RiboSimR: A Tool for Simulation and Power Analysis of Ribo-seq Data

Patrick Perkins[(✉)], Anna Stepanova, Jose Alonso, and Steffen Heber

North Carolina State University, Raleigh, NC 27502, USA
Pjperki2@ncsu.com

Abstract. RNA-seq and Ribo-seq are widespread quantitative methods for assessing transcription and translation. They can be used to detect differential expression, differential translation, and differential translation efficiency between conditions. The statistical power to detect differential genes is affected by multiple factors, such as the number of replicates, sequencing depth, magnitude of differential expression and translation, distribution of gene counts, and method for estimating biological variance. As power estimation of translational efficiency involves the combination of both RNA-seq measurements and Ribo-seq measurements, this task is particularly challenging. Here we propose a power assessment tool, called RiboSimR, based purely on data simulation. RiboSimR, produces semi-parametric simulations that generate data based on real RNA and Ribo-seq experiments, with customizable choices on baseline parameters and tool configurations. We demonstrate the usefulness of our tool by simulating data based on two published Ribo-seq datasets and analyzing various aspects of experimental design.

Keywords: Ribosome profiling · RNA-seq · Simulation · Power analysis · Experimental design · Differential analysis

1 Introduction

RNA-seq and Ribo-seq are popular techniques for investigating the transcriptional and translational landscape [1, 2]. The techniques use next-generation sequencing to produce genome-wide high-resolution snapshots of the total populations of mRNAs and translating ribosomes, respectively. Although it is a less widely used assay, Ribo-seq has shown that it can be used to assess the dynamics of ribosome activity during translation at a nucleotide specific resolution [1]. These techniques generally produce count tables, which quantify transcription and translation for each gene. These tables can be used to measure changes in transcription and translation across biological conditions, treatments, or timepoints by calculating differential expression (DE), differential translation (DT), and differential translation efficiency (DTE). Differential translation measures changes in translation without taking changes in transcript abundance into account, while differential translational efficiency assesses changes in translation after accounting for changes in transcript abundance [3].

In order to determine statistical significance when testing for differential genes, sample replication is used to account for biological and technical variability [4].

The amount of replication researchers should employ is a complicated but important aspect of experimental design in and sequencing experiment. Additionally, for both Ribo-seq and RNA-seq, there exists a minimum sequencing depth threshold for identifying biologically important genes [5]. Sequencing under this threshold reduces the power of the test to identify differential genes, while sequencing above this threshold has diminishing returns in terms of the yield of differential genes per amount of sequencing. As we are dealing with read count data, which is commonly modeled using a negative binomial distribution, counting error plays a role in gene determining optimal sequencing depth [6, 7]. Genes with low counts are more greatly affected by this counting error than those with high counts, and therefore, our ability to detect a differential gene is dependent on the count value of the gene. Determining the appropriate levels of sample replication and sequencing depth for Ribo-seq and RNA-seq are critically important aspects of experimental design and should to be assessed prior to conducting any experiment.

Our ability to detect differential genes can also be affected by which methods we choose to use when performing tests for statistical significance. As sample sizes are often limited, empirical Bayes methods are commonly used to estimate gene-specific biological variation by shrinking variance estimates towards the average trend across all genes [6, 7]. These procedures are typically effective at providing variance estimates which are closer to the biological variance, but can also introduce dependencies amongst genes that violate some of the assumptions of the error-controlling procedures during multiple testing correction. It is essential that researchers understand the differences in methodologies between tools for differential analysis, as they often contain different variance estimation procedures and multiple testing correction methods, both of which can have significant effects on the results of the tests. Other seemingly simple decisions, like choosing a statistical significance threshold to use when determining differential genes, can be of great importance. The level of change that is considered biologically significant can vary based on the purpose of an experiment, as can trends in the relationship between p-values and effect sizes [8].

Researchers have previously shown that simulations can be used to examine different experimental design setups for RNA-seq, including scenarios with varying sequencing depth and replicate number [9–11]. Wu et al. proposed a prospective power analysis setup to visualize power in multiple forms and across various circumstances [9]. Similarly to other studies, they simulate datasets using the negative binomial distribution, and draw parameters from real experiments. In Ribo-seq, the complexity of these experimental design questions only increases, as we must ask ourselves if our decisions might effect each experiment type differently. Researchers also need to assess scenarios where there are potentially large discrepancies in the level of sampling depth and replication between RNA-seq and Ribo-seq experiments. Issues can also arise when testing for statistical significance, as evident by the increased amount of discrepancy amongst methods which employ more complex two factor tests to calculate differential translation efficiency [12–15].

We therefore propose a comprehensive simulation strategy for the purpose of performing dynamic power analyses of Ribo-seq experiments. We apply this strategy to assess various questions in experimental design of Ribo-seq and RNA-seq experiments. These questions include the optimal level of sample replication and sequencing depth,

the tradeoff in value of increasing depth versus replication, the effect of a gene's expression has on our ability to identify it as differential, the effect of adding differing amounts of depth and replication to pre-existing experiments, and the effect of various tools and parameters for performing differential analysis. We show the effects of this analysis using two datasets in Arabidopsis thaliana and Toxoplasma gondii. Furthermore, we present a R Shiny web application which allows users to perform simulations and power analyses based on their own data or pilot datasets. This tool can be used to decide which experimental parameters a user wants to use in a prospective project, or to aid them to add to a preexisting study.

2 Methods

We propose the web application RiboSimR, which uses simulations to evaluate the effect of experimental design on our power to identify differential genes in Ribo-seq experiments. Our simulation methodology is largely adapted from Wu et al.'s work on RNA-seq simulation, and our tool is composed of two similar steps [9]. Initially, we use a semi-parametric simulation scheme to produce count tables for Ribo-seq and RNA-seq experiments. These tables are made by drawing parameters from experimental count tables provided within the tool, or optionally uploaded by the user. The simulations therefore borrow multiple aspects from the real data, such as the distribution of biological dispersion within genes, the distribution of average transcription and translation levels, the negative relationship between dispersion and count magnitude, and the positive association between average transcripts and footprints. Secondly, we assess power and false discovery rates (FDRs) for the simulations using a large number of custom visualizations. Within the output of RiboSimR, we highlight features like false discovery cost, stratified power, and sequencing value. The tool allows users to customize various parameters within both the simulation setup and output generation steps.

2.1 Simulation Strategy

In order to perform downstream power analyses, we first use a negative binomial model to simulate both RNA-seq and Ribo-seq count tables. Researchers commonly use negative binomial models to generate RNA-seq and Ribo-seq count data, because the flexibility of these models allows for accurate representation of the mean-variance relationship found in these data [6, 7]. The NB distribution corresponds to a gamma-Poisson mixture, with the gamma distribution representing the biological variation, and the Poisson layer modeling the variability in sequencing read counts [9]. If Y_{gi} is the count value for gene g in replicate i, then $Y_{gi} \sim NB(s_i \mu_g, \phi_g)$, where μ_g is the mean count for gene g, ϕ_g is the dispersion for gene g, and s_i is the normalization factor for the library size of replicate i. We use this technique to model RNA-seq and Ribo-seq counts separately, i.e. each gene has a different mean count for Ribo-seq (Y_{giRibo}) and RNA-seq (Y_{giRNA}). Our method uses non-parametric resampling of mean count and dispersion parameters empirically from existing datasets, as we lack a valid justification for drawing the parameters parametrically. The dispersion parameter corresponds to the squared biological coefficient of variance, and can be thought of as representing the biological variation of gene

expression between replicates [6, 7]. We choose to sample mean count and dispersion parameters for individual genes as pairs in order to preserve their relationship, which has been previously described [6, 7]. Additionally, we sample parameters for our simulated RNA-seq and Ribo-seq counts in pairs, as there is a positive relationship between the level of transcription and translation within genes. For example, for a simulated gene g, we will sample parameters $[\mu_{gRNA}, \Phi_{gRNA}, \mu_{gRibo}, \Phi_{gRibo}]$ from a single row of the empirical count table.

Another important step in our simulations is to set effect sizes. As we seldom know the precise amount of DE, DT, and DTE that is biological significant within an experiment, this can be a risky assumption to make. These effect sizes are applied to mean count parameters across the two conditions, in effect generating artificial differential genes. For Ribo-seq experiments, we must assume separate effect sizes across the two different experiment types. Several techniques have been used to set effect sizes for simulated RNA-seq experiments, but we will focus on using a mixture model [9–11]. For a one factor test, such as a test for DE in RNA-seq, we let zg be the indicator that gene g is differentially expressed across a given treatment, and the proportion of differential genes be $P(z_g = 1) = \pi_1$. We have the effect size β_g satisfying $\beta_g | z_g = 0 = 0$ and $\beta_g | z_g = 1 \sim N(0, \sigma^2)$. This would be designated as a zero-inflated normal distribution for β_g. For Ribo-seq, because we are dealing with two experimental types, we require both $\beta_{gRibo}, \beta_{gRNA}$ and z_{gDT}, z_{gDE}, which describe the effect size and differential indicators for changes across conditions in either experiment type. We also define the differential effect size between experiments as β_{gE} where $\beta_{gE} = |\beta_{gRibo} - \beta_{gRNA}|$, and z_{gDTE} as the indicator for differential translation efficiency. Therefore, we have four possible types of differential genes:

1. $\beta_{gRNA} \neq 0, \beta_{gRibo} = 0, \beta_{gE} \neq 0 \rightarrow z_{gDE} = 1, z_{gDT} = 0, z_{gDTE} = 1$
2. $\beta_{gRNA} = 0, \beta_{gRibo} \neq 0, \beta_{gE} \neq 0 \rightarrow z_{gDE} = 0, z_{gDT} = 1, z_{gDTE} = 1$
3. $\beta_{gRNA} \neq 0, \beta_{gRibo} \neq 0, \beta_{gE} = 0 \rightarrow z_{gDE} = 1, z_{gDT} = 1, z_{gDTE} = 0$
4. $\beta_{gRNA} \neq 0, \beta_{gRibo} \neq 0, \beta_{gE} \neq 0 \rightarrow z_{gDE} = 1, z_{gDT} = 1, z_{gDTE} = 1$

Up to this point, we have described genes with zg = 1 as differentially expressed, but it may be the case that these genes are not biologically interesting, as $|\beta_{gRNA}|$ or $|\beta_{gRibo}|$ may be extremely low but non-zero. We would expect to have little power to detect these genes. Thus, we may be interested in defining differential genes of interest with an indicator $z_g* = 1$ if $|\beta_g| \geq \Delta$. This allows us to investigate the power of finding genes which we think are more likely to be biologically relevant. We can let the user decide the 'meaningful effect size', Δ. The meaningful effect size can also be different between experiment types, as users may expect larger overall differences to occur in either RNA-seq or Ribo-seq, or want to relax constraints on either experiment type.

2.2 Differential Analysis

Once the count tables for both experiment types have been simulated, we can assess differential expression, differential translation, and differential translation efficiency. Table 1 depicts the scenarios for each type of differential test, assuming a meaningful effect size. We have implemented four tools for analysis of differential genes: edgeR,

DESeq 2, limma-voom, and DSS for analysis of DE, DT, and DTE [6, 7, 17, 18]. Tests for DE and DT essentially ignore counts from the opposite experiment type, while tests for DTE perform a two factor test which investigates changes in translation levels after accounting for changes in levels of transcription. This can be seen as conducting an overall test for interaction using the formula~Experiment + Condition + Experiment: Condition, where we are looking for genes which respond differently to the treatment relative to the experiment type [6].

Table 1. Differential scenarios, assuming effect size Δ

	$|\beta_{gRibo}| < \Delta$		$|\beta_{gRibo}| \geq \Delta$	
	$|\beta_{gE}| < \Delta$	$|\beta_{gE}| \geq \Delta$	$|\beta_{gE}| < \Delta$	$|\beta_{gE}| \geq \Delta$
$|\beta_{gRNA}| < \Delta$	$z^*_{g\text{DTE}} = 0$	$z^*_{g\text{DTE}} = 1$	$z^*_{g\text{DTE}} = 0$	$z^*_{g\text{DTE}} = 1$
	$z^*_{g\text{DE}} = 0$	$z^*_{g\text{DE}} = 0$	$z^*_{g\text{DE}} = 0$	$z^*_{g\text{DE}} = 0$
	$z^*_{g\text{DT}} = 0$	$z^*_{g\text{DT}} = 0$	$z^*_{g\text{DT}} = 1$	$z^*_{g\text{DT}} = 1$
$|\beta_{gRNA}| \geq \Delta$	$z^*_{g\text{DTE}} = 0$	$z^*_{g\text{DTE}} = 1$	$z^*_{g\text{DTE}} = 0$	$z^*_{g\text{DTE}} = 1$
	$z^*_{g\text{DE}} = 1$	$z^*_{g\text{DE}} = 1$	$z^*_{g\text{DE}} = 1$	$z^*_{g\text{DE}} = 1$
	$z^*_{g\text{DT}} = 0$	$z^*_{g\text{DT}} = 0$	$z^*_{g\text{DT}} = 1$	$z^*_{g\text{DT}} = 1$

2.3 Power Assessment

For our purposes, we are not interested in finding genes with variation between conditions which is less than the effect size Δ, and can therefore restrict our genes to two categories: $|\beta_g| < \Delta$ and $|\beta_g| \geq \Delta$, where βg represents the effect size for the test of interest, be it DE, DT, or DTE. Table 2, which has been adopted from Wu et al. and simplified, shows the possible scenarios for any test [9].

Let G be the total number of genes tested, where G_1 are genes which exhibit an effect size of at least Δ, and G_0 are genes which do not. D_g is the decision on any gene, where $D_g = 1$ is a discovery and $D_g = 0$ is a non-discovery. Here V represents the total number of type I errors, or false discoveries, S the number of true positive, and R the total number of discoveries. The type 1 error is therefore P($V > 0$) and the FDR is E[V/R]. The FDC, described by Wu et al., can be defined as the number of false discoveries made for every true discovery, and represented as E[V/S] [9].

Here we are labeling genes which have $0 < |\beta_g| < \Delta$ as false discoveries, meaning that we only want to measure our ability to find genes with a meaningful effects size, and penalize all other discoveries.

Therefore, when we talk about power, we are actually discussing a targeted power, instead of a family-wise power, which is the probability of detecting all true differential genes [8]. This is because we both doubt our ability to find genes with very small effect sizes, and also question the biological importance of such genes. The targeted power, which from now on we simply refer to as power, is therefore E[S/G_1]. For each

simulation, we calculate the power, false discovery cost, and true positive rate of the discoveries made using a differential analysis tool and nominal FDR provided by the user. Averages across a number of repeated simulations are reported.

Table 2. Differential scenarios, assuming Δ

	Differential?	Discovery?		Total
	z_g^*	$D_g = 1$	$D_g = 0$	
$\lvert \beta_g \rvert < \Delta$	0	V	$G_0 - V$	G_0
$\lvert \beta_g \rvert \geq \Delta$	1	S	$G_1 - S$	G_1
Total		R	$G - R$	G

2.4 Implementation

We have implemented the proposed methods for simulation and power analysis in a R Shiny web application called RiboSimR, available at http://pjperki2.shinyapps.io/power. The tool allows users to upload their own count files for RNA-seq and Ribo-seq and create simulated data sets based on the parameters of their experiments. Users may also test different experimental scenarios for prospective experiments using provided pilot data sets. The app allows multiple options to customize the simulations, by choosing the size of the datasets they wish to simulate, the number of repetitions, and the meaningful effect size. Users may also choose between a large number of different visualizations, which include simulated variations of sample replication and sequencing depth, gene count stratification, comparisons between differential analysis tools, etc.… The tool also provides the choice of which type of test they want to conduct (DE, DT, DTE), and which power metrics they want to assess (power, true positive rate, false discovery cost). A description of the various tools and functionalities included in RiboSimR can be found at the apps home page. Runtime for the simulations depends on the size of the simulated data and the number of repetitions conducted, but a typical simulation with 20,000 genes and 20 repetitions takes approximately 10–30 min.

3 Results

3.1 Data Preparation

To display the utility of RiboSimR, we performed simulations and power analyses using two published datasets. The Merchante et al. data is from Arabidopsis thaliana seedlings, and was used to study of the role of the phytohormone ethylene as a gene-specific regulator of translation [19]. The data are split into two conditions, control samples treated with air, and ethylene treated seedlings. The Hassan et al. data are from intracellular and extracellular Toxoplasma gondii parasites [20]. This data was used to quantify the differences in translation between the intracellular and extracellular stages of the parasites.

Both datasets contain two replicates from two conditions for both RNA-seq and Ribo-seq. These datasets were chosen because they originate from studies which focus on using both RNA-seq and Ribo-seq data to identify genes which show differential translation efficiency between two treatments, and because they are from different organisms, and show varying levels of biological variability.

For all of the following simulations, we generate 20,000 genes and repeat each of the simulations 20 times. We assume that 2% of all genes fall into each of the four types of differential genes described in Sect. 2.1. Therefore, accounting for overlap, 6% of the total set of genes are DE, 6% are DT, and 6% are DTE. These simulation parameters were chosen to mirror the conditions of similar RNA-seq experiments [9]. The effect size for non-differential genes is set to 0, while the effect sizes of DE and DT genes are drawn from a normal distribution~$N(0, 1.52)$. Means and dispersions for simulated genes are drawn from the real data in paired sets with dependency. Unless stated otherwise, we employ the quasi-likelihood F-test in edgeR to calculate genes with DE, DT, and DTE. A meaningful effect size of 0.5 is used to determine biologically meaningful genes. All of the following results and figures represent output from riboSimR.

3.2   Simulation Results

Figure 1 examines how the power of our tests for DTE, DE and DT are affected by artificially altering the number of sample replicates and the sequencing depth. The x-axis represents the factor by which the empirical depth is multiplied. For DTE and DT, it seems as though, for a given replicate number, sequencing depth increases cease to yield increases in power at around 75% of the real dataset size. This indicates that increasing sequencing depth beyond this point would have diminishing returns.

However, the power to identify DE genes continues to increase as sequencing depth increases. For each type of test, it is clear that higher power can be achieved by increasing the number of sample replicates. This increase in power does not seem to have diminishing returns, as increases are observed in up to ten replicates. Our power to identify genes which show DE seems more largely reliant on both sequencing depth and replicate number than tests for DT and DTE.

To further quantify the differing effects of increasing the sequencing depth and replicate number, we can compare experiments in which the same number of total reads are added, but added by increasing only either depth or replicate number. Figure 2A shows the results of such an analysis, where the blue line indicates increasing the total number of reads by increasing depth, and the orange line indicates increasing the total number of reads by the same amount, but via increased replication. These results once again highlight that, for the Merchante dataset, power can only be increased by adding replicates.

We also investigated increasing replication and depth at differing levels in the two experiment types. The x-axis of Fig. 2B represents increasing levels of Ribo-seq replicates, while the different colored lines represent the different numbers of RNA-seq replicates. These results indicate that increasing the number of RNA-seq replicates generally has a larger positive effect on power than increasing Ribo-seq replicates. Figure 2C shows the result of a similar experiment, but increasing depth instead of replication. Here
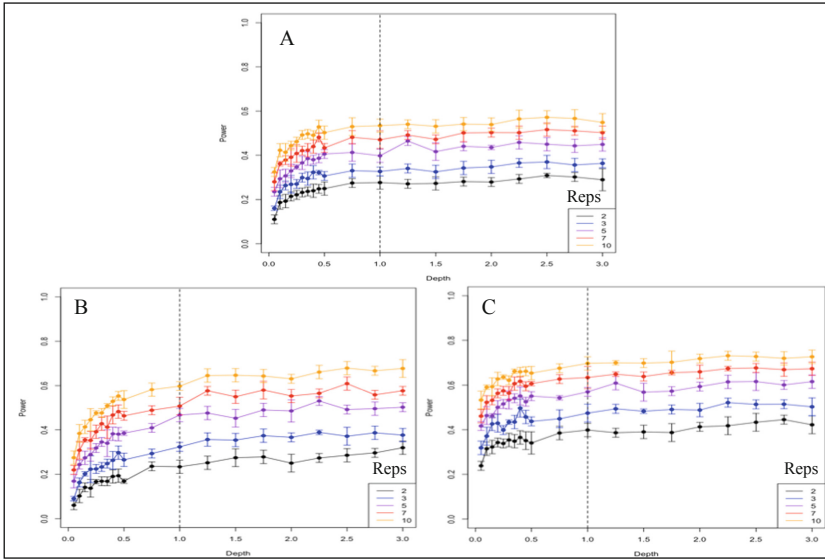
Fig. 1. Simulations based on Merchante et al. data for (A) DTE, (B) DE, and (C) DT. X-axis represents the factor by which depth is changed relative to the complete dataset. Different colored lines represent number of simulated replicates. (Color figure online)
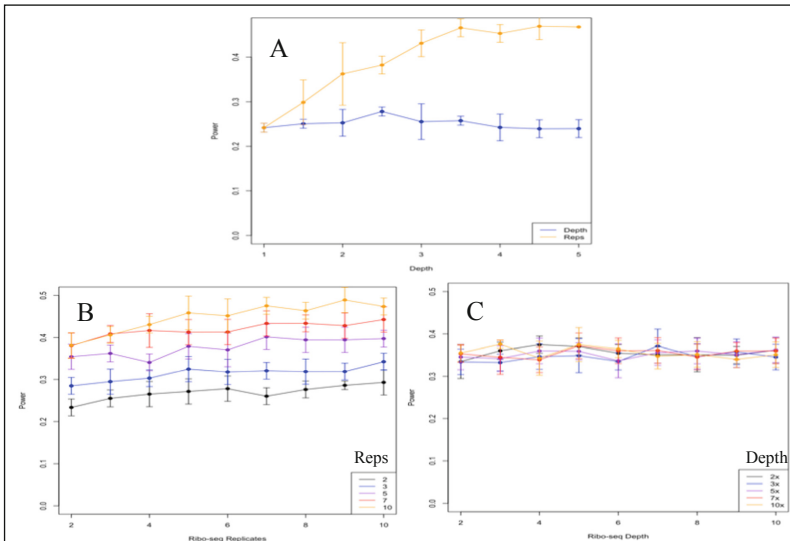

Fig. 2. (A) Value comparison between increasing reads via replications and sequencing depth. (B) Effect of increasing replication at differing levels across experiment types. (C) Effect of increasing sequencing depth at differing levels across experiment types. (Color figure online)

we are simulating the addition of reads to existing datasets, and therefore we begin simulating at 200% of empirical read levels. As previously noted, adding additional depth to the Merchante dataset has proven to have little effect on power, and these results indicate that addition of varying levels of RNA-seq or Ribo-seq have little effect on these results.

Figure 3 shows a quantification of the power to identify genes which show DTE, stratified by average RNA-seq read (3A) count and Ribo-seq read count (3B). These results indicate that, for a given replicate number, our power to detect genes with DTE is largely reliant on both the RNA-seq and Ribo-seq count for the genes. Our power to predict genes with lower than 50 average reads across replicates is significantly lower. For the RNA-seq reads, the difference in power between different replicate values is more pronounced in low read count genes, indicating that increasing the number of replicates can have a significant effect on our ability to find low read count genes which truly show DTE.
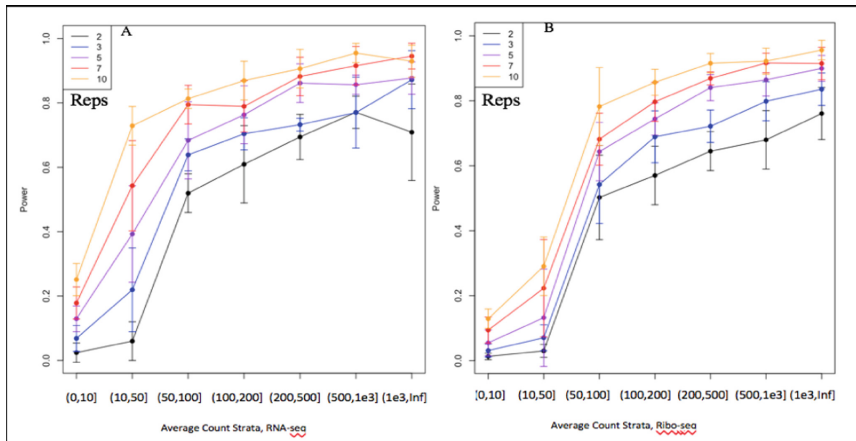

Fig. 3. Power stratification for DTE by average count value in (A) RNA-seq and (B) Ribo-seq. Simulated based on data from Hassan et al.

We also look to measure the effects that our choice of tools and parameters have when testing for differential genes. Figure 4 shows the use of four different tools, edgeR, DESeq 2, limma-voom, and DSS, for detection of DTE. The tools are evaluated by measuring power and FDR from simulations based on both the Merchante and Hassan datasets. The simulations are performed for both 2 and 5 replicates. The results in Fig. 4A and C indicate that DESeq 2 and edgeR generally achieve the highest power for each of the experimental setups, while limma-voom and DSS yield lower power. Figures 4B and D show that, in terms of FDC, limma-voom appears to outperform DESeq 2 and edgeR for both datasets, while DSS performs well for the Hassan data and poorly for the Merchante data. These results indicate that nuanced differences between program may effect results differently on an experiment-by-experiment basis. It is therefore valuable to have a way which we can evaluate each program for a user's specified dataset.
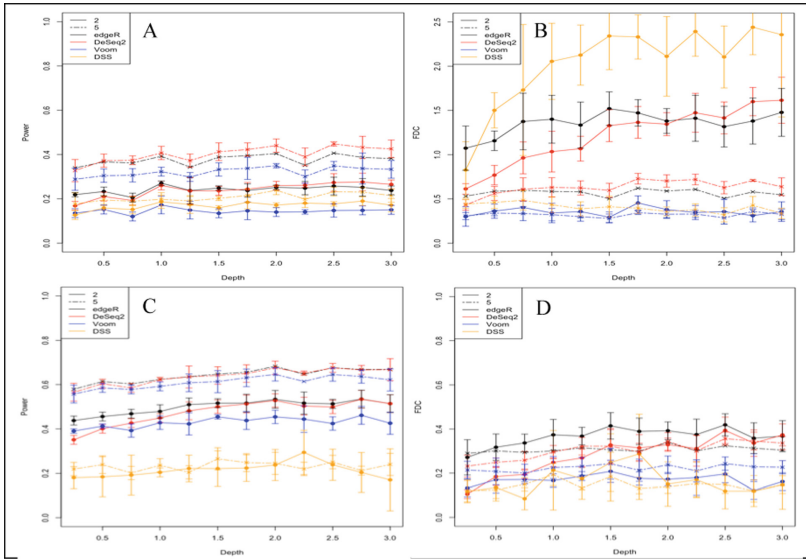
Fig. 4. Tool comparison results for power and FDC of (A, B) Merchante et al. data and (C, D) Hassan et al. data. Line types differentiate between using 2 and 5 replicates.

Figure 5 show the effects of three different FDR significance thresholds on power, true positive rate, and false discovery cost for the Merchante data. The simulations are performed for 1,5 and 8 replicates. As can be expected, the use of less stringent significance thresholds leads to a larger number of true positives and higher power, as the number of genes which pass the threshold increases. However, as can be seen in Fig. 5C, the false discovery cost when using less stringent thresholds can be significantly higher. In most cases, researchers should choose parameters values by balancing their ability to identify truly significant genes with the cost of making false discoveries. Interestingly, these results also depict a relationship between the significant threshold, replicate number, and FDC. The differences in power and FDC between the three FDR thresholds changes seems to be more drastic for two replicates than five and eight. This implies that the choice of FDR threshold has a larger effect on false discovery for experiments with a limited number of replicates.
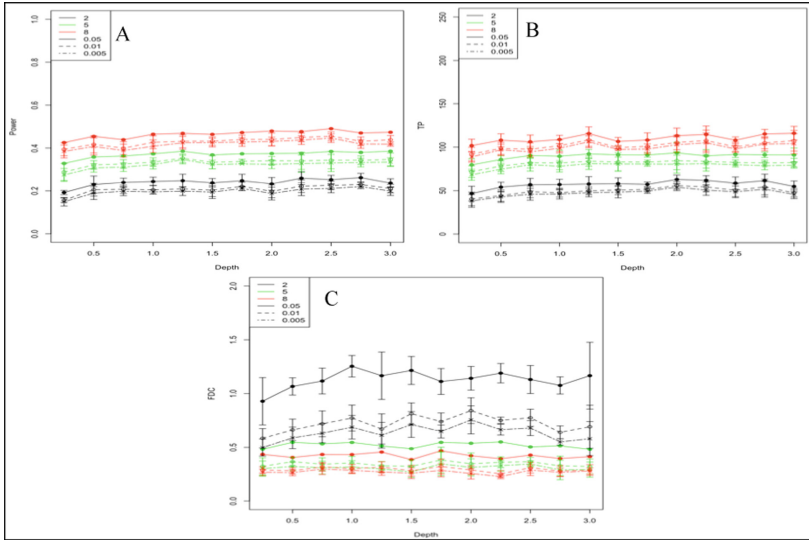
Fig. 5. The effect of using different FDR significance thresholds on (A) power, (B) true positive rate, and (C) false discovery cost. Colors represent differences in number of replicates. (Color figure online)

4   Conclusion

In this paper, we describe RiboSimR, an R Shiny web application for the analysis of experimental design parameters in Ribo-seq and RNA-seq experiments. This tool can be used prior to performing sequencing in order to test experimental design parameters using provided pilot data, or after the fact, to test the effects of adding additional sequencing information to an existing study. Using two published Ribo-seq and RNA-seq datasets, we demonstrate how RiboSimR can be used to investigate the effects of these various factors on the power to detect genes which exhibit differential expression, differential translation, and differential translation efficiency.

Firstly we showed that we can assess the effects that variations in replicate number and sequencing depth have on power to detect differential genes. We provided an example of an experiment which has reached a critical sequencing depth threshold in terms of identifying DTE and DT genes. For researchers, having the ability to predict the effects of increasing sequencing depth can save them valuable time and money. The results from our simulations also confirm the previously asserted notion that increasing sample replication yields larger increases in power than increasing sequencing depth.

We further looked and quantify the effects of increasing replication and depth differently between experiment types. This simulation setup can aid researchers in testing the experimental design parameters of an experiment they have already conducted, and help them predict how recreating the experiment with different factors, or adding additional information to their experiment, might affect their results.

We unexpectedly found that increasing the number of RNA-seq replicates has a larger positive effect on power than increasing Ribo-seq replicates. Additionally, RiboSimR can be used to analyze the effect that the magnitude of expression and translation have our ability to detect genes as differential. We confirmed that we have less power to detect genes with low counts, and found that increasing replication, especially for RNA-seq samples, had a strong influence on our ability to detect genes with less than 50 reads.

Finally, we tested the effects of more downstream parameters of differential analysis, including the choice of differential analysis tool and significance threshold. While it is not our goal to recommend a specific tool for identifying differential genes, we believe that a platform which allows the comparison of tools for individual experiments is of great use to researchers. We also highlight the importance of balancing power and false discovery cost when choosing tools and significance thresholds. These choices can have large effects that vary from experiment to experiment, and using an out-of-the-box tool or p-value threshold can lead to misleading results or loss of information. In conclusion, RiboSimR allows users to quantify the consequences of important experimental design choices at various points throughout the experimental process.

References

1. Ingolia, N., Ghaemmaghami, S., Newman, J., Weissman, J.: Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science **324**, 218–223 (2009)
2. Wang, Z., Gerstein, M., Snyder, M.: RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. **10**, 57 (2009)
3. Larsson, O., Sonenberg, N., Nadon, R.: Identification of differential translation in genome wide studies. Proc. Natl. Acad. Sci. **107**, 21487–21492 (2010)
4. Hansen, K., Wu, Z., Irizarry, R., Leek, J.: Sequencing technology does not eliminate biological variability. Nat. Biotechnol. **29**, 572 (2011)
5. Robles, J., Qureshi, S., Stephen, S., Wilson, S., Burden, C., Taylor, J.: Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing. BMC Genom. **13**, 484 (2012)
6. Anders, S., Huber, W.: Differential expression analysis for sequence count data. Genome Biol. **11**, R106 (2010)
7. Robinson, M., McCarthy, D., Smyth, G.: EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**, 139–140 (2010)
8. McCarthy, D., Smyth, G.: Testing significance relative to a fold-change threshold is a TREAT. Bioinformatics **25**, 765–771 (2009)
9. Wu, H., Wang, C., Wu, Z.: PROPER: comprehensive power evaluation for differential expression using RNA-seq. Bioinformatics **31**, 233–241 (2014)
10. Frazee, A., Jaffe, A., Langmead, B., Leek, J.: Polyester: simulating RNA-seq datasets with differential transcript expression. Bioinformatics **31**, 2778–2784 (2015)
11. Busby, M., Stewart, C., Miller, C., Grzeda, K., Marth, G.: Scotty: a web tool for designing RNA-Seq experiments to measure differential gene expression. Bioinformatics **29**, 656–657 (2013)
12. Eastman, G., Smircich, P., Sotelo-Silveira, J.: Following ribosome footprints to understand translation at a genome wide level. Comput. Struct. Biotechnol. J. **16**, 167–176 (2018)
13. Larsson, O., Sonenberg, N., Nadon, R.: Anota: Analysis of differential translation in genome-wide studies. Bioinformatics **27**, 1440–1441 (2011)

14. Olshen, A., Hsieh, A., Stumpf, C., Olshen, R., Ruggero, D., Taylor, B.: Assessing gene-level translational control from ribosome profiling. Bioinformatics **29**, 2995–3002 (2013)
15. Xiao, Z., Zou, Q., Liu, Y., Yang, X.: Genome-wide assessment of differential translations with ribosome profiling data. Nat. Commun. **7**, 11194 (2016)
16. Chang, W., Chang, J., Allaire, J., Xie, Y., McPherson, J.: Shiny: Web Application Framework for R. https://CRAN.R-project.org/package=shiny
17. Law, C., Chen, Y., Shi, W., Smyth, G.: Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. **15**, R29 (2014). https://doi.org/10.1186/gb-2014-15-2-r29
18. Wu, H., Wang, C., Wu, Z.: A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. Biostatistics **14**, 232–243 (2012)
19. Merchante, C., et al.: Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. Cell **163**, 684–697 (2015)
20. Hassan, M., Vasquez, J., Guo-Liang, C., Meissner, M., Siegel, T.: Comparative ribosome profiling uncovers a dominant role for translational control in Toxoplasma gondii. BMC Genom. **18**, 961 (2017)