nature structural & molecular biology

Efficiency and specificity in microRNA biogenesis

Omer Barad[1,5], Mati Mann[1,2,5], Elik Chapnik[1], Archana Shenoy[3,4], Robert Blelloch[3,4], Naama Barkai[1] & Eran Hornstein[1]

Primary microRNA cleavage by the Drosha–Dgcr8 'Microprocessor' complex is critical for microRNA biogenesis. Yet, the Microprocessor may also cleave other nuclear RNAs in a nonspecific manner. We studied Microprocessor function using mathematical modeling and experiments in mouse and human tissues. We found that the autoregulatory feedback on Microprocessor expression is instrumental for balancing the efficiency and specificity of its activity by effectively tuning Microprocessor levels to those of its pri-miRNA substrate.

Identification and cleavage of the primary miRNA (pri-miRNA) hairpin structures by the Drosha–Dgcr8 Microprocessor complex is a crucial step in miRNA biogenesis[1–4]. This cleavage distinguishes pri-miRNAs from any other RNA population in the nucleus. The Microprocessor has to comply with two main requirements. First, efficient pri-miRNA processing is critical because Microprocessor substrates, namely pri-miRNAs, can be directed to alternate routes (for example, transport to outside the nucleus or degradation). Second, undesired nonspecific targets ('off-targets') should be avoided, as most long RNAs in the nucleus can generate transient hairpin structures[5,6] but are not regularly cleaved[7]. Here we show that the previously described autoregulatory feedback mechanism, in which the Microprocessor cleaves *Dgcr8* mRNA[8–10], is crucial for regulating the interplay between efficient and specific cleavage of genuine pri-miRNA targets (**Fig. 1a**).

To characterize the interplay between the efficiency and specificity of Microprocessor activity, we used mathematical modeling. A canonical definition of enzyme activity compares the initial velocities for catalysis of two competing substrates, thus quantifying differences in specificity[11], whereas additional approaches account for both specificity and efficiency[12].

We defined a performance score S_p that increases with the probability, $P_{\text{pri-miRNA}}$, of cleaving a pri-miRNA substrate and decreases with the probability, $P_{\text{off-target}}$, of cleaving off target: $S_p = P_{\text{pri-miRNA}} \times (1 - P_{\text{off-target}})\ S_p$, which ranges from zero to one, is maximal when Microprocessor cleaves most of its true pri-miRNA targets and avoids most of its potential off-targets.

The cleavage probability, $P_{\text{substrate}}$, of a specific substrate depends on the unbound Microprocessor concentration [MP] and on the cleavage constant, $K_{\text{substrate}}$, which corresponds to the [MP] required to cleave 50% of a specific substrate (Online Methods): $P_{\text{substrate}} = \frac{[\text{MP}]}{K_{\text{substrate}} + [\text{MP}]}$. The cleavage constant of genuine substrates should be lower than that of off-targets, $K_{\text{pri-miRNA}} << K_{\text{off-target}}$, reflecting cleavage at lower Microprocessor concentration.

Therefore, when [MP] is low, specificity is maintained but cleavage efficiency is low, whereas high [MP] enables highly efficient yet poorly specific performance (**Fig. 1b**). Good performance, with high specificity and high efficiency, is obtained only for a narrow range of [MP] values. Therefore, in the absence of additional feedbacks, Microprocessor performance will be highly sensitive to most parameters in the system defining [MP] level, including the production and degradation rates of [MP] and the amount of miRNA substrates.

Recent studies have shown that the *Dgcr8* mRNA is cleaved by the Microprocessor itself. We hypothesized that this autoregulatory feedback[8–10] keeps Microprocessor performance robust to fluctuations under diverse biochemical parameters. Extending the model to account for this feedback, we discovered that S_p depends on two effective parameters: the total level of real targets, $[\text{Mir}]_T$, and the maximal level of Microprocessor, $[\text{MP}]_{max}$, obtained at the limit of negligible *Dgcr8* mRNA autoregulation (Online Methods).

To describe performance sensitivity to changes in $[\text{Mir}]_T$ and $[\text{MP}]_{max}$, we needed to first measure the cleavage constants. To estimate those constants, we compared RNA levels in wild-type mouse embryonic stem cells (ESCs) to RNA in *Drosha*-null and *Dgcr8*-null mouse ESCs, which lack Microprocessor activity[13–15]. Genome-wide data based on Affymetrix arrays[7] has shown that there is an overall elevation of pri-miRNAs in *Dgcr8*-null mouse ESCs. Furthermore, the expression of pri-miRNA remained unchanged in *Dicer1*-null mouse ESCs, ruling out potential unspecific secondary effects caused by miRNA depletion. Comparing different pri-miRNAs species, we found heterogeneity in pri-miRNAs cleavage rates (**Supplementary Fig. 1** and **Supplementary Table 1**), which can be partially explained by higher cleavage of pri-miRNAs that are embedded within polycistronic clusters (**Supplementary Fig. 2**). Because of the limited dynamic range of microarrays, we chose a set of representative monocistronic pri-miRNAs (miR-21 and miR-130a) or polycistronic pri-miRNAs (Let-7a and miR-16-1) for quantitative real-time PCR (qrtPCR) analysis. We found that RNA levels of these Microprocessor targets increased by 5- to 22-fold (average ~12-fold; **Fig. 1c**,**d**) in *Drosha*-null cells, corresponding to a relatively high average cleavage probability: $P_{\text{pri-miRNA}} \sim 0.9$. By contrast, the levels of mRNAs

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [2]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. [3]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, Center for Reproductive Sciences, University of California San Francisco, San Francisco, California, USA. [4]Department of Urology, University of California San Francisco, San Francisco, California, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to N.B. (naama.barkai@weizmann.ac.il) or E.H. (eran.hornstein@weizmann.ac.il).

Received 21 July 2011; accepted 4 April 2012; published online 13 May 2012; doi:10.1038/nsmb.2293

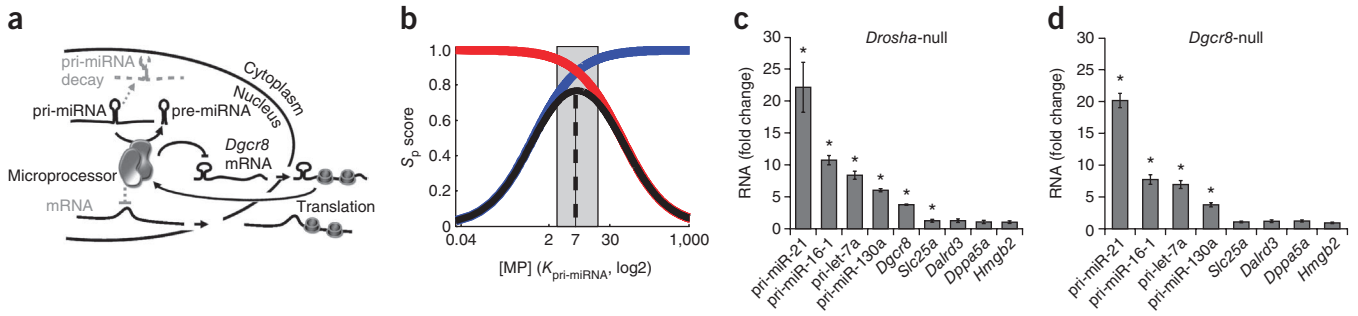

Figure 1 The interplay between efficient miRNA processing and minimal off-target cleavage. (**a**) A diagram for the kinetic model of Microprocessor activity. Efficient Microprocessor activity cleaves most of its real pri-miRNA targets, producing pre-miRNAs. Inefficient microprocessing may result in pri-miRNA decay. *Dgcr8* mRNA is translated in the cytoplasm, but a fraction of *Dgcr8* mRNA is degraded by Microprocessor autoregulatory feedback. Specific Microprocessor activity is defined by the avoidance of cleavage of undesired off-target mRNAs that should be translated in the cytoplasm. (**b**) The performance score S_p is defined by the probability that pri-miRNA substrates will be cleaved ($P_{pri\text{-}miRNA}$, blue) and the probability that nonspecific substrates will be avoided ($1 - P_{off\text{-}target}$, red). High specificity and efficiency are obtained only in a narrow range of [MP] values, allowing less than 2.5-fold deviation (gray box) from the values for maximal performance (dashed vertical line). Unbound Microprocessor levels are given in concentration units of $K_{pri\text{-}miRNA}$ (see Online Methods). (**c**) The expression of selected pri-miRNAs, *Dgcr8* mRNA and four additional Microprocessor mRNA targets in mouse ESCs depleted of *Drosha*, relative to wild-type control. Data are shown as mean ± s.d. ($n = 2$). (**d**) The expression of selected pri-miRNAs and four additional Microprocessor mRNA targets in mouse ESCs depleted of *Dgcr8*, relative to a wild-type control. Data are shown as mean ± s.d. ($n = 3$). Asterisks mark significant changes ($P < 0.05$).

previously reported to be cleaved by Microprocessor[13,14] increased by less than 25% ($P_{off\text{-}target} \sim 0.15$, **Fig. 1c**,**d** and **Supplementary Fig. 1**). This suggests that nuclear mRNAs make inferior Microprocessor substrates relative to pri-miRNAs, even though many of these off-target RNAs were abundantly expressed in mouse ESCs. The empirically calculated value for this ratio, $\frac{K_{off\text{-}target}}{K_{pri\text{-}miRNA}}$, is ~50. In contrast to other mRNAs, *Dgcr8* mRNA levels increased four-fold in *Drosha*-null mouse ESCs relative to control, consistent with reports that it is a genuine Microprocessor substrate[8–10] (**Fig. 1c**). This corresponds to cleavage probability $P_{Dgcr8} = 0.75$ and $\frac{K_{Dgcr8}}{K_{pri\text{-}miRNA}} \sim 3$. Thus, nuclear mRNAs make inferior Microprocessor substrates, whereas *Dgcr8* mRNA is at the lower end of the genuine substrate range.

We then numerically calculated the performance scores for different values of $[MP]_{max}$ and $[Mir]_T$ (**Fig. 2a**). This simulation revealed that a good performance score ($S_p > 0.7$) can be obtained for a wide range of parameters. Similarly, we derived the actual total Microprocessor level, $[MP]_T$, as a function of $[MP]_{max}$ and $[Mir]_T$ (**Fig. 2b** and Online Methods). It is noteworthy that we identified two modes of operation, which are both compatible with high Microprocessor performance ($S_p > 0.7$): mode 1, in which the total Microprocessor level is much higher than the total pri-miRNA level and Microprocessor is present in excess and mostly unbound to pri-miRNA substrate ($[MP]_T >> [Mir]_T$, **Fig. 2b**); and mode 2, in which the amount of Microprocessor approximates its substrate levels and it is mostly occupied by pri-miRNAs ($[MP]_T \sim [Mir]_T$, **Fig. 2b**). Notably, in both modes 1 and 2, high performance is maintained ($S_p > 0.7$) for a wide range of parameters—for example, if there were a two-fold change in $[MP]_{max}$ levels (**Fig. 2a** and **Supplementary Fig. 3**).

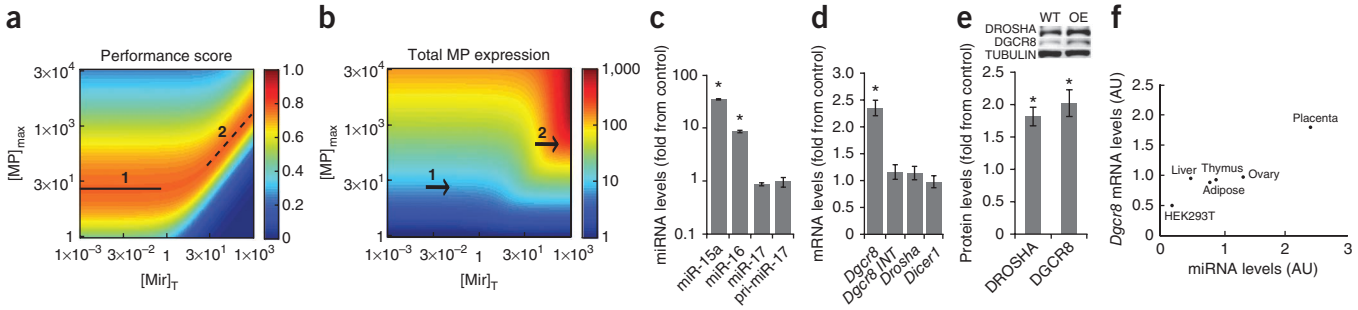

Figure 2 Microprocessor-dependent *Dgcr8* cleavage maintains robust balance of efficiency and specificity in Microprocessor activity. (**a**) Heat map exemplifying the performance score S_p as a function of normalized miRNA substrate level ($[Mir]_T$, *x* axis) and the maximal Microprocessor expression obtained in the absence of feedback ($[MP]_{max}$, *y* axis). S_p is high in two qualitatively different domains: mode 1, in which $[MP]_T >> [Mir]_T$ (black line), and mode 2, in which $[MP] \sim [Mir]_T$ (dashed line). (**b**) Heat map exemplifying the total Microprocessor level as a function of parameters, as in **a**. Changes in substrate concentration $[Mir]_T$ do not affect total Microprocessor levels in mode 1, as $[Mir]_T$ is in much excess over substrate levels. However, upregulation of pri-miRNA levels that are predicted to drive a proportional increase in Microprocessor is required in mode 2 (arrows). (**c**–**e**) mRNA and protein studies after overexpression of miR-15a and miR-16 in HEK293 cells. Shown are levels of mature miR-17 and pri-miR-17 (**c**) (data are shown as mean ± s.d., $n = 6$); levels of *Dgcr8*, a *Dgcr8* intron, *Drosha* and *Dicer1* mRNA (**d**) (data are shown as mean ± s.d., $n = 6$); and levels of DGCR8 and DROSHA protein (**e**) (data are shown as mean ± s.d., $n = 3$, $^*P < 0.05$). (**f**) *Dgcr8* mRNA expression level (*y* axis) is correlated with total miRNA expression[16] (*x* axis) in multiple human cell lines and tissues (each sample is a pool of 10–20 individual samples, Online Methods). AU, arbitrary units. MP, Microprocessor. Asterisks mark significant changes ($^*P < 0.05$).

We reasoned that the actual mode for activity can be revealed by examining experimentally whether total Microprocessor levels change in response to alterations in pri-miRNA levels (**Fig. 2b**). Our model provides testable predictions: if the Microprocessor function adheres to mode 1, then an increase in substrate levels will only have a marginal effect on $[MP]_T$. This is because Microprocessor levels in mode 1 greatly exceed substrate levels, so additional substrate will be buffered. This is in contrast to mode 2, for which substrate and Microprocessor levels are similar. An increase in $[Mir]_T$ will result in a corresponding increase in $[MP]_T$, to ensure that the level of unbound Microprocessor (and hence the performance score) remains robust. The bicistronic miR-15-16 represents one-twentieth of the total miRNA population in HEK293 cells[16]. We overexpressed miR-15-16 about 20-fold, leading to a two-fold increase in the total miRNA production (**Fig. 2c**). The processing of miR-17, an miRNA unrelated to miR-15-16, was not affected, supporting the notion that cleavage efficiency is robust to changes in substrate level. Quantifying *Dgcr8* by western blot analysis and by qrtPCR, we found that both protein and mRNA levels increased by ~2.2-fold (**Fig. 2d**,**e**). RNA levels of *Dicer1* and an intron of *Dgcr8* remained unchanged (**Fig. 2d**), indicating that the increase in *Dgcr8* was specific and driven by suppression of mRNA cleavage. *Drosha* protein levels were upregulated ~1.8-fold, whereas *Drosha* mRNA levels were unchanged (**Fig. 2e**), consistent with reported stabilization of *Drosha* by *Dgcr8* at the protein level[8].

Graded miR-15-16 overexpression correlated with stepwise upregulation of *Dgcr8* as well. Furthermore, overexpression of another miRNA, miR-17-18a, or overexpression of miRNAs in another cell line, namely HeLa, resulted in comparable upregulation of *Dgcr8* expression (**Supplementary Fig. 4**). In addition, in a set of human tissues, we discovered a strong correlation ($R^2 \sim 0.91$) of miRNA content[16,17] with *DGCR8* expression levels in these tissues (**Fig. 2f**), suggesting that *in vivo* Microprocessor levels are also attuned to pri-miRNA levels. We conclude that the Microprocessor is mostly bound to its substrate (mode 2 in **Fig. 2a**–**b**) and that Microprocessor robustness is plausibly maintained by autoregulatory feedback.

Our model suggests that Microprocessor activity will be robust to a two-fold change in the expression of its components. To test this, we quantified *Dgcr8* mRNA levels in a set of tissues harvested from a *Dgcr8* heterozygous mouse model and mouse ESCs[15] (Online Methods). The Microprocessor system is indeed robust, as *Dgcr8* expression changed by less than 20% (**Supplementary Fig. 5**), consistent with previous reports[18,19].

Based on the heterogeneity in pri-miRNA cleavage probability, we sought to uncover the limitations of Microprocessor robustness. We hypothesized that pri-miR-134 will be a poor substrate, as mature miR-134 was extremely sensitive to *Dgcr8* heterozygosity, in contrast to most other miRNAs[18,19]. Indeed, pri-miR-134 expression is upregulated only 2.2-fold upon Microprocessor knockout, corresponding to a relatively low cleavage probability ($P_{\text{pri-miR-134}} \sim 0.54$) (**Supplementary Fig. 5**).

In conclusion, we find that Microprocessor levels adjust to those of its pri-miRNA substrate, probably through the autoregulatory feedback on *Dgcr8* expression. We suggest that this ensures that the levels of unbound Microprocessor are kept at the range required for optimal activity, namely, at the range required for high efficiency and specificity.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank G.J. Hannon of the Cold Spring Harbor Laboratory (CSHL), F.V. Karginov (CSHL), D.R. Littman (New York University School of Medicine), R. Agami (Het Nederlands Kanker Instituut—Antoni van Leeuwenhoek) and N.V. Kim (Seoul National University) for reagents. We thank N. Shomron, D. Sprintzak, A. Horovitz, D.S. Tawfik and A. Eldar for comments on the manuscript. This work was supported by grants to E.H. from the Israel Science Foundation; the estates of F. Blau, L. Asseof, F. Sherr, C. Vener, J. Benattar, D. Francouer (Merck), Lilly Fulop, N. Baltor and M. Halphen; and the Wolfson family charitable trust. O.B. was supported by the Azrieli Foundation. E.H. is the incumbent of the Dr. Sydney Brenner Chair and the Helen and Milton A. Kimmelman Career Development Chair. N.B. is the incumbent of the Lorna Greenberg Scherzer Professorial Chair and was supported by a European Research Council Advanced Grant and a Helen and Martin Kimmel Award for Innovative Investigation. R.B. was supported by the California Institute for Regenerative Medicine (New Faculty Award RN2-00906-1) and the US National Institutes of Health (R01 NS057221).

AUTHOR CONTRIBUTIONS

Mathematical modeling was conducted by O.B. Experiments were conducted by M.M. Array data analysis and *Dgcr8* reagents were provided by A.S. and E.C. The study was designed and coordinated by O.B., M.M., N.B. and E.H. The paper was written by O.B., M.M., N.B. and E.H., with input from A.S. and R.B.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nsmb.2293.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Gregory, R.I. *et al. Nature* **432**, 235–240 (2004).
2. Han, J. *et al. Genes Dev.* **18**, 3016–3027 (2004).
3. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. & Hannon, G.J. *Nature* **432**, 231–235 (2004).
4. Bartel, D.P. *Cell* **116**, 281–297 (2004).
5. Bentwich, I. *et al. Nat. Genet.* **37**, 766–770 (2005).
6. Pedersen, J.S. *et al.* Rerer. *PLOS Comput. Biol.* **2**, e33 (2006).
7. Shenoy, A. & Blelloch, R. *PLoS ONE* **4**, e6971 (2009).
8. Han, J. *et al. Cell* **136**, 75–84 (2009).
9. Triboulet, R., Chang, H.M., Lapierre, R.J. & Gregory, R.I. *RNA* **15**, 1005–1011 (2009).
10. Kadener, S. *et al. RNA* **15**, 537–545 (2009).
11. Fersht, A.W.H. *Enzyme Structure and Mechanism* 2nd edn. (Freeman, San Francisco, 1985).
12. Savir, Y. & Tlusty, T. *Mol. Cell* **40**, 388–396 (2010).
13. Chong, M.M. *et al. Genes Dev.* **24**, 1951–1960 (2010).
14. Karginov, F.V. *et al. Mol. Cell* **38**, 781–788 (2010).
15. Wang, Y., Medvid, R., Melton, C., Jaenisch, R. & Blelloch, R. *Nat. Genet.* **39**, 380–385 (2007).
16. Lee, E.J. *et al. RNA* **14**, 35–42 (2008).
17. Liang, Y., Ridzon, D., Wong, L. & Chen, C. *BMC Genomics* **8**, 166 (2007).
18. Stark, K.L. *et al. Nat. Genet.* **40**, 751–760 (2008).
19. Schofield, C.M. *et al. Neural Dev.* **6**, 11 (2011).

ONLINE METHODS

Cell culture. HEK-293T and HeLa cells (American Type Culture Collection) were maintained in DMEM with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 U/ml streptomycin at 37 °C, in a 5% CO_2 humidified incubator.

miRVec pMSCV vectors for Pri-miR-15a-16 and pri-miR-17-18a are the kind gift of R. Agami. For overexpression, miRNAs were transiently transfected with Lipofectamine 2000 (Life technologies), according to the manufacturer's instructions.

Analysis of human and mouse tissues. Tissues harboring monoallelic expression of *Dgcr8* (ref. 15) were harvested from heterozygous offspring of *Dgcr8*$^{loxp/loxp}$;PGK-Cre transgenes. FirstChoice human total RNA human tissue RNA was purchased from Ambion. *DGCR8* mRNA in human tissues was measured by qrtPCR and normalized to *DROSHA* mRNA in the same tissue. Total microRNA in each tissue was estimated, based on qrtPCR for all mature microRNA[16], and total miRNA and *DGCR8* were normalized to their corresponding amounts in HEK293 and scaled to their average values.

RNA expression analysis. Total RNA extracted with miRNeasy Mini Kit and synthesis of cDNA was carried out using the miScript PCR System (Qiagen). All qrtPCR analyses were conducted for at least three technical repeats. Expression analysis of wild-type, Dgcr8 knockout and Dicer1 knockout mouse ESCs were carried out using 1.0 mouse gene ST microarrays[7] (Affymetrix) (GEO accession number GSE16923).

Primer sequences. Primer sequences were generated by Integrated DNA Technologies.

Primer name	Sequence
U6	GATGACACGCAAATTCGTGAA
B-actin for	ACAGAGCCTCGCCTTTGCCG
B-actin rev	ACCCATGCCCACCATCACGC
HPRT for	ACTTTGCTTTCCTTGGTCAGGCAGT
HPRT rev	CGTGGGGTCCTTTTCACCAGCA
Drosha for	TCGCCACCGCAGCTACGAAC
Drosha rev	CCAGCAGGTTCAGGAACAACCGA
Dicer1 for	GCAGAAGCTTGGGCATGCTGTGA
Dicer1 rev	CAGGCCTGCCATGCTGAGGG
Dppa5a for	AAACTTCCTGAACCTGGAGCTGTGGG
Dppa5a rev	CAGCACCAGCGACTGGACCTGGAATA
slc25a42 for	CCTCTGCCTCCCAAGTGCTGGGATTA
slc25a42 rev	CCTGCTCAGGTCCTGCATGTGAGATG
Dalrd3 for	TCAGTACTACAGGCTCAGACATGCTCAG
Dalrd3 rev	CAAGGCGGGCACAGTTATACATGACAA
hmgb2 for	TGAGATGTGGTCTGAGCAATCTGCCAA
hmgb2 rev	CTTTGAGCCTGTTGGCCTACCAGGAC
Dgcr8 for	GCTGAGTGCATTGTGATTTCCA
Dgcr8 probe	TAATTGAGGCAGTGGTTCT
Dgcr8 rev	TGGCCACATTGCTCTTTTCA
mmu-pri-miR-21 for	ATGGCTGTACCACCTTGTCGGATA
mmu-pri-miR-21 probe	AGCAGTCGATGGGCTGTCTGACATTT
mmu-pri-miR-21 rev	TTGACTGCAAACCATGATGCTGGG
mmu-pri-mir-16-1 for	AGCTCCTATGATAGCAATGTCAGCGG
mmu-pri-mir-16-1 probe	GCCTTAGCAGCACGTAAATATTGGCG
mmu-pri-mir-16-1 rev	TATTGCCAACCTTACTTCAGCAGCA
mmu - pri-let7a for	AAAGGAGAACGGCTTCCTGTGATG
mmu - pri-let7a probe	ACCCTGGATGTTCTCTTCACTGTGGGAT
mmu - pri-let7a rev	TAGTTATCTCCCAGTGGTGGGTGT
mmu-pri-miR-130a for	TTGCTGGGAAGGAAATGAGGACGA
mmu-pri-miR-130a probe	TGAGTGTGGCCAGGGACTGGGAGAAA
mmu-pri-miR-130a rev	TTGCACTGCTCGGTACACGTTAGA
hsa-Pri-miR-17 for	GCATCTACTGCAGTGAAGGCACTTGT
hsa-Pri-miR-17 probe	ATTATGGTGACAGCTGCCTCGGGAA
hsa-Pri-miR-17 rev	GCCCTGCACTTTAAAGCCCAACTT
hsa-miR-15a	TAGCAGCATAATGGTTTGTG
hsa-miR-16	TAGCAGCACGTAAATATTGGC
hsa-miR-17	CAAAGTGCTTACAGTGCAGGTAG

Western blotting. The following antibodies were used for western blotting of protein extract from whole-cell lysate: rabbit anti-DGCR8 (1:500, 10996-1-AP; Proteintech Group), rabbit anti-DROSHA (1:200, ab12286, Abcam Cambridge) and mouse anti-TUBULIN (1:1000, Sigma-Aldrich).

Software. Data processing was carried out using MATLAB 2010 (MathWorks). Images were created with MATLAB 2010, Adobe Illustrator and Microsoft Excel Office 2007.

Mathematical modeling. *Cleavage probability and performance score.* The cleavage probability $P_{substrate}$ of a substrate can be calculated by comparing the relative rates of the two independent and competing processes that act on the RNA substrate in the nucleus. On the one hand, RNA may be cleaved at a rate μ[MP] that depends (linearly) on the available free (unbound) MP and on some parameter, μ. This parameter describes the rate by which MP binds the specific RNA and is determined, among other ways, by the structure affinity to MP[20] and the number of embedded hairpins within the RNA. Alternatively, the RNA can also be degraded or exported from the nucleus at a rate λ. Therefore, the cleavage probability is given by

$$P_{\text{substrate}} = \frac{\mu[\text{MP}]}{\lambda + \mu[\text{MP}]} \quad (1)$$

We denote the [MP] required to cleave 50% of a particular substrate as the cleavage constant, $K = \lambda/\mu$. For simplicity, we scale the concentration units throughout the paper by the average cleavage constant of pri-miRNA targets, $K_{\text{pri-miRNA}}$. Thus, unbound MP is replaced by a normalized term $[\text{MP}]_{\text{norm}} = \frac{[\text{MP}]}{K_{\text{pri-mRNA}}}$.

Equation (1) is valid for all MP targets, with different targets varying only by the respective cleavage constant K. Therefore, the cleavage probabilities of pri-miRNAs can be written in terms of $[\text{MP}]_{\text{norm}}$ only

$$P_{\text{pri-mRNA}} = \frac{[\text{MP}]_{\text{norm}}}{1+[\text{MP}]_{\text{norm}}} \quad (2)$$

Off-target cleavage probability can be written in terms of $[\text{MP}]_{\text{norm}}$, given the ratio $\eta = \frac{K_{\text{off-target}}}{K_{\text{pri-mRNA}}}$ between the cleavage constant for the pri-miRNA target and the cleavage constant for the off-target

$$P_{\text{off-target}} = \frac{[\text{MP}]_{\text{norm}}}{\eta+[\text{MP}]_{\text{norm}}} \quad (3)$$

Similarly, the *Dgcr8* mRNA cleavage probability can be written in terms of $[\text{MP}]_{\text{norm}}$, given the ratio $\rho = \frac{K_{\text{Dgcr8}}}{K_{\text{pri-miRNA}}}$ between the cleavage constant for the pri-miRNA target and for *Dgcr8*

$$P_{\text{Dgcr8}} = \frac{[\text{MP}]_{\text{norm}}}{\rho+[\text{MP}]_{\text{norm}}} \quad (4)$$

With these definitions for the cleavage probabilities of pri-miRNAs and off-targets (equations (2) and (3)), the equation for defining the performance score, $S_p = P_{\text{pri-miRNA}}(1 - P_{\text{off-target}})$, can be rewritten as

$$S_p = \frac{[\text{MP}]_{\text{norm}}}{1+[\text{MP}]_{\text{norm}}} \times \frac{1}{1+[\text{MP}]_{\text{norm}}/\eta} \quad (5)$$

Empirical data were used to calculate the RNA cleavage probability and cleavage constant ratios.

At steady state, the amount of a specific RNA that is transcribed with a rate β and is eliminated with a rate λ is given by β/λ. Thus, the relative RNA expression

comparing two conditions with different elimination rates, with MP = μ[MP] + λ or without MP = λ, is given by

$$\frac{\text{RNA (no MP)}}{\text{RNA (with MP)}} = \frac{\mu[\text{MP}]+\lambda}{\lambda} \quad (6)$$

The first term in equation (6) can be empirically derived through qrtPCR measurements of RNA in the presence or absence of MP by comparing RNA levels in wild-type mouse ESCs to RNA in *Drosha*-null and/or to RNA in *Dgcr8*-null mouse ESCs, both of which lack MP activity.

Then, using equation (1), these values enable calculation of specific RNA cleavage probabilities

$$P = 1 - \frac{\text{RNA(WT)}}{\text{RNA(no MP)}} \quad (7)$$

The result is also obtained when the alternative process does not lead to elimination of the RNA but to its export to the cytoplasm. In this case, MP depletion will increase the RNA amount because all of its population will be exported instead of just the fraction that is not cleaved by the MP, $\left(1-P=\frac{\lambda}{\mu_i[\text{MP}]+\lambda}\right)$, leading to the same ratio as in equation (1).

By measuring the cleavage probabilities of pri-miRNAs, off-targets and *Dgcr8* mRNAs at the same time, under a certain condition, we can solve equations (2–4) together and extract the cleavage constant ratios

$$\eta = \frac{P_{\text{pri-miRNA}}}{P_{\text{off-target}}}\left(\frac{1-P_{\text{off-target}}}{1-P_{\text{pri-miRNA}}}\right) \quad (8)$$

$$\rho = \frac{P_{\text{pri-miRNA}}}{P_{\text{Dgcr8}}}\left(\frac{1-P_{\text{Dgcr8}}}{1-P_{\text{pri-miRNA}}}\right) \quad (9)$$

A model for the autoregulatory feedback on MP activity. DGCR8 protein is produced from Dgcr8 mRNA in the cytoplasm. D_{mc} is defined as the level of Dgcr8 mRNA in the cytoplasm and is derived from Dgcr8 mRNA transcripts that were not cleaved by the Microprocessor in the nucleus, using 1 – P_{Dgcrs} (see equation (4)). Thus, D_{mc} dynamics is described by the equation

$$\frac{dD_{\text{mc}}}{dt} = \beta_{D_{\text{m}}}\left(\frac{1}{1+\text{MP}_{\text{norm}}/\rho}\right) - \lambda_{D\text{mc}}D_{\text{mc}} \quad (10)$$

where $\beta_{D_{\text{m}}}$ is the transcription rate of the Dgcr8 mRNA (D_{m}) and $\lambda_{D_{\text{mc}}}$ is the degradation rate of the cytoplasmic Dgcr8 mRNA (D_{mc}).

MP levels are determined by *DGCR8*, which stabilizes *DROSHA* at the protein level[8], such that MP_{T} dynamics is given by

$$\frac{d(\text{MP}_{\text{T}})}{dt} = \beta_D D_{\text{mc}} - \lambda_{\text{MP}}\text{MP}_{\text{T}} \quad (11)$$

where β_D is the *Dgcr8* translation rate, and λ_{MP} is the degradation rate of MP.

At steady state, the level of *Dgcr8* mRNA is given by $D_{\text{mc}} = \frac{\beta_{D_{\text{m}}}}{\lambda_{D_{\text{mc}}}}\left(\frac{1}{1+\text{MP}_{\text{norm}}/\rho}\right)$, using equation (10), and from equation (11), we get the steady-state level of total MP expression

$$\text{MP}_{\text{T}} = \frac{\beta_D\beta_{D_{\text{m}}}}{\lambda_{\text{MP}}\lambda_{D_{\text{mc}}}}\left(\frac{1}{1+\text{MP}_{\text{norm}}/\rho}\right) \quad (12)$$

In general, the total expressed Microprocessor can be found in either unbound or free MP states. Bound MP can be associated with the following RNAs: miRNA hairpins, [[MP]M]; targets on mRNA (that is, off-target hairpins), [[MP]O]; or even nonhairpin RNA structures, which it cannot digest (that is, k_{cat} = 0), [[MP]R]. MP_{T} is calculated by summing these terms

$$\text{MP}_{\text{T}} = \text{MP} + [[\text{MP}]M] + [[\text{MP}]O] + [[\text{MP}]R] \quad (13)$$

We assume that the amount of nonhairpin RNA exceeds by far the hairpin-structured RNA in the nucleus. Under commonly accepted quasi steady-state conditions, we assume that the concentration of the intermediate complex does not change on the time scale of product formation . Therefore, we can approximate [[MP]R] ∝ R_{T}MP. Additionally, it is reasonable to assume that the majority of the hairpin structures that MP is bound to are pri-miRNA. Thus, we can further approximate the total MP as

$$\text{MP}_{\text{T}} \sim \text{MP}(1+\delta) + [[\text{MP}]M] \quad (14)$$

where $\delta > 0$ measures the fraction of MP that is bound to nonhairpin structures. Defining β_{M} as the rate by which miRNA hairpins are transcribed (total rate) and using equation (2), we can estimate the average number of processed miRNA hairpins per unit time as $\beta_M\frac{\text{MP}_{\text{norm}}}{1+\text{MP}_{\text{norm}}}$. Thus, at steady state, the MP fraction that is bound to miRNA hairpins is given by

$$[[\text{MP}]M] = \frac{\beta}{k_{\text{cat}}} \times \frac{\text{MP}_{\text{norm}}}{1+\text{MP}_{\text{norm}}} \quad (15)$$

Thus, using equations (12), (14) and (15), we obtain the steady-state equation for the unbound MP

$$\frac{\beta_D\beta_{D_{\text{m}}}}{\lambda_{\text{MP}}\lambda_{D_{\text{mc}}}}\left(\frac{1}{1+\text{MP}_{\text{norm}}/\rho}\right) = (1+\delta)\text{MP} + \frac{\beta_M}{k_{\text{cat}}} \times \frac{\text{MP}_{\text{norm}}}{1+\text{MP}_{\text{norm}}} \quad (16)$$

Next, we define two effective concentration parameters in the system. First, we define $[\text{MP}]_{\text{max}} = \frac{\beta_D\beta_{D_{\text{m}}}}{\lambda_{\text{MP}}\lambda_{D_{\text{mc}}}} \times \frac{1}{(1+\delta)}$, which represents (in the limit of $\delta \to 0$) the maximal level of [MP], obtained at the limit of no negative feedback on *Dgcr8* mRNA expression ($[\text{MP}]_{\text{max}}$). Second, we define $[\text{Mir}]_{\text{T}} = \frac{\beta_M}{k_{\text{cat}}} \times \frac{1}{(1+\delta)}$, which represents (in the limit of $\delta \to 0$) the total level of real targets ($[\text{Mir}]_{\text{T}}$). Using the scaled concentration (in the units of $K_{\text{pri-miRNA}}$), equation (16) is transformed to

$$[\text{MP}]_{\text{T max}}\left(\frac{1}{1+[\text{MP}]_{\text{norm}}/\rho}\right) = [\text{MP}]_{\text{norm}} + [\text{Mir}]_{\text{T}} \times \frac{[\text{MP}]_{\text{norm}}}{1+[\text{MP}]_{\text{norm}}} \quad (17)$$

In order to generate **Figure 2a,b**, equation (17) was numerically solved for varying $[\text{Mir}]_{\text{T}}$ and $[\text{MP}]_{\text{T max}}$ to calculate $[\text{MP}]_{\text{norm}}$, and hence the performance score S_{p}, using equation (5), and MP_{T} , using equation (12).

System robustness as a function of autoregulatory feedback stringency. To study in more detail the role of stringent feedback on *Dgcr8* expression, we considered a family of models with different numbers of equivalent hairpins, by modifying the feedback term $\left(\frac{1}{1+\frac{n[\text{MP}]_{\text{norm}}}{2\rho}}\right)_{n=2}$ in equation (17) to other n values. When n, denoting the hairpins, equals zero, then the MP system is modeled without feedback, whereas n = 1 describes weaker feedback with a single hairpin. Noting that there is experimental evidence for the dominance of one of the hairpins on *Dgcr8* mRNA, for simplicity, we consider all hairpins to be similarly accessible to the MP for $n > 1$. Notably, robust system response to variations in $[\text{MP}]_{\text{max}}$ and $[\text{Mir}]_{\text{T}}$, and hence to all other biochemical parameters, including substrate levels, substantially increased with n: 1.2-fold change in $[\text{MP}]_{\text{max}}$ for n = 0, 1.8-fold change for n = 1 and 2.2-fold change for n = 2, for operation mode 2 (**Supplementary Fig. 3b,c**). Increasing n further will slightly increase system robustness but will also require that most *Dgcr8* be cleaved (**Supplementary Fig. 3d**).

20. Han, J. *et al. Cell* **125**, 887–901 (2006).

doi:10.1038/nsmb.2293