Regulating the Many to Benefit the Few: Role of Weak Small RNA Targets

Daniel Jost, Andrzej Nowojewski, and Erel Levine*
Department of Physics and FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts

ABSTRACT   Small regulatory RNAs are central players in the regulation of many cellular processes across all kingdoms of life. Experiments in mouse and human have shown that a typical small RNA may regulate the expression of many different genes, suggesting that small RNAs act as global regulators. It is noted though that most targets respond only weakly to the presence of the small RNA. At the same time, evidence in bacteria and animals suggest that the phenotypes associated with small RNA mutants are only due to a few of their targets. Here we assume that targets regulated by a small RNA to control function is in fact small, and propose that the role of the many other weak targets is to confer robustness to the regulation of these few principal targets. Through mathematical modeling we show that auxiliary targets may significantly buffer both number and kinetic fluctuations of the principal targets, with only minor slowdown in the kinetics of response. Analysis of genomic data suggests that auxiliary targets experience a nonspecific evolutionary pressure, playing a role at the system level. Our work is of importance for studies on small RNA functions, and impacts on the understanding of small RNA evolution.

INTRODUCTION

Small RNA molecules play critical regulatory roles in many different organisms, from small regulatory RNAs regulating stress response in bacteria to microRNAs regulating development and homeostasis in mammals (1,2). Many small RNA families achieve target-specificity at the posttranscriptional level via basepairing of a very short (6–8 nucleotides) region with the targeted microRNA. Consequently, one predicts that a particular binding sequence could appear at random in multiple loci across the genome, which may explain why many genes carry binding sites in their 3′UTR. Sequence pairing, however, may not be sufficient for successful interaction, and one might expect that most of these putative sites are nonfunctional, and would be lost during evolution. Surprisingly, bioinformatic searches for microRNA targets that focus on highly conserved sites still predict a large number of target genes per microRNA (3). In addition, recent experiments in mouse (4) and in humans (5) demonstrate that the transfection of a single microRNA affects the expression of many genes that carry a seed match, although most of these genes respond to the microRNA very weakly. It is therefore reasonable to hypothesize that these many binding sites mediate real interactions and have some functionality within the cellular context. One tempting interpretation is that microRNAs act as global regulators, affecting directly genes across pathways or modules of the cellular network.

On the other hand, detailed studies focused on particular small regulatory RNAs (srRNAs) draw a very different picture. Studies in bacteria and animals (6–16) suggest that the phenotype associated with mutating srRNAs is due only to few of their targets. For example, the lethal phenotype associated with mutating the recognition sequence in the heterochronic microRNA let-7 of *Caenorhabditis elegans* (6,7) is rescued by a compensating mutation in only one gene (*lin-41*). Similarly, concurrent reexpression of only three targets of the human microRNA miR-31 (10) (out of the hundreds predicted) is enough to abrogate the metastasis suppression phenotype associated with overexpression of the microRNA. These and other examples suggest that only few targets are directly involved in the cellular response to the level of the srRNA.

As the functions of more and more srRNAs are being elucidated in a variety of organisms and tissues, the mystery remains: what is the role of the many weakly interacting conserved targets (17)? Recently, it has been proposed that the number of phenotypically relevant targets regulated by a srRNA is in fact small (18,19) and that the many other weak targets which we accordingly term auxiliary might be competitive inhibitors for the srRNA by preventing the srRNA binding to the few principal targets. In this article we propose that the role of these auxiliary targets is to confer robustness and suppress both number and kinetic fluctuations. Through dynamical analysis we show that auxiliary targets can strongly suppress intrinsic fluctuations in their target levels with only minor effect on the kinetics of response. We support these predicted functional role of auxiliary targets by analyzing their conservation across vertebrates, and show that auxiliary targets experience a nonspecific evolutionary pressure, suggesting that their role is at the system level. With small RNAs acting in development and response pathways, these features are expected to be of major physiological consequence.

CrossMark

MATERIALS AND METHODS

Unified model for srRNA regulation

General description of the model

In absence of auxiliary targets, the general picture of srRNA posttranscriptional regulation is well described by modeling the dynamics of the number of free srRNA s, of the number of principal target mRNAs m, and of the number of principal target proteins p, by a set of mass-action equations [20]:

$$\frac{ds}{dt} = \alpha_s - \beta_s s - ksm, \tag{1}$$

$$\frac{dm}{dt} = \alpha_m - \beta_m m - ksm, \tag{2}$$

$$\frac{dp}{dt} = \gamma m - \beta_p p. \tag{3}$$

Here α_s (α_m) refers to the transcriptional rate of srRNA (mRNA) and β_s (β_m) to its degradation or turnover rate. The value k represents the interaction rate between the srRNA and its target mRNA. The target proteins are produced at a rate γ per mRNA molecule and self-degrade at a rate β_p. Equations 1 and 2 assume that the pairing between the srRNA and the mRNA either leads to degradation of the complex or sequesters it (e.g., while blocking translation) for considerable time. Active codegradation is believed to occur for many prokaryotic small RNA-mRNA couples [1], while titration may be a more dominant mode of action of eukaryotic miRNAs [2]. Within the same framework, it is easy to generalize the model and allow a fraction of the srRNA to be recycled. This does not affect the main results of our work (see the Supporting Material).

We generalize this model by accounting for the interaction between the srRNA and the auxiliary targets, leading to the formation of a transient complex (Fig. 1). The kinetics of the number of free auxiliary target mRNAs n and of complexed srRNAs c follow the mass-action equations

$$\frac{dn}{dt} = \alpha_n - \beta_n n - k_a sn + k_d c, \tag{4}$$
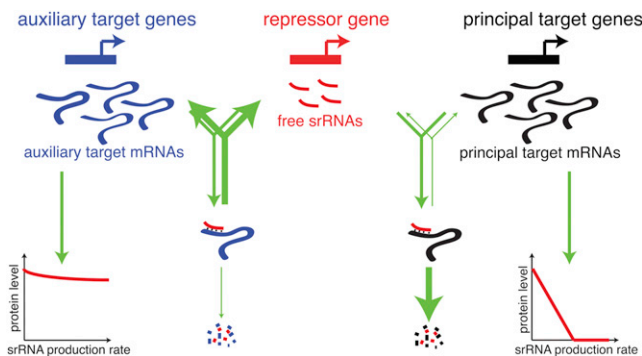

FIGURE 1 Interactions of a srRNA with its targets. In the pathway, the main difference between auxiliary and principal targets resides in the fate of the mRNA-srRNA complexes: for principal targets, the complexes are rapidly degraded leading to a strong effect on the overall principal protein level, whereas for auxiliary targets, the complexes preferentially dissociate, leading to only a weak effect on the auxiliary protein level.

$$\frac{dc}{dt} = -\beta_c c + k_a sn - k_d c, \tag{5}$$

where α_n is the total production rate over all auxiliary targets and β_n is their average turnover rate. The value k_a is the association rate between the srRNA and the auxiliary targets and k_d the dissociation rate of the transient complex which is degraded at a rate $\beta_c \ll k_d$. To account for the effect of the auxiliary targets on the srRNA level, one has to augment Eq. 1 by

$$[-k_a sn + k_d c + \beta_c (1 - p_d)c],$$

with p_d the probability that the srRNA is also eliminated during the complex degradation.

Accounting for stochasticity

Stochasticity of the underlying biochemical reactions, including the transcriptional burstiness [21,22] and the effect of transport by diffusion of the interacting molecules [23,24], is accounted for by augmenting the previous set of mass-action equations with Langevin terms [25,26] which captured the intrinsic fluctuations of each reaction (see the Supporting Material for a detailed description of the full stochastic model).

Analysis of steady-state properties

We analyzed the steady-state properties of srRNA regulation in absence or in presence of auxiliary targets within the limit of small noise, using the linear noise approximation [25]. The mean steady-state levels were estimated by solving the mass-action equations (Eqs. 1–5). Fluctuations were analyzed by solving the fluctuation-dissipation relation

$$JC + CJ^{\dagger} + N = 0, \tag{6}$$

with C as the covariance matrix of the system, J as the Jacobian of the set of mass-action equations (Eqs. 1–5), and N as a diagonal matrix whose entries are the amplitudes of the noise (Langevin) terms (see the Supporting Material).

Efficacy of the srRNA regulation and limit between slow and fast transport modes

As a robust measure of the efficacy of the srRNA regulation, we choose an information theory-based measure, which considers the interactions between a regulator and its target as a communication channel, and measure the information capacity of this channel [27,28]. The major advantage of this choice is that it does not require any specific knowledge of the function of the regulator or the precise signal it transduces. Moreover, this measure is insensitive to the particular form of the response function, and can therefore be applied across the entire range of parameters in our model.

In the limit of low noise, the information capacity for srRNA pathways can be approximated by (see the Supporting Material)

$$I_{\max} = \log_2 \left(\frac{1}{\sqrt{2\pi e}} \int d\alpha_s \left| \frac{d\langle p \rangle(\alpha_s)}{d\alpha_s} \right| \frac{1}{\sigma_p(\alpha_s)} \right), \tag{7}$$

where $\langle p \rangle$ and σ_p are, respectively, the steady-state mean and standard deviation of the principal target level. At a fixed set of parameters the impact of auxiliary targets on this information-based efficacy strongly depends on the mobility of the molecules: auxiliary targets increase the capacity of the channel if they diffuse rapidly, but decrease its capacity if diffusion is

slow. One can therefore define a critical diffusion rate that marks the boundary between these two behaviors (see the Supporting Material).

Analysis of kinetic properties

The kinetic response of the pathway was studied using stochastic Gillespie simulations (29), where, for simplicity, we neglect fluctuations in the total number of auxiliary targets. Each simulation starts with a configuration sampled from the steady-state distribution of principal targets (Poissonian distribution with mean α_m/β_m). At $t = 0$ the transcription of srRNA is switched on, and the kinetic response is assessed by measuring the distribution of the first-passage time when the number of principal mRNAs reaches zero.

Choice of parameter values

Typical values for parameters are only known for bacterial srRNA pathways (20,30). We choose to use these ranges of values (given in the Supporting Material) to plot our figures. However, Fig. S2 in the Supporting Material shows that the effects described in the article are robust over a wide range of parameter values that likely also include the eukaryotic pathways.

Analysis of genomic data

The conservation of microRNA targets was studied for a set of 87 microRNAs conserved among vertebrates (see Table S1 in the Supporting Material) and compared with an artificially generated set of 1000 mock-microRNAs (see Table S2). Seeds of the mock microRNAs were generated as 7-mers that obey the same dinucleotide statistics as the 87 conserved microRNAs in our sample (31). See the Supporting Material for details.

For each microRNA, a list of predicted targets in 12 vertebrates (*Homo sapiens*, *Pan troglodytes*, *Macaca mulatta*, *Mus musculus*, *Rattus norvegicus*, *Canis familiaris*, *Equus caballus*, *Bos taurus*, *Monodelphis domestica*, *Ornithorhynchus anatinus*, *Gallus gallus*, and *Xenopus tropicalis*) was obtained using the TargetScan algorithm and the aligned 3′UTRs dataset available at the TargetScan website (32). The number of targets $N_{i,s}$ per microRNA i in each species s was normalized by the total number of genes per species $N_{tot,s}$ in the alignment. Conservation of the number of targets was quantified by computing for each microRNA the relative fluctuations of $\{N_{i,s}/N_{tot,s}\}$ across species (see the Supporting Material).

Taking the list of human targets as reference, we represent TargetScan predictions by a binary array $x^t_{i,s}$, where $x^t_{i,s} = 1$ if gene t is a target of microRNA i in human and in species s, and 0 otherwise. The observed frequency f^t_i is then defined as the mean value of $x^t_{i,s}$ across species. The corresponding conservation score C^t_i is obtained by normalizing f^t_i by the number of species that carry a homolog of gene t (see the Supporting Material). For example, $C^t_i = 1$ means that gene t is a target of microRNA i in every species where t is present.

To investigate the conservation of principal and auxiliary targets, we use published experimental data that measured global proteome response to transfection of a microRNA (miR-1, miR-124, and miR-181) in (human) HeLa cells (4). For each transfected microRNA, we extract from the experimental dataset the protein level change of each of its target genes (as defined above), if it had been measured. Table S3 contains the list of these genes for miR-1 (382 genes), miR-124 (249 genes), and miR-181 (345 genes), as well as the corresponding fold-change in protein level (in $\log_2$ unit).

The distributions of fold-repression do not offer a natural separation between principal and auxiliary targets in HeLa cells. We therefore split the targets based on their fold-change into an subset of strong targets, representing targets with a fold-repression below a given arbitrary limit ($\log_2$ of the fold-change ≤ -0.6) and the complementary subset of weak targets,

We verified that our conclusions do not depend on the choice of this limit (see the Supporting Material).

RESULTS

Simple model for srRNAs with two classes of targets

Although it is still unclear which factors determine the efficiency of a given srRNA on a particular target, experiments on mammalian somatic tissues have revealed that most of the putative targets are only weakly affected by the transfection of a single microRNA, and only a small proportion is more significantly repressed (4,5). Below we identify the phenotypically relevant srRNA targets with this limited subset of strongly affected targets. This assumption may not always be justified, as in some cases small modulation may have strong physiological impact (see Table 2 in Seitz (18)), and conversely one could imagine an auxiliary role even for a strongly suppressed target (17,33) (see the Supporting Material).

Despite the fact that small RNA pathways differ in many details (including their biogenesis and mechanism of action), at an abstract level they can all be described by a unified model (20,26,30,34–38), which accounts for synthesis of all RNA species, interaction of the srRNA with its targets, and the consequential suppression of translation and/or promotion of degradation of the mRNA (and, perhaps, the small RNA molecule itself). This model can be translated into a simple mathematical framework (20) that we augment to describe the interactions with two classes of targets: the principal targets and the auxiliary targets (Fig. 1). Within a Langevin formalism, our model accounts for the stochastic nature of the underlying biochemical reactions, including the effect of transport by diffusion of the interacting molecules, by taking the limit of weak noise (for details, see Materials and Methods as well as the Supporting Material).

Auxiliary targets can finely tune the expression of principal targets

In absence of auxiliary targets, the effect of srRNAs on the mean level of principal targets falls into one of three categories (20), depending on the relative strength of the small RNA expression compared with that of the target (Fig. 2 A, *black line*). For efficient srRNA production (i.e., in cases where the synthesis rate of mature srRNAs α_s is much larger than that of mRNA, α_m), most of the mRNAs are targeted by the large srRNA pool, leading to complete silencing. Conversely, under conditions where the production of the srRNA is less efficient ($\alpha_s \ll \alpha_m$) the target is almost unaffected by the presence of the small RNA and is normally expressed. Intermediate scenarios, where the production rates of the srRNA and of the mRNA are comparable ($\alpha_s \sim \alpha_m$),
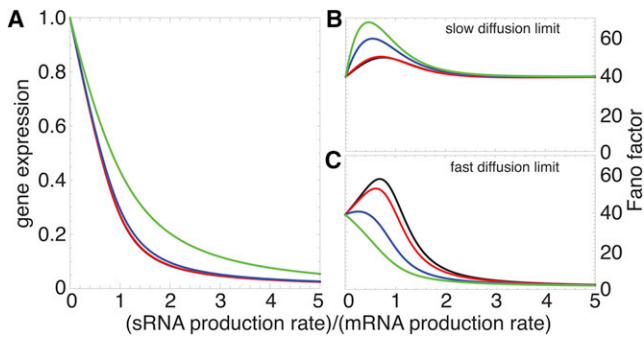
FIGURE 2   Steady-state properties of the principal targets in absence or presence of auxiliary targets. (*A*) Mean protein level of the principal targets as a function of the ratio α_s/α_m between the production rates of srRNA and principal mRNAs, and for different numbers n of auxiliary mRNA molecules (*black*, 0; *red*, 100; *blue*, 1000; *green*, 10,000). (*B* and *C*) Effect of the number of auxiliary targets on the noise/signal ratio (Fano factor) of the principal target proteins in the slow (*B*) and fast (*C*) transport limit.

allow fine-tuning of the target expression. Under these conditions, quantitative changes in the transcription rate of the srRNA correspond to quantitative changes in the target level. The sharpness of the transition between silenced and expressed regimes is controlled by the strength of the interaction between srRNA and mRNA. For strong interactions, the regulatory logic is a sharp linear-threshold response (see Fig. S1 *A*).

Additional targets may affect the level of srRNA available for regulating the principal targets in many ways, and one possibility is that auxiliary targets could significantly change this simple picture. At steady state, auxiliary targets can be interpreted as stoichiometric weak targets with an effective interaction constant $k_{eff} = k_a p_d (\beta_c/k_d)/(1 + \beta_c/k_d)$ (see the Supporting Material). The fact that the auxiliary targets are, by definition, weakly affected by the srRNA suggests that $k_{eff} \ll k$. The level of auxiliary targets would therefore play a measurable role on the principal one only at $k_{eff} \alpha_n \gtrsim k \alpha_m$ (20). In this situation, the auxiliary targets modify the steady-state level of the srRNA and then indirectly that of principal targets. This could help in finely tuning the position of the transition between the expressed and silent regimes (Fig. 2 *A*) but does not change the regulatory logic of the posttranscriptional regulation. Typical parameters (20,30) suggest that this effect starts to be significant for thousands of auxiliary target molecules.

This characteristic regulatory logic has been verified experimentally in vivo for bacterial small RNA pathways in *Escherichia coli* (20) and for the mammalian microRNA pathway in HeLa cells (38). These results allow us to characterize the effect of a small RNA on a target in terms of three qualitatively different operating regimes: the silencing regime, the tuning (or crossover) regime, and the expression regime. In this language, one may deduce that the effect of many mammalian microRNAs at their normal expression level lies within the tuning category. Indeed, on one hand,

overexpression experiments show at most a twofold repression compared to wild-type (4,5), suggesting that in normal tissues the targets are not completely silenced. On the other hand, knockdown of the same microRNA leads to, at most, a twofold upregulation (5), suggesting that targets are already somewhat suppressed in normal tissues. It is therefore possible that many microRNA-target pairs act in the crossover regime, which—as we discuss next—exposes them to amplified fluctuations.

Auxiliary targets reduce the intrinsic noise of the principal targets

The stochastic nature of the biochemical reactions composing gene regulation pathways leads to intrinsic fluctuations around the mean signal levels (39–41). A canonical way to appreciate the strength of protein fluctuations is to consider the Fano factor $\nu = \sigma^2_p/\langle p \rangle$, with σ^2_p the variance and $\langle p \rangle$ the mean level of the principal protein number p at steady-state. The value ν is a measure of the noise/signal ratio, which confers two advantages: in simple cases it does not depend on the transcription rate; and it allows comparison of the fluctuations in a particular system to a simple memoryless Poisson process, where $\nu = 1$. For example, in the absence of srRNA regulation, the noise/ signal ratio is mainly driven by the burstiness of the protein translation and $\nu \approx 1 + b$ (39) with $b = \gamma/\beta_m$ the protein burst size, which is the average number of protein translated from a single mRNA molecule.

Models of posttranscriptional regulation by srRNAs suggest that this mode of regulation is particularly effective in suppressing intrinsic fluctuations in the silenced regime (26,35) (Fig. 2 *C*, *black lines*). Indeed, in the limit of efficient transport, the effective lifetime of an active mRNA ($[\beta_m + k\langle s \rangle]^{-1}$) is dramatically reduced through the interaction with the srRNA and $\nu \approx 1 + b^*$ with $b^*/b = \langle p \rangle/\langle p \rangle_{max} = \beta_m /[\beta_m + k\langle s \rangle] \ll 1$ and $\langle p \rangle_{max} = (\gamma \alpha_m)/(\beta_m \beta_p)$, the expression level in absence of srRNA (see the Supporting Material). If transport is inefficient, the slow stochastic diffusion of a very low number of mRNA molecules can lead to relatively high fluctuations in the local mRNA concentration and diffusion noise may dominate (Fig. 2 *B*, *black lines*).

In the crossover regime, where synthesis of srRNA and mRNA are similarly efficient, the ultrasensitivity of the stoichiometric system leads to a high correlation between the abundance of the srRNA and that of its principal target (42). The cell state becomes broadly distributed and alternates between unrepressed and repressed states, yielding a large distribution for the protein level with a high noise/ signal ratio (26,35,42) (Fig. 2 *C*, *black lines*). These fluctuations are enhanced by strong RNA-RNA interaction (Fig. S1 *B*) and amplified by stochasticity in active gene copy number and transcriptional bursting (26,35). The presence of such large fluctuations deems small RNAs

unsuitable for fine-tuning gene expression, and has adverse impact on their suitability for patterning gene expression during development (43). As discussed above, it is likely that many mammalian microRNA-target pairs act within the tuning regime and are therefore exposed to these amplified fluctuations. In what follows we suggest that the role of auxiliary targets is to suppress these fluctuations, making the tuning regime operational.

Our key observation is that the presence of auxiliary targets serves to diminish the correlation between the free srRNA copy number and that of its principal target, thus reducing fluctuations significantly (Fig. 2 C). By following the propagation of fluctuations within the model, we find that large intrinsic fluctuations in the srRNA level are quickly absorbed by the dynamic equilibrium between free and complexed srRNAs. To appreciate the effect of auxiliary targets on the fluctuations of principal targets, we first simplify the model by assuming that auxiliary targets have no effect on the mean steady-state level of principal targets ($\beta_c/k_d \to 0$). We will return to this point at the end of this section.

Consider first the case of efficient transport. The presence of auxiliary targets has a weak impact on the steady-state level of free srRNA while creating a pool of complexed srRNA. This pool plays the role of particle reservoir that absorbs srRNA fluctuations (due, e.g., to transcriptional burstiness or ultrasensitivity of the interaction with the principal targets). In particular, for a high number of auxiliary targets, the noise/signal ratio of the principal targets is given by $\nu \approx 1 + b^*$ (see the Supporting Material). To appreciate this result we call the reader's attention to the fact that in the absence of auxiliary targets this expression is only valid in the repressed regime. The presence of auxiliary targets reduces the fluctuations for any level of the srRNA, even in the noise-sensitive crossover regime, thus permitting fine-tuning and better control.

Interaction between srRNAs and their targets occur in a complex environment in the cytoplasm, perhaps in dedicated bodies (44). We reason that the presence of auxiliary targets may alter the delivery of a srRNA to its principal targets and consequently increases fluctuations. We therefore depart from the limit of efficient diffusion, and study the model over a range of diffusion constants governing the srRNA mobility. As suspected, when transport is slower, the presence of auxiliary targets yields a notable increase in fluctuations (Fig. 2 B). Indeed, the stochastic diffusion of a small number of free srRNAs (compared with the high number of free auxiliary targets) impacts on the noisy formation of the srRNA-auxiliary target complexes, leading to an increase of the srRNA fluctuations which propagates into that of principal targets.

We conclude that auxiliary targets have two opposing effects on the fluctuations in the level of a principal targets: they buffer number fluctuations due, for example, to bursty transcription or chromatin fluctuations, while increasing

noise due to slow transport. Using the notions of efficacy and information capacity of the regulatory pathway (27) (see Materials and Methods and the Supporting Material), we show that the competition between these two effects is mainly settled by comparing the diffusibility of the RNA molecules and their half-lives (see Fig. 3 and Fig. S2). Based on these two parameters (which may be srRNA- and condition-specific), the principal targets of a srRNA may benefit or suffer from the presence of auxiliary targets. Interestingly, the boundary between these two situations does not depend significantly on the microscopic details of the various interactions in the model. For bursty promoters, the buffering of fluctuations is even stronger and auxiliary targets help in maintaining a low noise level even for slow transport.

Which regime characterizes the behavior of naturally occurring small RNAs? It has long been assumed (and in some cases verified experimentally) that srRNAs are stable (45,46), and that they are actively mobile (46). We therefore expected srRNA to be in the fast diffusion-slow decay regime, which is favorable for auxiliary targets. Indeed, estimating typical parameters for small RNAs in bacteria, yeast, and metazoan taken from the BioNumbers website (47) (see Table 1) suggests that they all sit safely in the region where auxiliary targets serve to suppress fluctuations. This observation can explain the ability of natural microRNAs present in somatic tissues to regulate their principal targets with low noise, even in the crossover regime.

To complete our discussion of noise attenuation we comment that the general picture presented above is not changed when allowing the srRNA-auxiliary target complexes to have a short but finite lifetime (see the Supporting Material). However, in this case, the presence
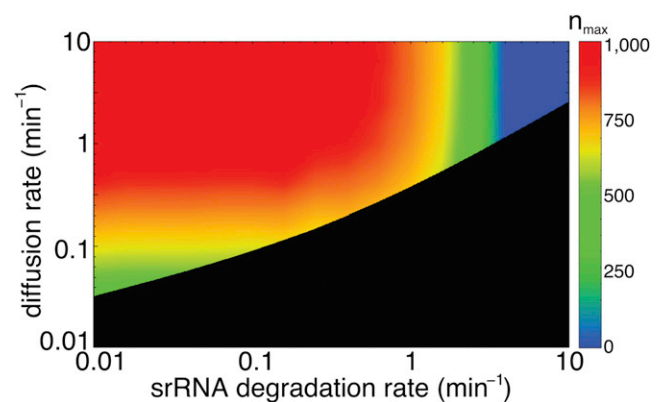

FIGURE 3 Impact of the srRNA degradation rate on the transition between fast and slow diffusion limit in the presence of auxiliary targets. In the fast transport region (*colored zone*), auxiliary targets help in maintaining a low noise level for the principal targets. In the slow transport region (*black zone*), diffusion noise dominates and the presence of auxiliary targets increases the internal fluctuations of the principal targets. (*Colors*) Upper bound of the number of auxiliary targets n_{max} imposed by limiting the slow-down in temporal response (see text).

TABLE 1   Diffusion mode in different organisms

	Escherichia coli	Budding yeast	HeLa cells
D ($\mu m^2\ s^{-1}$)	5	60	100
Ω (μm^3)	0.7	33	2000
β (min^{-1})	1/10	1/20	1/600
β/d	0.05	0.09	0.11

Estimation of typical parameters for different organisms. The transition between fast and slow diffusion modes is mainly controlled by the ratio between the RNA self-degradation rate β and the diffusion rate $d \equiv D\ell/\Omega$ with D the effective diffusion constant, Ω the volume of the cell (or of the cytoplasm), and $\ell \sim 5$ nm the length-scale of the reaction volume (see the Supporting Material). A small (high) ratio means fast (slow) transport mode. Numbers are taken from the BioNumbers database (47).

of auxiliary targets may affect the efficacy of the regulation. Indeed, under these conditions auxiliary targets effectively reduce the numbers of srRNA that are presented to their principal targets. In principle, one could have used this feature to impose a bound on the number of allowed auxiliary targets. However, as discussed in the next section, auxiliary targets act also as kinetic traps for the srRNAs and slow down the regulation process. As it turns out, this imposes an even stronger bound on the number of auxiliary targets.

Auxiliary targets maintain a robust kinetic response

In response to changes in environmental signals or development transitions, cells may need to adapt by stopping the expressions of specific proteins. Posttranscriptional regulation has been suggested as a mechanism to accelerate the response of gene expression to changes in the input signals (26,30). Indeed, in the absence of auxiliary targets newly produced srRNAs are directly used for depleting the pool of principal mRNA targets. However, it is natural to assume that auxiliary targets could act as kinetic traps and decelerate the adaptation of principal targets to signals that affect the srRNA production. The number of auxiliary targets exceeds significantly that of the principal target (4,5), and one may wonder if this would not result in an impossible hindrance of cellular response.

To address this question, we compare the kinetic response of two simple scenarios. In Scenario I, the srRNA bears sole responsibility for suppression of the target. Transcription of the mRNA is unaffected while the production of the srRNA is turned on at a high rate ($\alpha_s \gg \alpha_m$), thus switching from the expressed to the repressed regime of Fig. 2 A. In Scenario II, which is srRNA-free, transcription of mRNA is stopped (e.g., by binding of transcription factors (TF)) and the remaining mRNAs and proteins are allowed to self-degrade. As we just mentioned, in the absence of auxiliary targets the response time is much faster in the srRNA-based Scenario I than in the TF-based Scenario II (Fig. 4, *green arrow*), and the question is whether this effect is
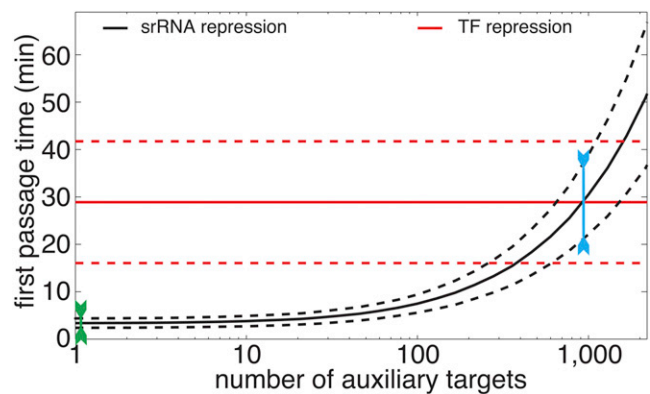

FIGURE 4   Impact of auxiliary targets on the temporal response of the principal targets. The temporal response is quantified by the first passage time, when the number of principal target mRNAs reaches zero. (*Solid lines*) Mean first passage time for the srRNA-based Scenario I and the TF-based Scenario II, as a function of the number of auxiliary targets. (*Dashed lines*) Standard deviation from the mean.

wiped out by the presence of auxiliary targets. To answer this question we extend the analysis by measuring the statistics of the first passage time when the number of principal target mRNAs reaches zero, using exact stochastic simulations (29) (for details, see the Supporting Material).

Our results suggest that the presence of auxiliary targets does not represent an overwhelming limitation to the reactivity of principal targets. As expected, the mean first passage time increases as the number of auxiliary targets is increased, imposing an upper bound on that number. To quantify it, we define n_{max} as the number of auxiliary target molecules that slow down the response time such that it is comparable with the response time in the TF-based Scenario II. Fig. 3 depicts this number for typical sets of parameters. We find that in a significant part of the parameter space this number can be very large (1000 auxiliary target molecules and more).

What determines the magnitude of n_{max} are the characteristics of the interaction between auxiliary targets and the srRNA. In particular, the rates of association and dissociation of srRNAs from their auxiliary targets turn out to be crucial for setting an upper bound on the number of auxiliary targets allowed without dramatically affecting the response time (Fig. 5 A). For example, if the srRNA-auxiliary target complex is very unstable (high dissociation constant $K_d \equiv k_d/k_a$), the effect of these targets on the response time is weak, and even a high number of targets can be present with no slow-down in response. On the other hand, if this complex is stable (low K_d), the dynamics is strongly affected even by a small number of auxiliary targets. Typical numbers for the association and dissociation rates ($k_a \approx k$ and $K_d \approx 10$–100 (30)) suggest that the number of auxiliary target molecules should be ~10^3–10^4 to significantly slow down the kinetic response of the principal targets.
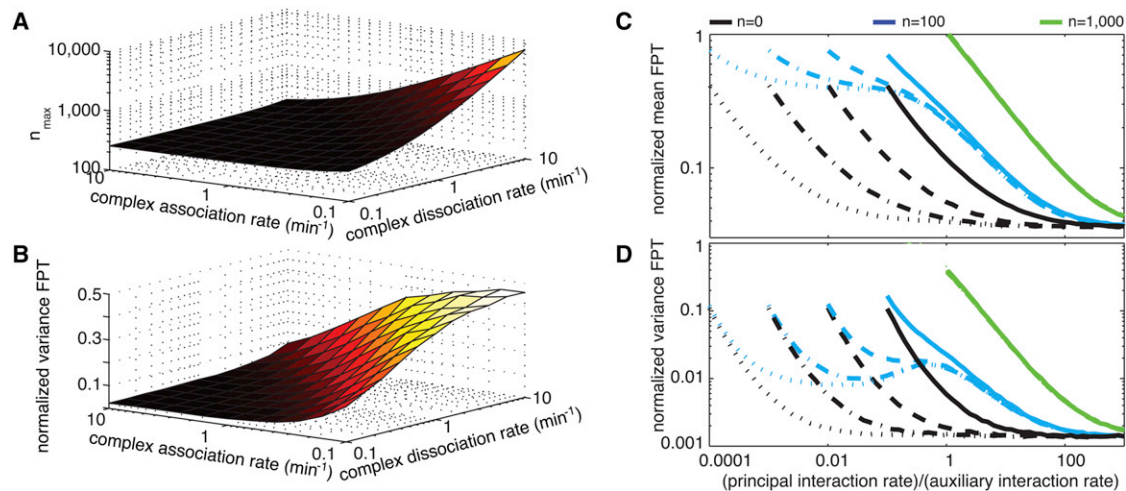
FIGURE 5 Impact of kinetic parameters on the efficiency of the temporal response of the posttranscriptional repression. (*A* and *B*) Number n_{max} of auxiliary targets (*A*) at which the mean time required for full suppression of the principal target by the srRNAs is equal to the corresponding time required by TF-based regulation, and the uncertainty in the response time (*B*) of the posttranscriptional regulation for $n = n_{max}$ (normalized by the corresponding value for the transcriptional regulation). (*C* and *D*) Mean and variance of the first passage time as a function of the ratio k/k_a between the interaction rate of the principal targets and the one of the auxiliary targets, for different numbers of auxiliary target molecules n and different values of the dissociation constant for the formation of the srRNA-auxiliary target complexes $K_d = 10$ (*full line*), 1 (*dashed line*), 0.1 (*dotted and dashed line*), and 0.01 (*dotted line*). Results are normalized by the corresponding values for the TF repression.

The statistics of first passage times allows us to consider another aspect of fluctuations that is typically ignored. In a fluctuating environment like the cell, the time at which all mRNA molecules initially present are degraded, bears some uncertainty: this time can be longer in one cell than in another. Looking back at the two scenarios described above, we notice that in the absence of auxiliary targets not only is the response time significantly reduced, but the uncertainty as well (Fig. 4, *green arrow*). This is yet another aspect of the robustness that small RNAs confer with their regulatory logic. Once again, one may be worried that the presence of auxiliary targets might reverse this outcome, and make the response time not only longer but also significantly less predictable.

Our model shows that this is not the case (Figs. 4 and 5 *B*). For example, consider the extreme case where the number of auxiliary targets is n_{max}, such that the mean response time is the same as in the TF-based Scenario II (Fig. 4, *blue arrow*). We find that the variance of this first passage time is still significantly smaller (at least twofold) in Scenario I than in Scenario II. Indeed, over the relevant range of parameters, and for any number of auxiliary targets below n_{max}, the uncertainty in the response time in Scenario I is significantly reduced compared to Scenario II.

Buffering of environmental noise by auxiliary targets

The cellular regulatory network is responsive to environmental changes, such that a change in the environment should lead to a robust response in gene expression. At the same time it is possible that a transient change in the environment should be ignored. We have previously suggested that small RNAs offer an inherent mechanism to buffer transient environmental changes that would result in the activation of a target gene (35). In this mechanism, the accumulation of target mRNAs is delayed with respect to the activation of transcription. Accumulation of the mRNA only starts after the pool of free srRNAs, which exists when the mRNA transcription is reactivated, is depleted. The size of the srRNA pool in the repressed state dictates the time window during which the changes in environmental signals are buffered.

Here we propose that the existence of auxiliary targets offers a similar mechanism in the opposite direction, i.e., when a change in environmental signals should result in suppression of the target (Scenario I above). We observe that under certain choices of parameters, the auxiliary targets provide the same kind of buffer for the srRNAs (Fig. 5, *C* and *D*). When the formation of auxiliary complexes is kinetically favored compared to the formation of principal complexes, srRNAs are first drawn to the pool of free auxiliary targets. Only after a significant fraction of these have been integrated into complexes can the srRNAs efficiently interact with the principal targets and silence their expression. This introduces a delay in the response time, followed by an abrupt repression of the principal targets. Both parts of this kinetic process are relatively noise-free, and together do not take longer than the more continuous (but slower) degradation occurring in Scenario II. This highlights a potential role for the auxiliary targets to delay the onset of the cell response without significantly affecting its efficiency.

Not surprisingly, the buffering effect exists when the rate of complex formation with auxiliary targets k_a is somewhat larger than the corresponding rate for principal targets k. The relevant range of parameters can be identified by observing the plateaus in the mean and the variance of the response time when plotted as a function of k. Such a range can be found for intermediate values of the association rate k_a and of the number of auxiliary targets n, such that $(k_a n)$ is comparable with k, but no fine-tuning is required (Fig. 5, C and D, cyan lines).

Nonspecific evolutionary pressure on auxiliary targets

In our model, the precise identity of the auxiliary targets is not significant for their role in suppression of noise. Indeed, the only requirement is that these targets would not be sensitive to a minor downregulation by their ineffective interactions with the srRNA, and that they would be coexpressed (at least partially) with the principal targets. The one important property for the functionality of the auxiliary targets is at the system level: the number of auxiliary targets should be large enough to create a noise-suppressing reservoir, and on the other hand not too large to titrate the srRNA away from its principal targets or to significantly slow down their temporal response. A prediction of this model is that the evolutionary force to maintain the interaction between auxiliary targets and the srRNA is nonspecific, and permits replacement of one target by another. We therefore ask if one could observe this evolutionary signature in available genomic data.

We focus on the conservation of microRNA targets across vertebrates for a set of 87 conserved microRNAs (see Table S1). Using the bioinformatic target prediction tool TargetScan (32), we generated a list of predicted targets for each microRNA in a set of 12 genomes. As a control, we employed the same procedure for a set of 1000 mock-microRNAs that were generated to mimic real microRNAs in many ways, but carry a random seed (see Table S2).

To explore the conservation of target identities, we focused on microRNA-target pairs in human. For each such pair we computed the probability that the target gene interacts with the microRNA in the other 11 species. Fig. 6 B shows the cumulative distribution of conservation scores. As expected, predicted targets of real microRNAs are slightly but significantly more conserved than those of mock microRNAs, suggesting a selection pressure acting on the interaction between natural microRNAs and their targets. Importantly, we observe that the total number of targets of each microRNA (but not their identity) remains strongly conserved among the 12 species (Fig. 6 A).

To classify the predicted targets as principal and auxiliary, we used published experimental data following the change in the proteome in human HeLa cells after transfection of three different microRNAs (4) (see Table S3).
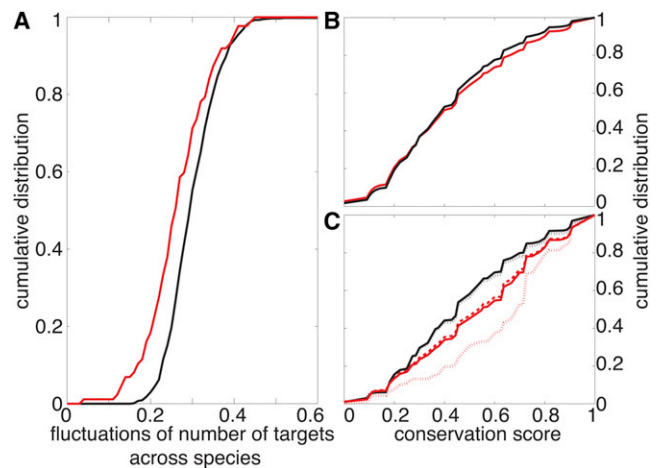

FIGURE 6 Conservation of microRNA targets across vertebrates. (A) Cumulative distributions for the relative fluctuations of the number of targets per microRNA across species for natural (red) and fake (black) microRNAs. Differences between the two distributions are significant (p-value $< 10^{-4}$ for the Kolmogorov-Smirnov test). Number of targets is more conserved for natural microRNAs. (B) Cumulative distributions for the conservation scores of predicted targets of natural (red) and fake (black) microRNAs. Differences between the two distributions are small but significant (p-value $< 10^{-10}$, Kolmogorov-Smirnov test). Proportion of conserved targets is slightly more important for natural microRNAs. (C) Cumulative distributions for the conservation scores of a subset of genes that are weak (red dashed line) or strong (red dotted line) targets of three natural microRNAs (miR-1, miR-124, and miR-181) in human (see Materials and Methods). (Red solid line) Distribution for the merged ensemble (weak and strong). (Black lines) Corresponding cumulative distributions for the same genes but when associated with fake microRNAs. Differences between natural and random microRNAs are always significant (p-values $< 10^{-10}$, Kolmogorov-Smirnov test). Proportion of conserved targets is larger for strong targets of natural microRNAs, then for weak targets of natural microRNAs, then for targets of fake microRNAs.

Targets that were strongly repressed (based on some arbitrary threshold) were classified as principal targets, whereas the many predicted targets that showed little or no response were classified as auxiliary. We note that in this approximation we do not use—or have—any information about the physiological significance of the different targets.

As observed previously (48), our data suggests that targets in the principal group are more conserved than those in the auxiliary group. Indeed, the data of Fig. 6 C, comparing the conservation of these two subsets of targets to those of mock microRNA, supports this claim. This reflects the strong evolutionary pressure that acts to conserve the interactions between microRNAs and their principal targets.

Taken together, we find that the number of targets for each microRNA is highly conserved, and that for principal targets this probably comes from the fact that the targets themselves are conserved. For the auxiliary targets, however, this target-number conservation is stronger than the conservation of target identity, supporting the notion

that it is their collective effect that is important, not their individual regulation.

DISCUSSION

Single microRNAs have been shown to regulate the expression of many different target genes (hundreds if not thousands), although most of them very weakly (4,5). This seems to be a universal feature, applicable to many small RNA pathways (including those acting in bacteria) that rely on a very short seed for specificity. These findings stimulated the contemporary view in the field, that small RNAs act as global modulators of gene expression at a network scale, rather than gene-by-gene.

On the other hand, in all cases we are aware of, where the functionality of a small RNA has been carefully studied, only a small number of target genes were identified as being phenotypically relevant (6–15). If this is the rule rather than the exception, one is led to ask what is the function of the many other weak targets of a srRNA. One possibility is that seed matching between most of the targets and a srRNA are completely coincidental and neutral (49). However, it has been reported that most of these interactions are evolutionarily conserved, suggesting possible functionality (3,18).

Our results support an elegant viewpoint of the relationships among srRNA targets, initially formulated by Seitz (18), which is unique to srRNA regulation, but not limited to a particular pathway or organism. Even if the precise molecular mechanisms that control the actual strength of the srRNA regulation on a specific target remain unclear (50), evidence suggest a hierarchy among srRNA targets. Principal targets are the ones behind the physiological role of the srRNA. Examples of such targets include the cancer-related proteins Suz12 for miR-200 (13), ANP32A and SMARCA4 for miR-21 (11), or integrin-α5, radixin, and RhoA for miR-31 (10). In contrast, auxiliary targets are not functionally regulated by the srRNA. In addition to the sponge-like effect (51) which allows a fine-tuning of the expression of principal targets (18), our model suggests functional roles for auxiliary targets in conferring robustness to the regulation and the kinetics of the principal targets. At an affordable price of slowing down the (already accelerated) reactivity of regulation, auxiliary targets significantly reduce fluctuations in the level of the principal targets, while maintaining the sensitivity of the regulatory logic. This observation is robust over the large range of in vivo parameters and does not depend on the details of the model. The buffer effect induced by the presence of auxiliary targets may not be limited to srRNA regulation and has also been suggested as a functional role for the many decoy binding sites of transcription factors along a genome (52).

These features may explain the wide spread of srRNAs in development and stress response pathways (1,2) where a precise but reactive response is needed. Specificity of each posttranscriptional pathway may have governed the evolution of the number of auxiliary targets for a given srRNA to find a compromise between the loss in reactivity and the gain in fidelity.

What are the auxiliary targets? The emerging picture suggests that auxiliary targets should be coexpressed with the principal targets of a srRNA. For a srRNA with multiple principal targets that are not coexpressed, we expect a set of auxiliary targets associated with each principal target. Any gene that share a pattern of expression with a principal target can be recruited as an auxiliary target, because the expression of these genes is hardly affected by the srRNA. Our results indicate that the exact number of auxiliary targets and their exact level of expression is of little consequence.

The role of most srRNA targets as auxiliary targets predicts that the evolutionary pressure in place to maintain the interaction between a srRNA and its auxiliary targets is nonspecific (18), and cannot be explained in evolutionary models that account for targets one at a time (53). The pressure to maintain auxiliary targets should also be contingent on the conservation of the principal targets. This type of evolutionary pressure should have unique fingerprints, and we expect it to be detectable as more whole-genome sequences become available.

SUPPORTING MATERIAL

REFERENCES

1. Gottesman, S. 2005. Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet.* 21:399–404.

2. Ghildiyal, M., and P. D. Zamore. 2009. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* 10:94–108.

3. Thomas, M., J. Lieberman, and A. Lal. 2010. Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* 17:1169–1174.

4. Baek, D., J. Villén, …, D. P. Bartel. 2008. The impact of microRNAs on protein output. *Nature.* 455:64–71.

5. Selbach, M., B. Schwanhäusser, …, N. Rajewsky. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature.* 455: 58–63.

6. Reinhart, B. J., F. J. Slack, …, G. Ruvkun. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans.* *Nature.* 403:901–906.

7. Slack, F. J., M. Basson, …, G. Ruvkun. 2000. The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol. Cell.* 5:659–669.

8. Massé, E., C. K. Vanderpool, and S. Gottesman. 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli.* *J. Bacteriol.* 187:6962–6971.

9. Liu, K., Y. Liu, …, R. He. 2011. MiR-124 regulates early neurogenesis in the optic vesicle and forebrain, targeting NeuroD1. *Nucleic Acids Res.* 39:2869–2879.

10. Valastyan, S., N. Benaich, …, R. A. Weinberg. 2009. Concomitant suppression of three target genes can explain the impact of a microRNA on metastasis. *Genes Dev.* 23:2592–2597.

11. Schramedei, K., N. Mörbt, …, K. Brocke-Heidrich. 2011. MicroRNA-21 targets tumor suppressor genes ANP32A and SMARCA4. *Oncogene.* 30:2975–2985.

12. Hyun, S., J. H. Lee, …, V. N. Kim. 2009. Conserved MicroRNA miR-8/miR-200 and its target USH/FOG2 control growth by regulating PI3K. *Cell.* 139:1096–1108.

13. Iliopoulos, D., M. Lindahl-Allen, …, K. Struhl. 2010. Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Mol. Cell.* 39:761–772.

14. Melton, C., R. L. Judson, and R. Blelloch. 2010. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature.* 463:621–626.

15. Zhao, C., G. Sun, …, Y. Shi. 2010. MicroRNA let-7b regulates neural stem cell proliferation and differentiation by targeting nuclear receptor TLX signaling. *Proc. Natl. Acad. Sci. USA.* 107:1876–1881.

16. Mavrakis, K. J., A. L. Wolfe, …, H. G. Wendel. 2010. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukemia. *Nat. Cell Biol.* 12:372–379.

17. Flynt, A. S., and E. C. Lai. 2008. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.* 9:831–842.

18. Seitz, H. 2009. Redefining microRNA targets. *Curr. Biol.* 19:870–873.

19. Swami, M. 2010. Small RNAs: pseudogenes act as microRNA decoys. *Nat. Rev. Genet.* 11:530–531.

20. Levine, E., Z. Zhang, …, T. Hwa. 2007. Quantitative characteristics of gene regulation by small RNA. *PLoS Biol.* 5:e229.

21. Chubb, J. R., T. Trcek, …, R. H. Singer. 2006. Transcriptional pulsing of a developmental gene. *Curr. Biol.* 16:1018–1025.

22. So, L.-H., A. Ghosh, …, I. Golding. 2011. General properties of transcriptional time series in *Escherichia coli*. *Nat. Genet.* 43:554–560.

23. Berg, H. C., and E. M. Purcell. 1977. Physics of chemoreception. *Biophys. J.* 20:193–219.

24. Bialek, W., and S. Setayeshgar. 2005. Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. USA.* 102:10040–10045.

25. van Kampen, N. 2001. Stochastic Processes in Physics and Chemistry. North-Holland, Amsterdam, The Netherlands.

26. Mehta, P., S. Goyal, and N. S. Wingreen. 2008. A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol. Syst. Biol.* 4:221.

27. Tkacik, G., C. G. Callan, Jr., and W. Bialek. 2008. Information flow and optimization in transcriptional regulation. *Proc. Natl. Acad. Sci. USA.* 105:12265–12270.

28. Cover, T., and J. Thomas. 1991. Elements of Information Theory. John Wiley and Sons, New York.

29. Gillespie, D. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81:2340–2361.

30. Shimoni, Y., G. Friedlander, …, H. Margalit. 2007. Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol. Syst. Biol.* 3:138.

31. Gentles, A. J., and S. Karlin. 2001. Genome-scale compositional comparisons in eukaryotes. *Genome Res.* 11:540–546.

32. Lewis, B. P., C. B. Burge, and D. P. Bartel. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell.* 120:15–20.

33. Bartel, D. P., and C.-Z. Chen. 2004. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* 5:396–400.

34. Mitarai, N., A. M. C. Andersson, …, K. Sneppen. 2007. Efficient degradation and expression prioritization with small RNAs. *Phys. Biol.* 4:164–171.

35. Levine, E., M. Huang, …, T. Hwa. 2008. On noise and silence in gene regulation by small RNA. *http://matisse.ucsd.edu/hwa/pub/sRNA-noise.pdf*.

36. Mitarai, N., J.-A. M. Benjamin, …, K. Sneppen. 2009. Dynamic features of gene expression control by small regulatory RNAs. *Proc. Natl. Acad. Sci. USA.* 106:10655–10659.

37. Jost, D., A. Nowojewski, and E. Levine. 2011. Small RNA biology is systems biology. *BMB Rep.* 44:11–21.

38. Mukherji, S., M. S. Ebert, …, A. van Oudenaarden. 2011. MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.* 43: 854–859.

39. Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA.* 98:8614–8619.

40. Elowitz, M. B., A. J. Levine, …, P. S. Swain. 2002. Stochastic gene expression in a single cell. *Science.* 297:1183–1186.

41. Paulsson, J. 2004. Summing up the noise in gene networks. *Nature.* 427:415–418.

42. Elf, J., J. Paulsson, …, M. Ehrenberg. 2003. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.* 84:154–170.

43. Levine, E., P. McHale, and H. Levine. 2007. Small regulatory RNAs may sharpen spatial expression patterns. *PLOS Comput. Biol.* 3:e233.

44. Goldie, B. J., and M. J. Cairns. 2012. Post-transcriptional trafficking and regulation of neuronal gene expression. *Mol. Neurobiol.* 45: 99–108.

45. Le Derout, J., I. V. Boni, …, E. Hajnsdorf. 2010. Hfq affects mRNA levels independently of degradation. *BMC Mol. Biol.* 11:17.

46. Krol, J., I. Loedige, and W. Filipowicz. 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* 11:597–610.

47. Milo, R., P. Jorgensen, …, M. Springer. 2010. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* 38(Database issue):D750–D753.

48. Friedman, R. C., K. K.-H. Farh, …, D. P. Bartel. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19: 92–105.

49. Chen, K., and N. Rajewsky. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* 8:93–103.

50. Grimson, A., K. K.-H. Farh, …, D. P. Bartel. 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell.* 27:91–105.

51. Ebert, M. S., and P. A. Sharp. 2010. Emerging roles for natural microRNA sponges. *Curr. Biol.* 20:R858–R861.

52. Burger, A., A. M. Walczak, and P. G. Wolynes. 2010. Abduction and asylum in the lives of transcription factors. *Proc. Natl. Acad. Sci. USA.* 107:4016–4021.

53. Shomron, N., D. Golan, and E. Hornstein. 2009. An evolutionary perspective of animal microRNAs and their targets. *J. Biomed. Biotechnol.* 2009:594738.

54. Berg, O. G. 1978. On diffusion-controlled dissociation. *Chem. Phys.* 31:47–57.

55. Hao, Y., L. Xu, and H. Shi. 2011. Theoretical analysis of catalytic-sRNA-mediated gene silencing. *J. Mol. Biol.* 406:195–204.

56. Walczak, A. M., G. Tkacik, and W. Bialek. 2010. Optimizing information flow in small genetic networks. II. Feed-forward interactions. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 81:041905.

57. Mugler, A., B. Grinshpun, …, C. H. Wiggins. 2011. Statistical method for revealing form-function relations in biological networks. *Proc. Natl. Acad. Sci. USA.* 108:446–451.