Bioenergetic costs and the evolution of noise regulation by microRNAs

Efe Ilker[1] and Michael Hinczewski[2]

[1] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[2] *Department of Physics, Case Western Reserve University, Cleveland OH 44106*

Noise control, together with other regulatory functions facilitated by microRNAs (miRNAs), is believed to have played important roles in the evolution of multicellular eukaryotic organisms. miRNAs can dampen protein fluctuations via enhanced degradation of mRNAs, but this requires compensation by increased mRNA transcription to maintain the same expression levels. The overall mechanism is metabolically expensive, leading to questions about how it might have evolved in the first place. We develop a stochastic model of miRNA noise regulation, coupled with a detailed analysis of the associated metabolic costs. Additionally we calculate binding free energies for a range of miRNA seeds, the short sequences which govern target recognition. We argue that natural selection may have fine-tuned the Michaelis-Menten constant K_{M} describing miRNA-mRNA affinity. K_{M} is constrained by seed length, and we show that optimal noise control (minimum protein variance at a given energy cost) is achievable for seeds of 6-7 nucleotides in length, the most commonly observed types. Moreover, at optimality the degree of noise reduction approaches the theoretical bound set by the Wiener-Kolmogorov linear filter. The results illustrate how selective pressure toward energy efficiency has potentially shaped a crucial regulatory pathway in eukaryotes.

Nonequilibrium processes within living systems exact a high price: the constant maintenance of fuel molecules and raw materials at sufficient concentrations to provide thermodynamic driving potentials for biological function [1]. Optimizing that function with respect to thermodynamic costs is a factor constraining evolution, and would have been particularly important at the very earliest stages of life where the metabolic chemistry responsible for maintaining those potentials was necessarily primitive and relatively inefficient. Yet thermodynamic costs are not the only factor that matters, and biology is full of counter-intuitively complex chemical mechanisms whose evolutionary predecessors, perhaps arising out of the randomness of genetic drift, may have consumed energy resources without any clear fitness benefit.

The discovery of microRNAs (miRNAs) along with their counterparts, small non-coding RNAs, raised many open questions about their functional purposes and evolution [2, 3]. These short endogenous RNAs, around 22 nucleotides (nt) in length, exist in many eukaryotic cells. They constitute the core of the RNA-induced silencing complex (RISC) that interacts with target messenger RNA (mRNA), leading to translational repression and the accelerated degradation of their target by a mechanism known as the RNA interference. One possible functional role for this interference, which is the focus of our work, is fine-tuning noise in protein populations by reducing the variance of protein copy numbers [4, 5], conferring robustness to cellular functions [6]. Such noise control, together with other regulatory functions facilitated by microRNAs, is believed to have played important roles in the evolution of complex multi-cellular life [7–10]. Yet it is a considerable expenditure of resources, similar to setting up a factory production line for a valuable good, funding gangs of thieves to constantly raid the factory, and compensating for losses by increasing the production rate. *So how would such a regulation scheme arise, and has evolution actually optimized it?* Using a combination of statistical physics, information theory, biochemistry, and population genetics, we arrive at some tentative answers to these questions.

Our results give insights into one peculiar feature of this system: why the job requires such short RNA molecules—the significance of the "micro" in microRNA. Specific interactions between the miRNA and its target in fact largely depend on only a 6-8 nt sequence known as the miRNA seed region, which forms Watson-Crick base pairs with a complementary sequence on the target mRNA. We argue that the observed length of the seed region lies in a metabolic sweet spot, giving just enough affinity between the miRNA and mRNA (as measured by their Michaelis-Menten constant K_{M}) to optimally control noise at a given level of energetic expenditure. Longer seed sequences (higher affinity, smaller K_{M}) would increase rather than decrease noise. On the other hand, shorter seed sequences (lower affinity, larger K_{M}), while allowing interactions with a wider range of targets, would also require significantly more energy to achieve the same level of control. By estimating these energy expenditures, we also show that effective noise control is costly enough to be under selection pressure in eukaryotic cells. A novel miRNA may initially appear during the course of evolution as a random product of genetic drift with non-optimal parameters, and then get gradually repurposed as a noise control mechanism that confers a fitness advantage. Once this starts to occur, our theory predicts that natural selection would hone K_{M} toward an optimal value.

miRNA regulation as a noise filter for protein expression. We start our theoretical description by introducing the architecture of the miRNA-regulated system

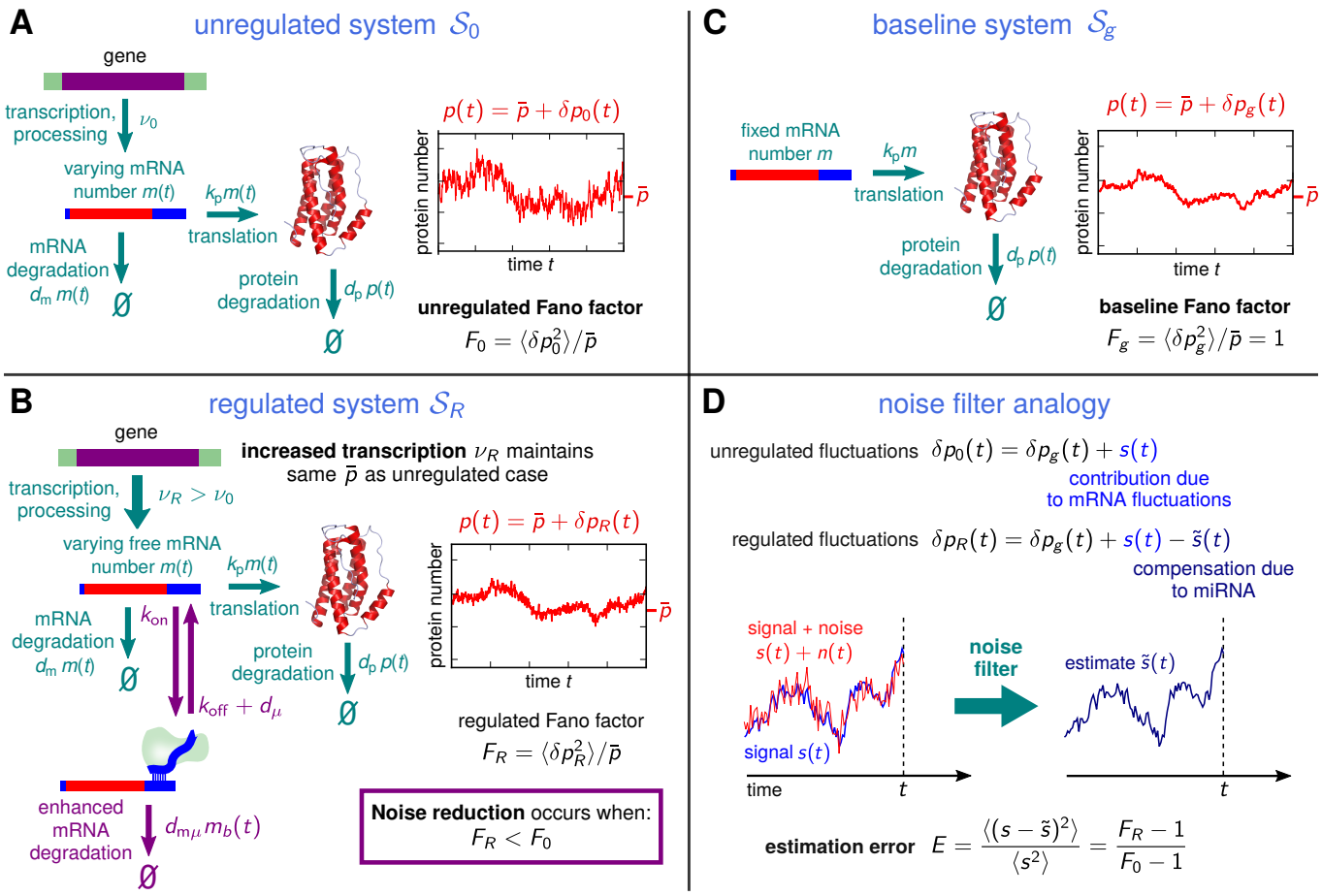

FIG. 1. **Overview of the miRNA noise regulation model:** (**A**) In the unregulated system $\mathcal{S}_0$ with no miRNA, noise in protein numbers has contributions from both varying mRNA population and intrinsic noise in the translation process, leading to a Fano factor F_0 (protein variance divided by the mean protein number $\bar{p}$). (**B**) For the regulated system $\mathcal{S}_R$, enhanced mRNA degradation due to targeting by miRNA can reduce protein noise, leading to a Fano factor $F_R < F_0$. To compensate for the loss of mRNA and achieve the same $\bar{p}$, the transcription rate ν_R must be increased relative to its value ν_0 in $\mathcal{S}_0$. (**C**) To set up the noise filter analogy, we introduce an imaginary basline system $\mathcal{S}_g$, which has fixed mRNA population and hence the protein noise is solely due to translation. The resulting Fano factor $F_g = 1$ is a lower bound on the two systems above: $1 \leq F_0$, F_R. (**D**) Decomposing the unregulated/regulated protein fluctuations $\delta p_0(t)$ and $\delta p_R(t)$ into a baseline contribution $\delta p_g(t)$ and an additional contribution allows us to define the signal $s(t)$ and estimate $\tilde{s}(t)$ in the noise filter analogy. Their normalized mean-squared difference E is the estimation error of the filter, which can be expressed in terms of F_R and F_0.

$\mathcal{S}_R$ (Fig. 1B) in comparison to the unregulated one $\mathcal{S}_0$ (Fig. 1A). Our description builds off the experimentally-validated model of Ref. [5], and takes the form of stochastic biochemical reaction networks with dynamics obeying linearized chemical Langevin (CL) equations [11, 12], as detailed in Supplementary Information (SI) Sec. S.I A. The CL results agree closely with available experimental data [5], and give us analytical expressions for correlation functions describing copy number fluctuations of chemical species in each system.

The fluctuating output numbers are denoted as $x(t) = \bar{x} + \delta x(t)$ where $\delta x(t)$ is the variation from the mean level $\bar{x}$. The main species of interest are total (bound + unbound) miRNA, mRNA and protein copy numbers, denoted by $x = \mu_{\text{tot}}$, m or p respectively. In the unregulated system $\mathcal{S}_0$ (Fig. 1A) we have transcription from a gene at rate ν_0 producing an mRNA population $m(t)$, which is then translated at rate k_p to a produce a protein population $p(t)$. The mRNA and proteins are degraded at rates d_m and d_p respectively.

To meaningfully compare $\mathcal{S}_R$ to $\mathcal{S}_0$, we assume parameter sets such that the mean protein output $\bar{p}$ is the same for both regulated and unregulated systems. This maintains the functional effectiveness of the protein in the regulated system, but with the potential added benefit of noise reduction. Indeed miRNA are often upregulated along with their target mRNA via feedforward loops [13, 14]. For $\mathcal{S}_R$ (Fig. 1B) there is an additional RNA interference mechanism: free miRNAs with population $\mu(t)$ can bind to the mRNA with rate k_{on} to form a bound complex with population $m_b(t)$. Considering a single miRNA binding site on the target mRNA, the

total number of miRNA is $\mu_{\rm tot}(t) = \mu(t) + m_b(t)$. The mRNA in the complex has an enhanced degradation rate $d_{m\mu} > d_m$ relative to the regular mRNA rate d_m. The miRNA unbinds with rate $k_{\rm off}$, and degrades with rate d_μ. For simplicity we assume the miRNA degradation rate d_μ is the same when both bound and unbound to mRNA [5]. miRNA-mRNA affinity can be characterized through the Michaelis-Menten constant,

$$K_{\rm M} = K_{\rm D} + k_{\rm cat}/k_{\rm on}, \tag{1}$$

where the dissociation constant $K_{\rm D} = k_{\rm off}/k_{\rm on}$ and $k_{\rm cat} = d_{m\mu} + d_\mu$ is an effective catalytic rate for the miRNA-catalyzed degradation reaction. K_M approximately relates $m_b(t)$ to $\mu_{\rm tot}(t)$ through $m_b(t) \approx m(t)\mu_{\rm tot}(t)/(K_M + m(t))$. Note that experimental values of $K_{\rm M}$ and $K_{\rm D}$ are reported in units of concentration (molars) while our CL formalism uses copy numbers of chemical species. We assume a typical eukaryotic cell volume $V = 2000\ \mu{\rm m}^3$ to convert between concentrations and copy numbers as needed.

Because of interference from the miRNA, the transcription rate ν_R of mRNA in $\mathcal{S}_R$ must be larger than the rate ν_0 in $\mathcal{S}_0$, if both systems maintain the same mean protein level $\bar{p}$. In the limit of no miRNA ($\mu_{\rm tot} \to 0$) or vanishing affinity ($K_{\rm M} \to \infty$) the regulated system approaches the same behavior as the unregulated one, with $\nu_R \to \nu_0$. The strength of regulation can be characterized by a parameter $R \equiv 1 - \nu_0/\nu_R$, where R ranges between 0 (no regulation) to 1 (maximum possible regulation). For known miRNA-mediated regulation networks, R typically lies between 0.05-0.95 [15, 16].

To quantify the effect of miRNA on noise, we look at the Fano factor $F_\lambda = \langle(\delta p_\lambda)^2\rangle/\bar{p}$ with $\lambda = R, 0$ labeling the system $\mathcal{S}_\lambda$ in which the quantity is calculated. Successful noise reduction implies that $F_R < F_0$: for the same mean protein output level, there is less protein variance in the presence of miRNA. Interestingly, miRNA actually increases fluctuations in mRNA as a side effect, but these fluctuations have less of an influence on protein numbers, since sudden jumps in the mRNA population are targeted by the miRNA. To understand the role of miRNA more precisely, it is helpful to use a noise filter analogy.

In order to motivate this mathematical analogy, we define an imaginary baseline system $\mathcal{S}_g$ (Fig. 1C), where we have fixed the mRNA population at a constant level $m(t) = \bar{m} = (d_p/k_p)\bar{p}$, which agrees with the mean $\bar{m}$ in $\mathcal{S}_R$ and $\mathcal{S}_0$ and hence gives the same $\bar{p}$. This removes the contribution of mRNA fluctuations to the noise, so $p(t) = \bar{p} + \delta p_g(t)$, where $\delta p_g(t)$ are the ground-level (baseline) fluctuations that come from protein translation and degradation (and cannot be mitigated by RNA interference). As summarized in Fig. 1D, the protein fluctuations in both $\mathcal{S}_0$ and $\mathcal{S}_R$ can then be compared to this baseline. For $\mathcal{S}_0$ we write $\delta p_0(t) = \delta p_g(t) + s(t)$, where $s(t)$ represents the added noise due to varying mRNA population $m(t)$. For $\mathcal{S}_R$ this added noise is partially compensated, $\delta p_R(t) = \delta p_g(t) + s(t) - \tilde{s}(t)$. Both $s(t)$ and $\tilde{s}(t)$ can be explicitly calculated using the CL formalism (SI Sec. S.II) and it turns out they are correlated: $\tilde{s}(t)$ takes the form of a convolution,

$$\tilde{s}(t) = \int_{-\infty}^{t} dt' H(t-t')(s(t') + n(t')), \tag{2}$$

where $H(t)$ and $n(t)$ are functions of the biochemical parameters. The above equation has a clear noise filter interpretation (depicted schematically in Fig. 1D): $s(t)$ is the "signal", $n(t)$ a "noise" that corrupts the signal, $H(t)$ is a linear filter function that acts via convolution on the past history of the corrupted signal $s(t) + n(t)$, and $\tilde{s}(t)$ is the "estimate" of the signal. It turns out the problem of reducing protein noise via mRNA interference (making F_R as small as possible) is equivalent to making $\tilde{s}(t)$ as close as possible to $s(t)$. We can see this directly in the error of estimation, which is defined as $E = \langle(s - \tilde{s})^2\rangle/\langle s^2\rangle$. For our case, E can be expressed in terms of the Fano factors, $E = (F_R - 1)/(F_0 - 1)$. Fine-tuning miRNA parameters only affects F_R, leaving F_0 fixed, so E can be minimized by decreasing F_R. Both F_R and F_0 are bounded from below by the Fano factor $F_g = 1$ of the baseline system, so perfect filtering, $E \to 0$, would correspond to $F_R \to 1$. The noise filter interpretation, which has been earlier applied to a variety of other biological networks (for a review see Ref. [17]), has an important payoff which we will return to later: it allows us to find the conditions for optimal noise reduction, and calculate tighter bounds on F_R, which in general will be greater than 1 at optimality.

Bioenergetic costs of miRNA regulation. The second major component of our model is an estimate of the costs for miRNA regulation, which we adapt from experimental data on eukaryotic transcription energetics collected in Ref. [18]. In general this includes energy expenditures channeled to the synthesis of new molecules, as well as maintenance, the recycling/repair of molecules to maintain steady-state levels (i.e. assembling mRNA from existing nucleotides to counterbalance degradation). Eukaryotic cells typically have long enough generation (cell division) times t_r, that maintenance is the dominant contribution to metabolic expenditures over a generation. Focusing on the maintenance costs, we can estimate the total metabolic consumption C_T (in units of phosphate [P] bonds hydrolyzed, namely ATP or ATP-equivalents) for the unregulated system per generation: $C_T \simeq t_r M_\nu$, where $M_\nu = \nu_0 \epsilon_m$ is the consumption rate. Here ν_0 is the mRNA transcription rate in $\mathcal{S}_0$, and ϵ_m is the energy cost in terms of P for assembling the mRNA (which will depend on the length of the transcript). For the regulated case, the total cost will be $C_T + \delta C_T$, where the extra contribution $\delta C_T = t_r \Delta M_\nu$. The difference in consumption rate $\Delta M_\nu = (\nu_R - \nu_0)\epsilon_m + \nu_\mu \epsilon_\mu$. The first term

accounts for the added costs of increased transcription to maintain the same $\bar{p}$, while the second term is the rate of miRNA assembly ν_μ times the cost ϵ_μ of that assembly in units of P (including potentially any related costs of the RISC complex). As shown in SI Sec. S.I C, ΔM_ν can be expressed in terms of the biochemical parameters of the system as

$$\Delta M_\nu = \frac{R}{1-R}\left(1+\gamma_\mu\left(1+\frac{K_{\rm M}}{\bar{m}}\right)\sigma_\epsilon\right)M_\nu, \quad (3)$$

with $\gamma_\mu \equiv d_\mu/d_{m\mu}$ and $\sigma_\epsilon \equiv \epsilon_\mu/\epsilon_m$. Based on experimental estimates (summarized in SI Table S1) we know that $\gamma_\mu \ll 1$ and $\sigma_\epsilon \ll 1$. Given the experimental range of $R = 0.05-0.95$, the term $R/(1-R)$ can vary by a factor of 361 between the smallest and largest observed regulation magnitudes, highlighting the strong dependence of ΔM_ν on R. The remaining terms in Eq. (3) encapsulate the modification to the costs due the parameters governing miRNA-mRNA interactions, particularly $K_{\rm M}$ and the degradation enhancement $d_{m\mu}$.

As discussed later in the section on evolutionary pressure, it is convenient to define a non-dimensional measure for the extra cost due to regulation: the extra cost as a fraction of the total metabolic expenditure of a cell per generation, $\delta C_T/C_T$. Here $C_T \simeq t_r M_{\rm tot}$, where $M_{\rm tot}$ is the total maintenance ATP consumption rate. Based on the data from Ref. [18], $M_{\rm tot}$ approximately scales with cell volume, and we use a value $M_{\rm tot} = 3\times10^{11}$ P/hr characteristic of eukaryotic cells. Thus we will report costs in terms of $\delta C_T/C_T = \Delta M_\nu/M_{\rm tot}$, with ΔM_ν given by Eq. (3).

Our model so far has assumed one mRNA target, but in general a single miRNA can target up to hundreds of mRNAs, which will also change the energetic costs of regulation. As a rough estimate of this multi-target scenario, we can assume similar biochemical parameters among different targets. This allows us to use the single-target theory but scaling up $\bar{m}$ (and hence $\bar{p}$) to reflect the total mRNA numbers when accounting for all targets. Note that the dependence of ΔM_ν on $\bar{m}$ is non-trivial, since both R and M_ν depend on $\bar{m}$. However when we demand a certain level of overall noise control (a specific value of E) the extra cost ΔM_ν will increase with $\bar{m}$ as the number of targets gets larger: the larger the system, the more expensive it is to control.

Seeds of length $6-7$ nt are most energetically efficient for noise reduction. With all the components of the model defined, we can now investigate the interplay between noise reduction and energetic costs. The error E (or equivalently the Fano factor F_R) can be expressed as a function of ΔM_ν, γ_μ, σ_ϵ, $\bar{m}$ and K_M (details in the SI Secs. S.I B,C). For a given cost ΔM_ν and fixed degradation/energy parameters γ_μ and σ_ϵ, we can ask what value of $K_{\rm M}$ minimizes E. $K_{\rm M}$ is an interesting tuning parameter because it is related to the binding strength between the miRNA and the mRNA, which in turn depends on the number of complementary interactions between the seed and the target region of the mRNA. From Eq. (1), $K_{\rm M} \geq K_{\rm D}$, and the dissociation constant $K_{\rm D}$ is related to the free energy of binding via $\Delta G = k_B T \ln(K_{\rm D}/[1{\rm M}])$. Longer seeds should allow for more negative ΔG (stronger binding) and hence smaller values of $K_{\rm D}$. This in turn gives access to smaller values of $K_{\rm M}$. For RNA interference systems, experimentally measured $K_{\rm M}$ values have ranged from comparable to $K_{\rm D}$ to about two orders of magnitude larger than $K_{\rm D}$ [19].

In Fig. 2A, we show the contour diagram of $\log_{10} E$ as a function of $K_{\rm M}$ and fractional metabolic cost $\delta C_T/C_T$, assuming a single mRNA target and using the experimentally-derived parameters of SI Table S1. The blue curve denotes the optimal value $K^*_{\rm M}$, which achieves the minimum E for a given cost $\delta C_T/C_T$. The red contour line marks the value $K^d_{\rm M}$ where $E = 1$, which corresponds to $F_R = F_0$. This is the boundary between the noise control region to the right, where $E < 1$ ($F_R < F_0$) and a "dud" region to the left, where $E > 1$ ($F_R > F_0$). In the latter region regulation adds protein noise to the system rather than mitigating it. For a fixed $\delta C_T/C_T$, as we scan K_M from small to large values, we cross from dud to noise control at $K^d_{\rm M}$, improve the filter performance until we reach $K^*_{\rm M}$, and then get progressively worse filtering for $K_M > K^*_{\rm M}$. Thus extremely strong affinities between miRNA and mRNA (small K_M) can be counterproductive and increase protein variance, and extremely weak affinities are ineffective at dampening the noise. K^*_M represents an optimal middle point that balances these two effects.

In the biologically relevant parameter regime γ_μ, σ_ϵ, $\phi \ll 1$, where $\phi \equiv d_p/d_m$, we can derive analytical approximations for both $K^d_{\rm M}$ amd $K^*_{\rm M}$. For small costs $\delta C_T/C_T$ we have $K^d_{\rm M} \approx \bar{m}(1+\gamma_\mu^{-1})$, and with increasing cost the boundary begins to decrease as $K^d_{\rm M} \approx \bar{m}((1-R)/\gamma_\mu)^{1/2}$. On the other hand, $K^*_{\rm M}$ is remarkably stable as $\delta C_T/C_T$ is varied. In the large cost limit it approaches

$$K^*_{\rm M} \approx \bar{m}\gamma_\mu^{-1}\sigma_\epsilon^{-1/2}. \quad (4)$$

Thus the optimal affinity depends on two non-dimensional biochemical ratios, γ_μ and σ_ϵ, and the mean mRNA target number $\bar{m}$.

Fig. 2B shows what happens when we scale up $\bar{m}$ by a factor of 100, roughly mimicking the case of 100 similar mRNA targets. As predicted by Eq. (4), $K^*_{\rm M}$ increases by a factor of 100, but the contour levels are pushed up by a similar factor: as expected, it costs more achieve the same level of noise control when compared to the single-target case. Interestingly, the $K^*_{\rm M}$ values in the single and multi-target case (about 0.65 nM and 65 nM respectively at high $\delta C_T/C_T$), are comparable to $K_{\rm M}$ measured in a fruit fly RNA interference pathway [19].

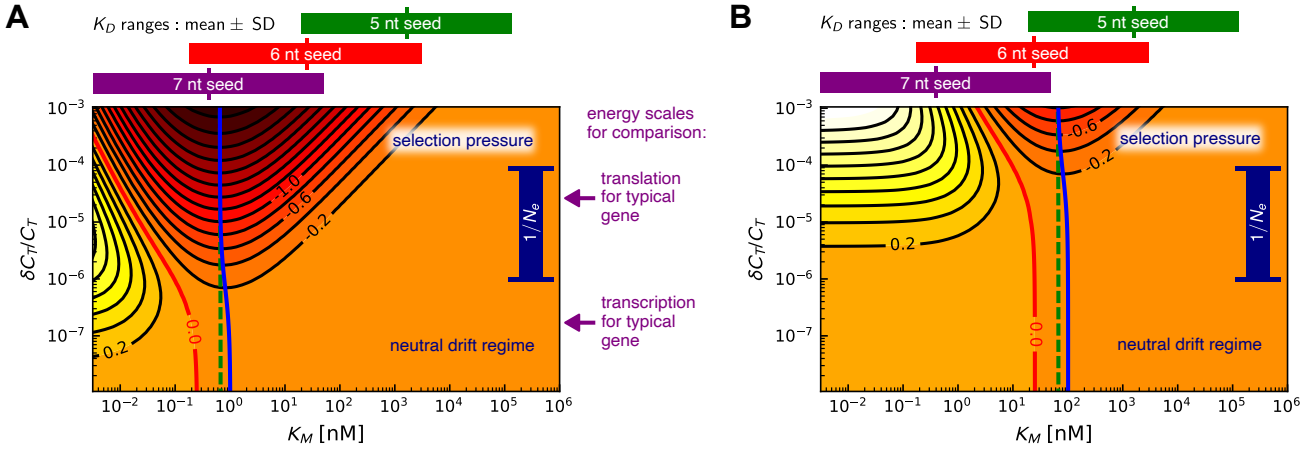

FIG. 2. Contour diagrams of noise filter error, $\log_{10} E$, in terms of Michaelis-Menten constant $K_{\rm M}$ and fractional metabolic cost $\delta C_T/C_T$ for a fixed protein output level $\bar{p}$ in the (**A**) single target and (**B**) 100 target cases. The spacing between contour values is 0.2. The minimum contour line for a given $\delta C_T/C_T$ corresponds to the most energetically efficient noise reduction, and this is achieved by K_M^* values along the blue curves. Similarly, the green dashed curves show the most efficient $K_{\rm M}^{\rm wk}$ values predicted by the WK optimal filtering theory. The red line is the boundary of the "dud" region on the left, inside which the miRNA regulation adds noise ($E > 1$) rather than mitigating it. To make a connection to physiological values of $K_{\rm M}$, we plot estimated K_D ranges for known miRNA-mRNA interactions above the plots for 5, 6, and 7 nt seed lengths. $K_{\rm D}$ sets a lower bound on $K_{\rm M}$ from Eq. (1). Altogether, this shows that for a single typical target gene, the most economical noise reduction is likely to occur for 7 nt seeds, while 6 nt seeds become favorable for miRNAs with many targets. To better illustrate the evolutionary relevance, we also plot the dark blue bar showing inverse effective population sizes $N_e^{-1} = 10^{-6} - 10^{-4}$ typical for metazoans. As the fitness disadvantage due to metabolic costs become significant, $\delta C_T/C_T \gtrsim N_e^{-1}$, there is selective pressure on the organism driving it towards the $K_{\rm M}^*$ value.

In the experiments, $K_{\rm M} = 1 \pm 0.2$ nM was found for a fully complementary interaction with a 7 nt seed, while single nucleotide mismatches in the seed binding (which in principle allow for a larger range of possible targets) boosted $K_{\rm M}$ by up to a factor of 82. While we do not yet have extensive experimental surveys of $K_{\rm M}$ values in miRNA (or related siRNA [small interfering RNA]) systems, it would be intriguing to check whether $K_{\rm M}$ tends to scale with the target population, as predicted by the optimal theory.

We can make the connection between the number of complementary matches (or seed length) and K_M more explicit. In SI Sec. S.IV we used the ViennaRNA server [20] to predict the ΔG values of $\approx 10^4$ human miRNA seed sequences of length 7 nt, resulting in a distribution of $K_{\rm D}$ values which covered around ten decades on a logarithmic scale. The mean and standard deviation of $\log_{10}(K_{\rm D}/[1M]) = -9.5 \pm 2.2$ is shown as a purple bar above the plots in Fig. 2A,B. To mimic shorter seeds, we deleted 1 or 2 nucleotides from the sequence, to give the ranges $\log_{10}(K_{\rm D}/[1M]) = -7.6 \pm 2.0$ (6 nt, red bar) and -5.8 ± 1.9 (5 nt, green bar). As validation, the calculation was able to correctly reproduce measured $K_{\rm D}$ ranges for fully matched 7 nt seeds in fruit fly and mouse siRNA-target complexes. Since $K_{\rm D}$ sets the floor for $K_{\rm M}$, we see from Fig. 2A that the optimal $K_M^* \approx 1$ nM for the single-target case is unlikely to be accessible for 5 nt seeds. It becomes plausible for 6 nt seeds, and even more so for 7 nt seeds. Since $K_{\rm M}$ can be up to two decades larger than $K_{\rm D}$ [19], it is notable that 7 nt seeds densely cover the range of $K_{\rm D}$ values one or two decades smaller than K_M^*. Extrapolating this pattern, longer seeds (≥ 8 nt) will be less likely than 7 nt to achieve K_M^*. For the 100 target case (Fig. 2B) we see that the ideal seed length is shifted to 6 nt as a result of the increase in K_M^*. Seeds of length 6 nt constitute 67% of a dataset of human and mouse miRNA seed sequences, with 7 nt seeds forming another 23% [21]. The preponderance of 6 nt seeds is in line with expectations if K_M was optimized, particularly since miRNAs will typically have many different targets.

Noise reduction can approach optimal linear filter performance. The filter analogy in Eq. (2) allows us to make an interesting comparison. For a given signal $s(t)$ and $n(t)$, we know that there is a filter function $H_{\rm wk}(t)$, the Wiener-Kolmogorov (WK) solution [22–24], which gives the best performance (smallest E) of all possible functions $H(t)$ for this type of linear noise filtering system (SI Sec. S.II B). We denote the corresponding value of error, $E^*_{\rm wk}$, and it serves as the overall lower bound on E. In our system, E reaches a minimum E^* at K_M^* for a given energetic cost, but is this minimum E^* comparable to $E^*_{\rm wk}$? In other words, can miRNA noise regulation approach an optimal WK filter? This type of comparison has recently proven fruitful in a variety of biological contexts [17, 25–31], for example yielding

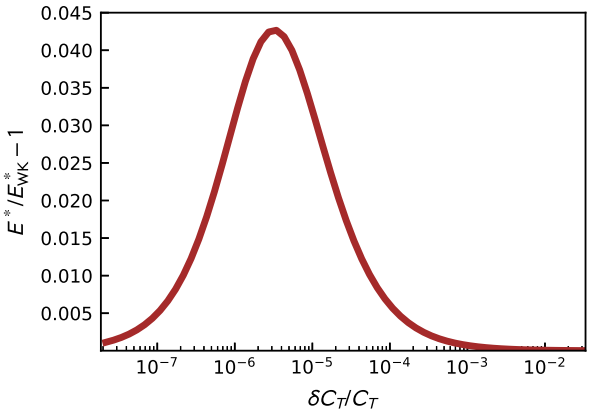

FIG. 3. The discrepancy between E^* and $E^*_{\rm wk}$, the error of the actual system and the WK optimal system respectively, at the most energetically efficient $K_{\rm M}$ for the single target case (Fig. 2A), at different values of the fractional cost $\delta C_T/C_T$. $E^*/E^*_{\rm wk} - 1 < 0.045$, so the miRNA system exhibits a performance close to WK optimality for this network motif.

tight bounds on the fidelity of information transmission in signaling networks.

The miRNA system does not exactly realize true WK optimality, because the optimal filter function $H_{\rm wk}(t)$ cannot be precisely implemented by the miRNA regulation network. However, the affinity $K_M^{\rm wk}$ predicted to be optimal by the WK theory for a given $\delta C_T/C_T$ (green dashed curve in Figs. 2A,B) is extremely close to K_M^* (blue curve). Fig. 3 shows the difference between the respective errors E^* and $E^*_{\rm wk}$ along these curves, as a function of $\delta C_T/C_T$. While $E^*/E^*_{\rm wk} - 1 > 0$, as expected, the difference is always smaller than 0.045, peaking at moderate cost values. As the cost $\delta C_T/C_T$ increases, E^* converges to $E^*_{\rm wk}$, and K_M^* similarly converges to $K_M^{\rm wk}$. Notably, despite its constraints, the miRNA system can get quite close to WK optimal performance.

Evolutionary pressure on miRNA noise regulation. The previous two sections have argued that optimality in noise reduction (in the broader WK sense) is in principle approachable, and 6-7 nt seeds put miRNA systems within reach of achieving it. The final question we would like to consider is whether there would be any pressure from natural selection actually driving miRNA regulation toward optimality. From a population genetics perspective, let us say that the fitness of an organism with a particular miRNA regulatory network is f_R, while the same organism missing the network has fitness f_0. The selection coefficient $s = f_R/f_0 - 1$, quantifying the relative fitness, can be decomposed into two contributions, $s = s_a + s_c$ [18]. Here s_a is the adaptive advantage due to the regulation, for example resulting from protein noise reduction. The remaining part, $s_c < 0$, comes from the added metabolic cost of implementing the regulation. In order for s to be overall positive (and hence $f_R > f_0$) we would need an $s_a > 0$ that is larger in magnitude than the cost, $s_a > |s_c|$.

The evolution of miRNA systems, however, poses a conundrum: a newly arisen miRNA, before any selective fine-tuning of its seed sequence, could potentially target hundreds of mRNA in a random manner, making $s_a < 0$ due to deleterious effects on existing genetic networks. So how would advantageous miRNA regulation eventually emerge? The solution to this problem, as argued by Chen and Rajewsky [2], is for new miRNA systems to be expressed at very low levels, such that s_a, even if negative, has a negligible magnitude. In this regime however $s \approx s_c < 0$, which is still an overall fitness disadvantage. Would such a deleterious variant survive in a population long enough for further mutations to confer a significant positive s_a? The answer depends on the magnitude of $|s_c|$ relative to a threshold known as the "drift barrier" [32]. This threshold is set by N_e^{-1}, where N_e is a measure of genetic diversity known as the effective population size [33]. N_e is the size of an idealized population that shows the same changes in genetic diversity per generation (due to genetic drift, or random sampling of variants) as the actual population. N_e is generally smaller than the real population size, and among more complex eukaryotes (like vertebrates) can be as small as $10^4 - 10^6$ [18, 34]. If $|s_c| \ll N_e^{-1}$, selection against the variant is weak, even when $s_c < 0$, so it can survive in a population via genetic drift as an effectively neutral mutant and even eventually take over (with a fixation probability roughly given by its initial fraction in the population [35]). On the other hand, if $|s_c| \gg N_e^{-1}$, then the costs are sufficiently high that selection would efficiently weed out the variant from the population, unless there were compensating advantages, $s_a \gtrsim |s_c|$.

A recent derivation based on a general bioenergetic growth model for organisms allows us to make the above discussion more quantitative: it showed that $s_c \approx -\ln(R_b)\delta C_T/C_T$, where R_b is the mean number of offspring per individual [36]. Assuming that $\ln(R_b)$ does not change the order of magnitude, we can thus use the fractional cost $\delta C_T/C_T$ as a proxy for $|s_c|$, and compare it to N_e^{-1}. The blue range bars on the right in Fig. 2A,B show possible N_e^{-1} values 10^{-6}–10^{-4} for higher-order eukaryotes, separating a selection pressure regime at large $\delta C_T/C_T$ from a neutral drift regime at low $\delta C_T/C_T$. For comparison we also show the cost scales for transcription and translation of a typical gene [18], which indicate that transcription is not generally not under selective pressure in these organisms, while translation may be.

At the lowest expression levels, a newly evolved miRNA system could survive in the neutral drift regime, even with a non-optimal $K_{\rm M}$. There is limited noise reduction achievable in this regime, since the contours indicating E significantly smaller than 1 require larger $\delta C_T/C_T$, particularly for the multi-target case (Fig. 2B).

Thus the initial evolution could be imagined as a random walk near the bottom edge of the diagram. Mutations that led to greater expression of an miRNA, moving up on the cost scale, would hit against the drift barrier, and would more likely survive if they came with compensating fitness advantages. In a context where noise control was beneficial, this would mean being funneled up the region where K_M is close to K_M^*, the path that confers the largest noise reduction as expression levels rise. Once the costs are above the drift barrier, there would be significant selective pressure to optimize K_M. In this high cost regime there is a direct tradeoff between extra metabolic expenditure and noise filtering, with $\Delta M_\nu \sim M_\nu/E$, and hence $s_c \sim -M_\nu/E$. Our theory predicts that any compensating fitness advantage $s_a > 0$ would have to grow like E^{-1} or faster as E decreases, in order for the regulation to be viable in the long term.

Conclusions. Putting all the elements of our theory together, we show that if noise control confers a fitness advantage for a particular miRNA system, there is selective pressure driving it toward an optimal miRNA-mRNA affinity, as described by the Michaelis-Menten constant K_M. Remarkably, the optimal value K_M^* in Eq. (4) can be achieved by a fairly narrow range of seed lengths (6-7 nt), which happen to make up the vast majority of miRNA seeds. While we argue the plausibility of this key result based on realistic ranges of biological parameters, Eq. (4) opens the way for future experimental validation in specific, fully characterized systems. Such validation should be practical, since the equation only involves a small number of biochemical parameters. The noise reduction at optimality approaches the performance of a Wiener-Kolmogorov filter, the best possible linear noise filter. If true, such optimization would be a striking example of metabolic costs directly shaping the course of evolution for a biochemical network in eukaryotes. This is unusual in itself, because eukaryotes are generally less likely to prioritize energy efficiency relative to prokaryotes, which have higher effective population sizes and thus lower drift barriers [18]. Beyond the implications for miRNA evolution, the theory could also find applications in the design of synthetic circuits with 3' UTR engineering [37] and artificial miRNAs [38].

Though our work focuses on noise control via miRNA regulation, it is also important to keep in mind that noise control can be implemented via other mechanisms, and that miRNA themselves have other functional roles. Protein noise can be reduced (maintaining the same expression level) by increasing transcription and either decreasing translation rates or enhancing mRNA degradation [39, 40], and this would avoid noise added due to miRNA interactions [41]. Degradation rates could for example be fine-tuned by the length of mRNA poly(A) tails [42]. However this mechanism is non-specific, while miRNAs allow for selective noise control via seed recognition. Though our results suggest a formative role for noise reduction in shaping miRNA-mRNA affinity, the filtering capacity likely co-evolved with other miRNA functions such gene silencing and crosstalk [43]. Perhaps in some cases noise regulation is simply a complimentary benefit of a far more complex utility scheme. Ultimately a deeper understanding of miRNA evolution will require larger-scale models of its full regulatory context, coupled with *in vivo* experiments to explore the tangled effects of function, metabolic costs, and fitness.

Acknowledgements. We thank Amit Singh Vishen for helpful discussions and suggestions.

[1] X. Yang, M. Heinemann, J. Howard, G. Huber, S. Iyer-Biswas, G. Le Treut, M. Lynch, K. L. Montooth, D. J. Needleman, S. Pigolotti, *et al.*, Proc. Natl. Acad. Sci. **118**, e2026786118 (2021).
[2] K. Chen and N. Rajewsky, Nat. Rev. Genet. **8**, 93 (2007).
[3] D. P. Bartel, Cell **173**, 20 (2018).
[4] M. Osella, C. Bosia, D. Corá, and M. Caselle, PLoS Comput. Biol. **7**, e1001101 (2011).
[5] J. M. Schmiedel, S. L. Klemm, Y. Zheng, A. Sahay, N. Blüthgen, D. S. Marks, and A. van Oudenaarden, Science **348**, 128 (2015).
[6] V. Siciliano, I. Garzilli, C. Fracassi, S. Criscuolo, S. Ventre, and D. Di Bernardo, Nat. Comm. **4**, 2364 (2013).
[7] E. Hornstein and N. Shomron, Nat. Genet. **38**, S20 (2006).
[8] M. S. Ebert and P. A. Sharp, Cell **149**, 515 (2012).
[9] K. J. Peterson, M. R. Dietrich, and M. A. McPeek, Bioessays **31**, 736 (2009).
[10] C. Alberti and L. Cochella, Development **144**, 2548 (2017).
[11] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).
[12] G. Tkačik and A. M. Walczak, Journal of Physics: Condensed Matter **23**, 153102 (2011).
[13] J. Tsang, J. Zhu, and A. Van Oudenaarden, Mol. Cell **26**, 753 (2007).
[14] N. J. Martinez and A. J. Walhout, Bioessays **31**, 435 (2009).
[15] D. Baek, J. Villén, C. Shin, F. D. Camargo, S. P. Gygi, and D. P. Bartel, Nature **455**, 64 (2008).
[16] M. Selbach, B. Schwanhäusser, N. Thierfelder, Z. Fang, R. Khanin, and N. Rajewsky, Nature **455**, 58 (2008).
[17] D. Hathcock, J. Sheehy, C. Weisenberger, E. Ilker, and M. Hinczewski, IEEE Trans. Mol. Biol. Multi-Scale Commun. **2**, 16 (2016).
[18] M. Lynch and G. K. Marinov, Proc. Natl. Acad. Sci. **112**, 15690 (2015).
[19] L. M. Wee, C. F. Flores-Jasso, W. E. Salomon, and P. D. Zamore, Cell **151**, 1055 (2012).
[20] R. Lorenz, S. H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, and I. L. Hofacker, Algorithms Mol. Biol. **6**, 1 (2011).
[21] D. C. Ellwanger, F. A. Büttner, H.-W. Mewes, and V. Stümpflen, Bioinformatics **27**, 1346 (2011).
[22] N. Wiener, *Extrapolation, interpolation, and smoothing*

of stationary time series, Vol. 2 (MIT press Cambridge, 1949).
[23] A. N. Kolmogorov, Izv. Akad. Nauk SSSR Ser. Mat. **5**, 3 (1941).
[24] H. W. Bode and C. E. Shannon, Proc. Inst. Radio. Engin. **38**, 417 (1950).
[25] M. Hinczewski and D. Thirumalai, Phys. Rev. X **4**, 041017 (2014).
[26] N. B. Becker, A. Mugler, and P. R. ten Wolde, Phys. Rev. Lett. **115**, 258103 (2015).
[27] M. Hinczewski and D. Thirumalai, J. Phys. Chem. B (2016).
[28] D. Hartich and U. Seifert, Phys. Rev. E **94**, 042416 (2016).
[29] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, J. Stat. Phys. **162**, 1395 (2016).
[30] C. Zechner, G. Seelig, M. Rullan, and M. Khammash, Proc. Natl. Acad. Sci. U. S. A. **113**, 4729 (2016).
[31] T.-L. Wang, B. Kuznets-Speck, J. Broderick, and M. Hinczewski, bioRxiv , 2020.10.06.327700 (2020).
[32] W. Sung, M. S. Ackerman, S. F. Miller, T. G. Doak, and M. Lynch, Proc. Natl. Acad. Sci. **109**, 18488 (2012).
[33] J. H. Gillespie, *Population Genetics: A Concise Guide* (JHU Press, 2010).
[34] B. Charlesworth, Nat. Rev. Genet. **10**, 195 (2009).
[35] M. Kimura, *The neutral theory of molecular evolution* (Cambridge University Press, 1983).
[36] E. Ilker and M. Hinczewski, Phys. Rev. Lett. **122**, 238101 (2019).
[37] F. J. Navarro and D. C. Baulcombe, ACS Synth. Biol. **8**, 358 (2019).
[38] A. Kotowska-Zimmer, M. Pewinska, and M. Olejniczak, Wiley Interdiscip. Rev. RNA **12**, e1640 (2021).
[39] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. Van Oudenaarden, Nat. Genet. **31**, 69 (2002).
[40] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[41] R. Fan and A. Hilfinger, Biophys. J. **122** (2023).
[42] O. S. Rissland, A. O. Subtelny, M. Wang, A. Lugowski, B. Nicholson, J. D. Laver, S. S. Sidhu, C. A. Smibert, H. D. Lipshitz, and D. P. Bartel, Genome Biol. **18**, 1 (2017).
[43] M. Jens and N. Rajewsky, Nat. Rev. Genet. **16**, 113 (2015).

Supplemental Material: Bioenergetic costs and the evolution of noise regulation by microRNAs

Efe Ilker[1] and Michael Hinczewski[2]

[1] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*

[2] *Department of Physics, Case Western Reserve University, Cleveland OH 44106*

CONTENTS

S.I. Analysis of microRNA-mediated noise regulation 1
A. Chemical Langevin formalism of the microRNA-mediated gene expression 1
B. Protein expression noise 3
C. Bioenergetic costs of miRNA regulation 5

S.II. The noise-filter formulation of the microRNA regulated system 5
A. Realizing miRNA regulation as a noise filtering mechanism 5
B. Wiener-Kolmogorov (WK) optimality of protein output noise level 6

S.III. Calculation of phase diagrams 8
A. Contour plots, fitness costs, and Michaelis-Menten constant 8
B. Minimal noise for the same metabolic cost 8

S.IV. Estimating dissociation constants K_{D} for seed sequences of varying length 9

References 11

S.I. ANALYSIS OF MICRORNA-MEDIATED NOISE REGULATION

A. Chemical Langevin formalism of the microRNA-mediated gene expression

Here, we present the continuum description of gene expression dynamics described by chemical Langevin (CL) equations.

a. General birth-death processes: We consider that the rate of change of each biomolecule $dn_j(t)/dt$ is related to a set of stochastic "reactions", which we will label with an index $\rho = 1, \ldots, r$, where r is the total number of possible reactions. The CL dynamics of the birth-death process for biomolecule j can be written as [1]:

$$\frac{dn_j(t)}{dt} = \sum_{\rho=1}^{r} \Delta_{\rho j} \pi_\rho(t) + \sum_{\rho=1}^{r} \Delta_{\rho j} \sqrt{\pi_\rho(t)} \eta_\rho(t) \tag{S1}$$

where $\Delta_{\rho j}$ are the stoichiometric coefficients representing the change in the population of j in reaction ρ, $\pi_\rho(t)$ are the rates of reactions and $\eta_\alpha(t)$ are independent Gaussian noise

terms at time t which satisfy $\langle \eta_\rho(t)\eta_\beta(t')\rangle = \delta_{\rho\beta}\delta(t-t')$ where $\delta_{\rho\beta}$ and $\delta(t-t')$ are Kronecker delta and Dirac delta functions. As the reactions for the production and degradation of j are uncorrelated, we can add up the uncorrelated noise terms, and obtain:

$$\frac{dn_j(t)}{dt} = \sum_{\rho=1}^{r} \Delta_{\rho j}\pi_\rho(t) + \sqrt{\alpha_j(t)}\eta_j(t) \tag{S2}$$

where $\alpha_j(t) = \sum_{\rho=1}^{r} \Delta_{\rho j}^2 \pi_\rho(t)$. Note that η_j's are uncorrelated at different times yet there can be non-zero correlations between the dynamics of different populations at time t if they share the same reactions.

b. Protein production: The reaction set of interest in the gene expression is given as [2]:

$$\begin{aligned}
\emptyset &\underset{d_m}{\overset{\nu_R}{\rightleftarrows}} m \\
m &\xrightarrow{k_p} m + p \\
p &\xrightarrow{d_p} \emptyset \\
\emptyset &\underset{d_\mu}{\overset{\nu_\mu}{\rightleftarrows}} \mu \\
m + \mu &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftarrows}} m_b \\
m_b &\xrightarrow{d_{m\mu}} \mu \\
m_b &\xrightarrow{d_\mu} m
\end{aligned}$$

where m, p, μ, m_b represent respectively mRNA, protein, miRNA and miRNA-mRNA complex. These reactions include birth-death processes of species as well as complex formation and degradation. As the miRNA-mRNA complex m_b degrades, the miRNAs are recycled. For convenience, we assume a similar property when miRNAs spontaneously degrade however in general $d_\mu \ll d_{m\mu}$. This model has been shown to agree well with experimental systems of miRNA-mediated gene expression and noise reduction [2, 3].

We may express the dynamics of biomolecular concentrations in law of mass action form:

$$\frac{d[\mu]}{dt} = \nu_\mu - d_\mu[\mu] - k_{\text{on}}[m][\mu] + (k_{\text{off}} + d_{m\mu})[m_b] \tag{S3a}$$

$$\frac{d[m]}{dt} = \nu_R - d_m[m] - k_{on}[m][\mu] + (k_{\text{off}} + d_\mu)[m_b] \tag{S3b}$$

$$\frac{d[m_b]}{dt} = k_{\text{on}}[m][\mu] - (k_{\text{off}} + d_{m\mu} + d_\mu)[m_b] \tag{S3c}$$

$$\frac{d[p]}{dt} = k_p[m] - d_p[p]. \tag{S3d}$$

Now, introducing $[\mu_{tot}] = [\mu] + [m_b]$ and assuming binding/unbinding rate of mRNA-microRNA complex is much faster than other dynamics, we obtain a quasi steady state

approximation indicating $[m_b] = \frac{[m][\mu_{tot}]}{K_M+[m]}$ with $K_\mathrm{M} = \frac{k_\mathrm{off}+d_{m\mu}+d_\mu}{k_\mathrm{on}}$. With these simplifications, we obtain

$$\frac{d[\mu_{tot}]}{dt} = \nu_\mu - d_\mu[\mu_{tot}] \tag{S4a}$$

$$\frac{d[m]}{dt} = \nu_R - d_m[m] - \frac{d_{m\mu}}{\kappa_M + [m]}[\mu_{tot}][m] \tag{S4b}$$

$$\frac{d[p]}{dt} = k_p[m] - d_p[p]. \tag{S4c}$$

Multiplying concentrations with the cell volume V, and using (S1) and (S2), we can now introduce the Chemical Langevin in terms of copy numbers (expressed without brackets) form as:

$$\frac{d\mu_{tot}(t)}{dt} = \nu_\mu - d_\mu\mu_{tot}(t) + \sqrt{\alpha_\mu}\eta_\mu(t) \tag{S5a}$$

$$\frac{dm(t)}{dt} = \nu_R - d_m m(t) - \frac{d_{m\mu}}{\kappa_M + m(t)}\mu_{tot}(t)m(t) + \sqrt{\alpha_m^R}\eta_m^R(t) \tag{S5b}$$

$$\frac{dp(t)}{dt} = k_p m(t) - d_p p(t) + \sqrt{\alpha_p^R}\eta_p^R(t). \tag{S5c}$$

c. Steady-state properties, regulation, bound fraction of miRNAs: At the steady-state, the mean populations are $\bar{p} = \frac{k_p}{d_p}\bar{m}$, $\bar{\mu}_\mathrm{tot} = \nu_\mu/d_\mu$, $\bar{m} = \frac{\nu_R}{d_m}(1-R)$ where the regulation strength is given by:

$$R = \frac{\theta_\mu\bar{\mu}_\mathrm{tot}}{\theta_\mu\bar{\mu}_\mathrm{tot} + \gamma_m\bar{m}} \tag{S6}$$

where $\gamma_m \equiv d_m/d_{m\mu}$ and θ_μ is the bound fraction of miRNAs, i.e.,

$$\theta_\mu = \frac{\bar{m}}{K_M + \bar{m}}. \tag{S7}$$

Fixing the mean protein population for all regulation levels requires transcriptional increase for compensation. Accordingly, we can express the transcription rate in regulated system ν_R in terms of the transcription in unregulated system ν_0:

$$\nu_R = \frac{\nu_0}{1-R} = \frac{d_m\bar{m}}{1-R}. \tag{S8}$$

B. Protein expression noise

The fluctuating output numbers are defined in the main text as $x(t) = \bar{x} + \delta x(t)$ where $\delta x(t)$ is the variation from the mean level $\bar{x}$ and the species of interest are $x(t) \equiv \{\mu_{tot}(t), m(t), p(t)\}$. As we focus on the stationary-state fluctuations, we may treat the jump noise amplitudes α_j's to be time-independent. Plugging these in Equation series (S5), using

linear-noise approximation and taking the Fourier transform leads us to

$$\delta\mu(\omega) = \frac{\sqrt{\alpha_\mu}\eta_\mu(\omega)}{d_\mu - i\omega} \tag{S9a}$$

$$\delta m_R(\omega) = \frac{\sqrt{\alpha_m}\eta_m(\omega) + \sqrt{\alpha_m R/(1-R)}\eta_m^R(\omega) - d_{m\mu}\theta_\mu\delta\mu(\omega)}{c_m - i\omega} \tag{S9b}$$

$$\delta p_R(\omega) = \frac{k_p\delta m_R(\omega)}{d_p - i\omega} + \frac{\sqrt{\alpha_p}\eta_p(\omega)}{d_p - i\omega}. \tag{S9c}$$

where $c_m = d_m(1 - R\theta_\mu)/(1 - R)$. The limit of $R \to 0$ corresponds to unregulated system as $c_m \to d_m$ and $d_{m\mu}\theta_\mu\delta\mu(\omega) \to 0$ in (S9b).

The main observable for noise in output population is the variance $\sigma_p^2 \equiv \langle(\delta p(t))^2\rangle = \langle(\delta p(0))^2\rangle$. In order to relate the kinetic parameters to noise spectrum, we first need to plug in the values of jump noise amplitudes α_j's write them in terms of equilibrium populations. These are given by we have $\alpha_\mu = 2d_\mu\bar{\mu}_{\text{tot}}$, $\alpha_m = 2d_m\bar{m}$, and $\alpha_p = 2d_p\bar{p}$ where $\bar{\mu}_{\text{tot}}$, $\bar{m}$, and $\bar{p}$ are mean steady-state population levels for miRNA, mRNA and protein respectively. By using the correlation property of Gaussian noise in Fourier space, the variance can be obtained by taking the inverse Fourier transform of $\langle(\delta p_R(\omega)\delta p_R(-\omega))\rangle$. Here, we quantify the amount of variance σ_p^2 at a given output level $\bar{p}$ by using the Fano factor, i.e., $F \equiv \sigma_p^2/\bar{p}$, and define F_0, F_R to describe the Fano factor respectively in unregulated and regulated systems. These give:

$$F_0 = 1 + \frac{k_p}{d_m + d_p}, \tag{S10a}$$

$$F_R = 1 + \frac{k_p}{(c_m + d_p)(1 - R\theta_\mu)} + \frac{d_{m\mu}\theta_\mu(c_m + d_\mu + d_p)k_pR}{(c_m + d_\mu)(c_m + d_p)(d_\mu + dp)(1 - R\theta_\mu)} \tag{S10b}$$

in which the third term in F_R is a result of coupling with microRNA pool noise (extrinsic noise). We can quantify the relative noise levels in miRNA regulated system with respect to the unregulated systems, with the function

$$E = \frac{F_R - 1}{F_0 - 1}. \tag{S11}$$

To simplify our analysis, we define the following nondimensional parameters:

$$\gamma_m \equiv \frac{d_m}{d_{m\mu}},\ \ \gamma_\mu \equiv \frac{d_\mu}{d_{m\mu}},\ \ \phi \equiv \frac{d_p}{d_m}. \tag{S12}$$

Assuming degradation rate of proteins is slower than of RNAs, i.e., $\phi \ll 1$, we get

$$E \approx \frac{(1 - R)(\gamma_\mu + R\theta_\mu)}{\gamma_\mu(1 - R\theta_\mu)^2} \tag{S13}$$

C. Bioenergetic costs of miRNA regulation

Following the Lynch & Marinov analysis [4] which is based on the biochemistry of transcription, the transcriptional cost of a typical gene in eukoryatic cells is given by the sum of maintenance and synthesis (growth) costs, $c^{\nu}_{\rm tot} \equiv M_\nu t_r = c_g + \nu_0 \epsilon_m t_r$ where ϵ_m is the cost of recycling mRNAs. For long enough cell-division times, the maintenance term dominates, and hence $M_\nu \approx \nu_0 \epsilon_m$. By contrast, the costs of miRNA regulation consists of two contributions: increased transcription of mRNAs in order to maintain $\bar{p}(R) = \bar{p}$ and transcription of miRNAs. This can be expressed as $\Delta M_\nu = (\nu_R - \nu_0)\epsilon_m + \nu_\mu \epsilon_\mu$. By substituting $\nu_R = \nu_0/(1-R)$, $\nu_\mu = d_\mu \bar{\mu}_{\rm tot}$, $\nu_0 = d_m \bar{m}$, and inverting (S6) for $\bar{\mu}_{\rm tot}$, we obtain Eq. (3) in the main text,

$$\Delta M_\nu = \frac{R}{1-R}\left(1 + \frac{\gamma_\mu \sigma_\epsilon}{\theta_\mu}\right) M_\nu \tag{S14}$$

$\sigma_\epsilon \equiv \epsilon_\mu / \epsilon_m$.

In order to determine the noise levels for a given cost ΔM_ν, we can invert the above equation for R, to plug in Eq.(S15). Hence, in the limit $\phi \ll 1$, we have

$$E(\mathcal{M}, \theta_\mu) \approx \frac{(\gamma_\mu \sigma_\epsilon + \theta_\mu)(\mathcal{M}\theta_\mu^2 + \gamma_\mu^2 \sigma_\epsilon + \gamma_\mu \theta_\mu (1+\mathcal{M}))}{\gamma_\mu \left(\gamma_\mu \sigma_\epsilon + \theta_\mu (1 + \mathcal{M}(1-\theta_\mu))\right)^2} \tag{S15}$$

where $\mathcal{M} = \Delta M_\nu / M_\nu$.

S.II. THE NOISE-FILTER FORMULATION OF THE MICRORNA REGULATED SYSTEM

A. Realizing miRNA regulation as a noise filtering mechanism

The microRNA regulation acts directly on mRNA related fluctuations on protein output production. We can then formulate the noise regulation as a filtering process by decomposing the protein fluctuations given in (S9c) as

$$\delta p_R(\omega) = \underbrace{\frac{k_p \delta m_R(\omega)}{d_p - i\omega}}_{s(\omega) - \tilde{s}(\omega)} + \underbrace{\frac{\sqrt{\alpha_p}\eta_p(\omega)}{d_p - i\omega}}_{\delta p_g(\omega)} \tag{S16}$$

where $\delta p_g(\omega)$ is the ground noise in protein production, $s(\omega) = \frac{k_p \delta m_R(\omega)}{d_p - i\omega}|_{R=0}$. This satisfies the signal estimation problem where $\tilde{s}(\omega)$ is the estimate of $s(\omega)$ which is the actual signal. Moreover, $\tilde{s}(\omega), s(\omega)$ have non-zero correlations, whereas $\langle(\delta p_g(\omega) s(-\omega))\rangle = \langle(\delta p_g(\omega)\tilde{s}(-\omega))\rangle = 0$. Accordingly, the normalized error in estimation is defined in time domain:

$$E = \frac{\langle (s(t) - \tilde{s}(t))^2 \rangle}{\langle s(t)^2 \rangle} \tag{S17}$$

which is then equivalent to (S11) as $\langle (s(t) - \tilde{s}(t))^2 \rangle = (F_R - 1)\bar{p}$ and $\langle s(t)^2 \rangle = (F_0 - 1)\bar{p}$. For linear filtering problem, we can write the estimate signal as:

$$\tilde{s}(\omega) = H(\omega)(s(\omega) + n(\omega)), \tag{S18}$$

where $s(\omega) + n(\omega) \equiv c(\omega)$ is the corrupted signal with noise $n(\omega)$ and $H(\omega)$ is the linear filter. Eq. (2) of the main text lays out this expression in the time domain, thus it is the inverse Fourier transform of (S18). Now, we can obtain each of the above terms using (S9b), (S16), and (S18). Thus, we get:

$$H(\omega) \equiv \frac{B - A}{A} \tag{S19}$$

$$s(\omega) \equiv \frac{k_p \sqrt{\alpha_m} \eta_m(\omega)}{DA} \tag{S20}$$

$$n(\omega) \equiv -\frac{k_p}{(B - A)D} \left(\sqrt{\alpha_m R/(1 - R)} \eta_m^R(\omega) - \frac{d_{m\mu}\theta_\mu}{C} \sqrt{\alpha_\mu} \eta_\mu(\omega) \right) \tag{S21}$$

$$c(\omega) \equiv s(\omega) + n(\omega) \tag{S22}$$

with complex functions

$$A \equiv d_m - i\omega, \quad B \equiv c_m - i\omega, \quad C \equiv d_\mu - i\omega, \quad D \equiv d_p - i\omega. \tag{S23}$$

The noise term $n(\omega)$ captures the additional fluctuations by the miRNA-mRNA interactions and it is uncorrelated with the actual signal $s(\omega)$.

B. Wiener-Kolmogorov (WK) optimality of protein output noise level

In order to determine the error function E in (S17), we need to calculate the variance given in the numerator:

$$\langle (s(t) - \tilde{s}(t))^2 \rangle|_{t=0} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left(|H(\omega)|^2 P_{nn}(\omega) + |1 - H(\omega)|^2 P_{ss} \right) \tag{S24}$$

where the power spectra (or cross-spectral density) functions are defined as:

$$2\pi P_{xy}(\omega)\delta(\omega + \omega') = \langle x(\omega) y(\omega') \rangle \tag{S25}$$

The inverse Fourier transform of this expression would give the cross correlations in time domain $C_{xy}(t) \equiv \langle x(t') y(t' + t) \rangle$. The Wiener-Kolmogorov (WK) optimal filter minimizing the error E for a given $s(t)$, i.e., minimizing (S24), satisfies

$$H_{\text{wk}}(\omega) = \frac{1}{P_{cc}^+} \left\{ \frac{P_{cs}}{P_{cc}^-} \right\}_+ \tag{S26}$$

which requires two types of decomposition for causal (+) and anti-causal (−) parts, i.e., multiplicative $P_{cc} = P_{cc}^{+} P_{cc}^{-}$ and additive $\{P_{cs}/P_{cc}^{-}\}_{+}$. Here, the "casual" means that the function in Fourier domain extended in complex ω plane must not have zeros or poles in the upper half-plane (Im $\omega > 0$).

The relevant power spectra are obtained by inserting (S20), (S21), (S22) in (S25):

$$P_{cs} = \frac{{k_p}^2}{|D|^2|A|^2}\alpha_m, \qquad P_{cc} = \frac{P_{cs}}{|C|^2 d_m^2}\mathcal{L}\left(\frac{d_m^2|C|^2}{\mathcal{L}} + |A|^2|C|^2 + d_m^2 b_\mu |A|^2\right) \tag{S27}$$

and $P_{ss} = P_{cs}$, $P_{nn} = P_{cc} - P_{cs}$ with dimensionless parameters $b_\mu = \frac{\alpha_\mu(1-R)}{\alpha_m R}(\theta_\mu/\gamma_m)^2$ and $\mathcal{L} = \frac{1-R}{R(1-\theta_\mu)^2}$. To achieve the causal decompositions required for WK equation (S26), P_{cc} can be rewritten as,

$$P_{cc} = \frac{P_{cs}}{|C|^2 d_m^2}\mathcal{L}(\lambda_1 + i\omega)(\lambda_1 - i\omega)(\lambda_2 + i\omega)(\lambda_2 - i\omega) \tag{S28}$$

where $\pm i\lambda_{1,2}$ are the roots of the numerator and given by

$$\lambda_{1,2} = d_m\left(\frac{\Gamma \pm \sqrt{\Gamma^2 - 4\beta}}{2}\right), \ \Gamma = b_\mu + \left(\frac{\gamma_\mu}{\gamma_m}\right)^2 + \frac{1+\mathcal{L}}{\mathcal{L}}, \ \beta = b_\mu + \frac{1+\mathcal{L}}{\mathcal{L}}\left(\frac{\gamma_\mu}{\gamma_m}\right)^2 \tag{S29}$$

Thus, we obtain

$$P_{cc}^{+} = \frac{P_{cs}^{+}}{C d_m}\sqrt{\mathcal{L}}(\lambda_1 - i\omega)(\lambda_2 - i\omega) \tag{S30a}$$

$$\frac{P_{cs}}{P_{cc}^{-}} = \frac{P_{cs}^{+} C^* d_m}{\sqrt{\mathcal{L}}(\lambda_1 + i\omega)(\lambda_2 + i\omega)}. \tag{S30b}$$

where $P_{cs}^{+} = \frac{k_p\sqrt{\alpha_m}}{DA}$. Next, we need to first determine the additive decomposition of P_{cs}/P_{cc}^{-} and then remove the poles in the upper half-plane (Im $\omega > 0$). This yields the casual part

$$\left\{\frac{P_{cs}}{P_{cc}^{-}}\right\}_{+} = \sqrt{\frac{\alpha_m}{\mathcal{L}}}\frac{d_m k_p}{(d_m - d_p)}\left(\frac{d_p + d_\mu}{(d_p + \lambda_1)(d_p + \lambda_2)(d_p - i\omega)} - \frac{d_m + d_\mu}{(d_m + \lambda_1)(d_m + \lambda_2)(d_m - i\omega)}\right). \tag{S31}$$

The optimal WK filter can be calculated from (S26) by inserting (S30a) and (S31) which results

$$H_{\mathrm{wk}}(\omega) = \frac{d_m^2(d_m s_m - d_p s_p)}{\mathcal{L} d_m d_p (d_m - d_p)}\frac{(d_\mu - i\omega)(\frac{d_m d_p(s_m - s_p)}{d_m s_m - d_p s_p} - i\omega)}{(\lambda_1 - i\omega)(\lambda_2 - i\omega)} \tag{S32a}$$

$$s_m = \frac{d_p(d_m + d_\mu)}{(d_m + \lambda_1)(d_m + \lambda_2)}, \tag{S32b}$$

$$s_p = \frac{d_m(d_p + d_\mu)}{(d_p + \lambda_1)(d_p + \lambda_2)}. \tag{S32c}$$

The filter $H_{\rm wk}(\omega)$ contains two characteristic frequencies λ_1, λ_2. This differs from the functional form of $H(\omega)$ in (S19) which has a single characteristic frequency c_m. To evaluate how close the miRNA-mediated noise regulation to the WK optimality, we numerically compare the errors $E, E_{\rm wk}$ of both cases respectively. We get the error in estimation in WK optimality using (S24) in (S17) with $H(\omega) \to H_{\rm wk}(\omega)$

$$E_{\rm wk} = \frac{F_{\rm wk} - 1}{F_0 - 1} = \frac{d_m^2}{\mathcal{L}(d_m - d_p)^2}\left((d_m + d_p)\left(\frac{s_m^2}{d_p} + \frac{s_p^2}{d_m}\right) - 4 s_m s_p\right). \tag{S33a}$$

S.III. CALCULATION OF PHASE DIAGRAMS

A. Contour plots, fitness costs, and Michaelis-Menten constant

In the main text, we plot the contour diagram of $\log_{10} E$ from Eq. (S11) which informs about the noise reduction of the miRNA-regulated gene expression at a fixed protein output $\bar{p}$. The gene expression parameters are given in Table S1, whereas Michaelis-Menten constant $K_{\rm M}$ and miRNA levels $\mu_{\rm tot}$ are free parameters which control both noise reduction and vary metabolic costs of regulation. The metabolic costs of regulation are calculated by using (S14), and the definitions of energetic costs per mRNA and per miRNA respectively,

$$\epsilon_m = \zeta_P L_{m,\rm pre}, \qquad \epsilon_\mu = \zeta_P L_{\mu,\rm pri} \tag{S34}$$

where $L_{m,\rm pre}$ and $L_{\mu,\rm pri}$ are the nucleotide numbers of precursor mRNA and primary miRNA. In principle, the costs associated with the miRISC can be larger than ϵ_μ. Thus, our estimations can be seen as a lower bound. Finally, we relate metabolic costs to a fitness disadvantage, $s_c \sim \Delta M_\nu / M_{\rm tot}$ where $M_{\rm tot}$ is the (mean) total metabolic rate of the cells.

B. Minimal noise for the same metabolic cost

The contour plots indicate a minimum E for a given metabolic cost $\sim \Delta M_\nu$ at a precise $K_{\rm M}$ value which remain in a narrow range as a function of metabolic costs as shown in the Fig.2 of the main text. Taking the limit of $\phi \ll 1$, we can estimate this value by minimizing Eq. (S15) with respect to $\theta_\mu(K_{\rm M})$. Among multiple solutions, there is a single physical one, i.e., $0 \le \theta_\mu(K_{\rm M}^*) \le 1$ from which we can infer $K_{\rm M}^*$ using Eq.(S7). For high $\mathcal{M}$ and $\gamma_\mu \ll 1$, this yields the Eq.4 of the main text, i.e.,

$$K_{\rm M}^*(\Delta M_\nu, \bar{p}) \approx \bar{m}\gamma_\mu^{-1}\sigma_\epsilon^{-1/2}. \tag{S35}$$

This approximation holds well for the parameters in Table S1, with only 0.6% discrepancy at high $\mathcal{M}$.

In this regime and as $\sigma_\epsilon \ll 1$, we observe that $E(K_{\rm M}^*) \sim \mathcal{M}^{-1}$ which relates noise reduction to metabolic and hence to the fitness costs, i.e., $s_c \propto -M_\nu / E(K_{\rm M}^*)$.

Parameter	Definition	Value	Reference
d_m	mRNA degradation rate	0.11 hr^{-1} per mRNA	[5]
d_p	protein degradation rate	0.022 hr^{-1} per protein	[5]
d_μ	microRNA degradation rate	0.05 hr^{-1} per miRNA	[6, 7]
$d_{m\mu}$	mRNA degradation through miRISC	0.44 hr^{-1} per complex	
K_M	Michaelis-Menten const. for target degradation	solved for	
$L_{m,\mathrm{pre}}$	precursor mRNA nucleotide number	20,000	[8]
$L_{\mu,\mathrm{pri}}$	primary miRNA nucleotide number	2,500	[9, 10]
ζ_P	P bonds hydrolyzed per nucleotide recycling	2.17 P hr^{-1}	[4]
M_tot	total metabolic budget per cell	3×10^{11} P hr^{-1}	[4]
V	volume of cells	2,000 $\mu\mathrm{m}^3$	[11]
ϕ	d_p/d_m	0.2	
γ_m	$d_m/d_{m\mu}$	0.25	
γ_μ	$d_\mu/d_{m\mu}$	0.114	[12, 13]
$\bar{m}$	mean mRNA copy number	31	[5, 11]
$\bar{p}$	mean protein copy number	200,000	[5, 11]
σ_ϵ	ϵ_μ/ϵ_m	0.125	

TABLE S1. Gene expression parameters (metazoan) used in calculating the phase diagrams. Typical (or estimated) values from the corresponding references are taken such as median values over a distribution of genes. When an extensive distribution is not present, a reasonable estimation is made compatible with the references. We have not found a value for $d_{m\mu}$ in the literature, however, target mRNAs degrade faster in the complex, and hence we assumed the ratio to be within an order of magnitude. In order to convert the copy number units to molar units of mRNA numbers and K_m, we divide these values by the cell volume V and Avagadro constant.

S.IV. ESTIMATING DISSOCIATION CONSTANTS K_D FOR SEED SEQUENCES OF VARYING LENGTH

To estimate a range of possible K_D values for miRNAs with different seed sequences, we started with a collection of 9990 human miRNA seed sequences of length 7 nucleotides (nt) taken from the TargetScanHuman database, ver. 7.1 [15]. For each sequence, we used the RNAcofold algorithm from the ViennaRNA package (ver. 2.3) to calculate the free energy ΔG at $T = 298°\mathrm{K}$ for binding to a complementary target sequence [14]. To mimic the effects of unpaired nucleotides flanking the seed-target complex, we added two random unpaired nucleotides at both the beginning and end of the seed/target to make the total length 11 nt. The algorithm was run with the default dangling end energy option. Once ΔG is known, K_D can be found using the relation $\Delta G = k_B T \ln(K_\mathrm{D}/[1M])$.

To validate the predictions of the algorithm, we checked it on the seed sequence GAGGUAG, for which experimental K_D values are available for fruit fly and mouse siRNA-target complexes [16]. Depending on the dangling ends, the algorithm predicted a range of K_D with mean 63 pM (95% confidence interval: 7 to 238 pM). This compares well with the

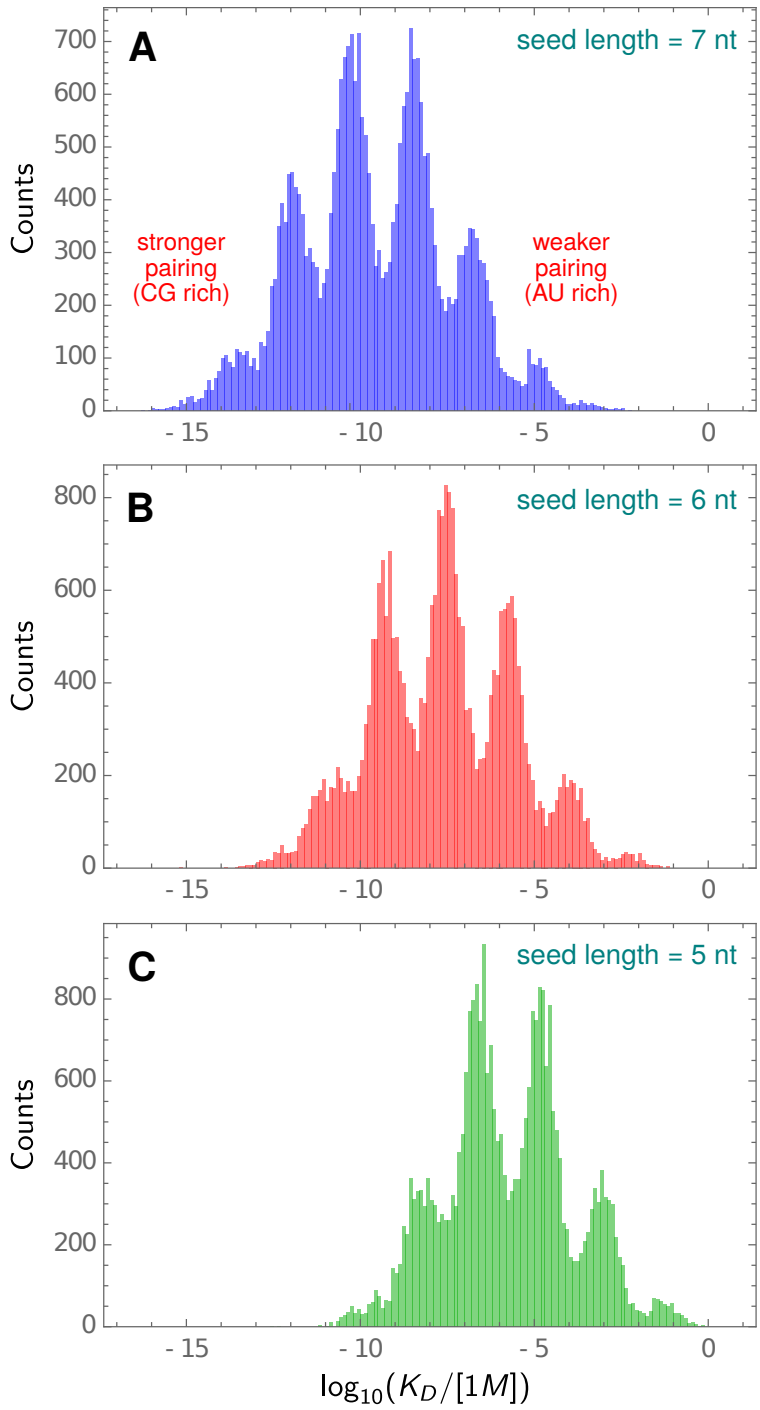

FIG. S1. Distribution of K_{D} values estimated for a different miRNA seed sequences with length: A) 7 nt; B) 6 nt; C) 5 nt. The calculation used the RNAcofold algorithm of the ViennaRNA package [14].

experimentally measured ranges for different seed-matched targets: 4 to 210 pM for fly, 13 to 26 pM for mouse [16].

A histogram of the results for all the seed sequences in the dataset is shown in Fig. S1A. The distribution has 8 peaks, roughly corresponding to the fact that the seed-target complex can have between 0 and 7 CG pairs. The more CG pairs relative to AU pairs, the stronger the binding and the lower the K_{D}. To model the effect of having a 6 nt seed, in Fig. S1B the calculation was redone after randomly deleting 1 nucleotide from each original seed. Similarly Fig. S1C shows the results for 5 nt seeds, after randomly deleting 2 nucleotides per seed. As expected, shorter seeds lead to overall weaker binding, shifting the distribution to larger K_{D} values.

---

[1] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).

[2] J. M. Schmiedel, S. L. Klemm, Y. Zheng, A. Sahay, N. Blüthgen, D. S. Marks, and A. van Oudenaarden, Science **348**, 128 (2015).

[3] S. Mukherji, M. S. Ebert, G. X. Zheng, J. S. Tsang, P. A. Sharp, and A. Van Oudenaarden, Nat. Genet. **43**, 854 (2011).

[4] M. Lynch and G. K. Marinov, Proc. Natl. Acad. Sci. **112**, 15690 (2015).

[5] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Nature **473**, 337 (2011).

[6] M. J. Marzi, F. Ghini, B. Cerruti, S. De Pretis, P. Bonetti, C. Giacomelli, M. M. Gorski, T. Kress, M. Pelizzola, H. Muller, *et al.*, Genome research **26**, 554 (2016).

[7] E. R. Kingston and D. P. Bartel, Genome Res. **29**, 1777 (2019).

[8] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell* (Garland Science, 2015).

[9] G. Song and L. Wang, PloS One **3**, e3574 (2008).

[10] H. K. Saini, A. J. Enright, and S. Griffiths-Jones, BMC Genom. **9**, 1 (2008).

[11] R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, Nucleic Acids Res. **38**, D750 (2010), http://bionumbers.hms.harvard.edu/.

[12] M. Jung, A. Schaefer, I. Steiner, C. Kempkensteffen, C. Stephan, A. Erbersdobler, and K. Jung, Clinical Chemistry **56**, 998 (2010).

[13] Y. K. Ho, W. T. Xu, and H. P. Too, PloS one **8**, e72463 (2013).

[14] R. Lorenz, S. H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, and I. L. Hofacker, Algorithms Mol. Biol. **6**, 1 (2011).

[15] V. Agarwal, G. W. Bell, J.-W. Nam, and D. P. Bartel, eLife **4**, e05005 (2015).

[16] L. M. Wee, C. F. Flores-Jasso, W. E. Salomon, and P. D. Zamore, Cell **151**, 1055 (2012).