

mRNABtab—a web application for mRNA stability analysis

Andrei Alic¹, José E. Pérez-Ortín², Joaquín Moreno² and Vicente Arnau^{1,*}¹Department of Informatics and ²Department of Biochemistry and Molecular Biology, Universitat de València, 46100 Burjassot, Spain

Associate Editor: Ivo Hofacker

ABSTRACT

Eukaryotic gene expression is regulated both at the transcription and the mRNA degradation levels. The implementation of functional genomics methods that allow the simultaneous measurement of transcription (TR) and degradation (DR) rates for thousands of mRNAs is a huge improvement in this field. One of the best established methods for mRNA stability determination is genomic run-on (GRO). It allows the measurement of DR, TR and mRNA levels during cell dynamic responses. Here, we offer a software package that provides improved algorithms for determination of mRNA stability during dynamic GRO experiments.

Availability and implementation: The program mRNABtab is freely accessible at <http://mRNABtab.uv.es/>. mRNABtab is written in C, PHP and R.

Contact: vicente.arnau@uv.es

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on October 17, 2012; revised on January 16, 2013; accepted on January 21, 2013

1 INTRODUCTION

mRNA is the messenger molecule transferring information from genes to proteins. It is by its own nature and function unstable (see Parker 2012; Pérez-Ortín *et al.*, 2007 for discussion). There is a variable turnover that depends on synthesis (transcription rate, TR) conducted by RNA polymerase II and degradation rate (DR) conducted by a complex system of RNases and other proteins or RNAs (Parker, 2012). The importance of TR in determining the levels of mRNAs is widely documented; however, the parallel importance of DR in it is only starting to be recognized (Pérez-Ortín *et al.*, 2007).

There are several methods currently available for DR determination. Some of them are able to measure mRNA stabilities genome-wide (see Chen *et al.*, 2008; Schwalb *et al.*, 2012 and Supplementary Data for a critical description and comparison). The most popular, although prone to mistakes and biases (Marín-Navarro *et al.*, 2011; Wang *et al.*, 2002), is the transcription shut-off and subsequent determination of mRNA decay: assuming first-order kinetics, the degradation constant (k_d) and the mRNA half-life (HL = $\ln 2/k_d$) can be calculated.

The TR and mRNA concentrations can be determined genome-wide by using a genomic run-on protocol (GRO), which allows to estimate the HLs of mRNAs even under non steady-state conditions (Fan *et al.*, 2012; García-Martínez *et al.*,

2004). This possibility is extremely interesting because it allows to determine TR, mRNA levels and HLs in a single experiment during the fast transcriptional response to an imposed environmental change (Castells-Roca *et al.*, 2011 and references therein). This is the simplest protocol currently available for monitoring mRNA turnover in dynamic situations. However, it involves complex numerical analysis and costly computing. Now we describe here the mRNABtab package that offers two alternative algorithms (programs 'StepK' and 'RampK') for determining mRNA HLs from data generated by GRO experiments. Additionally, it also includes a program ('Shutoff') to calculate k_d from simple decay experiments (Wang *et al.*, 2002). All three programs are devised to cope with genome-wide data [TR and/or (mRNA) time sequences for thousands of genes].

2 WHAT THE PROGRAMS DO

Typically, the GRO experiment delivers values of TR (TR_0 , TR_1 , $TR_2 \dots$) and mRNA concentration ($[mRNA]_0$, $[mRNA]_1$, $[mRNA]_2 \dots$) simultaneously determined at discrete times (t_0 , t_1 , $t_2 \dots$). By assuming a linear variation of TR for a particular gene between consecutive pointwise determinations (e.g. values of TR and $[mRNA]$ at t_0 and t_1) with slope p :

$$p = (TR_1 - TR_0)/(t_1 - t_0) \quad (1)$$

A mean value for the first-order degradation constant (k_d) throughout the considered time interval can be determined (Pérez-Ortín *et al.*, 2007; see Supplementary Data). This calculation is sequentially translated to subsequent intervals and afterwards to all genes. A program following this procedure has been already available for spreadsheets (Marín-Navarro *et al.*, 2011), but a faster version (in C) is now included in mRNABtab with the name of StepK.

An inconvenient feature of StepK is that mean values for k_d jump stepwise from one interval to the next. A more realistic change would be to suppose that k_d varies also linearly in between sample points as TR is assumed to do. Under these conditions, the rate of mRNA change for a gene would be

$$d[mRNA]/dt = TR_0 + p(t - t_0) - [(k_d)_0 + q(t - t_0)] [mRNA] \quad (2)$$

where q is the (unknown) rate of k_d change with time, assumed constant for the time interval under consideration. For the first interval (between t_0 and t_1), the initial value of k_d [i.e. $(k_d)_0$] is calculated supposing an initial steady state (as time 0 is usually taken before the onset of the environmental change triggering the

*To whom correspondence should be addressed.

transcriptional response) as TR_0 divided by $[mRNA]_0$. The RampK program integrates numerically Equation (2) for the considered time interval by Euler's method taking a random initial q value. This value is contrasted by checking the difference between the calculated and experimental value of mRNA concentration at the end of the interval and iteratively refined through a bisection algorithm. Once the value of q has been determined to a chosen precision, the final value of $(k_d)_1$ (i.e. the value of k_d at t_1) is calculated as:

$$(k_d)_1 = (k_d)_0 + q(t_1 - t_0) \quad (3)$$

and this is used as initial value of k_d for integrating the next interval (between t_1 and t_2). Similarly, the procedure is extended to subsequent intervals and further to all genes.

Therefore, under slightly different assumptions, both programs (StepK and RampK) deliver a time course for k_d from experimental determinations of TR and mRNA concentration sampled at a sequence of time points. StepK estimates a mean value of k_d for each interval, whereas RampK gives the k_d values at the interval extremes assuming a linear time course between them. Shutoff is an additional program included in this website to treat data of mRNA decay after transcriptional arrest (Wang *et al.*, 2002). Thus, data will be mRNA concentrations at different time points. Shutoff calculates k_d fitting these data to an exponential decay by minimizing square deviations.

3 STRUCTURE OF THE WEBSITE

We offer a novel implementation of a website holding the algorithms described previously. The user has the chance to work with a set of well-known tools, all gathered in a user-friendly web application. The algorithms run on our server. It introduces the notion of *Experiment*, which contains one or more *Operations*, including *Input*, *Output* and *Settings* as child notions for the latter.

The input file must have a TR and/or an RA set of columns, each one having the time in the header. If one of the types is missing, neither StepK nor RampK can run. Furthermore, the first column always contains the unique names of the genes analysed. The application consists of three layers: the UI, the algorithms and the storage, built using the Yii framework. It can be viewed as three-tier architecture (Fig. 1), using different languages at each level. The first one uses PHP with JavaScript/HTML/CSS. It controls the algorithms execution written in R (UPGMA, Sota, Kmeans—not implemented by us—and Shutoff), C (StepK and RampK) and PHP (Filtering). The last layer consists of the file system (storing of the inputs/outputs of each algorithm) and the database (store the successfully executed operations for a registered user).

4 USAGE AND APPLICATION

One can create an account or to use the application anonymously. Both cases offer the same features when it comes to capabilities. The difference between them lies in the possibility to save the experiments with their history, inputs and results when an account is used. A user must upload his/her file, and then

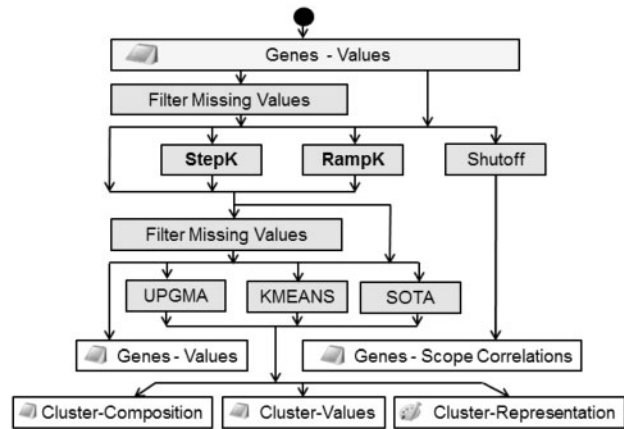


Fig. 1. Application working schema of the mRNAstab site

he/she can use the algorithms to process it. The results (visual or text) can be downloaded onto the user's computer. There is also an FAQ section where more details about the algorithms and the workflow can be found. This website was tested under Mozilla Firefox, Google Chrome and Microsoft Internet Explorer.

5 CONCLUSIONS

We offer a web application to analyse the stability of the genes from microarray data. It contains two new applications capable of estimating the degradation constant (k_d) and, therefore, the mRNA stability for a list of time points of TR and mRNA.

Funding: Spanish MCINN (BFU2010-21975-C03-01) and Regional Valencian Government (PROMETEO 2011/088) (to J.E.P.O.); MCEI-DGI (BFU2009-11965 to J.M.). Erasmus Grant offered by the Technical University of Cluj-Napoca, Romania (to A.A.).

Conflict of Interest: none declared.

REFERENCES

- Castells-Roca, J. *et al.* (2011) Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities. *PLoS One*, **6**, e17272.
- Chen, C.Y. *et al.* (2008) Messenger RNA half-life measurements in mammalian cells. *Methods Enzymol.*, **448**, 335–357.
- Fan, J. *et al.* (2012) Array-based nuclear run-on analysis. *Methods Mol. Biol.*, **809**, 505–517.
- García-Martínez, J. *et al.* (2004) Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol. Cell*, **15**, 303–313.
- Marín-Navarro, J. *et al.* (2011) Global estimation of mRNA stability in yeast. *Methods Mol. Biol.*, **734**, 3–23.
- Parker, R. (2012) RNA degradation in *Saccharomyces cerevisiae*. *Genetics*, **191**, 671–702.
- Pérez-Ortín, J.E. *et al.* (2007) Genomics and gene transcription kinetics in yeast. *Trends Genet.*, **23**, 250–257.
- Schwalb, B. *et al.* (2012) Measurement of genome-wide RNA synthesis and decay rates with Dynamic Transcriptome Analysis (DTA). *Bioinformatics*, **28**, 884–885.
- Wang, Y. *et al.* (2002) Precision and functional specificity in mRNA decay. *Proc. Natl Acad. Sci. USA*, **99**, 5860–5865.