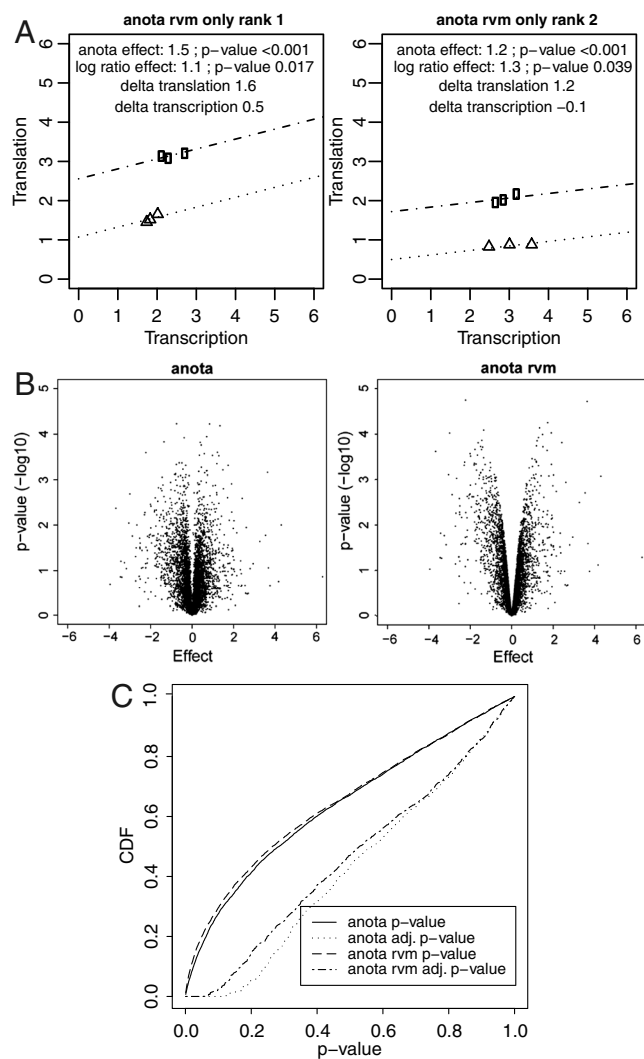


Fig. 4. RVM improves the performance of APV within anota. (A). Translational activity data (indicated as translation on each y-axis) and cytosolic mRNA data (indicated as transcription on each x-axis) from the top two genes (ranked by p -value) that were identified among the top 5% differentially translated (ranked by p -value) by anota RVM but not by the log-ratio approach are shown. The two sample categories are indicated by squares and triangles. The same scale as in Fig. 2 is used. Effects calculated by anota RVM and by the log-ratio approach are indicated together with their corresponding p -values. The differences in means between the sample classes for the translational activity data only and the cytosolic mRNA data only are indicated as delta translation and delta transcription, respectively. Lines represent the regression lines estimated by anota for the two sample categories. (B). Volcano plots for anota and anota RVM. The $-\log_{10}$ p -value is plotted as a function of the effect. (C). A comparison of statistical power between anota and anota RVM. Shown is the cumulative distribution function (CDF) of all p -values from anota and anota RVM. CDFs for both the raw p -values and p -values adjusted for multiple testing as described in the text are shown.

ciated with high significance when anota RVM is used (Fig. 4B). Finally, we compared the cumulative distribution function of the raw p -values and p -values that had been adjusted for multiple testing (32) from anota and anota RVM (Fig. 4C). This analysis indicated that anota RVM generated smaller p -values than anota without RVM adjustment [similar results were obtained for the other two datasets from Fig. 1F (Figs. S4 and S5)].

We searched for enrichment of genes with shared biological functions among genes identified by anota RVM and the log-ratio approach to assess whether the former enabled new biological insights. Genes identified as differentially translated by anota

RVM shared functions to a larger extent than genes identified as translationally regulated by the log-ratio approach in all three datasets shown in Fig. 1F (Figs. S6–S8). This suggests that anota RVM is associated with fewer biological false positive and false negative findings inasmuch as a completely random set of genes would not enable identification of any enriched biological functions. Thus, anota identifies differential translation that is more biologically coherent compared to log ratios and thereby offers greater promise for biological discoveries.

Conclusions

Regulation of translation is important for many biological processes and dysregulation is present in diseases such as cancer and fibrosis. Despite this, genome-wide studies of differential translation have been rare. Downstream of the technical challenges, data analysis has in the past relied on simple intuitive approaches that we show are ineffective for analysis of differential translation and have led to many false inferences that may have hampered the enthusiasm for genome wide studies of translational control. Our approach to analysis of translational activity provides an advancement that is necessary to understand how mRNA specific translation is regulated on a transcriptome wide scale and shows promise for generating reproducible and more biologically sound findings.

We show that our approach for analysis of differential translation is versatile and can be used both on count data from sequencing as well as microarray data. Moreover, our approach is not limited to studies of translational control but could be used in, for example, ribonucleoprotein immunoprecipitation–microarray (RIP-chip) studies where the observed differential level of immunoprecipitated mRNA will be partially dependent on the total mRNA amount. In general, anota can be used profitably with any datasets containing paired-controls.

Defining which genes are under translational control only represents the first step toward mechanistic understanding of translational regulation and its underlying organization. Efforts are currently being made toward identifying tools and approaches that will be necessary to discover mechanisms that define the common regulatory patterns as described in the posttranscriptional regulation theory and by the originally proposed informosomes (33–36). We believe that our methodological improvements for analysis of genome wide datasets are in this spirit and will help to understand how the transcriptome is regulated from a systems perspective. Such characterization will be necessary to understand how dysregulation of translational control leads to human diseases.

Materials and Methods

Sampling of Example Data. Data shown in Fig. 1 A–D were sampled from normal distributions. First, 5 cytosolic mRNA data points were sampled and then a set of corresponding translational activity data points were sampled using the values of the sampled cytosolic mRNA data as distribution means. A second sample class was generated in the same manner. The different examples in Fig. 1 A–D were generated by changing the means and the standard deviations of the normal distributions within the “*rnorm*” function in R.

A Survey of the Occurrence of Spurious Correlations in Published Datasets. We searched both the ArrayExpress (37) and the Gene Expression Omnibus (38) databases for a set of studies of translational activity. Datasets generated using Affymetrix GeneChips were normalized with the robust multiarray average method (39, 40) using updated probe set definitions when available (41, 42). For the cDNA dataset (11), we used the normalized data provided by the authors. For the sequencing dataset (24), we used the count data supplied by the authors, filtered for identifiers originating from the coding regions, and used quantile normalization and a transformation to stabilize the variance. For all studies, we identified paired translational activity and cytosolic mRNA data and calculated correlations between the \log_2 difference scores and the cytosolic mRNA data per gene across all samples in each study.

Analysis of Differential Translation Using Three Datasets. We identified two sample classes from each study to be used for the analysis of differential

translation. For the Kitamura et al. dataset (21), we compared “no treatment” to “LPS treatment 4 h.” For the Otulakowski et al. dataset (9), we compared “FD19” to “P1.” The sequencing-based dataset contained only two samples classes (“rich” and “starved”) and only two samples per class (24). Because anota requires at least 3 samples per sample class, we simulated a third sample for each class and type of data (actively translated and cytosolic mRNAs). The simulation was done using the transformed and normalized data and by sampling one new data point from a normal distribution with a mean and a standard deviation derived from the data points of that gene in the empirical data. Correlation between the simulated and empirical data was similar to the correlation between the two empirical data replicates. To simulate the presence of outliers, we also generated a dataset for which 5% of all data points were shifted by adding or subtracting 3 gene-

specific standard deviations to/from the mean of the normal distribution from which the data point was sampled. The two simulated datasets generated similar results; we present results from the dataset without outliers. We generated log ratios using data from actively translated mRNAs and cytosolic mRNAs that originated from the same sample. Student's *t* test (two-tailed unequal variance) was used to compare the log ratios between the sample categories.

ACKNOWLEDGMENTS. We thank Richard Simon and George Wright for permission to incorporate some of their original RVM code into anota. O.L. was initially supported by a fellowship from the Knut and Alice Wallenberg Foundation. Partial support was provided by Le Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) Grant 119258 (R.N.).

- Mathews M, Sonenberg N, Hershey JBW (2007) *Translational Control in Biology and Medicine* (CSHL Press, New York), p 1.
- Polunovsky VA, Bitterman PB (2006) The cap-dependent translation apparatus integrates and amplifies cancer pathways. *RNA Biol* 3:10–17.
- Petroulakis E, et al. (2009) p53-dependent translational control of senescence and transformation via 4E-BPs. *Cancer Cell* 16:439–446.
- Silvera D, et al. (2009) Essential role for eIF4G1 overexpression in the pathogenesis of inflammatory breast cancer. *Nat Cell Biol* 11:903–908.
- Larsson O, Nadon R (2008) Gene expression—Time to change point of view? *Biotechnol Genet Eng* 25:77–92.
- Arava Y, Boas FE, Brown PO, Herschlag D (2005) Dissecting eukaryotic translation and its control by ribosome density mapping. *Nucleic Acids Res* 33:2421–2432.
- Pearson K (1896) On a form of spurious correlation which may arise when indices are used in the measurement of organs. *P R Soc London* 60:489–498.
- Cohen J, Cohen P, West SG, Aiken L (2003) *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Erlbaum, London).
- Otulakowski G, Duan W, O'Brodovich H (2009) Global and gene-specific translational regulation in rat lung development. *Am J Respir Cell Mol Biol* 40:555–567.
- Sampath P, et al. (2008) A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2:448–460.
- Lu X, de la Pena L, Barker C, Camphausen K, Tofilon PJ (2006) Radiation-induced changes in gene expression involve recruitment of existing messenger RNAs to and away from polysomes. *Cancer Res* 66:1052–1061.
- Chen YW, et al. (2002) Response of rat muscle to acute resistance exercise defined by transcriptional and translational profiling. *J Physiol* 545:27–41.
- Rajasekhar VK, et al. (2003) Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol Cell* 12:889–901.
- Iguchi N, Tobias JW, Hecht NB (2006) Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. *Proc Natl Acad Sci USA* 103:7712–7717.
- Spence J, Duggan BM, Eckhardt C, McClelland M, Mercola D (2006) Messenger RNAs under differential translational control in Ki-ras-transformed cells. *Mol Cancer Res* 4:47–60.
- Larsson O, et al. (2007) Eukaryotic translation initiation factor 4E induced progression of primary human mammary epithelial cells along the cancer pathway is associated with targeted translational deregulation of oncogenic drivers and inhibitors. *Cancer Res* 67:6814–6824.
- Parent R, Kolippakkam D, Booth G, Beretta L (2007) Mammalian target of rapamycin activation impairs hepatocytic differentiation and targets genes moderating lipid homeostasis and hepatocellular growth. *Cancer Res* 67:4337–4345.
- Genolet R, Araud T, Maillard L, Jaquier-Gubler P, Curran J (2008) An approach to analyze the specific impact of rapamycin on mRNA-ribosome association. *BMC Med Genomics* 1:33.
- Larsson O, et al. (2008) Fibrotic myofibroblasts manifest genome-wide derangements of translational control. *PLoS One* 3:e3220.
- Grech G, et al. (2008) Igbp1 is part of a positive feedback loop in stem cell factor-dependent, selective mRNA translation initiation inhibiting erythroid differentiation. *Blood* 112:2750–2760.
- Kitamura H, et al. (2008) Genome-wide identification and characterization of transcripts translationally regulated by bacterial lipopolysaccharide in macrophage-like J774.1 cells. *Physiol Genomics* 33:121–132.
- Parent R, Beretta L (2008) Translational control plays a prominent role in the hepatocytic differentiation of HepaRG liver progenitor cells. *Genome Biol* 9:R19.
- Ramirez-Valle F, Braunstein S, Zavadil J, Formenti SC, Schneider RJ (2008) eIF4G1 links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. *J Cell Biol* 181:293–307.
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324:218–223.
- Kim YY, et al. (2009) Eukaryotic initiation factor 4E binding protein family of proteins: Sentinels at a translational control checkpoint in lung tumor defense. *Cancer Res* 69:8455–8462.
- Ceppe M, et al. (2009) Ribosomal protein mRNAs are translationally-regulated during human dendritic cells activation by LPS. *Immunome Res* 5:5.
- Dang Do AN, Kimball SR, Cavener DR, Jefferson LS (2009) eIF2alpha kinases GCN2 and PERK modulate transcription and translation of distinct sets of mRNAs in mouse liver. *Physiol Genomics* 38:328–341.
- Zid BM, et al. (2009) 4E-BP extends lifespan upon dietary restriction by enhancing mitochondrial activity in *Drosophila*. *Cell* 139:149–160.
- Schleifer SJ, Eckholdt HM, Cohen J, Keller SE (1993) Analysis of partial variance (Apv) as a statistical approach to control day-to-day variation in immune assays. *Brain Behav Immun* 7:243–252.
- Allison DB, Cui X, Page GP, Sabripour M (2006) Microarray data analysis: From disarray to consolidation and consensus. *Nat Rev Genet* 7:55–65.
- Wright GW, Simon RM (2003) A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* 19:2448–2455.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 57:289–300.
- Fan D, Bitterman PB, Larsson O (2009) Regulatory element identification in subsets of transcripts: Comparison and integration of current computational methods. *RNA* 15:1469–1482.
- Keene JD (2007) RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* 8:533–543.
- Spirin AS (1969) The second Sir Hans Krebs Lecture. Informosomes. *Eur J Biochem* 10:20–35.
- Foat BC, Stormo GD (2009) Discovering structural cis-regulatory elements by modeling the behaviors of mRNAs. *Mol Syst Biol* 5:268.
- Parkinson H, et al. (2009) ArrayExpress update—From an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res* 37:D868–D872.
- Barrett T, et al. (2009) NCBI GEO: Archive for high-throughput functional genomic data. *Nucleic Acids Res* 37:D885–D890.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185–193.
- Irizarry RA, et al. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31:e15.
- Dai M, et al. (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33:e175.
- Sandberg R, Larsson O (2007) Improved precision and accuracy for microarrays using updated probe set definitions. *BMC Bioinformatics* 8:48.