Quantitative design of cell type-specific mRNA stability from microRNA expression data

Lukas Oesinghaus[1]*, Sebastian Castillo-Hair[1], Nicole Ludwig[2], Andreas Keller[2], Georg Seelig[1,3]*

Affiliations:

[1]Department of Electrical & Computer Engineering, University of Washington, Seattle, WA, United States

[2]Department for Clinical Bioinformatics, Center for Bioinformatics, PharmaScienceHub, Saarland University, Germany

[3]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA

*Corresponding authors: loesing@uw.edu, gseelig@uw.edu

Abstract: Limiting expression to target cell types is a longstanding goal in gene therapy, which could be met by sensing endogenous microRNA. However, an unclear association between microRNA expression and activity currently hampers such an approach. Here, we probe this relationship by measuring the stability of synthetic microRNA-responsive 3'UTRs across 10 cell lines in a library format. By systematically addressing biases in microRNA expression data and confounding factors such as microRNA crosstalk, we demonstrate that a straightforward model can quantitatively predict reporter stability purely from expression data. We use this model to design constructs with previously unattainable response patterns across our cell lines. The rules we derive for microRNA expression data selection and processing should apply to microRNA-responsive devices for any environment with available expression data.

Main text: The ability to predictably sense and respond to specific cell states promises to increase the specificity of mRNA, gene, and cell therapies (*1–4*). In principle, cell state-sensing can be achieved by engineering cis-regulatory elements that respond to transcription factors, microRNA, or RNA binding proteins that are differentially expressed in the target cell state. Practically, we usually lack the ability to quantitatively predict sensor function from available prior information, which tends to be limited to expression data for trans-acting RNAs and proteins that serve as inputs to the sensor, making it difficult to generalize beyond cellular environments compatible with iterative experimental testing.

So far, maybe the most widespread approach to transgene targeting relies on microRNA (miRNA) (*5*, *6*). miRNA are short (~22nt) regulatory RNA bound by the RNA-induced silencing complex (RISC), which actively search for and degrade complementary targets, usually in the 3'UTR of mRNAs (*7*). miRNA are attractive for engineering applications because regulation can be achieved simply by inserting target sites complementary to a miRNA into the transgene. Although natural miRNA regulation uses partial target sites, fully complementary target sites confer a stronger regulatory effect, making them the more popular choice. Due to their often cell type-specific expression, miRNA are well-suited to classify cell state via synthetic gene circuits (*8–11*). Most commonly, multiple target sites for a highly and differentially expressed miRNA are used to exclude transgene expression in a specific tissue (*12*, *13*).

Still, it remains difficult to achieve complex expression patterns that require multiple target sites for miRNA that are expressed in several tissues because there is no method to quantitatively predict endogenous miRNA activity, i.e., the loss of expression due to target mRNA degradation, from expression data. Previous high-throughput studies have found low correlation between miRNA expression and target stability beyond a minimum expression threshold necessary for regulation (*14*, *15*). Smaller studies find low, intermediate, and high correlations (*5*, *16–18*). miRNA expression and activity are thus generally believed to be relatively weakly associated (*6*, *10*, *11*). How multiple miRNA targets combine to produce an overall regulatory effect is similarly contested (*9*, *15*, *18*). This necessitates individual validation of miRNAs and makes it difficult to anticipate behavior in novel environments (*10*).

To address these limitations, we measure the activity of all annotated high-confidence human miRNAs (*19*, *20*) across ten different cell lines using reporter libraries. We find that it is possible to quantitatively predict the stability of reporter genes containing arbitrary combinations of miRNA targets in their 3'UTR purely from readily collectible expression data, but only after identifying high-quality microRNA expression datasets and correcting them for systematic bias. Using this model, it is possible to quantitatively design mRNA with defined stability patterns across cell types without the need for individual verification or high-throughput activity measurements, which are infeasible in real tissues.

A universal transfer function predicts full miRNA target sites

To measure the regulatory activity of human miRNA, we created reporters for 1,382 miRNAs selected from two major miRNA databases (miRBase (*19*), MirGeneDB (*20*)) (**Fig. 1A**). Single, fully complementary miRNA target sites were inserted into the 3'UTR of a plasmid-based reporter gene (**Fig. S1**). The reporter library was transiently transfected into 10 cell lines, each of which is characterized by a unique miRNA expression profile. Stability of reporter transcripts in each cell line was inferred from the ratio of mRNA to input plasmid counts as determined by sequencing (*21*). The resulting relative stability values are normalized such that the stability for constructs with inactive target sites in the 3'UTR context used (the 'main context') is 1.

Our experiments reveal a nearly monotonic, universal relationship between miRNA expression obtained via microarray (*22*) and stability (**Fig. 1B**). As an example, measurements of the miR-100-5p reporter across cell lines reveal varying stability values, and aggregating these across all miRNAs enables mapping the full regulation function. Experimental replicates have a Pearson r^2 above 0.98, suggesting that the assay is highly reproducible (**Fig. S2-S4**). High correlation between mRNA abundance and reporter protein expression measured with flow cytometry is consistent with a model wherein full target sites in the 3'UTR primarily modulate mRNA stability rather than translation (r^2=0.95, **Fig. S5**). Different miRNAs have varying levels of experimental evidence in databases. The monotonic relationship only clearly holds for miRNAs that are listed in the manually curated MirGeneDB, while low-confidence miRNAs in miRBase sometimes display erroneously (*22*) high measured expression with no activity (**Fig. S6**). We therefore exclude them from the analysis.

The observed relationship between mRNA and miRNA abundance is well explained by a simple transfer function derived under the assumptions of mass action kinetics and instantaneous degradation of the target transcript by the RISC (r^2 = 0.44-0.69, **Fig. 1B inset, Supplementary Text 1**) (*23*, *24*). For low miRNA concentrations ($[x] \ll k_{deg}/k_{on}$) mRNA degradation is

dominated by the baseline degradation rate k_{deg}. In the opposite regime, the degradation rate increases linearly with miRNA concentration. Empirically, mRNA stability starts decreasing around 1,000 miRNA transcripts per million (tpm), in line with previous results (*14*). To confirm that our observations are not an artifact of the specific reporter context used, we embedded 10 miRNA targets in 30 additional 3'UTR sequences (Methods), yielding 3,000 measurements across all cell lines. The transfer function correctly predicts the relative stability due to miRNA target sites across the different contexts (**Fig. 1C**). Different context sequences could occlude target sites by secondary structure formation, thereby causing outliers. We engineered additional context sequences to test this hypothesis and found that, while strong occlusion of the target site strongly reduces miRNA activity (**Fig. S7**), secondary structure in natural 3'UTRs is usually too weak to create strong outliers (**Fig. S8**).

We observe mostly minor but systematic global deviations from the transfer function for some cell lines (**Fig. S9**), likely due to variations in total miRNA levels and the main context stability k_{deg} between cell lines (**Fig. S10**, Methods). A scaling factor of the total miRNA concentration for each cell line accounts for both of these issues and improves the model fit (**Fig. S11**).

Combining datasets corrects data bias

The universal relationship we observe between miRNA expression and reporter activity seems partially at odds with previous high-throughput experiments (*14*, *15*) and in fact, we only observe a universal relationship when using certain miRNA measurements. We hypothesized that the correlation between stability and expression is determined by the method used to collect the miRNA expression data. To test this hypothesis, we collected expression datasets of different origins, including short RNA sequencing and microarrays from Agilent and Affymetrix, for our cell lines and tested their mutual correlation (**Fig. S12A-C**). We additionally include datasets from what we refer to as "improved sequencing methods", which involve approaches to reduce ligation bias in library construction (*25–29*). High correlation between different collection methods is only observed between Agilent microarray data and improved sequencing methods. We hypothesize that cross-method correlation signifies correctness: We calculated r^2 values for our transfer function (**Fig. 1D** and **S12D**) and found consistently high correlation even for older publications using Agilent microarray data (r^2~0.62) and for improved sequencing methods (r^2~0.52). In contrast, standard small RNA sequencing datasets perform poorly (r^2 0.11 to 0.41).

Even the highest quality microarray and sequencing data still suffer from technology-specific biases in which certain miRNAs are over- or undercounted. To identify putative technology-specific outliers we combine microarray data (*22*) with improved sequencing data. Because the sample preparation shares little in common, technical biases between these two different types of measurements should not correlate. Potential biased data points are defined as outliers from the transfer function in one but not the other dataset for at least one cell line (**Fig. S13**). Because technology-specific bias depends only on miRNA sequence features, we only identify consistent outliers across multiple cell lines as genuine biased data points. Bias-aware dataset merging is performed by using the geometric mean except in cases of technology-specific bias, where the data from the other technology is used. Dataset merging removes around a third of outliers with a more than two-fold difference to the predicted stability value. Because biases are technology-specific, this identification of false positives and negatives in highly tractable systems such as cell lines should be transferable to more difficult systems such as tissues.

Crosstalk between miRNA family members explains outliers from ideal behavior

MicroRNAs occur in families with highly similar sequences, leading to potential crosstalk (**Fig. S14**). If a microRNA degrades a reporter transcript carrying targets that are not fully complementary, such crosstalk can result in apparent deviation from the transfer function model. To quantify crosstalk, we measured a reporter library of 1,096 partially mutated miRNA target sites (**Fig. S15A**). Results for individual mismatches qualitatively agree with previous observations (**Fig. 1E**, **S15B-C**): An adenine at position 1 does not act as a mismatch (*30*). The strongest effects are observed early in the seed and around the cleavage site at positions 10 and 11 (*31*, *32*). Wobble base pairs are generally more tolerated than genuine mismatches but are still damaging (*32*).

We then used these measurements to build a model to predict the impact of multiple mutations. We classify the impact of individual mutations as high, medium, low and none based on their position only and whether the mutation causes a mismatch or a wobble base pair with the miRNA. Targets with multiple mutations are then grouped by the number and impact class of their individual mutations (**Fig. S15D-E**). Activity rapidly falls off with increasing numbers of mutations and multiple mutations can abrogate cleavage even for combinations of mutations with no individual effect (**Fig. S15E, S16A**). We used the classification of individual mutations to build a regression tree model that predicts the effect of multiple mutations (**Fig. S16B**). The model predicts the impact of mutations on a held-out miRNA (**Fig. S16C**), although the effect of mutations is highly sequence-specific (**Fig. S16D**). We merge miRNAs that are identical for the first 18 bases and then use the mutation classification (Methods) to filter targets that are likely to experience crosstalk (**Fig. 1F**), removing many of the strongest outliers.

Bias correction yields an updated model

Several additional sources of bias and noise can largely be excluded: Low counts in the stability data and miRNA GC content have little to no effect (**Fig. S17A-D**). Some miRNA targets contain homopolymer stretches but except for miR-3613-3p, these do not impact the result (**Fig. S17E**). Each step of bias removal improves the model fit (**Fig. 1G**, **Fig. S18**). The effect of scaling miRNA concentrations is minor for all cell lines except HUH7 and PC3. Merging data sources and addressing crosstalk substantially improves the fit for almost all cell lines. Tera1 and Jeg3 improve the least, which likely indicates biological variation in the miRNA composition within these cell lines. Bias removal substantially improves the agreement between the transfer function and measured stabilities (**Fig. 1H**). We call the model without and with scaling factors and bias-aware merging the "baseline model" and the "updated model".

Multiple target sites are described by an additive model of miRNA concentrations

Next, we investigated how multiple target sites cooperate to control reporter stability. Multiple models of miRNA interactions have been suggested in the literature (*9*, *15*, *33*). Here, we focus on the additive model, which assumes that target sites are independent of each other, because it arises naturally from the same differential equations that yielded our transfer function (**Fig. 2A**, **Supplementary Text 1**). We chose 100 miRNAs spanning a range of expression values and created reporters by repeating their target sites two to six times. In the additive model, the behavior should be predicted by multiplying the miRNA expression data by the repeat number and using the transfer function to predict stability. This approach yields accurate predictions (r^2~0.8) with similar accuracy for all target copy numbers (**Fig. 2B**, **S19-20**). An alternative

model which assumes that regulation is independent of the number of repeats leads to systematic overprediction.

To investigate the behavior of multiple target sites for different miRNAs, we tested 500 combinations of 2 to 6 different miRNA targets chosen to best distinguish different models (Methods). We see an overall r^2 of 0.76 (**Fig. 2C**, left panel) for the correlation between the additive model and measured data, confirming that the additive model also works well for multiple different miRNAs. A prediction based simply on the strongest target site systematically overpredicts the stability (**Fig. S21**). Although stability prediction works well, there are still outliers. To distinguish between effects not considered in the additive model and inaccuracies in our expression data, we used the stability data for single target sites to estimate the true miRNA expression levels (inverted transfer function model, see methods). This approach reduces some of the deviation from ideal behavior in the prediction (**Fig. 2C**, right panel).

As a further test of the additive model, we chose 30 different sets of 5 targets and generated 15 randomly shuffled versions of their position in the 3'UTR. If secondary structure, cooperative effects, proximity to the coding region, or other factors matter, shuffling of target site position should change the measured stability. If the additive model is correct, the position should make no difference. We compare the mean stability of the 15 sites with the individual measured stabilities (**Fig. 2D**). With an r^2 of 0.95, we can conclude that the position along the UTR is not relevant. Overall, the updated model prediction accuracy from expression data comes close to matching and sometimes exceeds that of the inverted transfer function model, which requires difficult-to-collect activity data, across all cell lines for both repeats and combinations (**Fig. 2E**).

Although the additive model works well for most miRNA combinations, there are notable outliers, especially in the case of repeated target sites for the same miRNA. First, some miRNAs show a pattern where an even number of repeats has less activity than an odd number (**Fig. S22A-B**). This is likely due to self-complementarity of the target, which causes strong target-target interactions and prevents miRNA binding via secondary structure formation (**Fig. S22C-D**). Unbiased discovery of all occluded targets via $\Delta\Delta G$ calculations is difficult due to the low accuracy of secondary structure prediction for mRNAs in cells (**Fig. S22E-F**) (*34*). Second, some miRNA target sites lead to a nearly monotonic increase in stability with increasing target numbers (**Fig. S23A**). This effect occurs more strongly in cell lines where the cognate miRNA is not expressed, indicating that this is not an effect of non-canonical miRNA regulation (**Fig. S23B-C**). We suspect RBPs as a potential cause, although identification of concrete candidates is difficult using current prediction methods.

Model-based design of 3'UTRs with tailored stability profiles

We used our model to design 3'UTRs with defined stability patterns based purely on miRNA expression data (**Fig. 3A**). The chosen stability patterns represent tasks of varying difficulty given the repressive nature of miRNA regulation. We test “binary” designs aiming for low expression (stability 0) in one or more target cell lines and high expression (stability 1) everywhere else, binary designs resulting in high expression in the target cell lines (stability 1) and low expression in all others (stability 0), and graduated expression patterns, where the goal is to achieve stabilities between 0 and 1 in each cell line (**Fig. 3B**). The baseline model was used to generate the designs. This reflects the real-world challenge of creating designs for tissues and cell types for which miRNA expression data is available but for which we have no reporter

activity measurements. We used an evolutionary algorithm to select four to six target sites based on the weighted mean squared error (mse) between the target and the predicted pattern. We created designs for either a subset of six or all cell lines based on the notion that the design should be easier for a smaller number of target cell lines or for all 10 cell lines. In total, we generated 1,782 designs for six cell lines and 1,986 designs for all cell lines, respectively (**Fig. S24**).

For the binary designs, we experimentally tested four 3'UTRs per design target type using 4, 5, or 6 miRNA target sites (**Fig. S25-27**). The best-performing designs generally come close to achieving the desired pattern (**Fig. 3C**). Easier design tasks lead to better results: Most designs that suppress stability in one or two cell lines have relatively little off-target knockdown, but constraining expression to one or two cell lines or suppressing activities in three cell lines is much more challenging. Still, we observe deviations from the target pattern even for the best designs. We next asked whether these deviations are predicted by our models or whether they are purely random. First, we note that measurements and predictions generally agree (**Fig. S25-27**). Approximately 70% of predictions have a distance of 0.2 or less to the measured value (**Fig. 3D**).

The difficulty of the design types is anticipated by the target-prediction root mean square deviation (rmsd) (**Fig. S28A**). Designs with five or more target sites lead to significantly better predicted and measured results for binary designs constraining expression to a single cell line but not for binary designs inactivating expression (**Fig. S28B**). **Fig. 3E** shows an example prediction for restraining activity to a single cell type. Many of the most prominent deviations from the desired pattern, e.g., off-target activity in JEG3 for the second design, are predicted by the model. Two other model types further increase the prediction accuracy (updated and inverted transfer function, **Fig. S28C-E**). These other two models represent how well we can predict stability given knowledge of total miRNA concentrations and background stability (updated model) and with access to a prior high-throughput assay of miRNA activity in the same system (inverted transfer function model). Notably, all models predict performance not just for well-performing designs (**Fig. S29**) but also for poorly performing ones (**Fig. S28A, S30**).

We experimentally tested 1,116 graduated designs targeting all cell lines, achieving varying levels of success in generating the desired patterns (**Fig. S31**). Comparing across quartiles of design success, we find that for the first three quartiles, the predicted rmsd values track the measured rmsd values for all models. For the fourth quartile, the baseline model loses accuracy (**Fig. 3F**). Predictions and measurements in individual cell lines have an r^2 of 0.6 for the updated model (**Fig. S32A**) and 75% of predictions deviate less than 0.2 from the measured value (**Fig. S32B**). 84% of designs have a mean absolute prediction error smaller than 0.2 across cell lines (**Fig. S32C**). As for the binary model, we can anticipate the feasibility of achieving a certain stability pattern and the success or failure of an individual design is often predicted in advance (**Fig. S32D**).

Thus, all models predict reasonably well whether a certain design goal will be fulfilled, although the two expression data-based models overpredict design success overall (**Fig. 3G**). Large unexpected failures also occur (**Fig. S33**). Strong target occlusion due to target-target interactions is again one potential cause. In the most extreme cases, sites from the 5p and 3p arm of related miRNAs are present on the same 3'UTR, leading to a systematic underprediction of the stability. Such designs can easily be excluded by prohibiting strong secondary structure or target-

target interactions. A few designs exhibit overall very high or very low stabilities for unknown reasons.

Targeting mRNA stabilities to human tissues

Next, we asked what sort of mRNA expression patterns are possible in real tissues (**Fig. 3H**). We chose a microarray (*35*) and an improved sequencing dataset (*36*) whose collection method matches our cell line expression data (**Fig. S34**). The two datasets show little correlation before consistent outliers are addressed (**Fig. S35**). We merged the two datasets similarly as for our cell line data (**Fig. S36**). We then used the merged dataset to generate designs with between 1 and 8 target sites that either exclude or constrict expression to a single tissue (**Fig. 3H, S37A-B**). While exclusion in some tissues with highly specific miRNAs, e.g. miR-122-5p in liver, can often be achieved by a single target site that could be chosen without a model, constraining expression to a single tissue requires many target sites and a careful balancing of on-target activity with unavoidable off-target activity that cannot be achieved without a quantitative model (**Fig. S37C-E**). Even for exclusion of a single tissue, the choice of miRNA is often not obvious and the optimal number of target sites is rarely the four repeats that are often used. Although we could not experimentally test these designs, the good agreement we observed between model and predictions in cell lines is encouraging.

Discussion

In summary, we find that a universal transfer function predicts the stability of miRNA-responsive 3'UTRs from total miRNA levels as calculated by an additive model. Universality here means that the regulatory effect is a function of the miRNA concentration but is largely independent of miRNA identity and of the cellular context. The choice of miRNA expression data is key and the universal function becomes apparent only when using high-quality miRNA expression datasets. Conversely, our reporter stability data provide an easy way of benchmarking collection methods by measuring expression for our used cell lines and seeing whether the universal relationship is retained. Systematic treatment of biases increases the prediction accuracy: The expression data should be scaled to total miRNA levels, strong secondary structure eliminated, outliers removed by combining data sources, and crosstalk excluded. Our results suggest that earlier conflicting results on the relationship between miRNA levels and activity can be resolved by careful selection and processing of the expression data. Given the ubiquitous use of miRNA regulation, this insight should be highly useful for gene therapy and cellular control circuits.

We designed hard-to-achieve patterns such as restriction to a single cell line or graduated expression. Importantly, we show that such patterns are achievable even with a limited number of target sites, resulting in compact 3’UTRs. We expect such designs to be practically important for increasing the specificity of future mRNA therapies where few other levers for achieving specificity are available. However, the simple nature of miRNA regulation, namely that it is repressive and non-cooperative, still constrains the achievable patterns. Incorporating our insights into the design of multi-gene circuits of miRNA-regulated mRNAs could address this limitation (*9*).

Going beyond cell lines, we have created designs for tissue datasets. Whether our model performs well for real tissues remains to be seen. The heterogeneity of cell types in a single tissue could pose a significant hurdle because tissue-averaged miRNA expression profiles might

not be representative of any single cell type within the tissue. Future single-cell miRNA sequencing techniques (*37*) could help overcome this limitation, though only if they avoid the biases of standard library construction workflows.

Further improvements in prediction accuracy seem possible. Biological variation in miRNA levels within the same cell type, sequence-specific RISC loading and activities (*38*), poorly predicted secondary structure (*34*), titration of miRNAs by other transcripts (*14*, *39*), subcellular localization (*14*), post-transcriptional miRNA modification (*40*), or miRNA-independent UTR stability variation (*41*) likely all cause residual prediction noise. Given the enormous impact of expression data quality, residual technical bias probably also plays a large role. We here focus on fully complementary target sites relevant for engineering applications. While biologically relevant seed target sites require more complex models that take into account the secondary structure and miRNA sequence features (*42*), any predictive model will also require debiased expression data as an accurate proxy for concentrations, making the elimination of residual biases equally relevant for basic biology.

In the longer term, we expect that the approach introduced here, i.e., learning quantitative relationships between the levels of trans-regulatory molecules and their targets that allow generalization from one cellular context to the other, can also be applied to RNA binding proteins or transcription factors, enabling the rational design of functional cis-regulatory elements given knowledge of only transregulator expression.

References and Notes:

1. G. M. Allen, W. A. Lim, Rethinking cancer targeting strategies in the era of smart cell therapeutics. *Nat. Rev. Cancer* **22**, 693–702 (2022).
2. I. I. Taskiran, K. I. Spanier, H. Dickmänken, N. Kempynck, A. Pančíková, E. C. Ekşi, G. Hulselmans, J. N. Ismail, K. Theunis, R. Vandepoel, V. Christiaens, D. Mauduit, S. Aerts, Cell-type-directed design of synthetic enhancers. *Nature* **626**, 212–220 (2024).
3. J. Zhang, K. Salaita, Smart Nucleic Acids as Future Therapeutics. *Trends Biotechnol.* **39**, 1289–1307 (2021).
4. A. P. Teixeira, M. Fussenegger, Synthetic macromolecular switches for precision control of therapeutic cell functions. *Nat. Rev. Bioeng.*, doi: 10.1038/s44222-024-00235-9 (2024).
5. B. D. Brown, B. Gentner, A. Cantore, S. Colleoni, M. Amendola, A. Zingale, A. Baccarini, G. Lazzari, C. Galli, L. Naldini, Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. *Nat. Biotechnol.* **25**, 1457–1467 (2007).
6. B. Dhungel, C. Ramlogan-Steel, J. Steel, MicroRNA-Regulated Gene Delivery Systems for Research and Therapeutic Purposes. *Molecules* **23**, 1500 (2018).
7. L. F. R. Gebert, I. J. MacRae, Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).
8. Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333**, 1307–1311 (2011).
9. P. Mohammadi, N. Beerenwinkel, Y. Benenson, Automated Design of Synthetic Cell Classifier Circuits Using a Two-Step Optimization Strategy. *Cell Syst.* **4**, 207-218.e14 (2017).
10. B. Angelici, L. Shen, J. Schreiber, A. Abraham, Y. Benenson, An AAV gene therapy

computes over multiple cellular inputs to enable precise targeting of multifocal hepatocellular carcinoma in mice. *Sci. Transl. Med.* **13**, eabh4456 (2021).
11. L. Wang, W. Xu, S. Zhang, G. C. Gundberg, C. R. Zheng, Z. Wan, K. Mustafina, F. Caliendo, H. Sandt, R. Kamm, R. Weiss, Sensing and guiding cell-state transitions by using genetically encoded endoribonuclease-mediated microRNA sensors. *Nat. Biomed. Eng.*, doi: 10.1038/s41551-024-01229-z (2024).
12. R. Jain, J. P. Frederick, E. Y. Huang, K. E. Burke, D. M. Mauger, E. A. Andrianova, S. J. Farlow, S. Siddiqui, J. Pimentel, K. Cheung-Ong, K. M. McKinney, C. Köhrer, M. J. Moore, T. Chakraborty, MicroRNAs Enable mRNA Therapeutics to Selectively Program Cancer Cells to Self-Destruct. *Nucleic Acid Ther.* **28**, 285–296 (2018).
13. S. E. Sinnett, E. Boyle, C. Lyons, S. J. Gray, Engineered microRNA-based regulatory element permits safe high-dose mini *MECP2* gene therapy in Rett mice. *Brain* **144**, 3005–3019 (2021).
14. G. Mullokandov, A. Baccarini, A. Ruzo, A. D. Jayaprakash, N. Tung, B. Israelow, M. J. Evans, R. Sachidanandam, B. D. Brown, High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. *Nat. Methods* **9**, 840–846 (2012).
15. J. J. Gam, J. Babb, R. Weiss, A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity. *Nat. Commun.* **9**, 2430 (2018).
16. O. Flores, E. M. Kennedy, R. L. Skalsky, B. R. Cullen, Differential RISC association of endogenous human microRNAs predicts their inhibitory potential. *Nucleic Acids Res.* **42**, 4629–4639 (2014).
17. A. Kozomara, S. Hunt, M. Ninova, S. Griffiths-Jones, M. Ronshaugen, Target Repression Induced by Endogenous microRNAs: Large Differences, Small Effects. *PLoS ONE* **9**, e104286 (2014).
18. I. Vainberg Slutskin, S. Weingarten-Gabbay, R. Nir, A. Weinberger, E. Segal, Unraveling the determinants of microRNA mediated regulation using a massively parallel reporter assay. *Nat. Commun.* **9**, 529 (2018).
19. A. Kozomara, M. Birgaoanu, S. Griffiths-Jones, miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
20. B. Fromm, D. Domanska, E. Høye, V. Ovchinnikov, W. Kang, E. Aparicio-Puerta, M. Johansen, K. Flatmark, A. Mathelier, E. Hovig, M. Hackenberg, M. R. Friedländer, K. J. Peterson, MirGeneDB 2.0: the metazoan microRNA complement. *Nucleic Acids Res.* **48**, D132–D141 (2020).
21. D. Griesemer, J. R. Xue, S. K. Reilly, J. C. Ulirsch, K. Kukreja, J. R. Davis, M. Kanai, D. K. Yang, J. C. Butts, M. H. Guney, J. Luban, S. B. Montgomery, H. K. Finucane, C. D. Novina, R. Tewhey, P. C. Sabeti, Genome-wide functional screen of 3′UTR variants uncovers causal variants for human disease and evolution. *Cell* **184**, 5247-5260.e19 (2021).
22. J. Alles, T. Fehlmann, U. Fischer, C. Backes, V. Galata, M. Minet, M. Hart, M. Abu-Halima, F. A. Grässer, H.-P. Lenhof, A. Keller, E. Meese, An estimate of the total number of true human miRNAs. *Nucleic Acids Res.* **47**, 3353–3364 (2019).
23. J. Martinez, T. Tuschl, RISC is a 5′ phosphomonoester-producing RNA endonuclease. *Genes Dev.* **18**, 975–980 (2004).
24. J. A. Broderick, W. E. Salomon, S. P. Ryder, N. Aronin, P. D. Zamore, Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. *RNA*

17, 1858–1869 (2011).
25. A. D. Jayaprakash, O. Jabado, B. D. Brown, R. Sachidanandam, Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res.* **39**, e141–e141 (2011).
26. R. T. Fuchs, Z. Sun, F. Zhuang, G. B. Robb, Bias in Ligation-Based Small RNA Sequencing Library Construction Is Determined by Adaptor and RNA Structure. *PLOS ONE* **10**, e0126049 (2015).
27. P. Mestdagh, N. Hartmann, L. Baeriswyl, D. Andreasen, N. Bernard, C. Chen, D. Cheo, P. D'Andrade, M. DeMayo, L. Dennis, S. Derveaux, Y. Feng, S. Fulmer-Smentek, B. Gerstmayer, J. Gouffon, C. Grimley, E. Lader, K. Y. Lee, S. Luo, P. Mouritzen, A. Narayanan, S. Patel, S. Peiffer, S. Rüberg, G. Schroth, D. Schuster, J. M. Shaffer, E. J. Shelton, S. Silveria, U. Ulmanella, V. Veeramachaneni, F. Staedtler, T. Peters, T. Guettouche, L. Wong, J. Vandesompele, Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat. Methods* **11**, 809–815 (2014).
28. M. D. Giraldez, R. M. Spengler, A. Etheridge, P. M. Godoy, A. J. Barczak, S. Srinivasan, P. L. De Hoff, K. Tanriverdi, A. Courtright, S. Lu, J. Khoory, R. Rubio, D. Baxter, T. A. P. Driedonks, H. P. J. Buermans, E. N. M. Nolte-'t Hoen, H. Jiang, K. Wang, I. Ghiran, Y. E. Wang, K. Van Keuren-Jensen, J. E. Freedman, P. G. Woodruff, L. C. Laurent, D. J. Erle, D. J. Galas, M. Tewari, Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. *Nat. Biotechnol.* **36**, 746–757 (2018).
29. P. M. Godoy, A. J. Barczak, P. DeHoff, S. Srinivasan, A. Etheridge, D. Galas, S. Das, D. J. Erle, L. C. Laurent, Comparison of Reproducibility, Accuracy, Sensitivity, and Specificity of miRNA Quantification Platforms. *Cell Rep.* **29**, 4212-4222.e5 (2019).
30. D. P. Bartel, MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136**, 215–233 (2009).
31. S. M. Elbashir, Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO J.* **20**, 6877–6888 (2001).
32. L. M. Wee, C. F. Flores-Jasso, W. E. Salomon, P. D. Zamore, Argonaute Divides Its RNA Guide into Domains with Distinct Functions and RNA-Binding Properties. *Cell* **151**, 1055–1067 (2012).
33. R. J. Bloom, S. M. Winkler, C. D. Smolke, A quantitative framework for the forward design of synthetic miRNA circuits. *Nat. Methods* **11**, 1147–1153 (2014).
34. J. Zhang, Y. Fei, L. Sun, Q. C. Zhang, Advances and opportunities in RNA structure experimental determination and computational modeling. *Nat. Methods* **19**, 1193–1207 (2022).
35. N. Ludwig, P. Leidinger, K. Becker, C. Backes, T. Fehlmann, C. Pallasch, S. Rheinheimer, B. Meder, C. Stähler, E. Meese, A. Keller, Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* **44**, 3865–3877 (2016).
36. A. Keller, L. Gröger, T. Tschernig, J. Solomon, O. Laham, N. Schaum, V. Wagner, F. Kern, G. P. Schmartz, Y. Li, A. Borcherding, C. Meier, T. Wyss-Coray, E. Meese, T. Fehlmann, N. Ludwig, miRNATissueAtlas2: an update to the human miRNA tissue atlas. *Nucleic Acids Res.* **50**, D211–D221 (2022).
37. S. M. Hücker, T. Fehlmann, C. Werno, K. Weidele, F. Lüke, A. Schlenska-Lange, C. A. Klein, A. Keller, S. Kirsch, Single-cell microRNA sequencing method comparison and application to cell lines and circulating lung tumor cells. *Nat. Commun.* **12**, 4316 (2021).

38. E. Goh, K. Okamura, Hidden sequence specificity in loading of single-stranded RNAs onto *Drosophila* Argonautes. *Nucleic Acids Res.* **47**, 3101–3116 (2019).
39. R. Denzler, S. E. McGeary, A. C. Title, V. Agarwal, D. P. Bartel, M. Stoffel, Impact of MicroRNA Levels, Target-Site Complementarity, and Cooperativity on Competing Endogenous RNA-Regulated Gene Expression. *Mol. Cell* **64**, 565–579 (2016).
40. L. Tomasello, R. Distefano, G. Nigita, C. M. Croce, The MicroRNA Family Gets Wider: The IsomiRs Classification and Role. *Front. Cell Dev. Biol.* **9**, 668648 (2021).
41. V. Agarwal, D. R. Kelley, The genetic and biochemical determinants of mRNA degradation rates in mammals. *Genome Biol.* **23**, 245 (2022).
42. S. E. McGeary, K. S. Lin, C. Y. Shi, T. M. Pham, N. Bisaria, G. M. Kelley, D. P. Bartel, The biochemical basis of microRNA targeting efficacy. *Science* **366**, eaav1741 (2019).
43. D. Briskin, P. Y. Wang, D. P. Bartel, The biochemical basis for the cooperative action of microRNAs. *Proc. Natl. Acad. Sci.* **117**, 17764–17774 (2020).
44. M. E. Fornace, J. Huang, C. T. Newman, N. J. Porubsky, M. B. Pierce, N. A. Pierce, NUPACK: Analysis and Design of Nucleic Acid Structures, Devices, and Systems. [Preprint] (2022). https://doi.org/10.26434/chemrxiv-2022-xv98l.
45. T. Fehlmann, F. Kern, O. Laham, C. Backes, J. Solomon, P. Hirsch, C. Volz, R. Müller, A. Keller, miRMaster 2.0: multi-species non-coding RNA sequencing analyses at scale. *Nucleic Acids Res.* **49**, W397–W408 (2021).
46. B. Muzellec, M. Teleńczuk, V. Cabeli, M. Andreux, PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *Bioinformatics* **39**, btad547 (2023).
47. H. Jin, C. Zhang, M. Zwahlen, K. Von Feilitzen, M. Karlsson, M. Shi, M. Yuan, X. Song, X. Li, H. Yang, H. Turkez, L. Fagerberg, M. Uhlén, A. Mardinoglu, Systematic transcriptional analysis of human cell lines for gene expression landscape and tumor representation. *Nat. Commun.* **14**, 5417 (2023).
48. T. Xu, N. Su, L. Liu, J. Zhang, H. Wang, W. Zhang, J. Gui, K. Yu, J. Li, T. D. Le, miRBaseConverter: an R/Bioconductor package for converting and retrieving miRNA name, accession, sequence and family information in different versions of miRBase. *BMC Bioinformatics* **19**, 514 (2018).
49. C. Backes, F. Sedaghat-Hamedani, K. Frese, M. Hart, N. Ludwig, B. Meder, E. Meese, A. Keller, Bias in High-Throughput Analysis of miRNAs and Implications for Biomarker Studies. *Anal. Chem.* **88**, 2088–2095 (2016).

Acknowledgements: We thank Jesse Bloom, Paul Nghiem, and Andrew Hsieh for providing cell lines. We thank Shusruto Rishik for assistance with miRNA expression data analysis. We thank the members of the Seelig lab for discussions on this manuscript.

Funding: EMBO Postdoctoral Fellowship Program to L.O., NSF Award 2021552 to G.S., NIH Award R56HG013312 to G.S., NIH Award R01GM149631 to G.S., DFG Award 469073465 to A.K..

Author contributions: L.O. and S.C.H. collected the high-throughput stability data. N. L. and A.K. collected the miRNA expression data generated for this study. A.K. gave advice on the analysis and interpretation of the results. L.O. designed the constructs and analyzed the data. L.O., S.C.H., and G.S. designed the project. L.O. and G.S. wrote the manuscript with input from the other authors.

Competing interests: G.S is a co-founder of Parse Biosciences and an advisor to Deep Genomics and Sanofi.

Data and materials availability: Raw sequencing reads and processed files for the measured cell line microRNA expression data will be made available on GEO upon publication of this manuscript. Raw stability sequencing reads, count data, and processed miRNA expression data will be made available on GEO upon publication of this manuscript. Files specifying library sequences, used microRNA target sites, processed stability data, and other relevant information necessary to reproduce the results are provided as **Supplementary Data**. The code used to analyze the data in this manuscript is available on GitHub at https://github.com/loesinghaus/miRNA.design.2024.

Supplementary Materials

Materials and Methods

Supplementary Text

Figs. S1 to S37

Tables S1 to S9

References (43-49)

Data S1 to S4

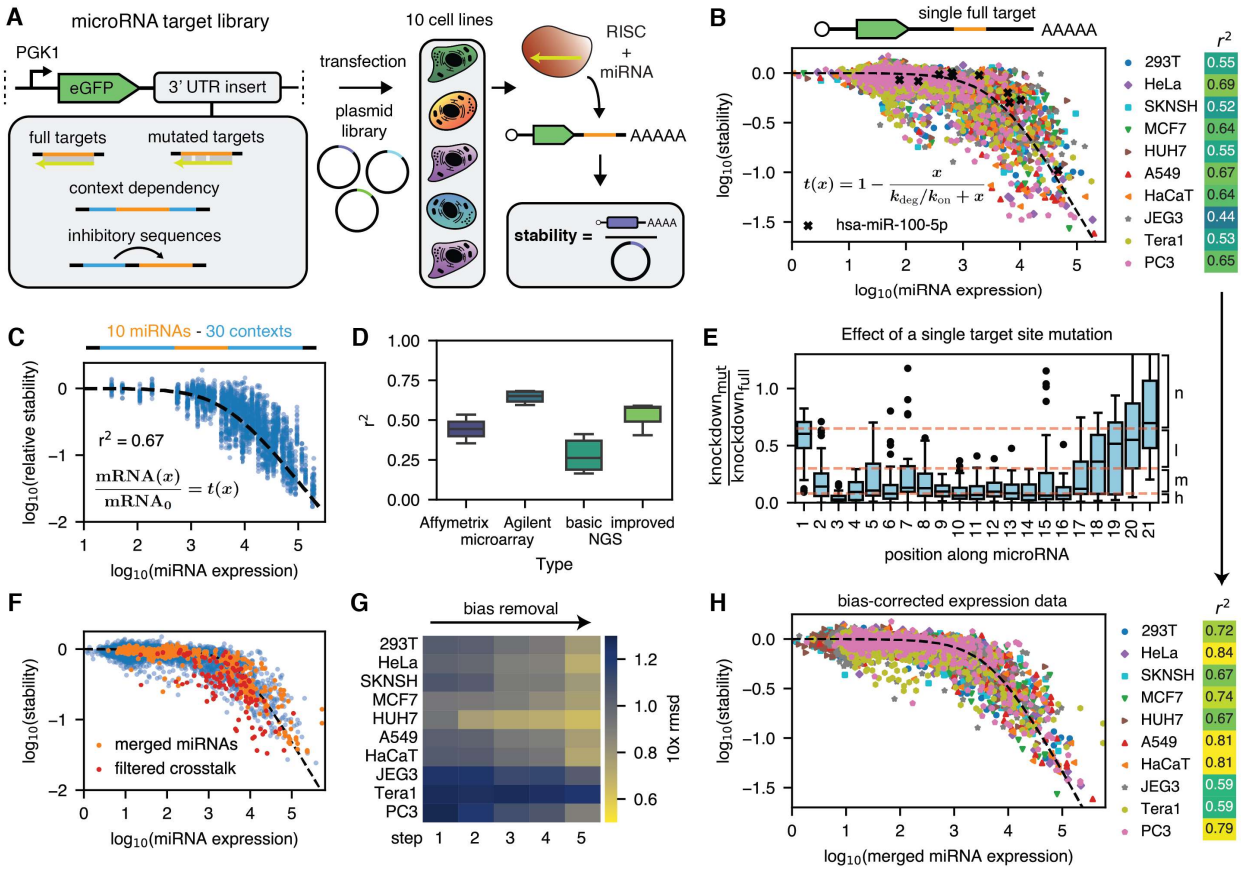


Fig. 1. A universal transfer function predicts miRNA activity from expression data. (**A**) Overview of the performed high-throughput stability assay. MicroRNA target sites are ordered as an oligo pool and cloned into a reporter plasmid. The plasmid library is transiently transfected into cell lines, where constructs are degraded by the RNA-induced silencing complex (RISC). Purified mRNA is sequenced. Sequencing counts for mRNA and the plasmid library are used to calculate stability values. (**B**) Stabilities for single full target sites across the different cell lines. The dashed line shows the best fit for the function shown in the inset. The x-axis uses unscaled microarray expression data (*22*). The heatmap shows Pearson r^2 values between the fit and the measured data. (**C**) The transfer function predicts the relative stability of 3'UTRs containing miRNA targets in different context sequences. (**D**) Pearson r^2 derived from fitting the transfer function for different miRNA data source types. (**E**) Effect of a single mutation on the relative knockdown between mutated and full target sites for miRNAs where the stability s_{full} of the non-mutated target site is less than 1/3. The knockdown k is given by $k = 1/s - 1$. n: no l: low, m: medium, h: high impact mutations (**F**) Potential crosstalk was identified based on the mutation classification in (E). The expression of near-identical miRNAs was merged (orange) and other miRNA targets that are likely to experience crosstalk (red) are filtered. (**G**) Changes in the root mean square deviation (rmsd) between measured data and the fit due to expression data bias removal. 1: unscaled microarray data, 2: scaled microarray data, 3: combination of microarray and sequencing data, 4: bias-aware merging, 5: removal of crosstalk. (**H**) Stability for single full target sites using the merged and cross-filtered expression data. The heatmap on the right demonstrates the improvements in the model fit.

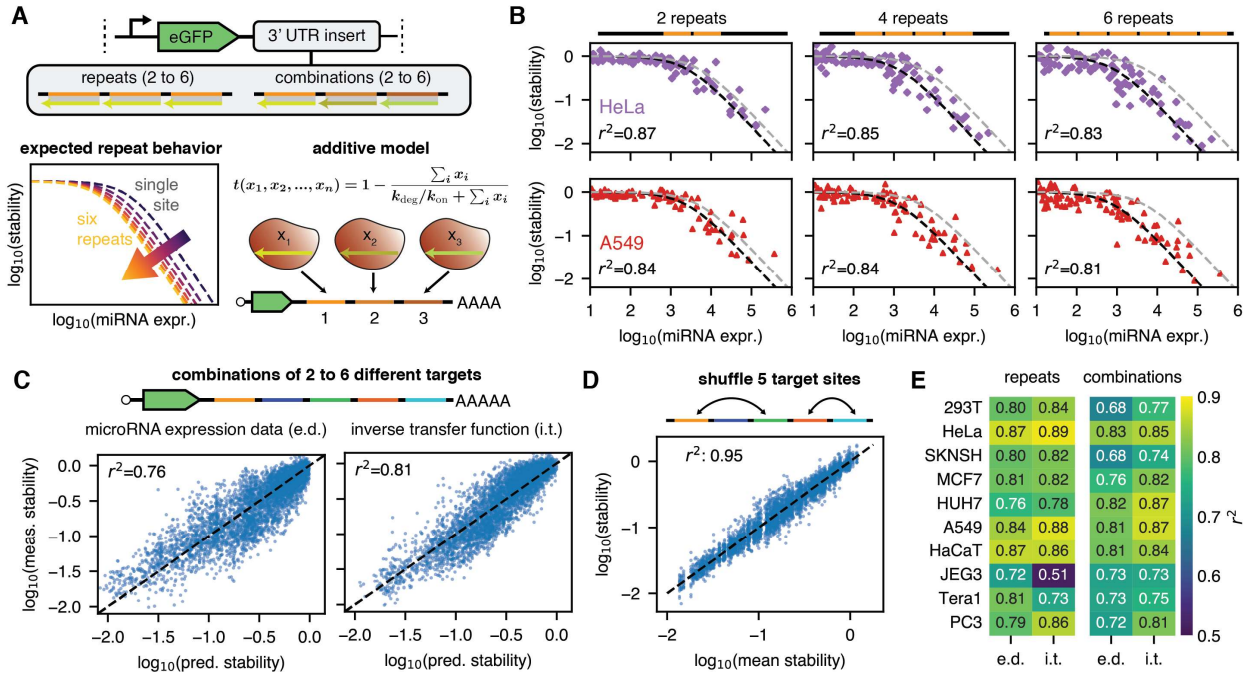


Fig. 2. The additive model predicts the behavior of multiple target sites. (**A**) We measure the stability of repeats and combinations of target sites. The additive model predicts that the concentration of individual target sites should be summed before the transfer function is applied. (**B**) Measurements and predictions using the additive model (dashed black line) for 2 to 6 target site repeats in two cell lines. We use scaled, merged and crosstalk-filtered microRNA expression data. The gray dashed line shows the prediction for a single target site. (**C**) Measurements and predictions using the additive model for combinations of 2 to 6 target sites across all our cell lines. We use either scaled, merged and crosstalk-filtered microRNA expression data (left) or expression data inferred from the stability data for individual target sites via inversion of the transfer function (right). (**D**) Mean and individual stability data for combinations of 5 target sites with shuffled ordering on the 3'UTR. (**E**) Comparison of model fit Pearson r^2 values for repeats and combinations of miRNA target sites for actual miRNA expression data (e.d.) or expression data derived from inverting the transfer function for individual sites (i.t).

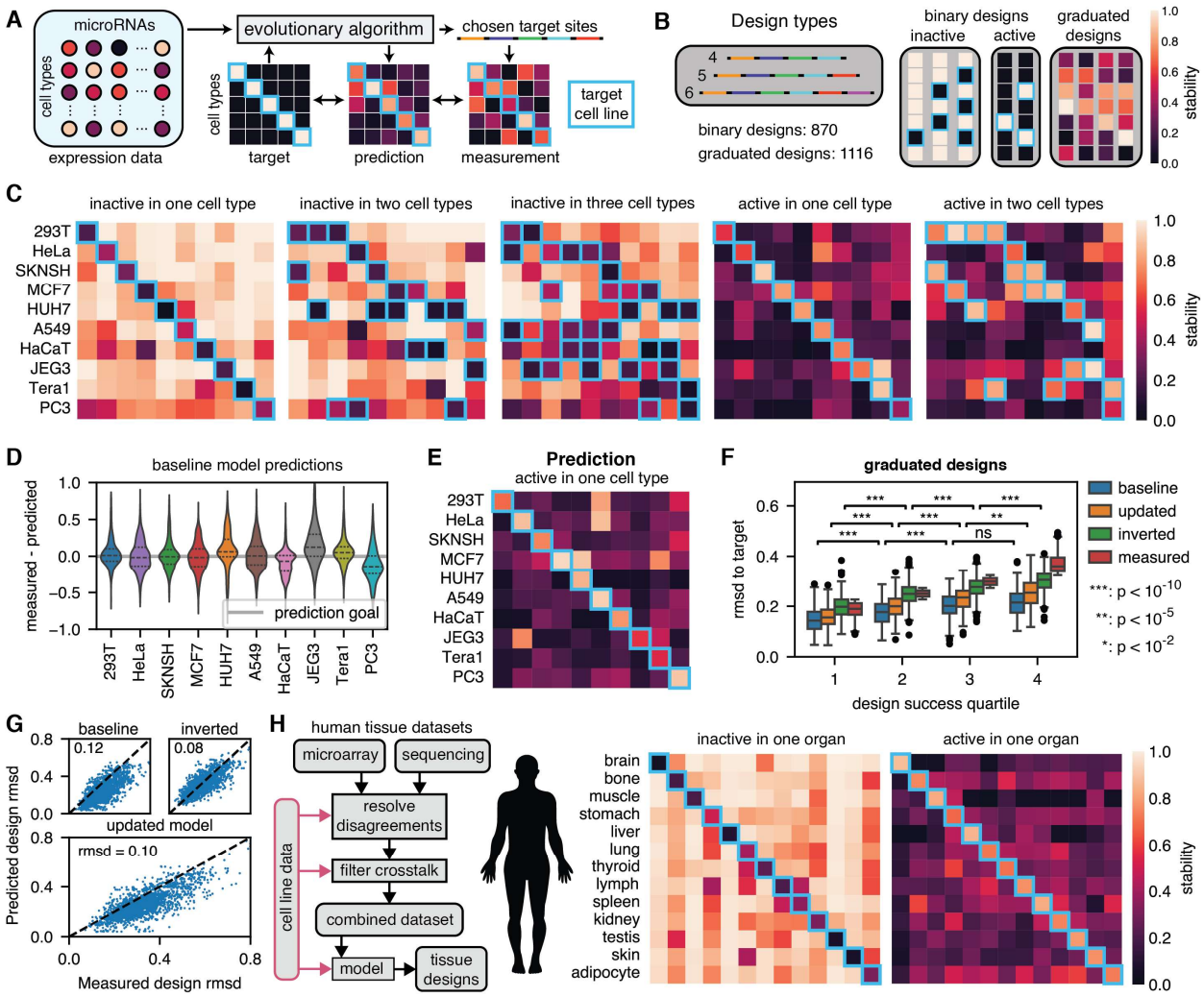


Fig. 3. Model-based design of tailored stability profiles from expression data. (**A**) An evolutionary algorithm uses expression data and our model to generate and predict 3'UTR designs. Designs are chosen based on a weighted root mean square deviation (RMSD) between predicted and target stability patterns. (**B**) Different stability patterns across cell lines are designed using 4,5 or 6 miRNA target sites. Designs either use binary expression patterns or graduated patterns. (**C**) Best-performing designs for either high or low stability in one or a few cell lines and the opposite in the others. The blue boxes show the cell lines in which high/low stability is desired. (**D**) Deviations of the measurement for binary designs from the predictions of the baseline model used for design. (**E**) Prediction by the baseline model for the best-performing designs with high stability in a single cell line. (**F**) RMSD of model predictions from target stability values for graduated designs across different quartiles of design success (measured RMSD to the target pattern). p-values were calculated using a two-sided Mann-Whitney U test. (**G**) Predicted and measured RMSD values from the target stabilities for the different models. (**H**) Two human tissue datasets collected using different methods were merged. We used our cell line data to help resolve disagreements and filter crosstalk. We used our model to create designs with up to six target sites that are predicted to be either active or inactive in a single organ.

Supplementary Materials

Materials and Methods

Library design

We designed and measured a first test library of 1,000 sequences and a second library of 10,021 sequences. The main text focuses on Library 2 because Library 1 was measured in fewer cell lines (HEK293T, HeLa, SKNSH, MCF7) and does not yield any additional insights. Not all sequences included in these two libraries are evaluated in the main text.

Library 1 design and purpose

Each sequence has a length of 142 nt, of which 15 nt were fixed at the 5' end (CGAGCTCGCTAGCCT) and 17 nt at the 3' end (AGATCGGAAGAGCGTCG), leaving a total of 110 variable nts. The purpose of Library 1 was for us to gain sufficient information to allow for the design of Library 2. We chose 145 miRNAs annotated as high confidence in miRBase v22 with high expression in the four cell lines. We measured them in two context sequences (data_high_fc_rs848_ref and data_high_fc_rs2303225_ref) of endogenous origin with neutral stability from Griesemer *et al.* (*21*). These measurements were used to fit the transfer function constant k_{on}/k_{deg} that was used to create designs with tailored stability profiles (**Fig. S4A**). Library 1 also contained two and three repeats of 41 miRNAs each, and 197 combinations of between two and five different miRNAs, which allowed for an initial verification of the additive model. It also included control sequences without miRNA target sites (*41*), which were used to generate further control sequences for Library 2. Other parts of Library 1 were not used in this manuscript. Sublibraries are listed in **Table S4** and sequences are provided in **Data S1**.

Library 2 design

Each sequence has a length of 200 nt, of which 18 nt were fixed at the 5' end (ACGACGCTCTTCCGATCT; part of the TruSeq Read 1 sequence) and 18 nt (CTCTGGATTTGCAACCGA) at the 3' end, leaving a total of 164 variable nts. Detailed library design, composition, and miRNA selection for Library 2 are described below (see **Table S5** for a list of sublibraries and **Data S3** for library sequences). Unless otherwise mentioned, we used the microarray data by Alles et al. (*22*) when expression data was required, e.g., to choose specific miRNA target sites to test.

Target site insertion

The vast majority of sequences were embedded into a single context sequence generated from a combination of two control sequences (high_fc_rs848_ref and high_fc_rs7539036_ref) with high stability and low variability in stability across cell lines from Library 1. Parts of high_fc_rs7539036_ref were added to the 5' and 3' end of high_fc_rs848_ref to generate a context sequence with a total length of 164 nts. MicroRNA targets are inserted into this 'main context' for all designs in Library 2 unless otherwise mentioned. MicroRNA target sites were inserted into context sequences as follows: Single target sites were inserted approximately in the middle of the variable context sequence. Multiple target sites were inserted at a distance of 6 nt to test potential cooperativity effects (*43*)**.** Start codons or poly(A) signals that were created by the insertion of the miRNA target were mutated to inactivate them.

Controls

In addition to the main context, we took all control sequences (with a length of 101 nt) from Library 1 that were measured as stable in all cell lines (mean stability of at least 1) and ranked

them by the variance of their stability across the measured cell lines. We appended the least variable sequences and then trimmed them from the 3' end to get a total length compatible with Library 2 (164 nt). We retained a total of 52 different control sequences.

Individual full miRNA target sequences

This sublibrary tests how a single full miRNA target site alters stability and the reliability of different data sources for miRNA sequences. Human miRNA sequences were taken from miRBase (version 22). MicroRNAs longer than 21 nt were truncated to 21 nt from the 3' end. Sequences shorter than 21 nt were padded with Us at their 3' end to reach a total length of 21 nt. A miRNA target is the reverse complement of these homogenized miRNA sequences. Targets containing a canonical poly(A) signal (AAUAAA) were filtered. We used almost all (856) high confidence human miRNAs in miRBase v22 with the exception of those containing canonical poly(A) signals in their targets. We added all sequences in MirGeneDB (166 out of 506 total) that were annotated as low confidence in miRBase. We also sampled 293 additional low confidence miRNAs from miRBase that are not in MirGeneDB, prioritizing miRNAs with large expression values in any of our chosen cell lines.

Assessing context effects and the impact of secondary structure

We chose ten miRNAs that cover a range of expression values in our cell lines (let-7a-5p, miR-16-5p, miR-19b-3p, miR-21-5p, miR-22-3p, miR-23a-3p, miR-24-3p, miR-31-5p, miR-107, miR-365a-3p). We embedded full targets for these miRNAs into two types of context sequences: First, we used the top 30 context sequences least variable in stability generated from our controls in Library 1 (see Controls). Second, we inserted 30 short sequences with varying amounts of complementarity to each of these miRNA targets starting 26 nt upstream of the 5' end of the miRNA target into the main context sequence. The degree of complementarity ranges from a block of 5 complementary bases to full complementarity to the miRNA target site. The distribution of $\Delta\Delta G$ values was calculated using NUPACK python package (version 4.0.0.20) (*44*) by deducting the free energy of the two individual strands (the 3'UTR and the miRNA) from the free energy of their complex.

Individual mutated miRNA sequences

We chose 11 miRNAs to mutate: 10 microRNAs (miR-16-5p, miR-19b-3p, miR-21-5p, miR-22-3p, miR-23a-3p, miR-24-3p, miR-31-3p, miR-31-5p, miR-107, miR-365a-3p) expected to have little crosstalk and one miRNA (let-7a-5p) with strong expected crosstalk. These microRNAs were chosen because they have highly variable expression values across our cell lines, allowing us to test the impact of mutations at different microRNA expression levels. We generated a total of 1092 mutated microRNA targets: a) 252 target sites with a single randomly chosen non-wobble mutation at all positions (252 total), b) all possible mutations leading to a single wobble base pair with the miRNA (132 total), c) insertion of an A at position 1 of the seed site (4 total), d) 704 target sites with multiple base changes distributed across the miRNA target for a total of two (231 total), three (187 total), four (132 total), five (99 total), and six (55 total) mutations.

Full miRNA target site repeats

We chose 100 miRNAs to generate constructs containing two to six repeats of a full target site. We included all miRNAs for which we tested mutations and all miRNAs of the 5p arm of the let-7 family. To make sure we cover the full range of miRNA expression values across our cell lines, we then classified miRNAs into buckets according to their maximum expression levels

across our cell lines: bucket 1 (max > 10^4 tpm), bucket 2 (10^4 tpm ≥ max > 10^3 tpm), bucket 3 (10^3 tpm ≥ max > 10^2 tpm), and bucket 4 (10^2 tpm ≥ max). We chose 14 miRNAs in bucket 1, 46 in bucket 2, 10 in bucket 3, and 5 in bucket 4. This choice ensures that we test all the most impactful miRNA targets across our cell lines.

Full miRNA target site combinations

To be able to distinguish different models of target site interactions, it is necessary to have constructs in which the different regulating miRNAs have similar expression levels. For example, in the additive model, if the ratio of expression between two cognate miRNAs for two target sites is 10, their impact on the stability fold change will have approximately the same ratio. Thus, it would be difficult to determine whether the less expressed miRNA has any impact at all. We therefore divided miRNAs into buckets depending on their mean expression across our cell lines: below $10^{2.5}$ tpm, between $10^{2.5}$ and 10^3 tpm, between 10^3 and $10^{3.5}$ tpm, and above $10^{3.5}$ tpm. When choosing multiple miRNA targets for a single UTR, we chose targets from the same bucket to maximize the odds of choosing miRNAs with similar expression levels. We started by generating 30, 30, and 20 combinations of two different miRNAs for buckets 2, 3, and 4, respectively. We then sampled an additional miRNA from the same bucket and added it to the previously created designs with one less target site, making sure to discard any duplicate designs with the same target sites. We repeated this procedure to generate constructs with between two and six different microRNA target sequences.

Design objectives for tailored stability profiles

We used five types of binary designs (active in one or two cell lines or inactive in one, two, or three cell lines) with target relative stability values of 0 or 1. For binary designs, we generally created 4 designs per cell line and miRNA target site number in two design rounds. In the first round, we generate 2 designs. In the second design round, 2 miRNA targets that were most often used in the first design round were excluded to create 2 additional designs. This prevents the algorithm from choosing nearly identical sites for every single design. We also used four types of designs with graduated target stability values: uniformly random values between 0 and 1 sampled independently for each cell line, a range of fixed uniformly spaced values between 0 and 1 randomly assigned to the different cell lines, values distributed uniformly on a log scale between 0.05 and 1 sampled independently for each cell line, and fixed logarithmically spaced values between 0.05 and 1 randomly assigned to the different cell lines. In the evaluation, these are uniformly treated as graduated designs. For graduated designs, we created 93 designs per design type and miRNA target site number with 1 or 2 designs per target pattern.

Model fitting for the design of tailored stability profiles

Sequences for tailored stability profiles were assayed as part of Library 2. Thus, all model design and fitting was done using Library 1 data and the insights gained from it. We derived the constant $\frac{k_{\text{deg}}}{k_{\text{on}}}$ for the transfer function by fitting the data of Library 1 for the four cell lines (HEK293T, HeLa, MCF7, SKNSH) we measured for that library. We heuristically filtered for crosstalk by first identifying all potentially crosstalking miRNAs as those with no mismatches in the core seed (positions 2 to 7) and up to 4 mismatches total. This crosstalk filtering is different from the one applied for the updated model as we did not have access to our measured crosstalk data at this point. A miRNA target was excluded when its crosstalking miRNAs were expressed to a level of at least 400 tpm and more than 1.35 times its own value in any cell line. We also

heuristically filtered false positives in the microarray data. For each cell line, we sorted all miRNA targets with a stability larger than $10^{-0.5}$ and positive deviation by their deviation from the model. miRNAs in the top 30% in at least three out of four cell lines (8 miRNAs total) were filtered. We fitted our transfer function to the filtered data, yielding $k_{deg}/k_{on}=10^{3.65}$, which agrees well with the value later derived from filtered Library 2 data ($k_{deg}/k_{on}=10^{3.68}$).

The genetic design algorithm

We used an evolutionary algorithm to create designs with 4, 5, and 6 miRNA target sites for each design type. We started with 300 sets of miRNA sites randomly sampled from all high-confidence miRNAs and ran the algorithm for 30 generations. In each generation, we first predict the stability pattern across our cell lines using the additive model and the transfer function derived from fitting the data for Library 1. We then evaluate the fitness of all designs. A set of 300 new designs was created by selecting 2 parents by tournament selection with a size of 3, merging the designs at a random position, and then randomly changing one of the miRNA targets with a 20% chance. We usually defined the fitness *f* of a design as the inverse of the weighted mean square error (wmse) between the target and the predicted stabilities s_{target} and s_{pred}: $f_{\mathrm{wmse}} = (\sum_{\text{cell lines } i=1}^{\mathrm{n}} w_i (s_{\mathrm{target},\, i} - s_{\mathrm{pred},\, i})^2/n)^{-1}$. For graduated expression patterns, the weighting was uniform across cell lines. For the binary designs, we weighted the error in the target cell lines 5 times for designs with one or two target cell lines and 3.33 times for designs with three target cell lines. For the designs meant to be active in a single cell type, we also created separate designs with a different fitness criterion f_{tsi} in which we multiply the tissue-specificity index (*35*) of the predicted stability with the stability in the target cell line s_{target}: $f_{\mathrm{tsi}} = s_{\mathrm{target}}(\sum_{\text{cell lines } i=1}^{n} 1 - s_i/s_{\mathrm{max},1\ldots n})/(n-1)$. This alternative criterion more strongly emphasizes high stability in the target cell line over reduction of stability in the other cell lines. For the graduated expression patterns with logarithmically distributed stability values, we also evaluated the mse on a log scale.

Library Cloning

Table S2 lists all oligos used for cloning and library preparation. Libraries were ordered as oligopools from Twist. We resupended the oligos to 5 ng/µl in 10 mM Tris-HCl, pH 8.0. We amplified the library using Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB, M0531L) using 0.25x EvaGreen (Biotium, #31000), 1 ng/µl resuspended library, and 0.5 µM of oligos 53 and 54 (Library 1) or oligos 51 and 52 (Library 2). We initially ran a test qPCR amplification for 25 cycles in a total of 10 µl to determine the optimal cycle number before the end of exponential amplification. Then we ran a larger reaction with this cycle number. The amplification protocol was 96°C for 40s, cycles of 96°C for 15s, 61°C for 20s, 72°C for 20s until the end of the exponential phase (10 cycles), then a final extension at 72°C for 8 min. The DNA was purified using 1.5x SPRIselect (Beckman Coulter, B23319) bead cleanup according to the manufacturer's instruction.

Table S6 lists the used plasmids. We used plasmid 1 (Addgene #176640 (*21*)) as the base plasmid for Library 1. For Library 2, BsaI cutting sites were inserted via overhang PCR to create plasmid 4 (oligos 49 and 50). The base plasmid was digested with BmtI and XbaI (Library 1) or BsaI (Library 2) and purified on a 1% agarose gel using Monarch® DNA Gel Extraction Kit (NEB, T1020L). We used a Gibson Assembly reaction (NEBuilder® HiFi DNA Assembly Master Mix, NEB, E2621L) using 800 ng of plasmid digest and 210 ng of amplified library at

50°C for 1h to assemble the library. The assembled plasmids were purified using a 1x SPRIselect bead cleanup.

We transformed 480 ng of assembled library into a vial of NEB® 10-beta Electrocompetent E. coli (NEB, C3020K). After resuspension in 1 ml of SOC and incubation at 37°C for 1h, we plated a 1:100,000 dilution to determine the transformation efficiency and grew the remaining cells in 200 ml of LB supplemented with 50 µg/ml kanamycin overnight. The libraries were purified using QIAGEN Plasmid Maxi Kit (Qiagen, 12162).

Individual construct cloning for flow cytometry

Starting from plasmid 1 digested with BmtI and XbaI, we inserted the main context sequence for Library 1 via Gibson assembly with oligo 26. The backbone plasmid for insertion of miRNA target sites was created from this plasmid by an overhang PCR with oligos 27 and oligo 28. This backbone plasmid was then digested with BsaI, and the individual constructs were created by ligation of the digested plasmid with oligos 29 to 48. The resulting constructs precisely match the associated sequences in Library 1.

Cell culture for high-throughput stability measurements

HEK293T, HeLa, HUH7, MCF7, K562, SKNSH, JEG3 and Tera1 cells were purchased from ATCC. A549 cells were a kind gift from Jesse Bloom. HaCaT cells were a kind gift from Paul Nghiem. PC3 cells were a kind gift from Andrew Hsieh. HEK293T, HeLa, MCF7, A549, HUH7, and HaCaT cells were cultured in DMEM supplemented with 10% FBS and 100 U/ml Penicillin/Streptomycin (ThermoFisher 15140122). SKNSH and JEG3 cells were cultured in EMEM supplemented with 10% FBS and 100 U/ml Penicillin-Streptomycin. K562 and PC3 cells were cultured in RPMI supplemented with 10% FBS and 100 U/ml Penicillin-Streptomycin. Tera1 cells were cultured in McCoy's 5A medium supplemented with 15% FBS and 100 U/ml Penicillin-Streptomycin.

MicroRNA expression measurements

The same RNA samples measured using microarrays in Alles et al. (*22*) were used for miRNA sequencing in this study. Cell culture conditions and the RNA isolation procedure have been described there. Small RNA expression libraries were generated using MGIEasy Small RNA Library Prep Kit (MGI Tech) according to the manufacturer's protocol. Briefly, 100 ng of total RNA were used for 3' adapter ligation. After removal of unused adapter by digestion, 5' adapters were ligated and the product was reverse transcribed using uniquely barcoded RT primers. The cDNA was PCR amplified for 21 cycles and PCR products derived from miRNAs were size selected by gel electrophoresis using Novex 6% TBE PAGE gels (Invitrogen). PCR products were pooled into a single library, circularized, and sequenced using DNBSEQ-G400RS High-throughput Sequencing Reagent Set (G400 sRNA FCL SE50) on a DNBSEQ-G400RS sequencer (both MGI Tech).

Fastq files were trimmed and quantified using the miRMaster 2.0 pipeline (*45*). The adapter sequence used for trimming was AGTCGGAGGCCAAGCGGTCTTAGG with a minimum overlap of 10 nt. We set a maximum edit distance of 1 nt for adapter matches and no ambiguous nucleotides were allowed. For each read 3 nucleotides were trimmed from the leading and trailing ends. Trimming was further performed with a sliding window of 4 nt with a phred score quality threshold of 20. Read lengths smaller than 17 nt were removed. Reads were collapsed and mapped against GRCh38 with a maximum mismatch of 1 using bowtie 1.1.2 using the

options –mm -v1 –m 100 –best –strata –fullref. MiRNA and isomiR quantification was performed against the *H. sapiens* miRNA set from miRBase v22.1.

Flow cytometry for control experiments

We seeded 100,000 HEK293T or HeLa cells in 24-well plates 24 h before transfection. We transfected 500 ng total of plasmid in each well. To screen the effect of plasmid concentration on miRNA activity, we used between 16.7 and 200 ng of reporter plasmid, 100 ng of a CMV-mCherry transfection control plasmid (plasmid 6), and enough pUC19 filler plasmid to reach 500 ng total. To measure the impact of different miRNA target sites, we used 30 ng (HEK293T) or 60 ng (HeLa) reporter plasmid, 100 ng of a CMV-mCherry plasmid (plasmid 6), and enough pUC19 filler plasmid to reach 500 ng total. Transfections were carried out using Lipofectamine™ 3000 Transfection Reagent (ThermoFisher, L3000001) according to the manufacturer's instructions. We exchanged media 4 h after transfection. Two days later, we detached cells with TrypLE, added media, then centrifuged for 5 min at 210 g. Cells were resuspended in 1 ml of PBS. We measured 200 µl of cell suspension either undiluted (HeLa) or diluted 1:1 in PBS (HEK293T) using an Attune NxT Flow Cytometer (ThermoFisher). We first gated cells on SSC-A and FSC-A (**Fig. S5B**), then on the mCherry fluorescence such that almost all non-transfected cells from a negative control were excluded (**Fig. S5C**). We calculated the median GFP fluorescence of the gated cells, subtracted the median GFP fluorescence of an untransfected negative control, then normalized to a positive control of the same plasmid without miRNA target sites (**Fig. S5E**).

Plasmid library transfection, RNA purification, and library amplification

We measured two replicates for each cell line. We seeded between 350,000 and 5 million cells two days (Tera1 cell line) or one day (all other cell lines) before transfection. The precise transfection parameters per cell line can be found in **Table S1**. Between 2.5 and 15 µg of plasmid library were transfected with Lipofectamine™ 3000 Transfection Reagent (ThermoFisher, L3000001) according to the manufacturer's instructions. We exchanged media 4 h after transfection. Two days later, we detached cells with TrypLE, added media, then spun down for 5 min at 210 g. We resuspended the cells in 1 ml of cold PBS and placed them on ice. We took 25 µl of each sample, added 175 µl of PBS, and measured the GFP fluorescence on the Attune NxT flow cytometer to determine cell numbers and transfection efficiencies. We used the Monarch Total RNA Miniprep Kit (NEB, T2010S) to purify total RNA. We eluted in 100 µl of nuclease-free water and determined concentrations using a NanoDrop 2000c (Thermo Fisher). mRNA was purified using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB, E7490L). We eluted in 18 µl of nuclease-free water and performed reverse transcription (RT) using 0.75 µM oligo 1 (library 1) or oligo 11 (library 2), 0.5 mM dNTPs (NEB), 0.5 U/µl SUPERase·In™ RNase Inhibitor (ThermoFisher, AM2696), and 10 U/µl Maxima H Minus Reverse Transcriptase (ThermoFisher, EP0752). Before adding RT buffer and enzyme, the mix was incubated at 65°C for 5 min. Reverse transcription was carried out at 50°C for 15 min, then 85°C for 5 min. We added 10 U RNase I (ThermoFisher, EN0601) and 5 U RNase H (NEB, M0297S) and incubated at 37°C for 30 min. cDNA was purified using DNA Clean & Concentrator-5 (Zymo, D4014) with 7x binding buffer, eluted in 10 µl nuclease-free water, and stored at -20°C until library amplification.

Libraries were amplified using KAPA HiFi HotStart ReadyMix (Roche, KK2602) with 0.5 µM of each index primer (Oligos 2 to and 8 and 12 to 25) and 1x EvaGreen (Biotium, #31000). The

protocol was 95°C for 3 min, a variable number of cycles of 96°C for 25s, 69°C for 15s, 72°C for 30s, and a final extension at 72°C for 5 min. We ran a pilot qPCR for 30 cycles with 1.09 µl of library in a total of 5 µl to determine the appropriate cycle number. The full library was amplified in 50 µl until the last cycle of exponential amplification. The PCR reaction was run on a 1.8% agarose gel at 120V for 35 min. The band matching the expected size was cut out, purified using Monarch® DNA Gel Extraction Kit (NEB, T1020L) and eluted in 20 µl of elution buffer. Concentrations were determined using a Qubit™ dsDNA Quantification Assay (ThermoFisher Q32851).

To introduce unique molecular identifiers (UMIs) for the DNA library, library amplification was preceded by two cycles of amplification with the reverse transcription primer (oligo 1 or 11) and the library amplification i5 primers (oligos 2 and 3 or 12 and 13). The DNA was purified using DNA Clean & Concentrator-5 (Zymo, D4014), eluted in 10 µl, and amplified as above.

Library sequencing and data processing

Library 1 was sequenced in-house on a NextSeq 500/550 Mid Output kit (150 cycles, Illumina) with custom primers (Oligos 9 and 10)**.** Library 2 was sequenced on a NovaSeq X Plus (Illumina) using a paired-end 300 cycle kit by Novogene. Cell line and replicate data was split according to the i5 and i7 indices. The UMI was extracted and added to the read name. Libraries were aligned to the reference sequences using BWA-MEM. Alignments were filtered using a custom python script. First, we removed ambiguous alignments: Reads were discarded when one of the two reads failed to align, when both reads had more than one read with the highest alignment score, or when one read aligned uniquely but did not match the top alignments of the other strand. We then filtered to at most one total insertion or deletion or four substitutions relative to the aligned reference sequence. The alignments were sorted and indexed using samtools. UMIs were deduplicated using UMI-tools with unique matching. Alignments for each reference were counted using a custom python script. Fold-change values were calculated from the count data of the two replicates using PyDESeq2 (*46*).

Stability data analysis

All analyses are performed with custom Python code.

Filtering and normalization

We filtered all sequences with less than 100 combined counts in the two plasmid DNA library replicates (one sequence out of 10,001 total unique sequences). The fold change values calculated by PyDESeq2 were normalized to make the average stability for non-regulated target sites in the main context equal to 1. Normalization was performed separately for each cell line. We initially divided all fold change values by the median of the 300 most stable constructs of containing individual high confidence miRNA target sites in the main context, most of which are inactive. For all UTRs using the main context sequence, we grouped designs by the number of miRNA target sites (1 to 6). Within each group, we divide fold change values by the median of the top 15% of most stable designs, then add the median of the median values across cell lines. The individual normalization for different target site numbers accounts for the fact that the average baseline stability can change with the number of miRNA target sites. This effect is usually small but significant for MCF7 and HUH7. Designs using other context sequences are only processed using the first normalization step. Normalized stability data for Library 1 and 2 are provided as **Data S2** and **Data S4**.

The choice of miRNA expression data

We used two miRNA expression datasets for most analyses: microarray data by Alles et al. (*22*) and BGISEQ-500-based sequencing data generated for this study. Both datasets cover all used cell lines. Although the latter was generated for this study, it was collected in a different laboratory from the one where the reporter experiments were performed. We deliberately made no effort to harmonize cell lines or culture conditions to avoid biasing our results in favor of this newly collected data. Of note, both datasets contain expression data for SH-SY5Y, while the stability data was collected in SKNSH cells. Since SH-SY5Y were created from the subclone of SKNSH, we expect their miRNA expression profiles to be sufficiently similar for this to be valid.

Fitting a function for a single target site in the main context

For the analysis of single full target sites, we only use high confidence miRNAs in miRBase and low confidence miRNAs in miRBase that are in MirGeneDB, but not low confidence miRNAs in miRBase that are not in MirGeneDB. First, we normalized the miRNA expression data to transcripts per million (tpm) on this subset of miRNAs. We then used SciPy to fit our transfer function either for fixed total miRNA levels or variable total miRNA levels, i.e., we introduced an additional fitting parameter for each cell line that scales total miRNA levels. For the latter, the total miRNA levels for HEK293T cells were kept constant to establish a reference point. We calculated deviation values for each miRNA target as the difference between the measured stabilities and the stabilities estimated by the transfer function on a logarithmic scale.

Single target sites in different context sequences

We used the fitted transfer function for target sites in the main context to predict the stability in all other contexts. Because the transfer function predicts relative stability changes due to the presence of miRNA target sites, measured stability values for miRNA targets $s_{\mathrm{miRNA,\,c}}$ in different context sequences c were divided by the measured baseline stabilities $s_{0,\,c}$ of the surrounding context sequences without a target site in the same cell line: $s_{\mathrm{miRNA,\,rel}} = s_{\mathrm{miRNA,\,c}}/s_{0,\,c}$.

Estimation of the scale factor

We hypothesized that systematic global deviations from the transfer function are the result of both genuine biological variations in total miRNA levels between cell lines and differences in baseline context stabilities k_{deg} between cell lines. Notably, Ago2 expression, which could serve as a proxy for the total slicing-competent miRNA level, is highest in PC3, where we do observe a stronger effect of miRNA (**Fig. S10A**). To estimate total miRNA levels, we therefore retrieved expression data (nTPM) for Ago proteins from the human protein atlas (*47*), which contains data for 8 of 10 of our cell lines. We then normalized these data to those in HEK293T cells. To estimate the relative stability of the main context, we took the geometric mean of 51 different measured context sequences and normalized this mean to the HEK293T value. For HUH7, stability of the main context relative to 51 other endogenously-derived context sequences is low, indicating that this effect might be dominant there (**Fig. S10B**). This approach assumes that the range of stability values covered by the different context sequences is similar across cell lines. The overall scale factor is then estimated by dividing the relative Ago2 expression by the relative mean context stability (**Fig. S10C**). We note that mathematically only a single scaling factor is necessary because the model only depends on the ratio of the miRNA-induced degradation and the baseline degradation (**Supplementary Text 1**).

Evaluation of the impact of secondary structure for single target sites

All ΔΔG values were calculated using the Python package for NUPACK version 4.0.0.20 (*44*) using the free_energy result of the complex_analysis function. When calculating these values, one must make a choice of how many bases to the 5' and 3' end of the miRNA target site to include. We used four different cutoffs (150, 100, 60, and 40 bases to the 5' and 3' end) and averaged the results for increased robustness. We calculated the free energy of the miRNA ΔG_{miRNA}), the target without context ($\Delta G_{target,\ ideal}$), the target with varying amounts of context ($\Delta G_{target,\ actual}$), and the miRNA in complex with the target both without ($\Delta G_{miRNA\text{-}target,\ ideal}$) and with context ($\Delta G_{miRNA\text{-}target,\ actual}$). We then subtracted the free energy of the miRNA and the target from the free energy of their complex to calculate the binding ΔΔG between the miRNA and its target: $\Delta\Delta G_{\text{ideal/actual}} = \Delta G_{\text{miRNA-target, ideal/actual}} - \Delta G_{\text{miRNA}} - \Delta G_{\text{target, ideal/actual}}$. The ΔΔG difference value that we use in figures is then calculated by subtracting the binding ΔΔG without context from the binding ΔΔG with context: $\Delta\Delta G_{\text{diff}} = \Delta\Delta G_{\text{actual}} - \Delta\Delta G_{\text{ideal}}$. For individual target sites, ΔΔG differences larger than 10 kcal/mole were heuristically classified as indicating strong secondary structure. P-values for the deviation differences between strong and weak secondary structure were calculated using a one-sided Mann–Whitney U test. We constrained the analysis of deviation values to miRNAs expressed to a level larger than $10^{3.5}$ tpm, i.e., those where we expect a measurable impact on stability in the absence of secondary structure. In the analysis of designed secondary structures, we excluded the two constructs containing the strongest designed secondary structure for each miRNA target. These constructs contain contiguous 21 bp and 24 bp RNA duplexes, which cause a decrease in measured stability that we attribute either to a destabilizing effect of having a very strong hairpin in the 3'UTR or to the negative impact of such hairpins on library preparation, rather than an unexpected increase in miRNA activity at stronger secondary structures.

Processing and comparison of miRNA expression datasets

Since we compare microarray and sequencing datasets, we use transcripts per million (tpm) instead of reads per million (rpm). We used miRNA expression values provided by the authors either in the original publication or deposited on the GEO database. The details of the used datasets are available in **Table S3**. The first step of dataset processing depended heavily on the individual dataset and was performed with a custom python script. The authors' files were converted into a CSV file containing miRNA names, cell line identifiers, and non-normalized expression values. We used a custom R script and miRBaseConverter (*48*) to map older miRNA names to a MIMAT ID and to check whether the miRNA in question is still considered real in the current version of miRBase. Invalid miRNAs were discarded. We then converted the MIMAT ID back to the miRNA identifier in miRBase v22. Datasets were normalized to tpm, a baseline value of 1 was added to the expression level of each miRNA, then the dataset was normalized to tpm again and converted to log10 values. To compare datasets, we limited the miRNAs in each dataset to high confidence miRNAs in miRBase present in all compared datasets (242 miRNAs total). Microarray datasets were classified according to the manufacturer (Agilent, Affymetrix). Sequencing datasets were classified as either standard or improved based on their use of degenerate adapters to reduce ligation bias. To calculate correlation values between datasets, correlations were averaged across all cell lines common to both datasets. To check the agreement with our stability dataset, the transfer function was fitted individually for

each cell line. The resulting correlation values were averaged across cell lines for each dataset and then grouped by the dataset collection method.

Identification of false positives and negatives

For false positives, we first filtered miRNA targets for which we measured stabilities greater than $10^{-0.25}$ and expression levels larger than $10^{3.8}$ tpm. We then flagged those which display less activity than expected by the model (deviation > $10^{0.3}$) as potential false positives. For false negatives, we first filtered miRNA targets for which we measured stabilities lesser than $10^{-0.25}$ and expression levels smaller than $10^{3.5}$ tpm. We then identified those which display more activity than expected by the model (deviation > $10^{-0.3}$) as potential false negatives. We compared potential false positives and false negatives between the two expression datasets. When a miRNA expression measurement was identified as a potential false positive/negative in one dataset but not the other, it was retained as a potential false positive/negative. If a miRNA expression measurement was identified as a potential false positive/negative in both datasets, we concluded that it is not a genuine false positive/negative. We labeled miRNAs that appear as a potential false positive or negative in at least 4 cell lines after this filtering step as genuine false positives or negatives. **Table S7** contains a list of these identified false positives and negatives.

Analysis of mutated target sites and generation of a crosstalk model

Addressing crosstalk serves to increase the predictive accuracy of our model. In principle, one could simply filter all but the most highly expressed miRNA that share a core seed sequence (nts 2 to 7). However, this leads to a strong reduction in the available number of potential miRNA target sites. This problem gets worse as the number of target cell types increases because the highest expressed miRNAs of a given family can differ between tissues. Thus, crosstalk filtering is a balance between increasing predictive accuracy and retention of as many potential target sites as possible. We note that simply adding expression levels of the same family is often worse than not addressing crosstalk at all and not a viable option for quantitative prediction.

The measured stabilities s were converted to relative knockdown values, which we define as $\frac{s_{\text{full target}}}{1-s_{\text{full target}}}\frac{1-s_{\text{mut target}}}{s_{\text{mut target}}}$. This expression yields a value of 1 when the mutated target site is as strong as the full target site and 0 when the mutated target site is completely inactive. Analyses were generally performed on miRNA target sites where the stability value for the full target site is less than ⅓ in a given cell line to ensure that this value can be accurately determined. We determined the distribution of these relative knockdown values depending on where in the sequence mutations and wobbles occur ignoring the specific miRNA sequence and identity. We then classify individual mutations and wobbles according to the median relative knockdown at a given position with values less than 0.08, 0.3, and 0.65, or greater than 0.65 being classified as high, medium, low, and no impact mutations and wobbles, respectively. This classification was performed independently for true mismatches and for wobbles. The classification thresholds for the median were chosen heuristically such that each category contains multiple positions (10, 12, 16, and 4 for high, medium, low, and no impact mutations). When the classification for a wobble base pair at a specific location was stronger than for the corresponding mismatch, the classification for the wobble was lowered to the level of the mismatch. P-values between the mismatch and wobble relative knockdown distributions were calculated using a two-sided Mann-Whitney U test.

To train a tree model, we applied the previously derived mutation classification to targets with multiple mutations and counted how often each mutation type occurs. The mutation counts by mutation strength were the input values for the tree model. Training was performed on relative knockdown values constrained to mutated target sites where the non-mutated target site had a stability value less than ⅓. We held out all mutated hsa-miR-31-5p target sites as test data. We trained both an XGBoost model (xgboost version 1.7.3) and a DecisionTreeRegressor (sklearn version 1.3.0) and found nearly identical performance. The decision tree predicts relative knockdown values, which are converted to stabilities afterwards. The predictions shown **Fig. S16C** were performed on all measured mutated target sites irrespective of the stability of the non-mutated target.

Because the impact of mutations turned out to be highly sequence-specific, we decided to largely filter miRNA target sites where significant crosstalk is expected instead of trying to quantitatively adjust predictions for target sites for which crosstalk is expected: First, we added the expression levels for miRNAs that are identical for the first 18 bases, which we expect to cause nearly identical behavior (**Table S8**). We note that this is not necessarily fully accurate as mutations after base 18 can still have an impact (**Fig. S15C-E**). We nevertheless consider it preferable to filtering these target sites. We then identified miRNAs with potential crosstalk as all pairs of miRNAs and miRNA targets separated by fewer than 2 high impact mutations, 4 high and medium impact mutations, and 5 total mutations (**Table S9**). This is based on the maximum number of mutations of a given impact that still allows for crosstalk (**Fig. S15E**). These sites with potential crosstalk were filtered if the stability predicted by the transfer function for the full target site of the crosstalking miRNA is smaller than $10^{-0.5}$ and less than ⅓ of the predicted stability of the fully complementary miRNA for the target site in question.

Repeat and combination data

We used the bias-corrected expression data demonstrated in **Fig. 1H**. We predicted the expected stability using the transfer function and scale factors derived for single target sites and using either the additive model or an antagonistic model where only the miRNA target site with the highest associated expression level is used. For the inverted transfer function, expression levels are calculated from stabilities s via $t^{-1}(s) = \frac{k_{\mathrm{deg}}}{k_{\mathrm{on}}}\frac{1-s}{s}$. Since the transfer function only outputs values between 0 and 1, stability values larger than 1 were reduced to 0.999 for calculations using the inverted transfer function. We then input these calculated expression levels into the transfer function using the additive model. For the final calculation of Pearson correlation values by cell line, we excluded excessively stable sequences (stability values larger than 1.5).

Analysis of outliers in the repeat and combination data

Unusual patterns in the repeat data were identified as miRNAs for which two and four repeats were at least a factor of $10^{0.1}$ more stable than one and three repeats, or three and five repeats were at least a factor of $10^{0.1}$ more stable than two and four repeats. Constructs with a stability value of at least 1.5 were identified as highly stable. This yielded 13, 3, and 14 miRNAs where these conditions were fulfilled in at least one cell line.

To identify the origin of the unusual patterns in the repeat data, we calculated pairwise interaction strengths between the individual target sites as the free energy difference between a complex of two targets and the individual targets. To more generally detect whether secondary structure plays a role for any construct with multiple target sites, ΔΔG difference values between

miRNAs and 3'UTRs with multiple target sites were calculated as for individual target sites but with 150 or 100 nts on either side of the center of the variable 3'UTR region. Here, we heuristically classified sites with ΔΔG differences larger than 14 kcal/mole as strong secondary structure sites. Constructs were identified as likely subject to prediction deviations due to secondary structure when the site with the largest cognate miRNA expression value was classified as having strong secondary structure.

Performance prediction and evaluation of constructs with designed stabilities

We compare three prediction methods: First, the baseline prediction based purely on microarray data (*22*) and the transfer function fitted using the measurements from Library 1. Second, an updated model based on bias-aware merging of microarray and improved sequencing data as demonstrated in Figure 1 of this manuscript. The updated model also applies the scale factors derived from individual target sites. Third, predictions based on estimating expression levels via the inverted transfer function as described above for repeat and combination data. All predictions use the additive model. We filtered designs where the measured stability in any cell line is larger than 1.5 as this makes it difficult to estimate true relative stabilities (15 out of 3753 total designs). To evaluate design performance, stabilities larger than 1 were set to 1. When choosing the best-performing designs or when ranking design success into quantiles, we used the weighted mean square error as explained above in the design section. When comparing the design success across design types, we used the unweighted root mean square error between the target and measured stabilities.

Processing and merging of human tissue miRNA expression datasets

We used an Agilent microarray dataset from two human subjects by Ludwig et al. (*35*) and a BGISEQ-500-based sequencing dataset from six human subjects by Keller et al. (*36*). The data collection methods match the two main datasets used for the rest of this study, making it likely that biases behave at least similarly. We first filtered to miRNAs that are either high confidence in miRBase or in MirGeneDB. Data were then normalized to tpm. To calculate correlation values between subjects for the same tissue and dataset, we set expression values less than 100 tpm to 100 tpm. To integrate data from the different subjects and produce a single expression dataset across tissues, we then calculated the geometric mean of expression values for each tissue across all subjects and again normalized to tpm (**Fig. S34**). To compare the two datasets, we first harmonized tissue names and again set values less than 100 tpm to 100 tpm. Tissues that were not in both datasets were discarded. To calculate brain expression values for the NGS dataset, we averaged over temporal lobe, occipital lobe, frontal lobe, white matter, and gray matter values. We found that there are many miRNAs that behave differently in the two datasets. To identify these outlier miRNAs, we first renormalized the expression values because the microarray dataset has more highly expressed miRNAs overall. Regular normalization to tpm therefore leads to an underestimation of individual expression levels relative to the NGS data. We determined the top 30 most different miRNA by expression in each tissue. We identified miRNAs that are in the top 30 in at least three tissues, then renormalized the dataset to tpm while excluding them from the calculation. We again compared miRNA expression levels across tissues and datasets. We then determined outliers in a specific tissue as those miRNAs where the renormalized expression values differ by more than a factor of 10. miRNAs which are consistent outliers in 5 or more tissues were classified as consistently larger in NGS (5 total) or microarray (94 total)

data. We then calculated Pearson correlation values with and without these outliers on a log expression scale (**Fig. S35A-B**).

To merge the two datasets, we asked whether it is possible to guess which of the two datasets is correct for which outlier. First, we calculated deviation values in the cell line stability for each outlier miRNA for both our microarray and NGS data cell line expression data (**Fig. S36A**). We excluded cell lines (JEG3, Tera1) with poorer prediction performance from the analysis. We subtracted absolute deviation values for the two cell line datasets and considered a difference greater than $10^{0.2}$ as signifying substantially more correct prediction by one or the other method (**Fig. S36B-D)**. Second, we added false positive and negative information from our cell line data. Third, we inspected the sequence features of outliers. We noticed that many of the strongest outliers had a G-content greater than 50% (**Fig. S36E-F**). An earlier publication (*49*) found similar outliers between microarray and NGS data. There, RT-qPCR agreed with the NGS data over the microarray data. We therefore also labeled miRNAs with a G content of over 50% that were previously identified as consistent outliers as more likely to be correctly measured by improved NGS data. In total, we classified 12 miRNAs as incorrect in the NGS data and 16 miRNAs as incorrect in the microarray data. We then performed a bias-aware merging based on the geometric mean as for the cell line data (**Fig. S36G**). The resulting data was then further merged and filtered for crosstalk as for the cell line data.

Designs with tailored stability profiles in human tissues

We used the merged and crosstalk-filtered human tissue expression dataset. Because this dataset contains fewer miRNAs than the merged cell line data due to fewer input miRNAs in the microarray tissue data, we performed another round of fitting to the cell line stability data for this subset of miRNAs. The design process was generally performed as for cell lines with a few differences. First, we only generated a single design per design target and miRNA target number (1 to 8). Second, we filtered the expression dataset to miRNAs with a maximum expression greater than 3000 tpm across target tissues, which reduces computational complexity and focuses the algorithm on potentially relevant miRNAs. Third, we introduced an empty target site with zero expression in all tissues, which allows the algorithm to use fewer than the maximum number of target sites. We generated designs that are either active or inactive in a single tissue. The stability in that tissue was given a weight of 6.5 in calculating the mse. To compare the design quality, defined as the inverse of the weighted mse between the prediction and the target stability profile, between different target site numbers, we divided all quality values for designs targeting the same tissue by the maximum quality across the allowed target site numbers.

Supplementary Text

<u>The transfer function for predicting the stability of an mRNA containing a single microRNA target site from expression data</u>

Here, we derive the simple transfer function for predicting the stability of an mRNA with concentration m containing a single target site for a microRNA with concentration x. We assume that x is not appreciably depleted by target binding. In the absence of a microRNA target site, m is degraded at a rate k_{deg} and produced at a rate p:

$$\dot{m} = p - k_{\text{deg}} m$$

with a resulting steady-state concentration of

$$m_0 = \frac{p}{k_{\text{deg}}}.$$

Let m_{free} be the concentration of free mRNA and m_{bound} be the concentration of mRNA bound by the microRNA. Assuming irreversible binding this results in

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} m_{\text{free}} &= p - k_{\text{deg}} m_{\text{free}} - k_{\text{on}} x m_{\text{free}} \\ \frac{\mathrm{d}}{\mathrm{d}t} m_{\text{bound}} &= -k_{\text{deg}} m_{\text{bound}} + k_{\text{on}} x m_{\text{free}} - k_{\text{cat}} m_{\text{bound}} \end{aligned}$$

for a catalytic rate k_{cat} of the RISC. In the steady state and with $c_1 = k_{\text{deg}}/k_{\text{on}}$, this equation can be rewritten as

$$\begin{aligned} & 0 = m_0 - m_{\text{free}} - \frac{x}{c_1} m_{\text{free}} \\ \Leftrightarrow \quad & \frac{m_0}{m_{\text{free}}} = 1 + \frac{x}{c_1} \\ \Leftrightarrow \quad & m_{\text{free}} = m_0 \frac{1}{1 + x/c_1}. \end{aligned}$$

Under the simplifying assumption that degradation by the RISC is very fast compared to its binding ($m \approx m_{\text{free}}$), the stability of a microRNA-regulated mRNA relative to its baseline stability is given by

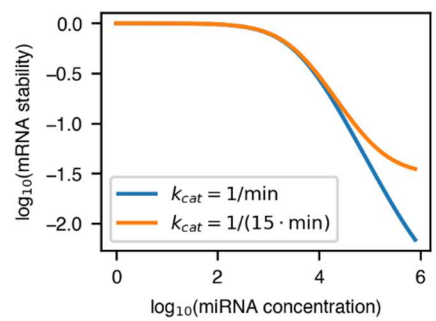
$$t(x) = \frac{m}{m_0} = 1 - \frac{x}{c_1 + x},$$

which is the transfer function used in the main text. When the finite catalytic rate k_{cat} of the RISC is taken into account, the bound mRNA fraction is given by

$$m_{\text{bound}} = \frac{k_{\text{on}}}{k_{\text{cat}} + k_{\text{deg}}} x m_{\text{free}},$$

which yields an expression for the total mRNA concentration:

$$\frac{m_{\text{total}}}{m_0} = \frac{m_{\text{bound}} + m_{\text{free}}}{m_0} = \left(1 - \frac{x}{c_1 + x}\right)\left(1 + \frac{x}{c_2}\right),$$
$$c_2 = \frac{k_{\text{cat}} + k_{\text{deg}}}{k_{\text{on}}}.$$



For large microRNA concentrations, the ratio between the mRNA concentrations without and with microRNA regulation thus tends to

$$\frac{m_0}{m_{\text{total}}} \overset{x \gg c_1, c_2}{\rightarrow} \frac{c_2}{c_1} = \frac{k_{\text{cat}} + k_{\text{deg}}}{k_{\text{deg}}} = 1 + \frac{t_{1/2,0}}{t_{1/2,\text{cat}}}.$$

This is the maximum loss of stability due to a single miRNA target site. In practice, we find that this term is not necessary even for repeats of six targets sites for highly expressed microRNAs. Martinez et al. (23) estimated the *in vitro* half-life of human RISC-bound RNA $t_{1/2,\text{cat}}$ to be approximately one minute, while $t_{1/2,0}$ is typically on the order of 8 hours. This would imply a maximum stability ratio of 480, which is substantially smaller than the smallest values measured in our libraries (cf. the illustration on the right). Note that we measure stability, so the maximum loss of fluorescence for a fluorescent reporter could be smaller if an mRNA with a cleaved tail can still be translated until it is fully degraded.

<u>The dependence of the transfer function on total microRNA concentrations and the baseline stability</u>

The transfer function

$$t(x) = \frac{m}{m_0} = 1 - \frac{x}{k_{\text{deg}}/k_{\text{on}} + x}$$

only has a single parameter, namely the ratio between the baseline degradation rate and the on-rate of the RISC. We initially fit $c_1 = k_{\text{deg}}/k_{\text{on}}$ as a global parameter for all cell lines. Different total microRNA concentration can be taken into account by a scaling factor s for the total microRNA concentration x, i.e., $x_{\text{scaled}} = s \cdot x$ for normalized concentrations x.

However, c_1 also depends on the baseline degradation rate $k_{\text{deg}} \propto 1/m_0$. When comparing context sequences with different baseline stabilities in one cell line, or a single context sequence with cell type-specific stability, c_1 also needs to be adjusted for the change in k_{deg}. A factor f change in stability $\tilde{m}_0 = f \cdot m_0$ thus leads to

$$\frac{\tilde{m}}{\tilde{m}_0} = \tilde{t}(x) = 1 - \frac{x}{\tilde{k}_{\text{deg}}/k_{\text{on}} + x} = 1 - \frac{x}{\frac{1}{f} k_{\text{deg}}/k_{\text{on}} + x}$$
$$= 1 - \frac{f \cdot x}{k_{\text{deg}}/k_{\text{on}} + f \cdot x} = 1 - \frac{\tilde{x}}{k_{\text{deg}}/k_{\text{on}} + \tilde{x}}.$$

An increase in baseline stability by a factor of f therefore has the same effect as an increase in the microRNA concentration by a factor of s.

The additive model for microRNA regulation of multiple target sites

We start with the same equation as above but for multiple independent miRNAs x_i for the target sites $t_1, t_2, \ldots, t_n$:

$$\frac{\mathrm{d}}{\mathrm{d}t} m_{\text{free}} \quad = p - k_{\text{deg}} m_{\text{free}} - \sum_{i=0}^{n} k_{\text{on}}\, x_i m_{\text{free}}$$

Assuming instant degradation by the RISC and using the same procedure as above, this yields the following transfer function

$$t(x_1, x_2, \ldots, x_n) = \frac{m}{m_0} = 1 - \frac{\sum_{i=0}^{n} x_i}{k_{\text{deg}}/k_{\text{on}} + \sum_{i=0}^{n} x_i},$$

which is the same transfer function as for a single target site except that the concentrations of attacking microRNAs are summed before the transfer function is applied.

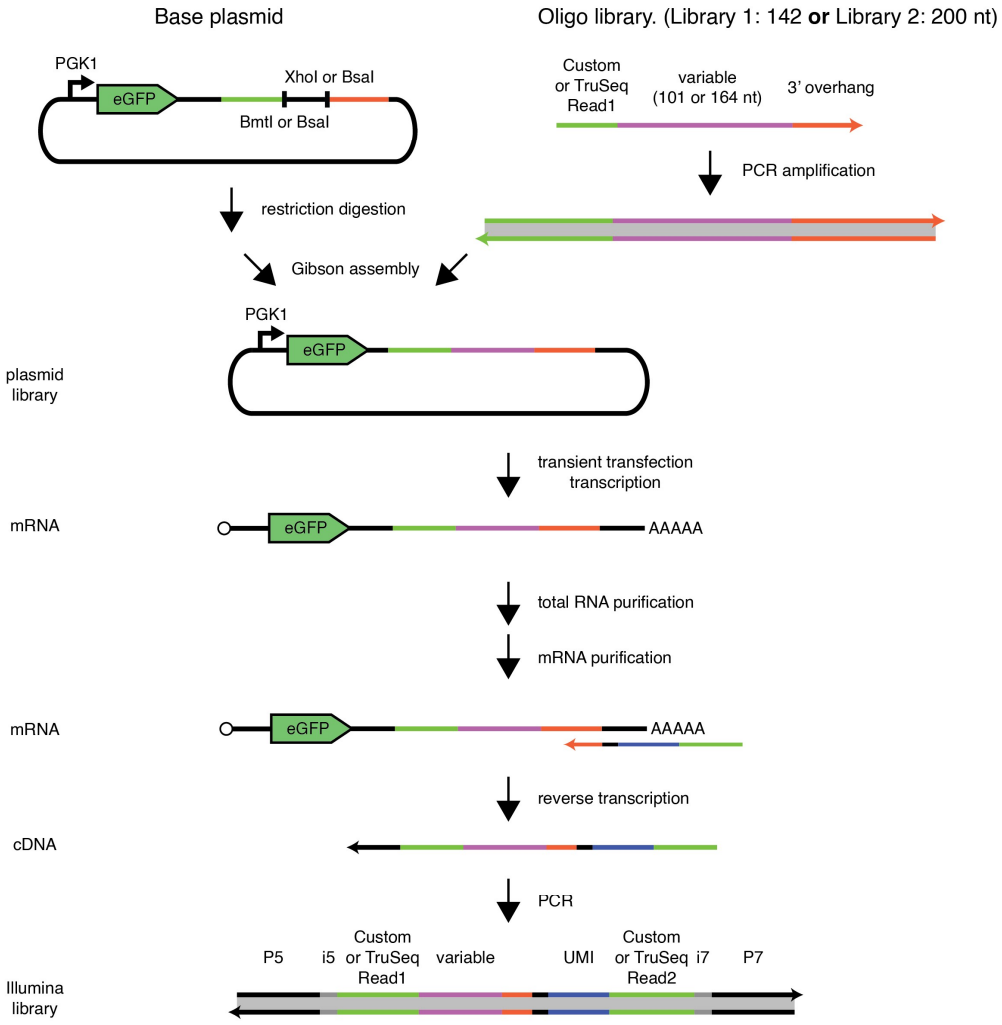


Fig. S1. Library construction and purification scheme. Assembly and processing of Library 1 and Library 2. Differences between the workflow for the two libraries are marked by an "or" (Library 1 or Library 2). Library oligo pools are amplified via PCR to introduce overhangs. The base cloning plasmid is digested with restriction enzymes and the amplified oligo pool is inserted via Gibson assembly to generate plasmid libraries. After transient transfection into the target cell lines, mRNAs transcribed from the library plasmids are first purified as total RNA, then with a poly(A)-based mRNA isolation kit. Reverse transcription is performed with a primer containing a unique molecular identifier (UMI). The cDNA is amplified with i5 and i7 index primers and sequenced. The primary differences between the library preparation for the two libraries are the restriction sites used for digestion of the base plasmid, the length of the variable region, and the use of custom read 1 and read 2 sequencing primers for Library 1.

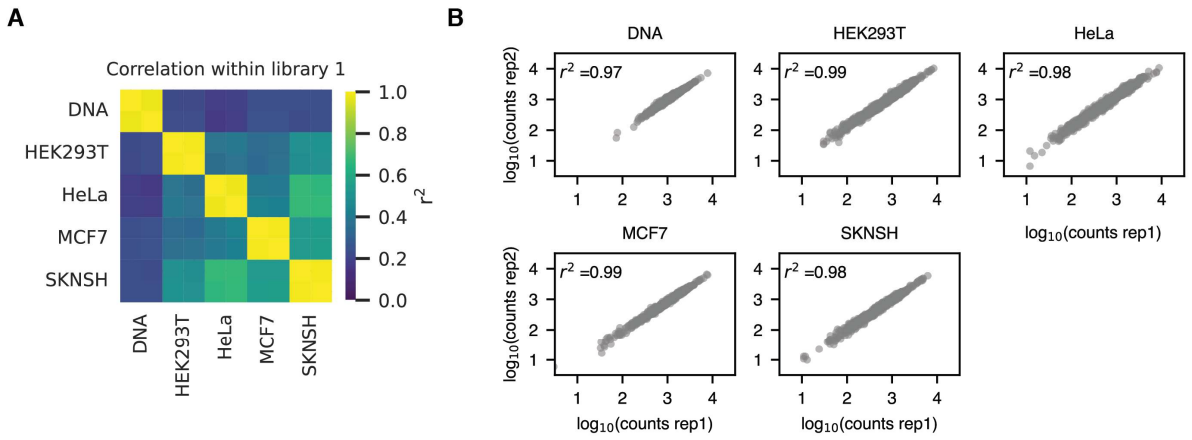


Fig. S2. Correlation between Library 1 replicates. (**A**) Correlation between counts across replicates and cell lines for Library 1. (**B**) Comparison of counts for individual designs between the two library replicates for different cell lines.

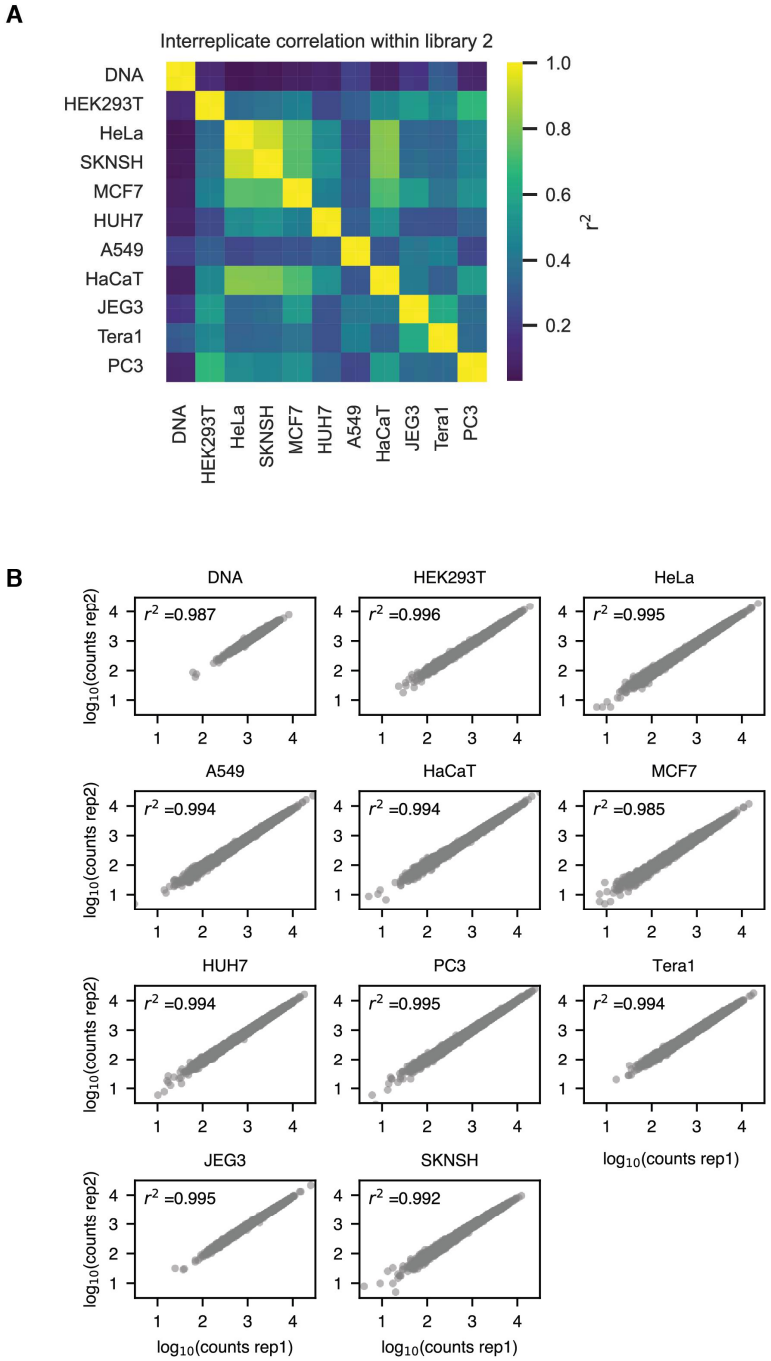


Fig. S3. Correlation between Library 2 replicates. (A) Correlation between counts across replicates and cell lines for Library 2. (**B**) Comparison of counts for individual designs between the two library replicates for different cell lines.

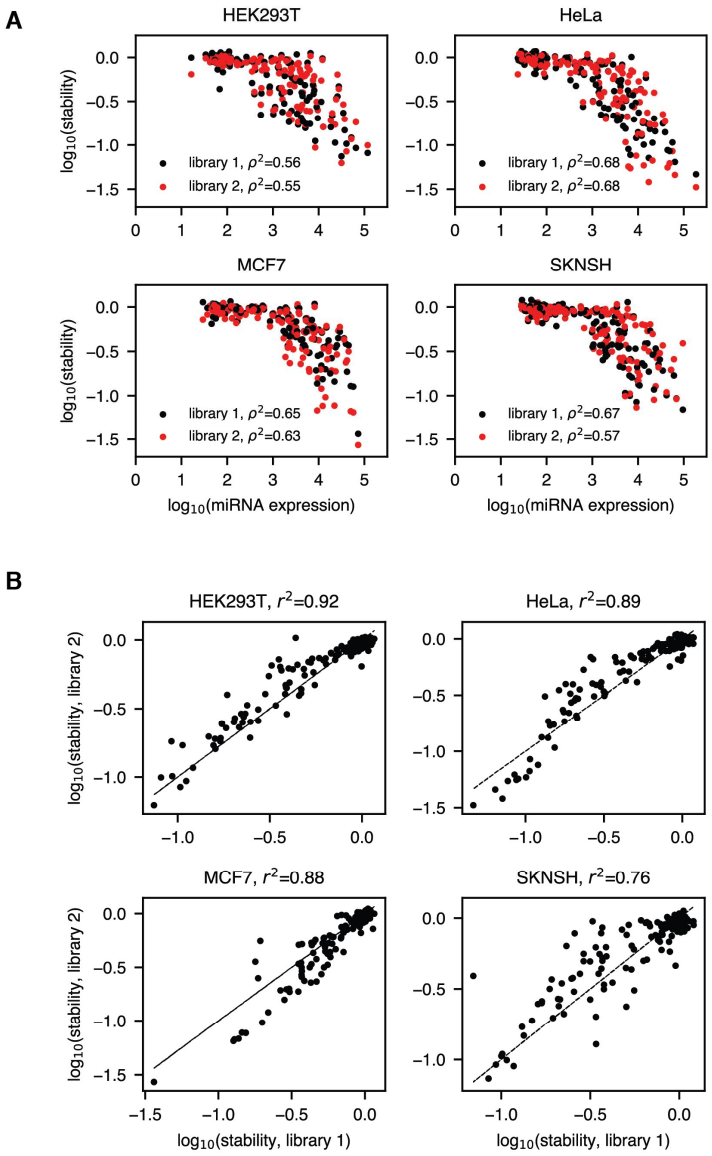


Fig. S4. Comparison of the results for Library 1 and Library 2. (A) Microarray-measured miRNA expression (*22*) versus the measured stability for miRNA targets and cell lines shared between Library 1 and 2. The distributions largely agree. The legend shows the Spearman ρ^2 between expression and stability. (**B**) The measured stabilities for both libraries. Note that the main context sequences are different for the two libraries. The relatively large variation for SKNSH could imply biological differences in the cell line miRNA concentrations at the time of measurement for the two libraries.

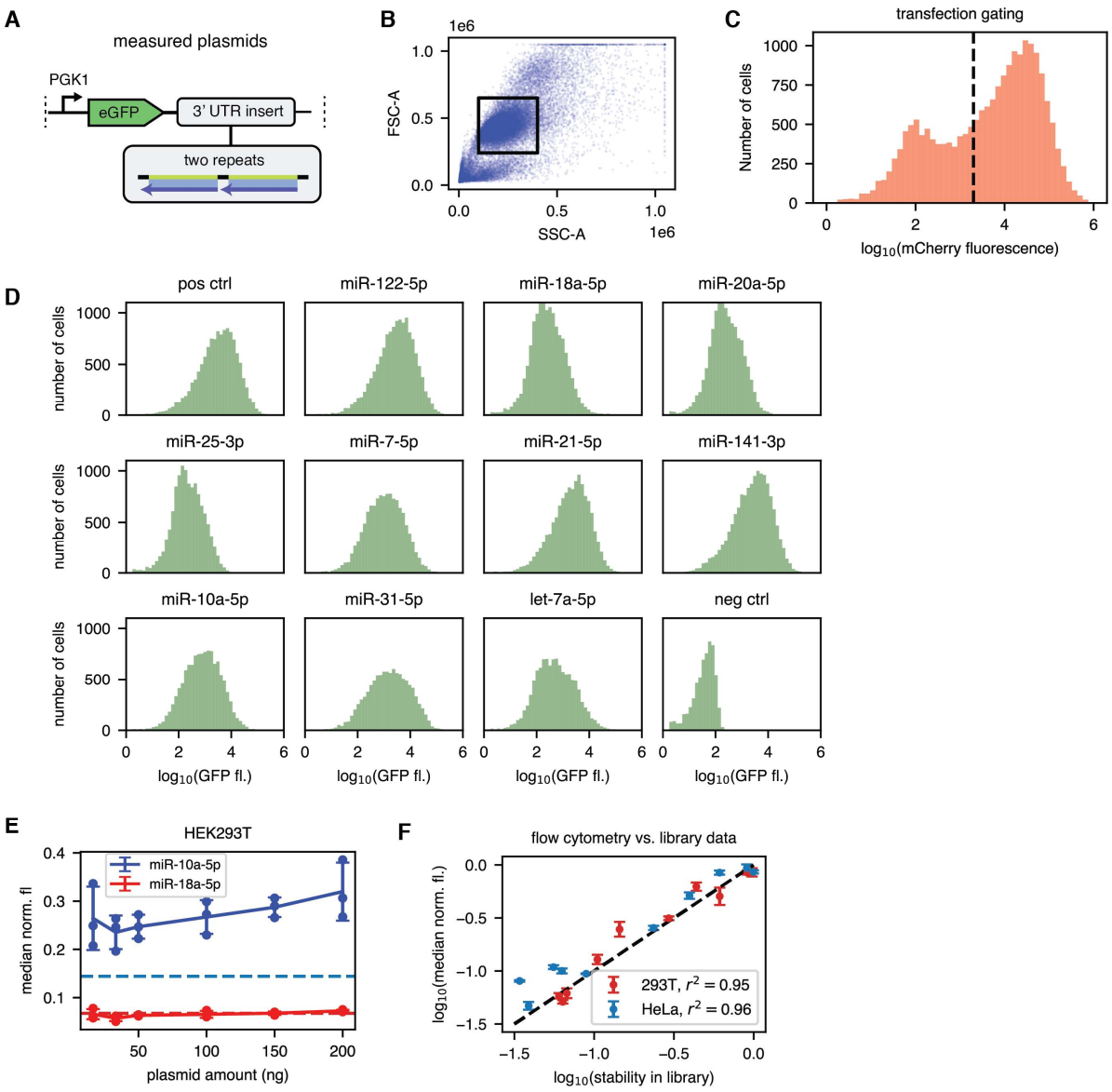


Fig. S5. Flow cytometry results for individual constructs from library 1. (**A**) We measured constructs containing two repeats of a microRNA target site. The sequence is identical to the same constructs in Library 1. (**B**) Gating strategy on SSC-A and FSC-A for HEK293T cells. (**C**) The gating for transfected cells using an mCherry-expressing transfection control. (**D**) GFP fluorescence distribution for one replicate in HEK293T after filtering for transfection using the mCherry signal. The negative control is shown without mCherry gating. (**E**) Median fluorescence normalized to the median fluorescence of the positive control for two different microRNA target types versus the transfected plasmid concentration. The result is largely independent of plasmid concentrations when using the median fluorescence. N=3 (**F**) Relationship between the stability measured for Library 1 for our high-throughput stability data and the median GFP fluorescence level normalized to the median fluorescence of the positive control measured using flow cytometry for two different cell lines (HEK293T, HeLa). N=3

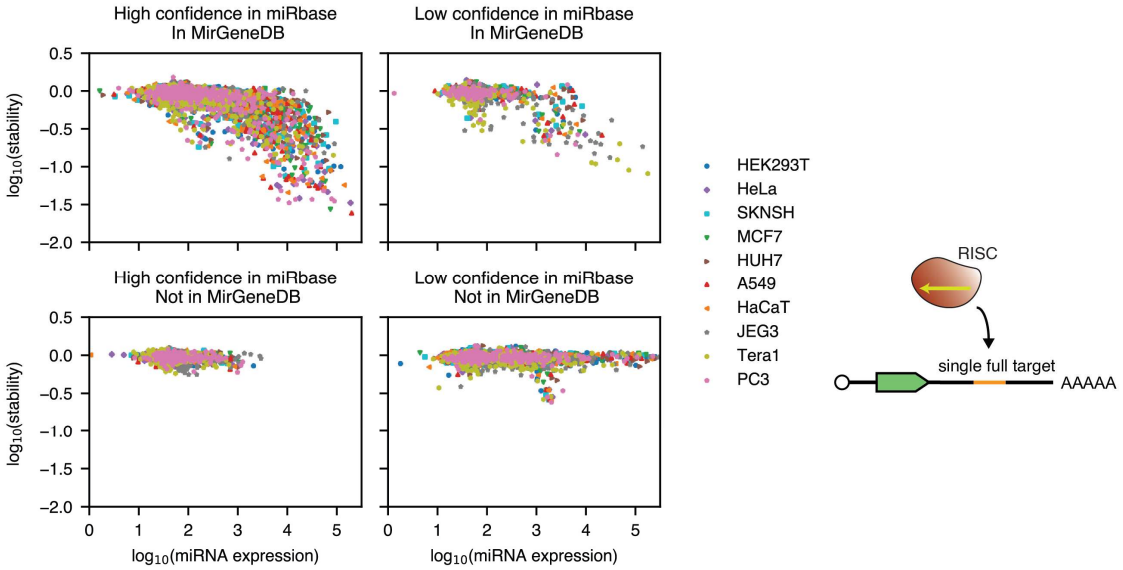


Fig. S6. The impact of the microRNA annotation data source on the observed behavior for single target sites. The microRNA expression data is from Alles et al. (*22*). MicroRNA targets in miRBase are annotated as high confidence or low confidence depending on the available level of experimental evidence for their existence. Later additions to miRBase are often low confidence sequences derived purely from sequencing data without any direct biological evidence. MirGeneDB is a manually curated microRNA database that more closely considers clear experimental evidence. Most, though not all, microRNAs that are in MirGeneDB are annotated as high confidence in miRBase. We find that the association between miRNA expression and target stability depends strongly on the annotation data source: All microRNAs in MirGeneDB approximately follow a monotonous relationship regardless of their status in miRBase, validating them as likely real microRNA. The microRNAs that are high confidence in miRBase but not in MirGeneDB have insufficient expression in our measured cell lines to tell either way. Low confidence microRNA that are not in MirGeneDB generally show no activity whatsoever even when (erroneously) measured as highly expressed. We therefore exclude them from the analysis.

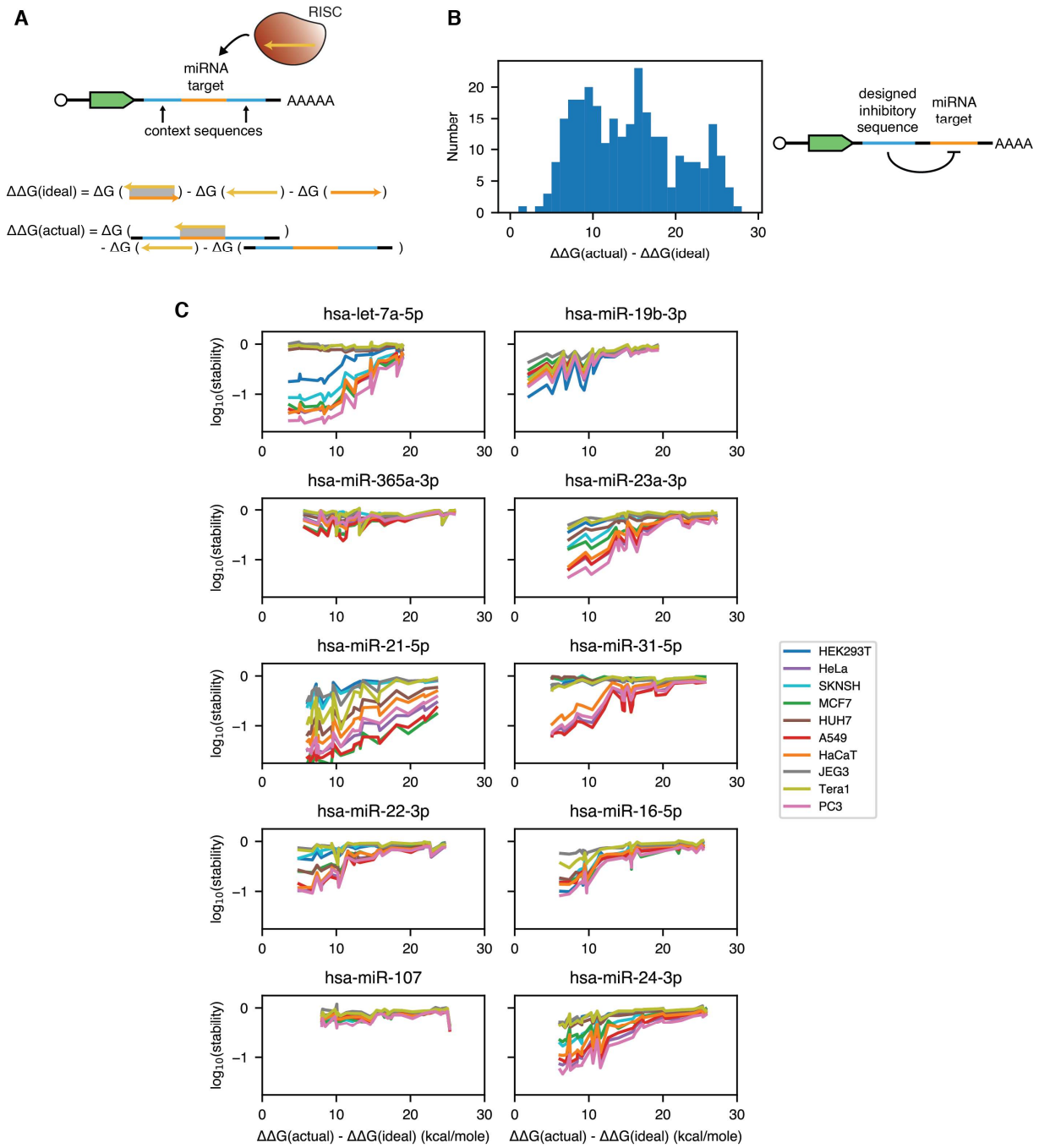


Fig. S7. Strong target secondary structure inhibits microRNA function. (**A**) We calculate the binding energy (ΔΔG) between the miRNA and its target without and with the surrounding context sequence (*44*). The ΔΔG value shown on the x-axis for all graphs investigating secondary structure in this publication is the difference between these two values. (**B**) Our engineered context sequences display a wide distribution of ΔΔG differences, allowing us to investigate the dependence of miRNA activity on target secondary structure. (**C**) The measured stabilities of identical miRNA targets depend on the ΔΔG difference between ideal and actual binding energies. The effect is relatively consistent across cell lines and microRNAs. A single connected line in this graph belongs to a single microRNA target site occluded by different amounts of secondary structure in a single cell line.

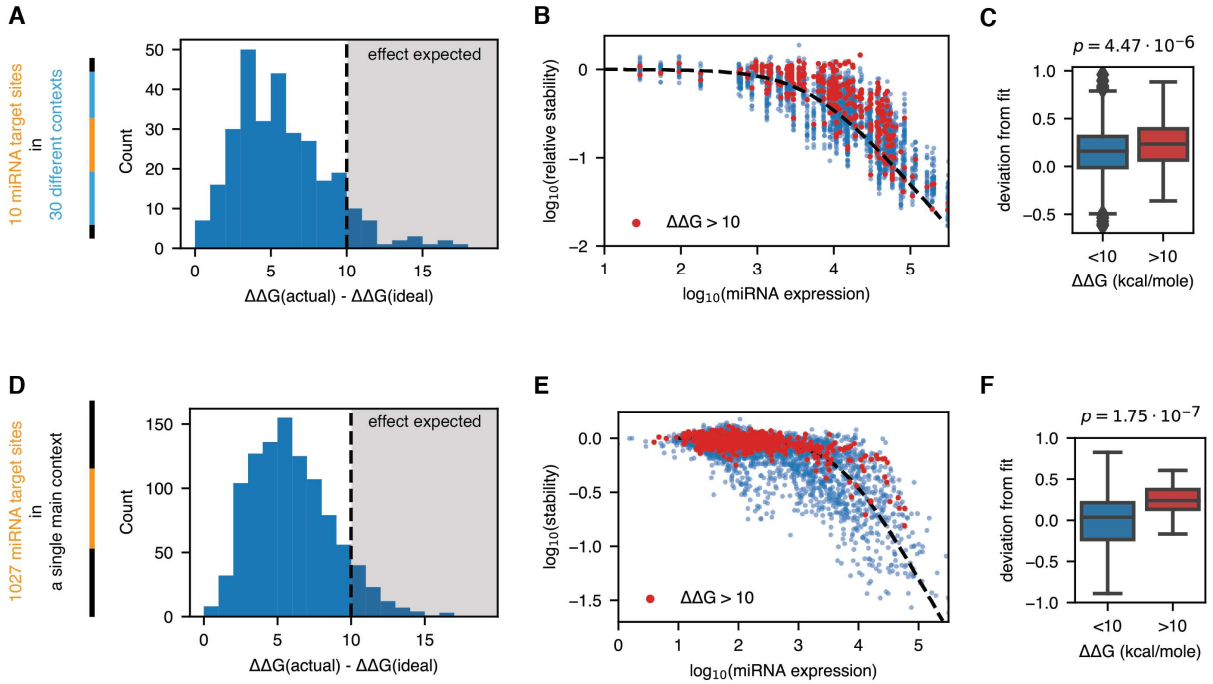


Fig. S8. The impact of secondary structure is generally small for natural context sequences. (**A**)-(**C**): Results for 10 miRNA target sites in 30 context sequences. (**D**)-(**F**): Results for 1027 miRNA target sites in the main context sequence. (**A**)/(**D**) Distribution of the binding energy differences. (**B**)/(**E**) MicroRNA expression and target transcript stabilities for low and high (>10 kcal/mole) binding energy differences. (**C**)/(**F**) Distribution of the fit deviation values for low and high binding energy differences. We constrained the calculation to miRNAs with an expression greater than $10^{3.5}$ tpm to exclude miRNA targets with no expected effect on stability. The p-value was calculated using a one-sided Mann-Whitney U test.

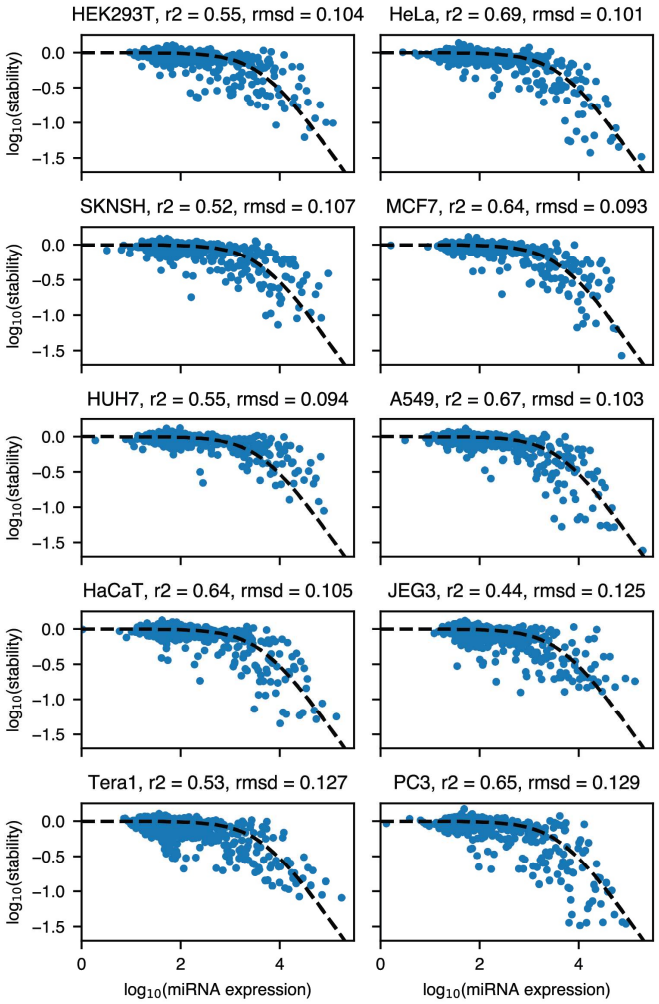


Fig. S9. Fit results for individual cell lines and normalized microRNA expression data. The x-axis contains microarray microRNA expression data (*22*). We fit a single universal transfer function for all cell lines. The titles show the Pearson r^2 values and the root mean square deviation (rmsd) between the measurements and the transfer function fit.

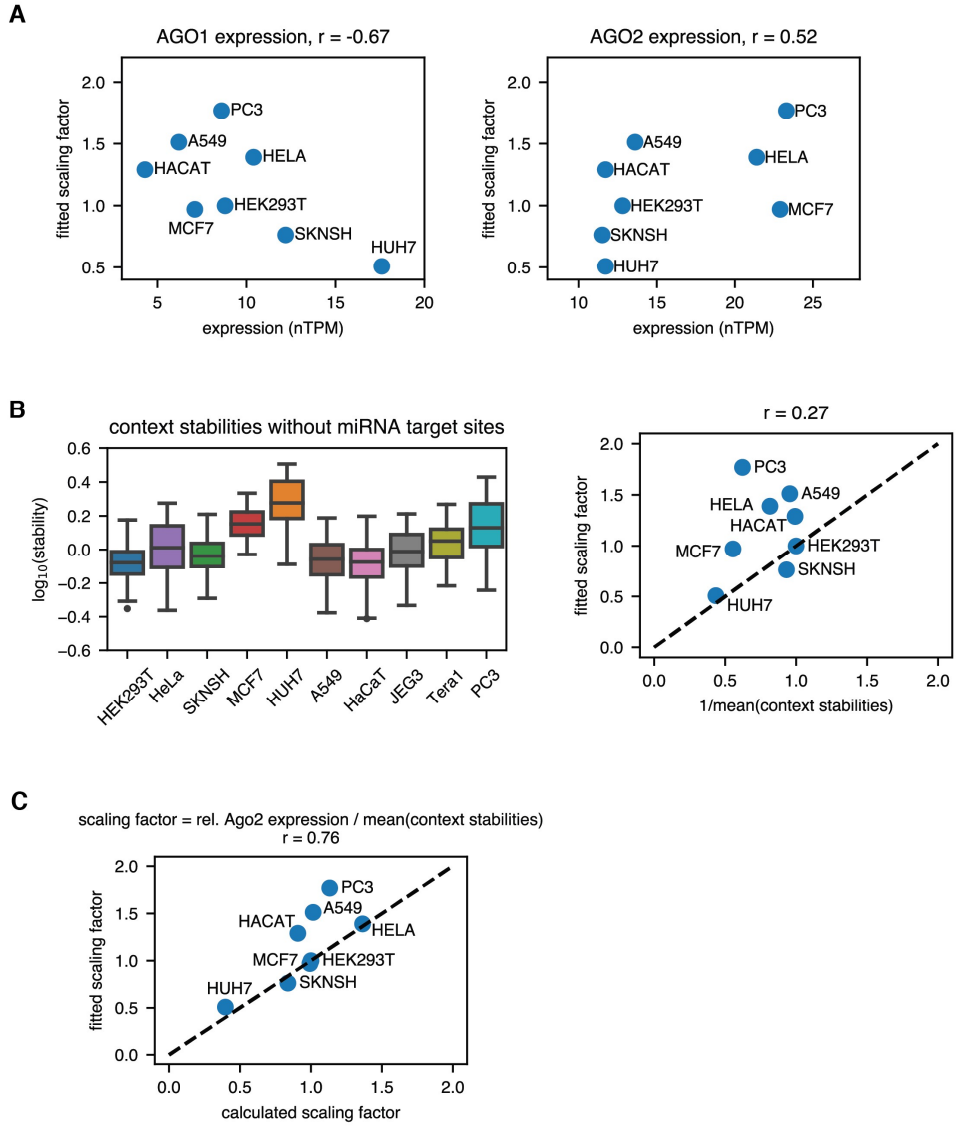


Fig. S10. Ago2 expression and context stability explain the scaling factor. (**A**) Ago1 and Ago2 expression according to protein atlas data versus the fitted scaling factor for individual target sites. Ago1 expression is negatively and Ago2 expression is positively correlated with the scaling factor. We used Ago2 expression as a proxy for the total amount of slicing-competent miRNA in a cell. (**B**) Left: The distribution of baseline stabilities for 51 different 3'UTRs (Methods) without miRNA target sites. Because normalization is performed relative to the stability of the main context, shown stabilities are also relative to the stability of the main context sequence. Right: The inverse of the geometric mean of the stabilities shown on the left versus the fitted scaling factor. This value approximates the relative stability of the main context across cell lines. (**C**) Dividing the relative Ago2 expression normalized by the HEK293T value by the geometric mean of the context stabilities yields a good estimate of the overall scaling factor.

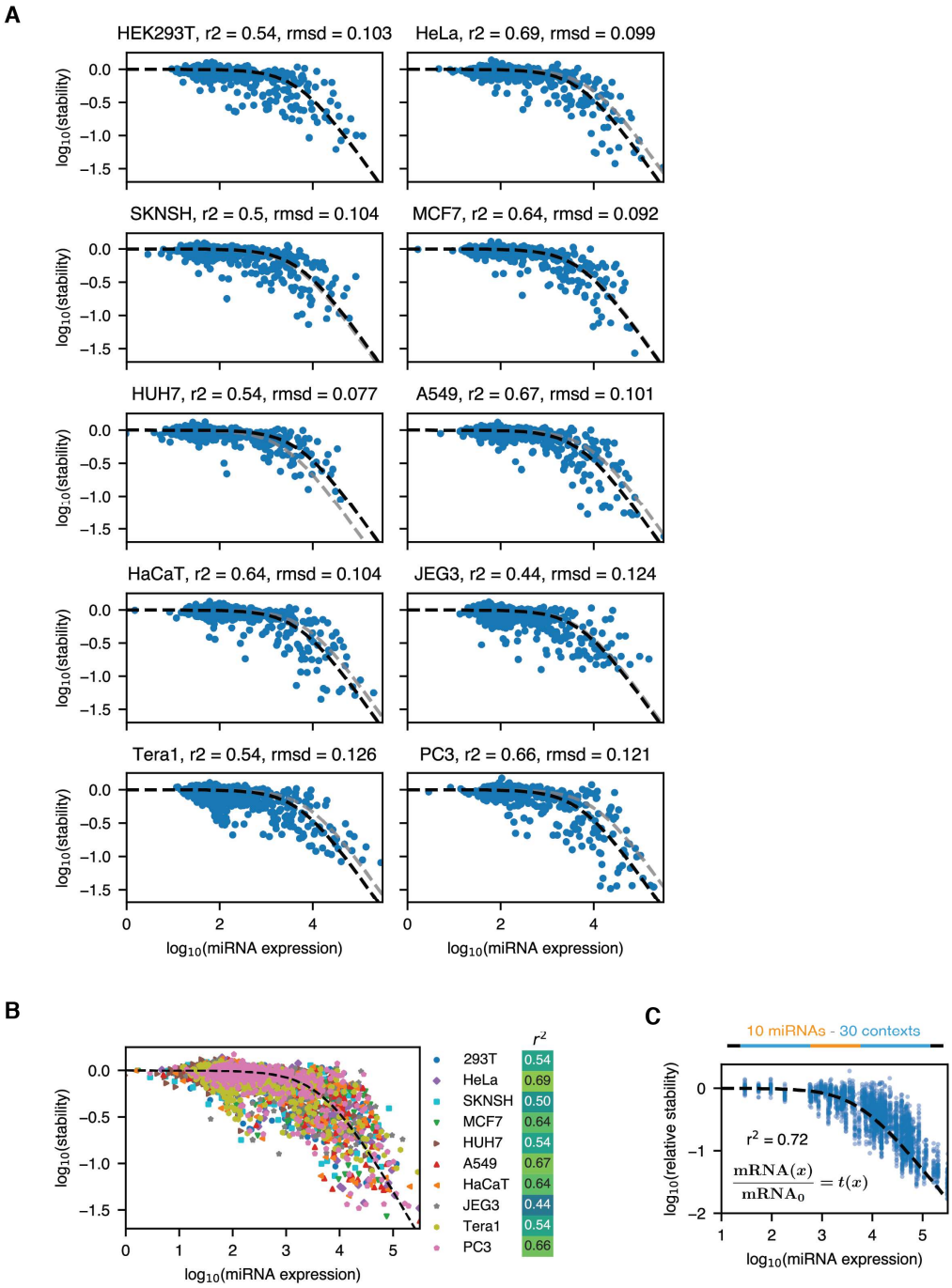


Fig. S11. A scaling factor for the total microRNA expression levels improves the model fit. (**A**)/(**B**) Scaled microRNA expression versus stability for target sites in the main context sequence. The titles and the heatmap show the Pearson r^2 values and the root mean square deviation (rmsd) between the measurements and the transfer function fit. The adjusted transfer function is shown in black and the unscaled transfer function in gray. (**C**) The scaling factor also improves the predictions of the relative stability of 3'UTRs containing miRNA targets in different context sequences.

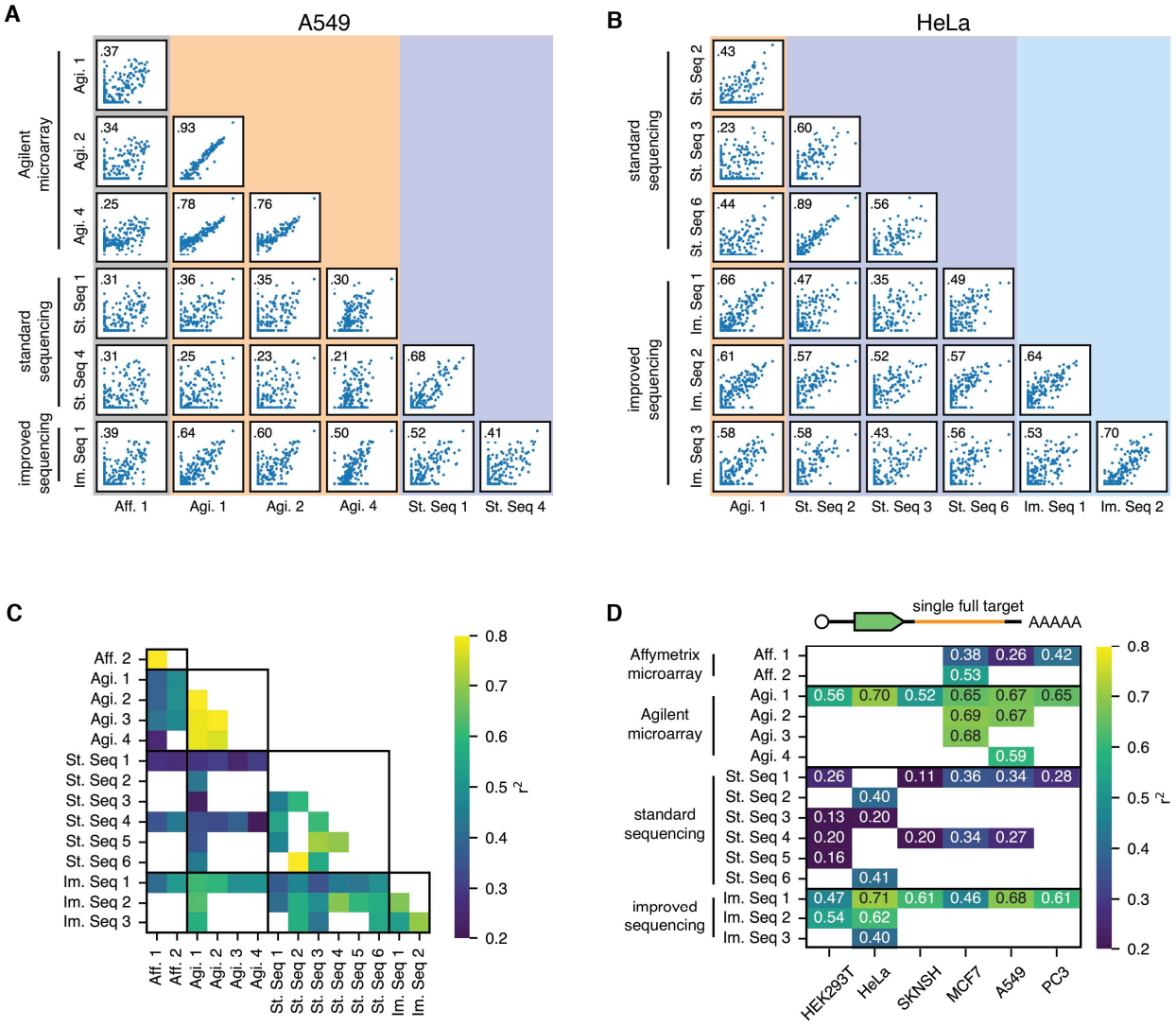


Fig. S12. Cross-dataset correlation predicts well-performing datasets. Cross-dataset correlation for (**A**) A549 and (**B**) HeLa cells. All values are set to a minimum of 100 tpm and the analysis is constrained to miRNAs common to all datasets. (**C**) Cross-dataset correlation averaged across all cell lines. (**D**) Correlation between the fitted transfer function using different microRNA expression datasets and measured stability for single microRNA target sites. Well-performing datasets have both high correlation with other datasets within their collection method and also with some other collection methods.

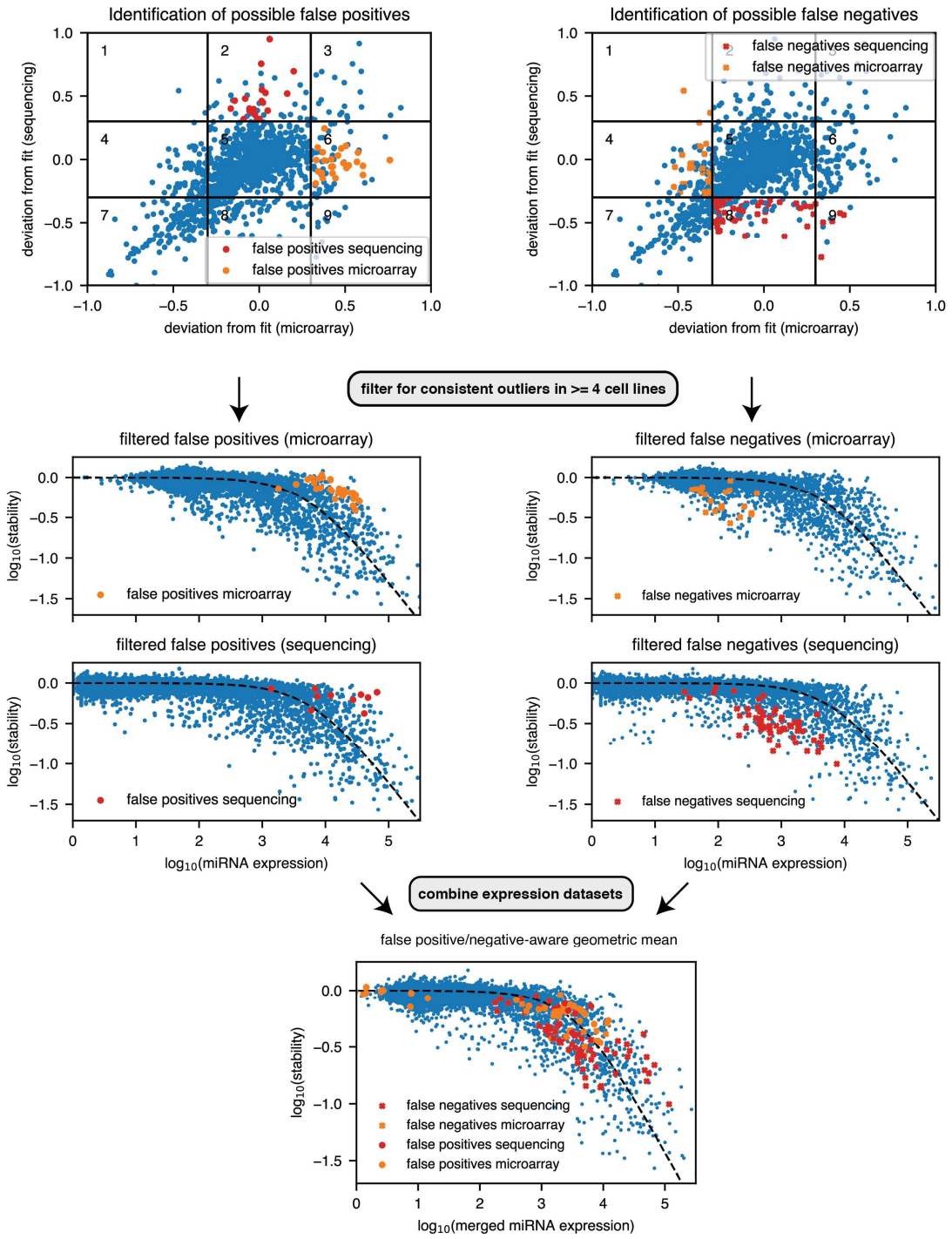


Fig. S13. A comparison of datasets identifies biased data points. We compare the transfer function fit for a microarray (*22*) dataset and the sequencing dataset collected for this study for all cell lines. First, we identified potential false positives and negatives in both datasets as outliers from the fit in only one of the two datasets (top plots, see methods for details). Then, these potential outliers were filtered for agreement across our cell lines: An outlier in a single cell line could represent measurement noise or genuine biological differences in the microRNA expression at measurement time. An outlier miRNA that is consistent across multiple cell lines is likely to be due to technical differences in the expression measurement method. The middle plots highlight miRNAs that were

identified as false positives and negatives in the two datasets in the expression versus stability plot. After merging, the previous false positives and negatives were predicted well by the transfer function, indicating that they were indeed incorrectly measured in the other dataset (bottom plot).

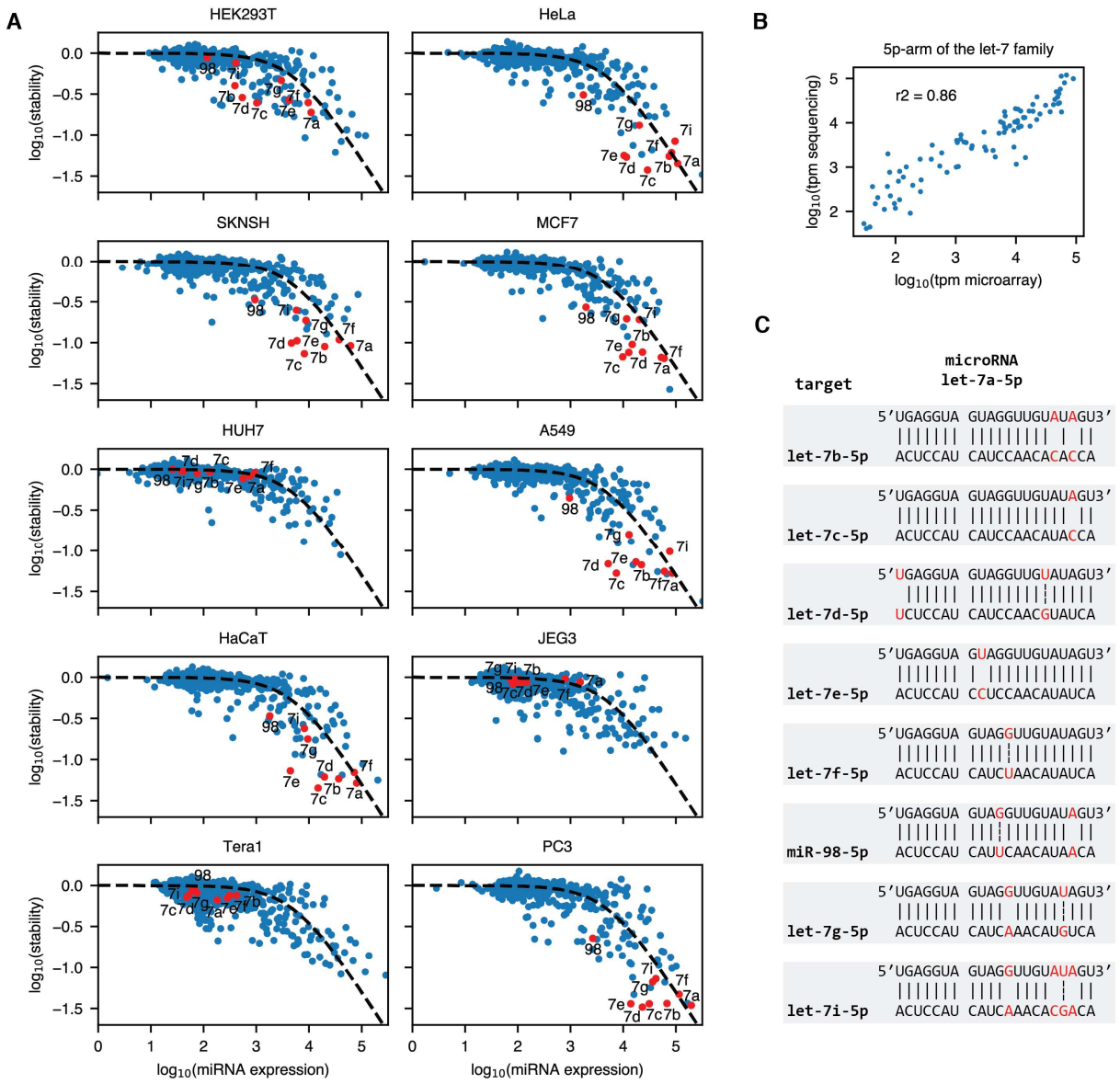


Fig. S14. Crosstalk between let-7 family members causes outliers from the fit. (**A**) In the stability versus expression plot, targets of the 5p arm of the let-7 family consistently appear as some of the strongest outliers with less stability than would be expected from the transfer function. The labeled miRNA targets are abbreviated from the full hsa-miR/let-x-5p name. (**B**) The let-7 family member expression is consistent between the microarray and sequencing measurements, indicating that the observed crosstalk is not a result of bias in the expression measurements themselves. (**C**) The mismatch pattern between an attacking hsa-let-7a-5p miRNA and target sites belonging to other let-7 family members generally only includes one or two mismatches (red, gap without dash) or wobble base pairs (red, dashed line), which explains qualitatively why such strong crosstalk is observed.

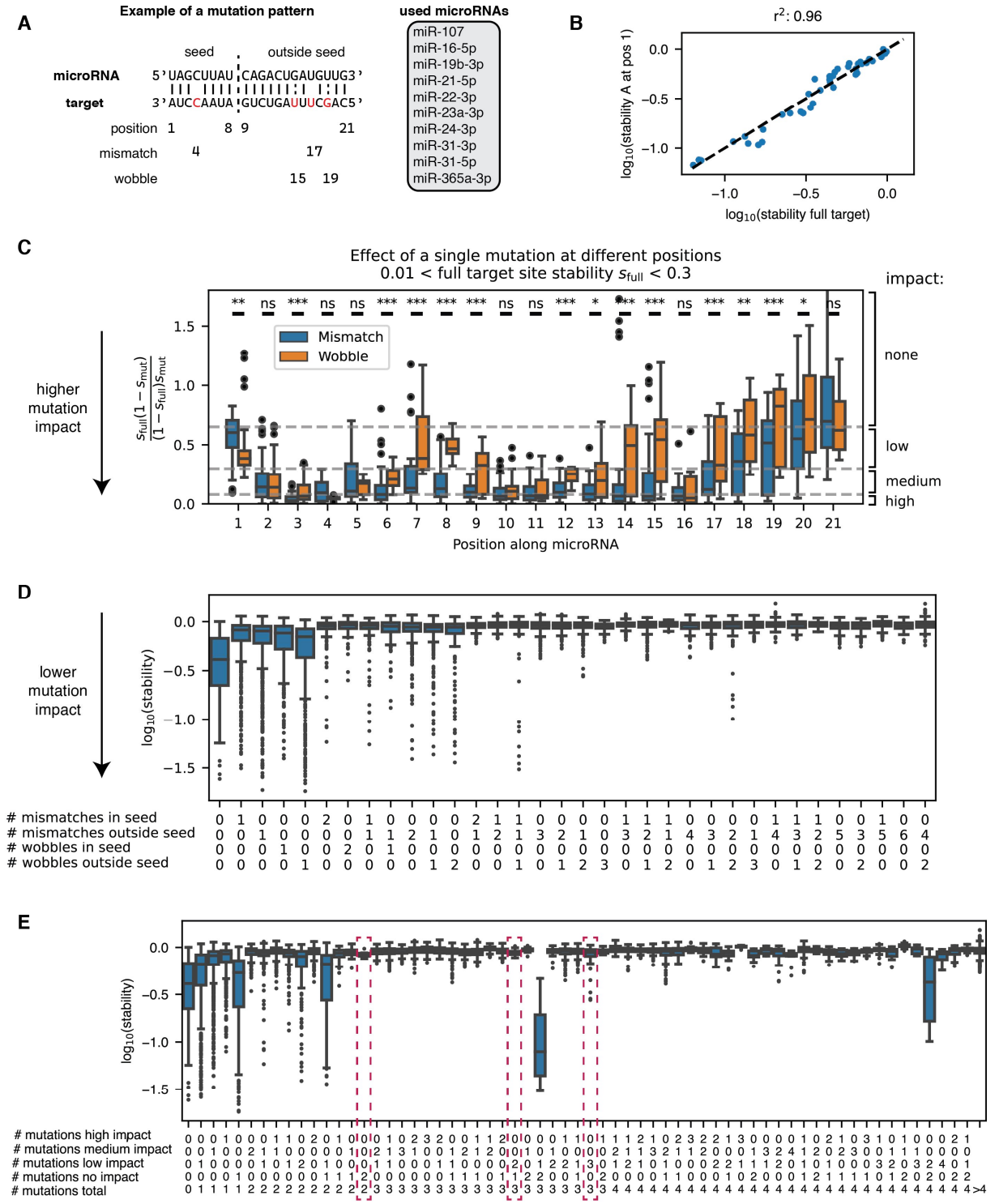


Fig. S15. The impact of mutations can be classified based on their position in the miRNA. (A) Example of a mutation pattern between a microRNA and a target. Note that the mutation position is denoted based on the position in the microRNA rather than in the target site although the mutations are introduced in the target site. (**B**) Stability for full target sites and those with an A mutation at position 1. A mutation to an A at the first seed-complementary nucleotide in the target does not reduce miRNA activity. (**C**) Relative loss of activity due to individual target site

mutations (mismatches and wobbles). The dashed lines denote classification lines based on the median mutation impact. P-values for the difference between mismatches and wobbles were calculated using a two-sided Mann-Whitney U test. (**D**) Distribution of 3'UTR stabilities for single target sites containing multiple mutations grouped by the number and position of mutations and wobbles. (**E**) Distribution of 3'UTR stabilities for single target sites containing multiple mutations grouped by the number of mutations with a specific impact. Classification based on the mutation impact leads to a much cleaner classification than one based on the location inside or outside the seed. Dashed red boxes show examples of combined mutations with little individual impact that nevertheless strongly reduce miRNA activity on the target in combination.

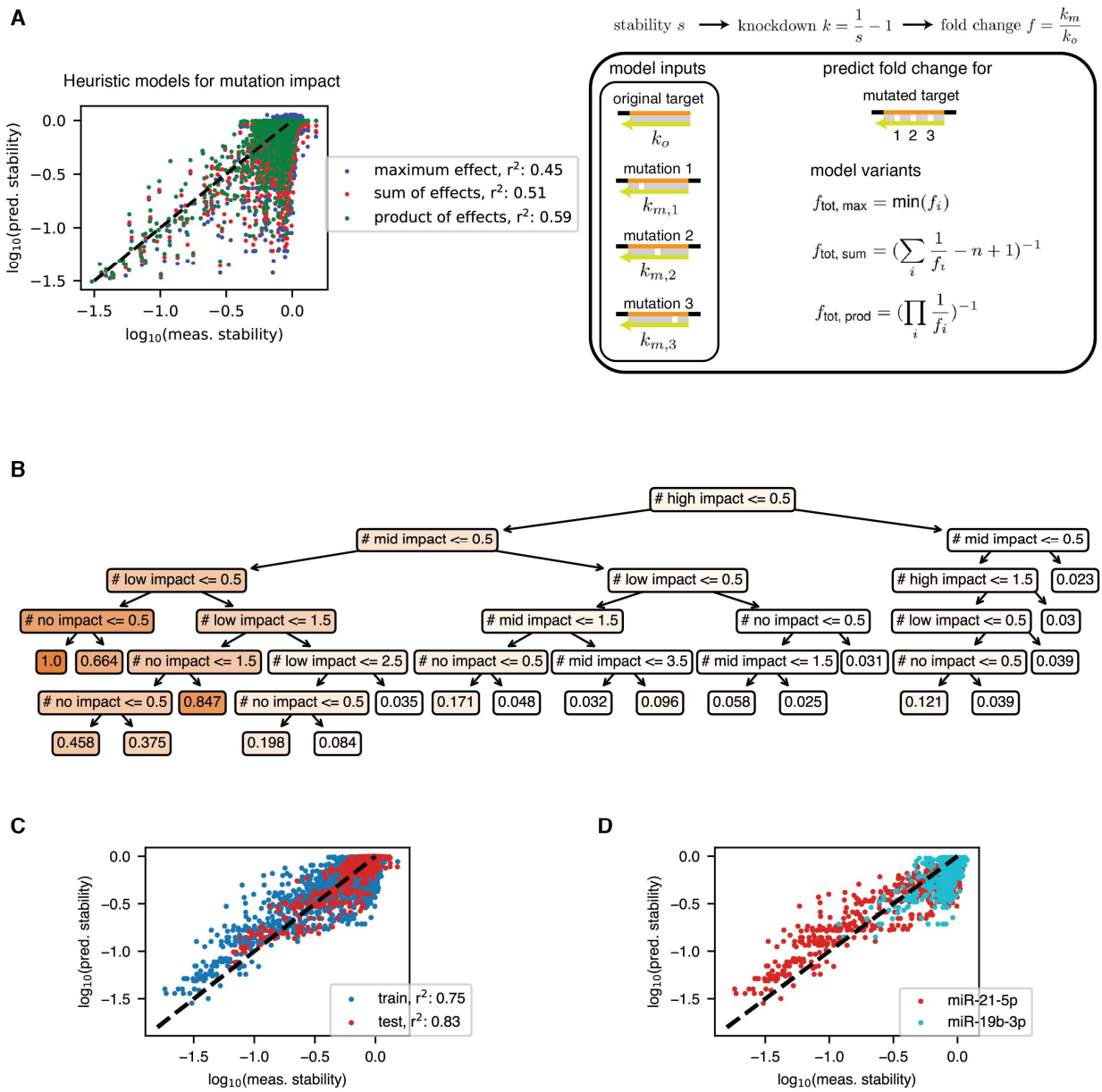


Fig. S16. A regression tree model predicts the impact of multiple target site mutations on miRNA activity. (**A**) The performance of three different heuristic models combining knowledge of the impact of individual mutations to calculate the overall mutation impact. None of the models work well because multiple individually inert mutations can combine to create a strong reduction in miRNA activity. (**B**) A regression tree that predicts relative knockdown values for a mutated target site from the presence and number of mutations classified as no, low, medium and high impact. (**C**) Performance of the regression tree on training and test data. The training was performed on data for all but one measured miRNA, the test data contains all mutation data for that specific miRNA (miR-31-5p). (**D**) Although the regression tree approximately captures the behavior of mutations, the behavior of mutations is strongly dependent on the individual miRNA sequence. While miR-21-5p is relatively resistant to mutations, miR-19b-3p quickly loses activity for mutated miRNAs.

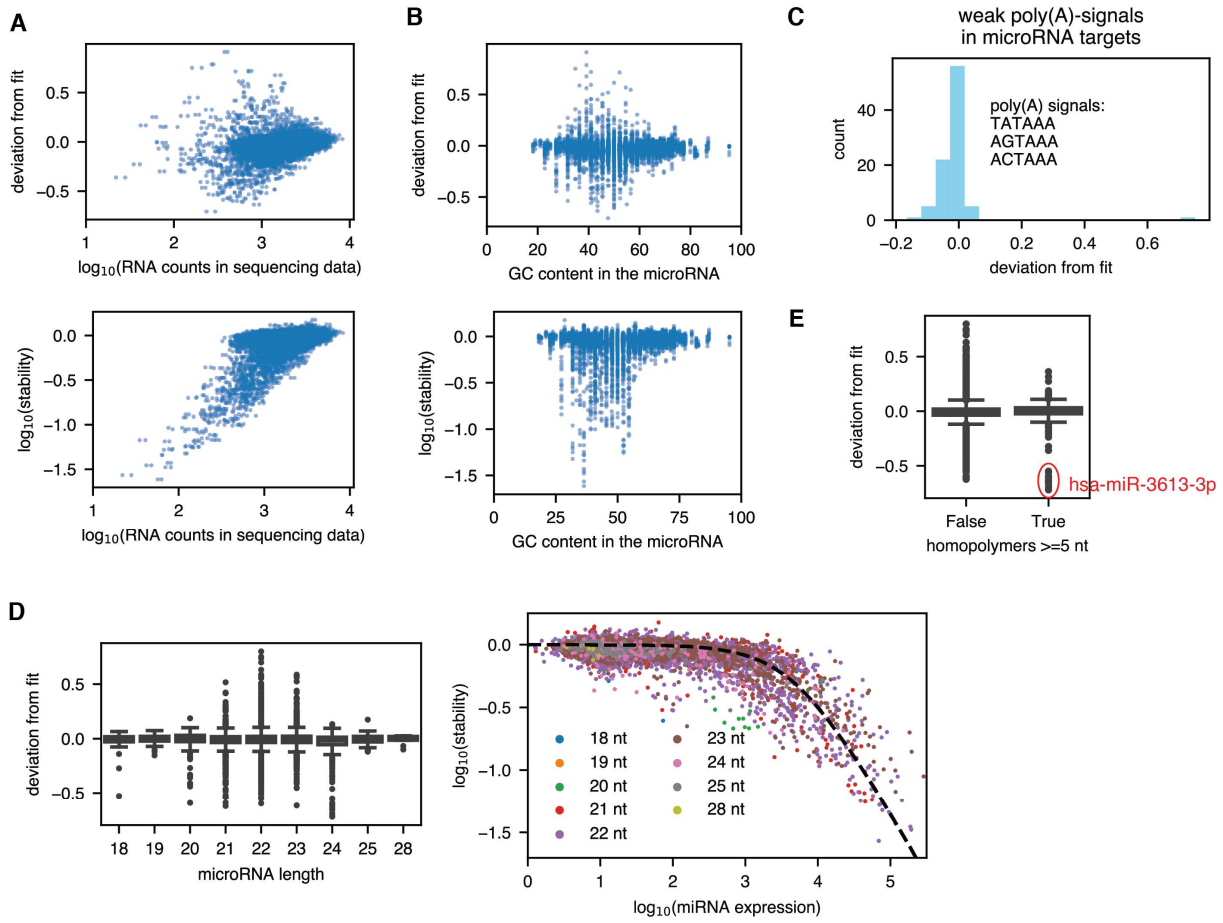


Fig. S17. RNA counts, miRNA GC content, weak poly(A) signals, miRNA length, and target site homopolymers have little effect on observed model deviations. (**A**) Counts in the RNA sequencing data (all cell lines) versus the observed deviation from the fit and the inferred stability. Strong deviation from the fit is not explained by undercounting. (**B**) MicroRNA GC content versus the observed deviation from the fit and the inferred stability. There is no clear association between GC content and observed deviation from the fit. (**C**) Weak non-canonical poly(A) signals in the target site do not lead to a lower measured transcript stability (negative deviation). (**D**) Deviation from the fit versus miRNA length and miRNA expression versus stability for the different miRNA lengths. In the design process, all miRNAs were standardized to a length of 21 nt by trimming from or adding uracils to the 3'end for the purpose of target site generation. We do not observe a pattern of reduced activity for short miRNAs. (**E**) Distribution of deviation values for microRNA targets depending on whether they contain contiguous homopolymers of a size of at least 5. Only hsa-miR-3613-3p stands out as an obvious outlier. It contains a long stretch of As: ACAAAAAAAAAAGCCCAACCCUUC. The resulting long stretch of Us in the 3'UTR probably destabilizes the transcript independent of miRNA regulation.

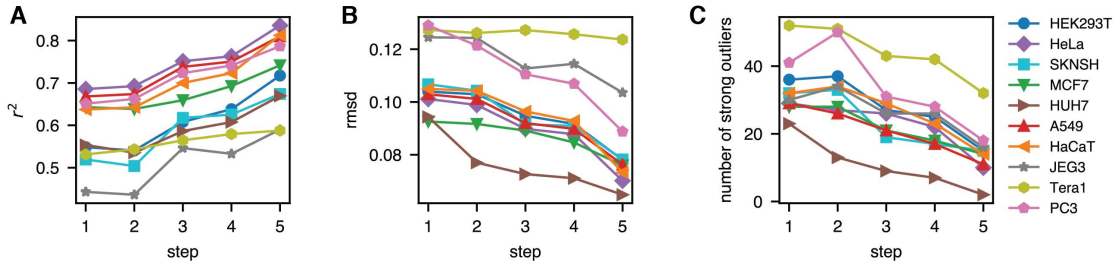


Fig. S18. Changes in the transfer function fit by processing step and cell line. (**A**) Pearson r^2 value. (**B**) Root mean square deviation (rmsd). (**C**) Number of strong outliers whose measured stability deviates by more than a factor of 2 from the predicted value. 1: unscaled microarray data, 2: scaled microarray data, 3: combination of microarray and sequencing data, 4: bias-aware merging, 5: removal of crosstalk.

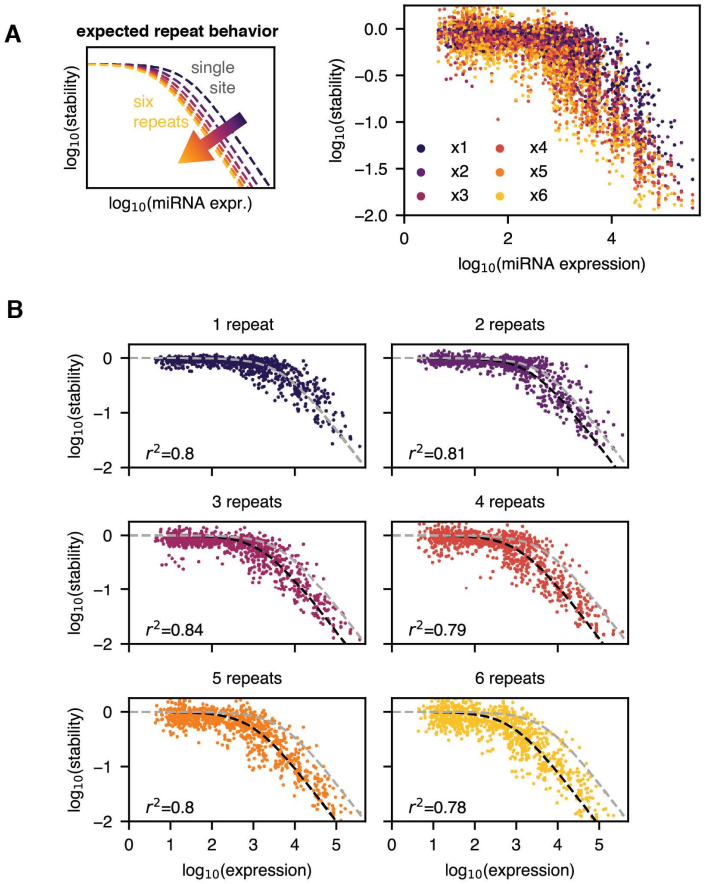


Fig. S19. The behavior of target repeats follows the additive model. (**A**) Stability versus miRNA expression for different repeat numbers across all measured cell lines. The plot on the left shows the expected behavior according to the additive model. (**B**) Stability versus miRNA expression for individual repeat numbers. The transfer function (dashed black line) uses the fitted parameters from Fig. 1. The dashed gray line shows the prediction for a single target site. The input miRNA expression is multiplied by the repeat number.

Fig. S20. Predictions of the additive model for two to six repeats of a single target site. The transfer function (dashed black line) uses the fitted parameters from Fig. 1. The dashed gray line shows the prediction for a single target site.

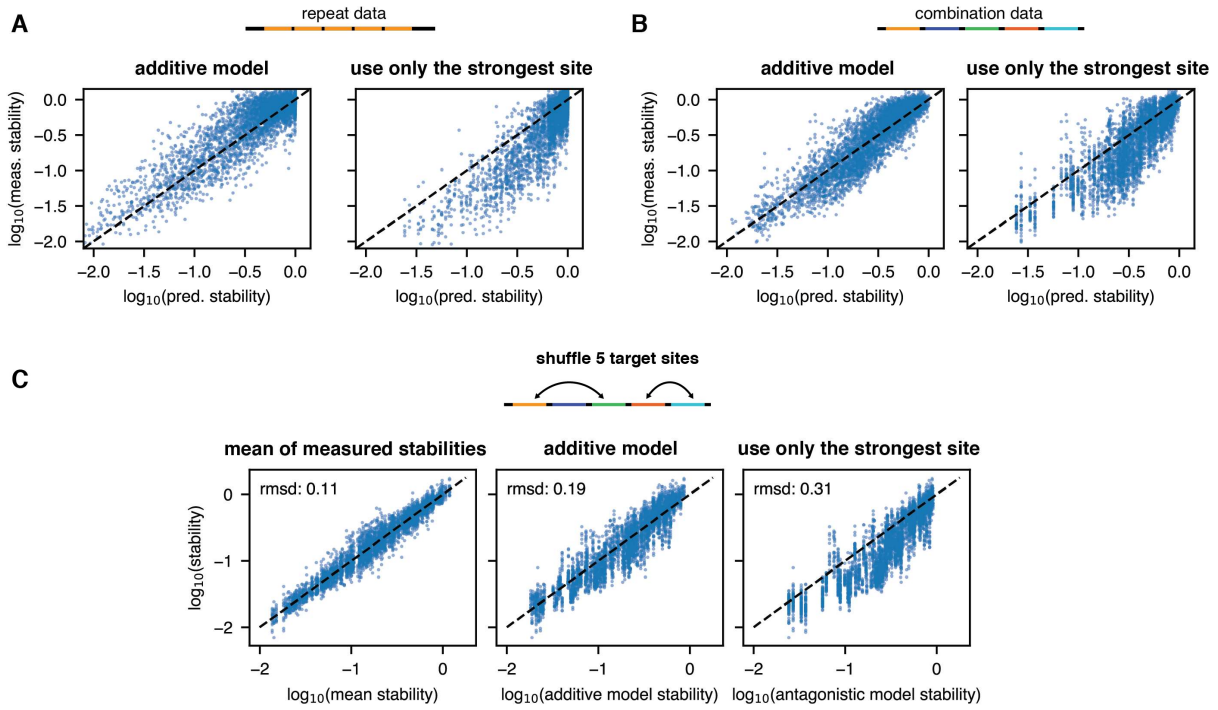


Fig. S21. Comparison of the additive model and antagonistic models using only the strongest site for repeat and combination data. All plots use inferred expression values derived by inversion of the transfer function. We compare the additive model discussed in the main text and an antagonistic model that only uses the strongest target site (i.e., the one whose cognate miRNA has the highest expression) to predict stability. (**A**) Predictions of the two models and measurements for target repeats. (**B**) Predictions of the two models and measurements or combinations of different targets. (**C**) Model comparison for shuffled target sites. On the left, we show predictions based on the measured mean stabilities of the 15 shuffled variants per set of target sites.

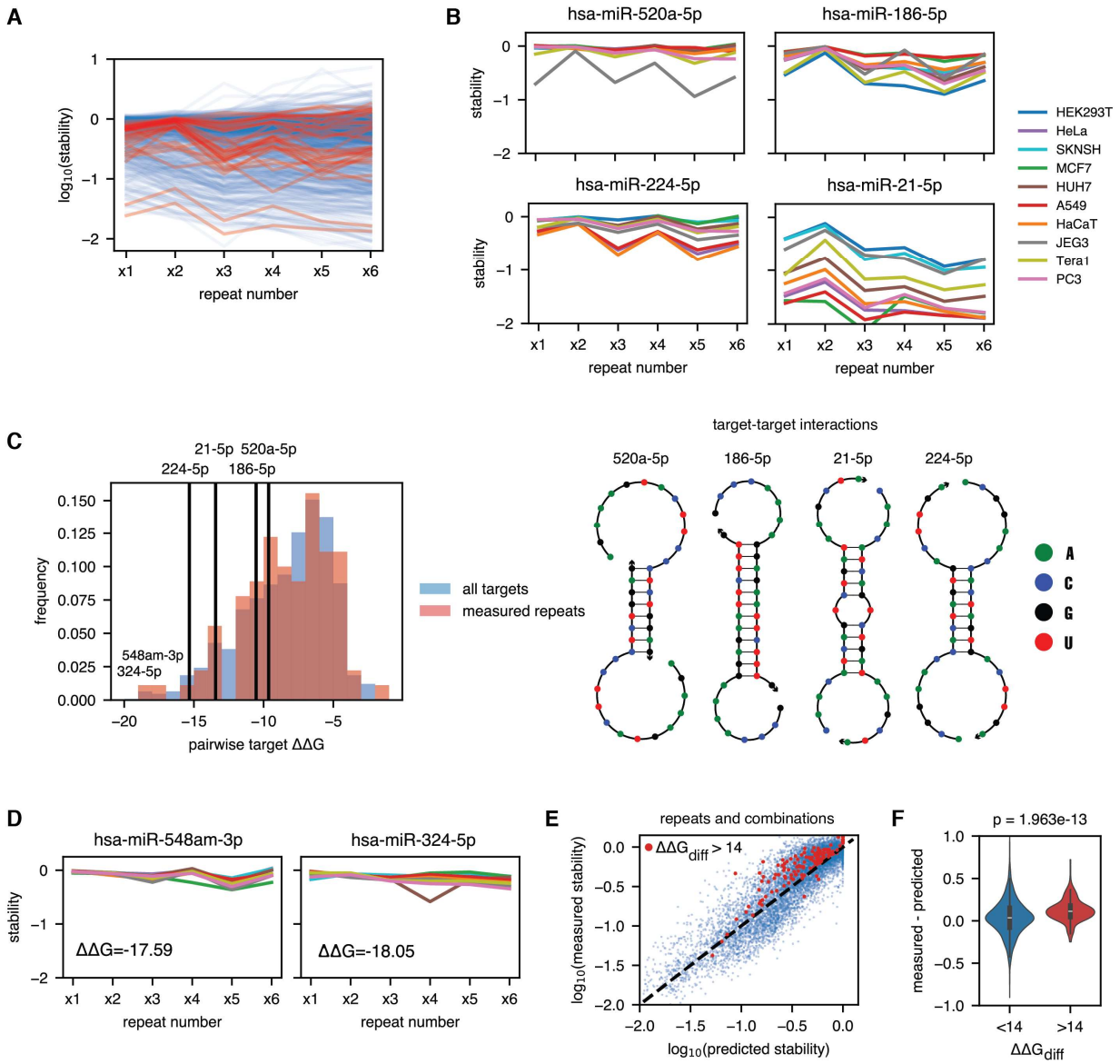


Fig. S22. Interactions between target sites explain outliers in the repeat stability data. (**A**). Stability of constructs containing one to six repeats of a single target site versus the number of repeats. Each solid line is for a single miRNA. Constructs with an unexpected pattern in which stabilities non-monotonically both increase and decrease at least twice with the number of repeats are highlighted in red. (**B**) Four example miRNA targets showing unexpected repeat stability patterns. There's an uptick in stability for an even number of repeats. (**C**) Left: Distribution of binding energies for two identical targets for all miRNAs (blue, high confidence in miRBase or in MirGeneDB) and of all miRNA targets we chose for measuring target repeats (Fig. 2B, S20). The interaction energies of the outliers shown in (B) and of the two most strongly self-interacting miRNA targets are highlighted. Right: Predicted secondary structures for target-target interactions for miRNAs shown in (B). (**D**) Observed stability patterns for the two most strongly self-interacting miRNA targets. These miRNAs do not show unusual stability patterns in our measured stability data because they are not significantly expressed in any of the cell lines. (**E**) Predicted versus measured stability for all measured repeats and combinations. Measurements where the dominant miRNA target is expected to be obscured by strong secondary structure are highlighted in red. (**F**) Difference between measured and predicted stability for constructs without and with strong secondary structure.

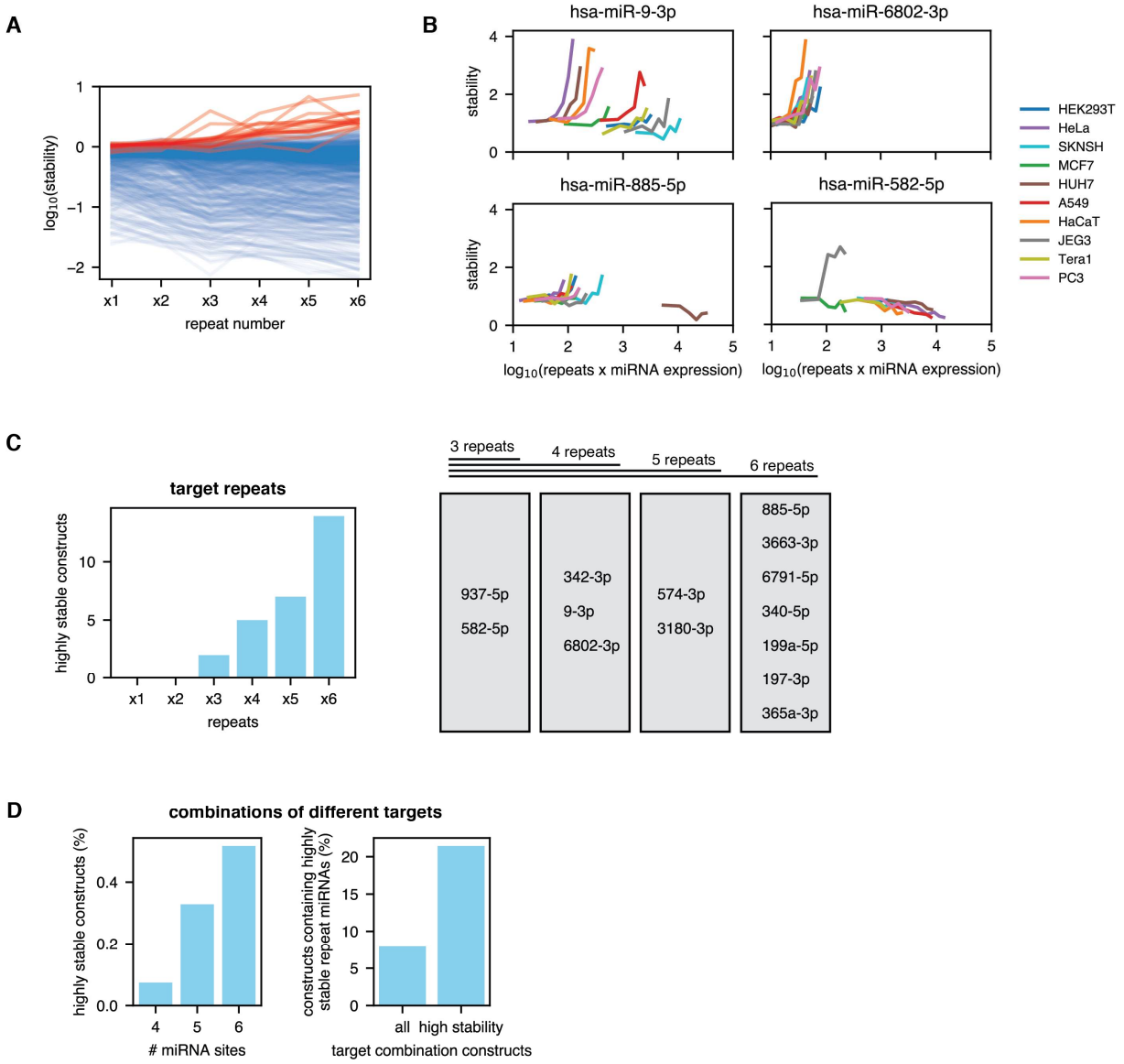


Fig. S23. Some miRNA target repeats cause strong increases in stability. (**A**) Stability of constructs containing one to six repeats of a single target site versus the miRNA expression. Each solid line is for a single miRNA. Constructs with an unexpected pattern in which there is a strong increase in stability with the repeat number are highlighted in red. (**B**) Four example miRNA targets that show a strong increase in stability in at least one cell line. Stability increases are more pronounced in cell lines where the cognate miRNA is not expressed. (**C**) The number of constructs with very high stability (>1.5) in at least one cell line increases with the number of repeats. All constructs that have high stability at a lower repeat number also show high stability at higher numbers of repeats. Stability-increasing miRNA targets behave consistently across repeat numbers. (**D**) For combinations of different miRNA targets, the number of constructs with very high stability also increases with the number of target sites (left plot). The fraction is far lower than for repeats of a single target site. Among those combinations that do have high stability, a greater fraction contains miRNA target sites that increase stability in constructs with four or more repeats (right plot).

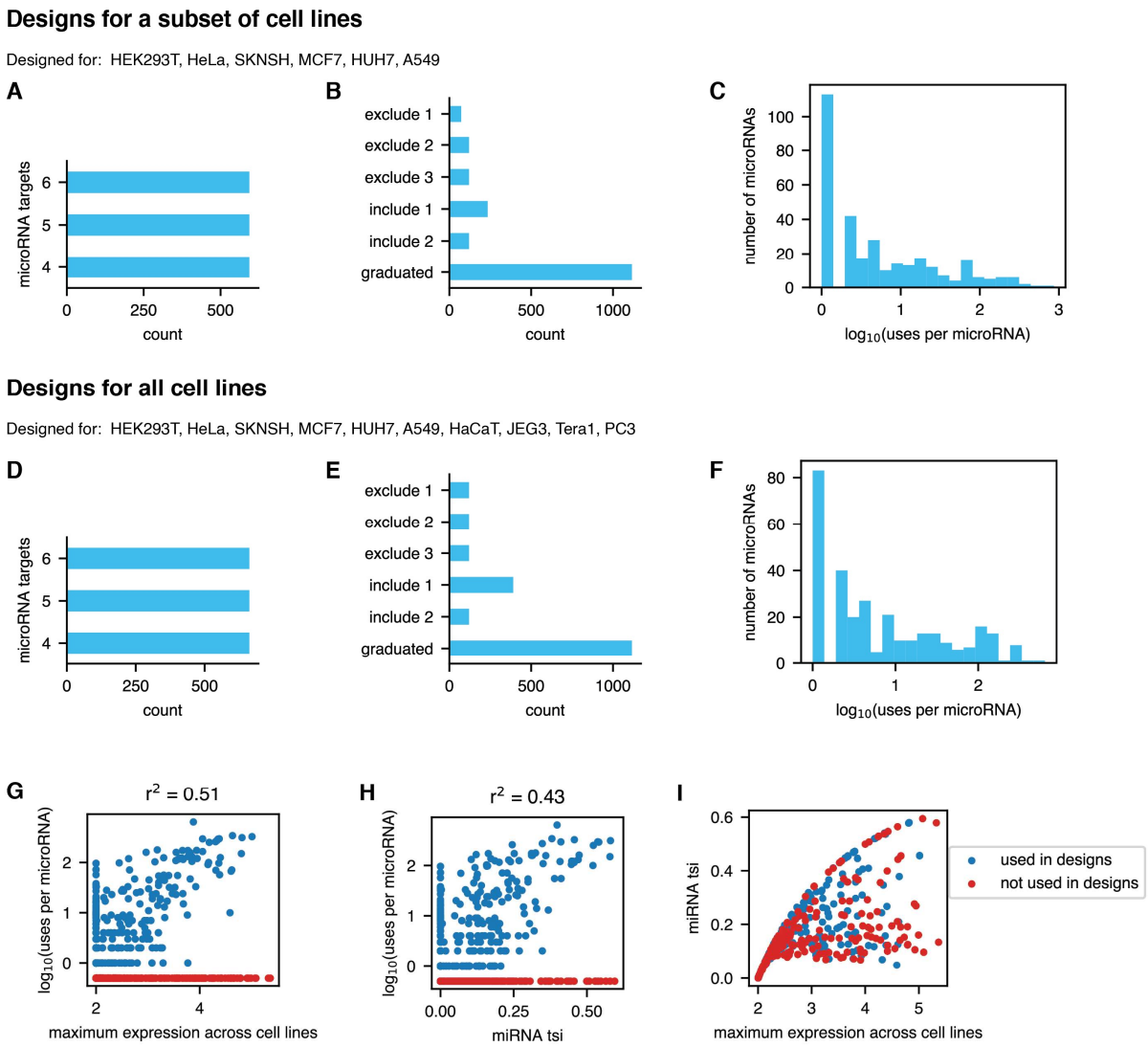


Fig. S24. Statistics on the generated designs. We created designs for either six of the ten cell lines (A)-(C) or all measured cell lines (D)-(I). (**A**)/(**D**). The number of designs with 4, 5, or 6 miRNA target sites. (**B**)/(**E**). The number of designs for different binary design targets including or excluding specific cell lines or graduated stability patterns. (**C**)/(**F**) Distribution of the number of times a specific miRNA target occurs across all designs. (**G**) The maximum expression level of a miRNA across cell lines and its usage frequency in our designs. More highly expressed miRNAs are used more often but many highly expressed miRNAs are nevertheless not used in any design. The correlation value is only calculated for used miRNAs. (**H**) Association of miRNA usage in our designs with the tissue-specificity index (tsi) (*35*) of the miRNA across our measured cell lines. More tissue-specific miRNAs are used more often but many tissue-specific miRNAs are not used in any of our designs. (**I**) Maximum expression versus tissue-specificity for miRNAs that were either used or not used for our designs.

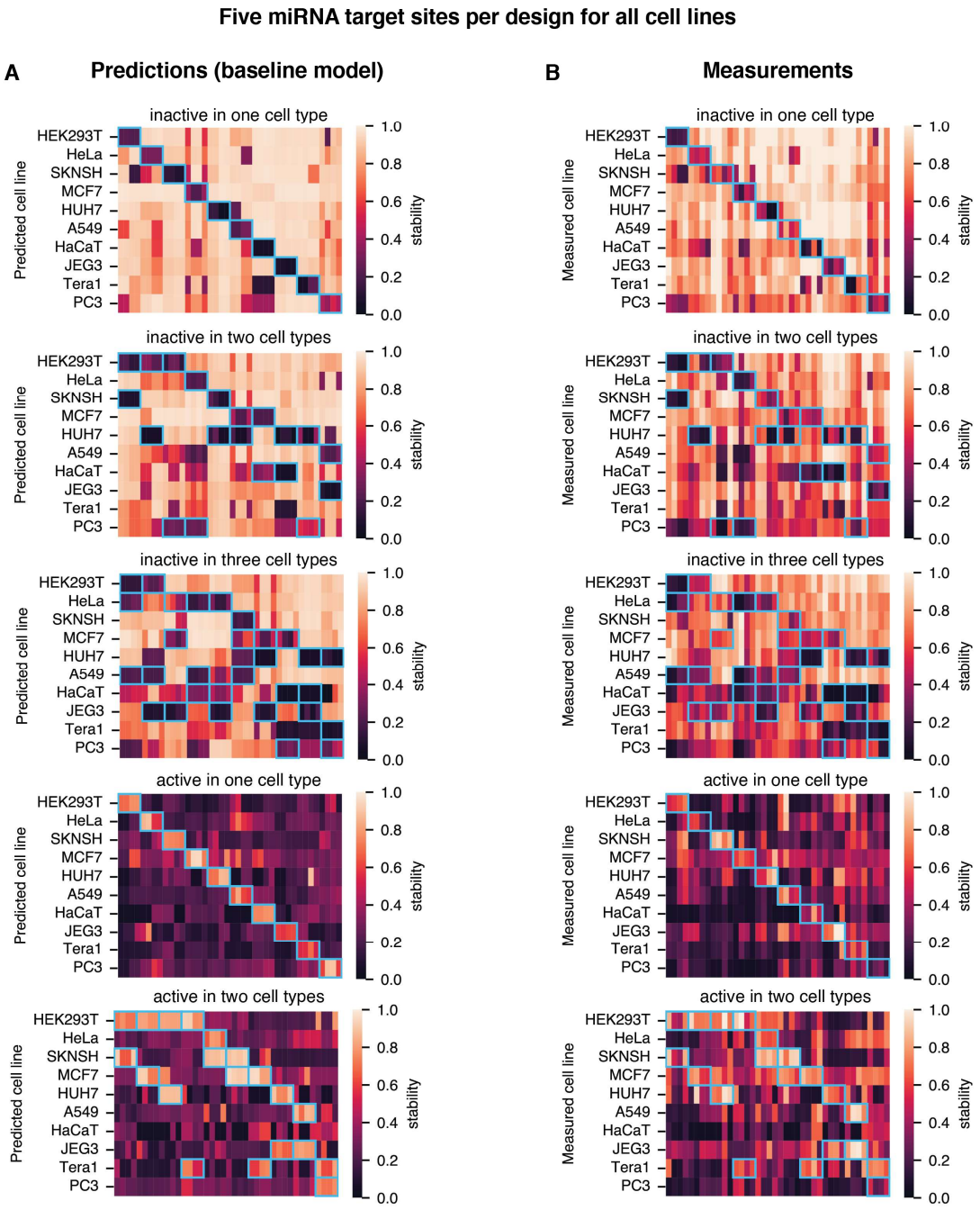


Fig. S25. Predictions and measurements for all mse-based binary designs for all cell lines with five microRNA target sites. We generated four designs per target pattern (e.g., inactivity in a single cell line) for each of the five design types. Each column shows one design. (**A**) Stabilities predicted by the baseline model and (**B**) measured stabilities across the cell lines. The blue boxes indicate the cell lines in which designs are meant to be active or inactive.

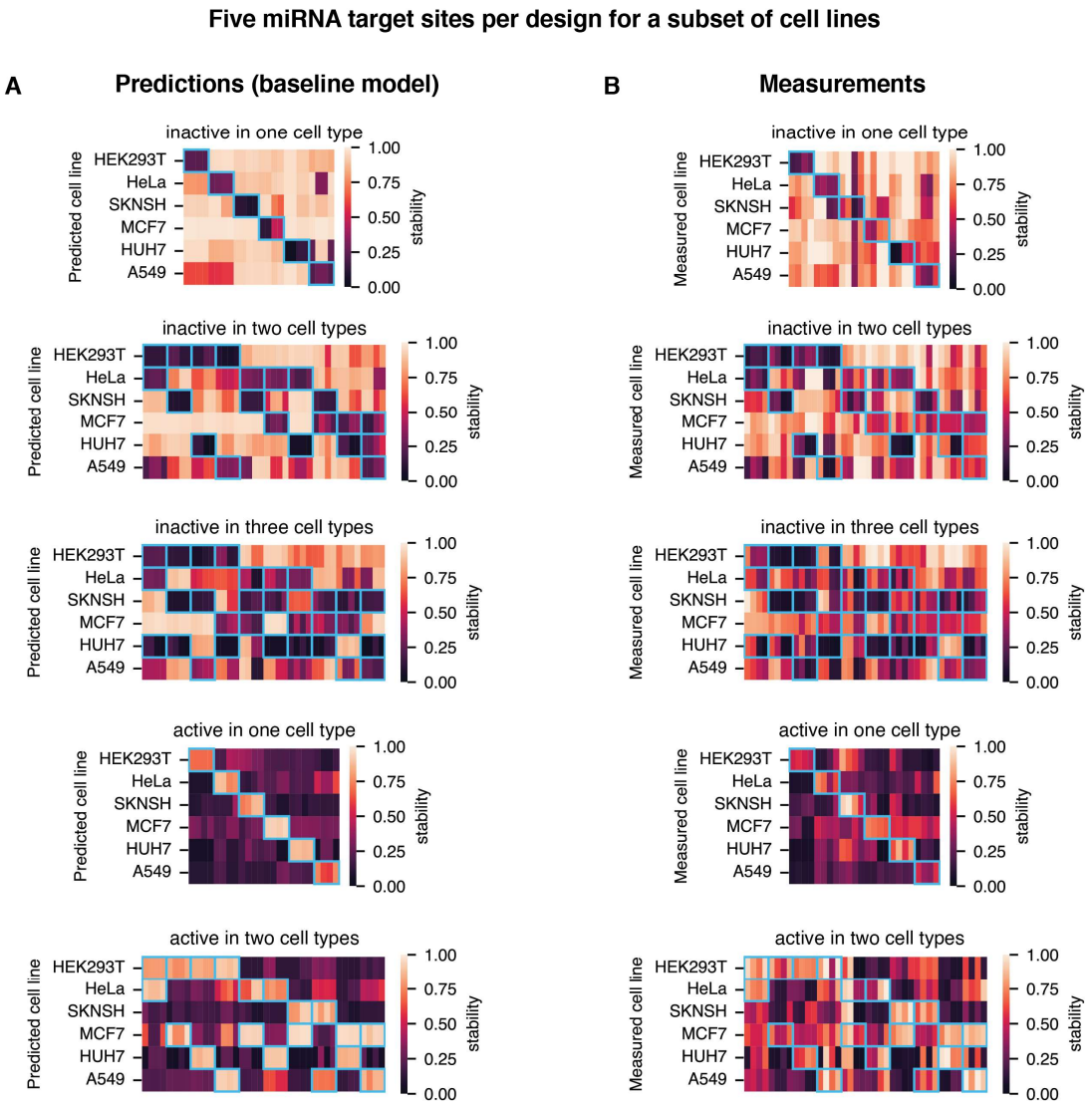


Fig. S26. Predictions and measurements for all mse-based binary designs for a subset of cell lines with five microRNA target sites. We generated four designs per target pattern (e.g., inactivity in a single cell line) for each of the five design types. Each column shows one design. (**A**) Stabilities predicted by the baseline model and (**B**) measured stabilities across the cell lines. The blue boxes indicate the cell lines in which designs are meant to be active or inactive.

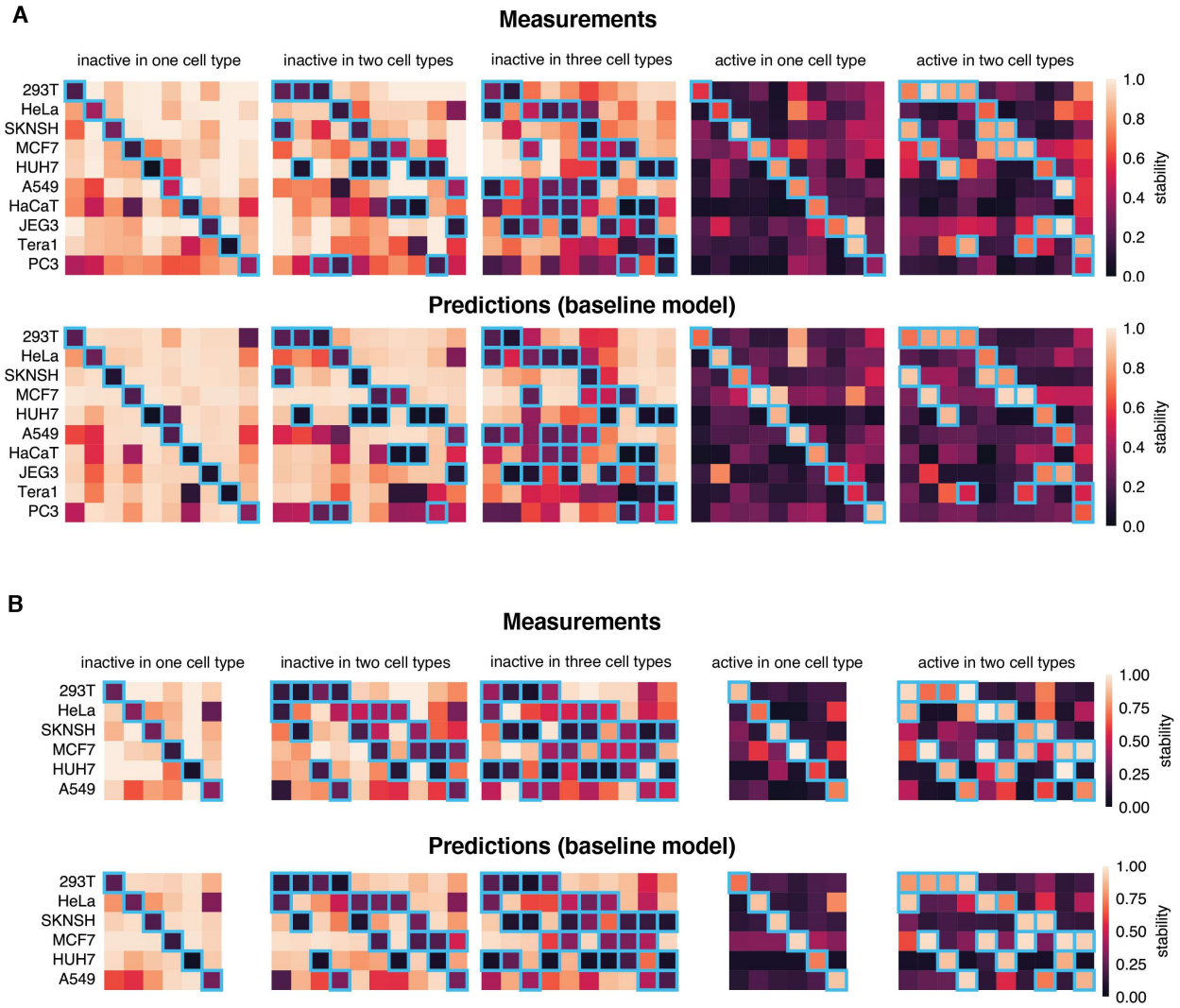


Fig. S27. Measurements and predictions of the baseline model for the best-performing binary designs. We show one design per target cell line and design type with the smallest measured weighted mean-squared error between the target stabilities and the measurement. Designs were created for (**A**) all cell lines or (**B**) a subset of 6 cell lines. Design failures are generally predicted in advance even by the baseline model. The blue boxes indicate the cell lines in which designs are meant to be active or inactive.

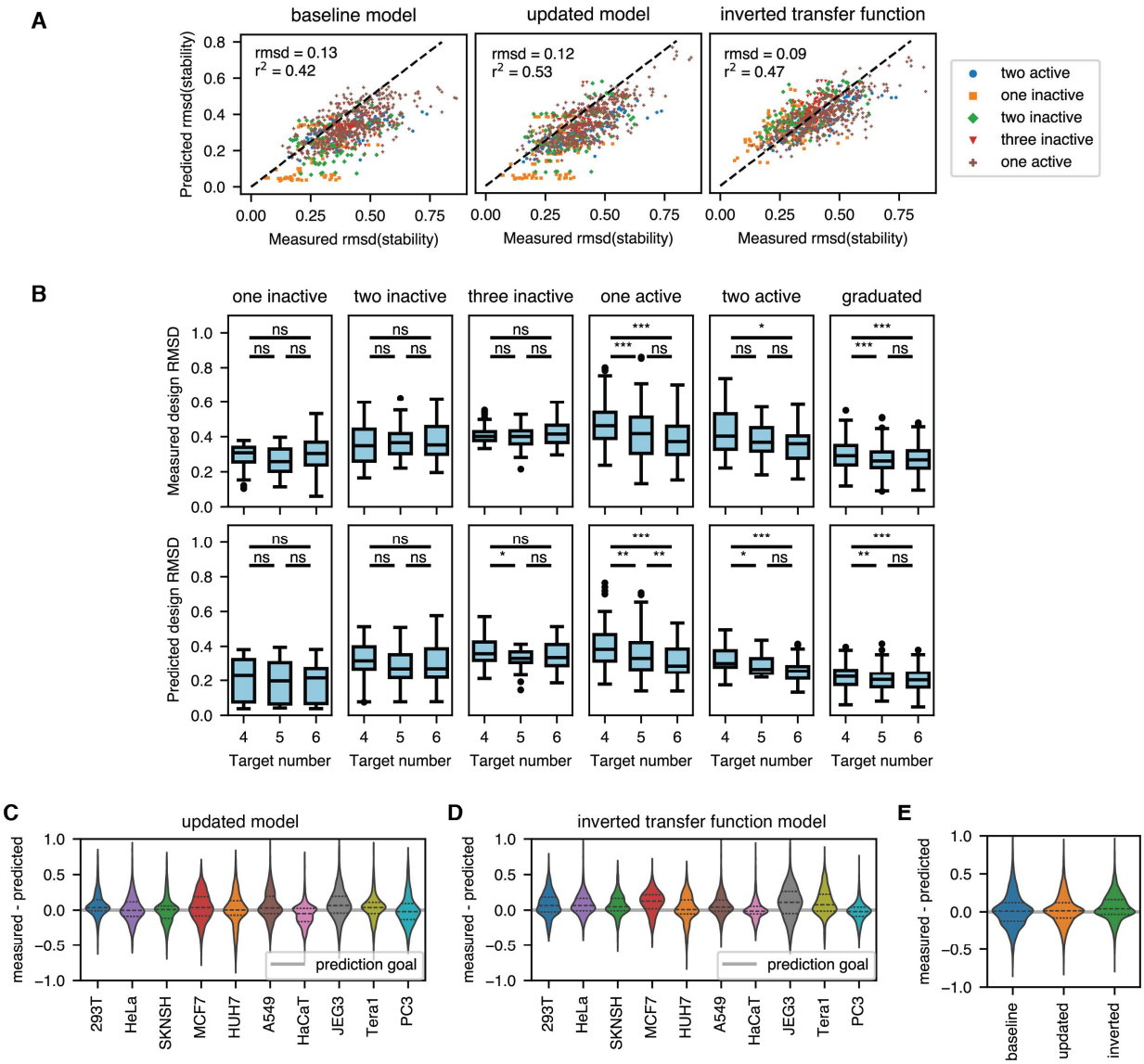


Fig. S28. The performance of different prediction models for binary designs. (**A**) Measured and predicted design rmsd values for the different binary designs. (**B**) Measured and predicted design rmsd (baseline model) for different numbers of target sites. The significance was calculated using a two-sided Mann-Whitney U test. *= p<0.05, **=p<0.01, ***=p<0.001, ns=not significant (**C**) Deviations of measurements from the predictions of the updated model. (**D**) Deviations of measurements from the predictions of the inverted transfer function model. (**E**) Deviations of measurements from predictions for the three models averaged across cell lines.

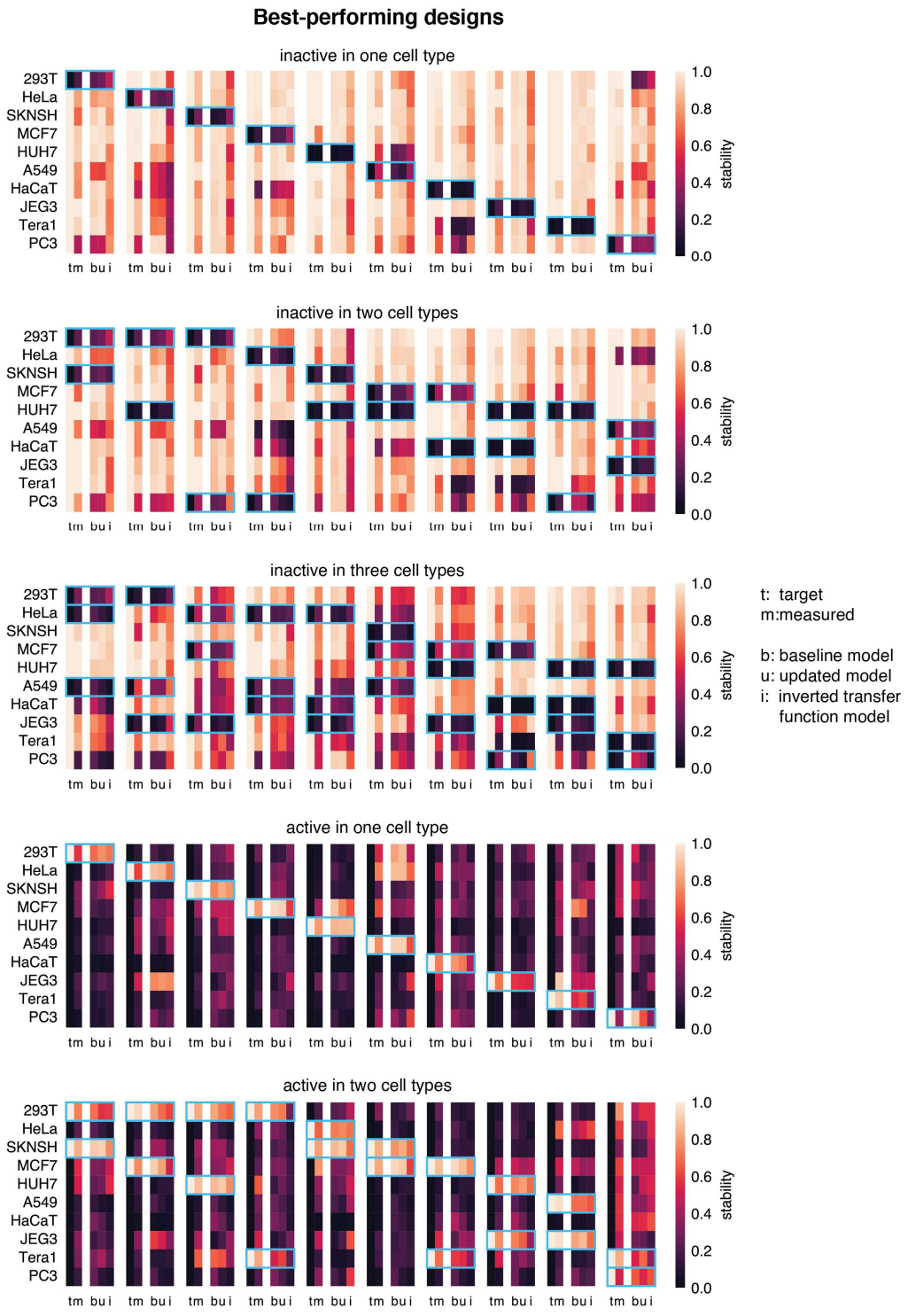


Fig. S29. Measurements and predictions by three different models for the best-performing binary designs targeting all cell lines. We show one design per target cell line and design type. We show the target and measured stability as well as predictions by the three different prediction models. The blue boxes indicate the cell lines in which designs are meant to be active or inactive. t: target stability, m: measured stability; b: baseline, u: updated, and i: inverted transfer function model.

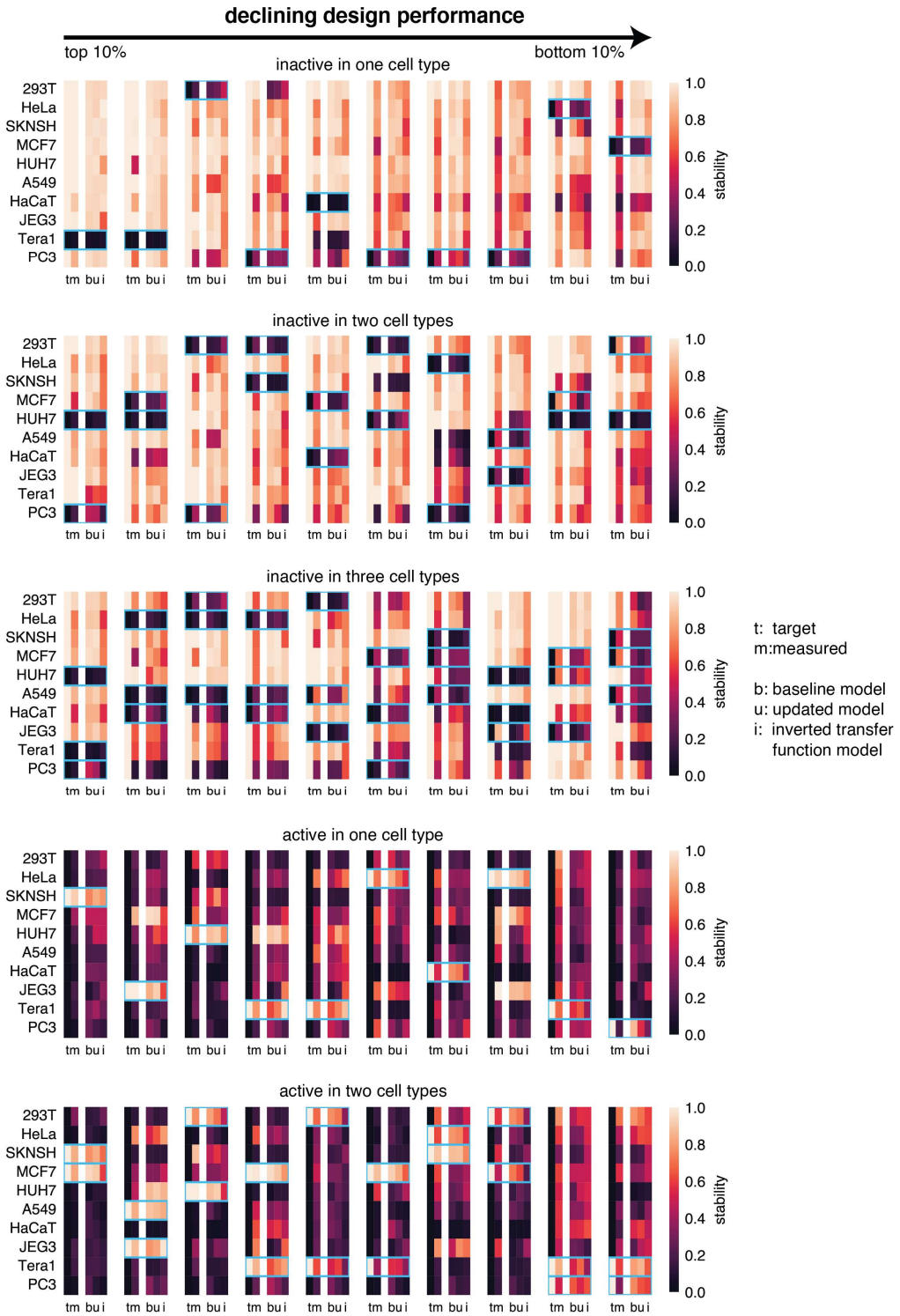


Fig. S30. Measurements and predictions by three different models for binary designs with a range of performances. The figure shows one design from each decile of design performance for each design type. We show the target and measured stability as well as predictions by the three prediction different models. The blue boxes indicate the cell lines in which designs are meant to be active or inactive. t: target stability, m: measured stability; b: baseline, u: updated, and i: inverted transfer function model.

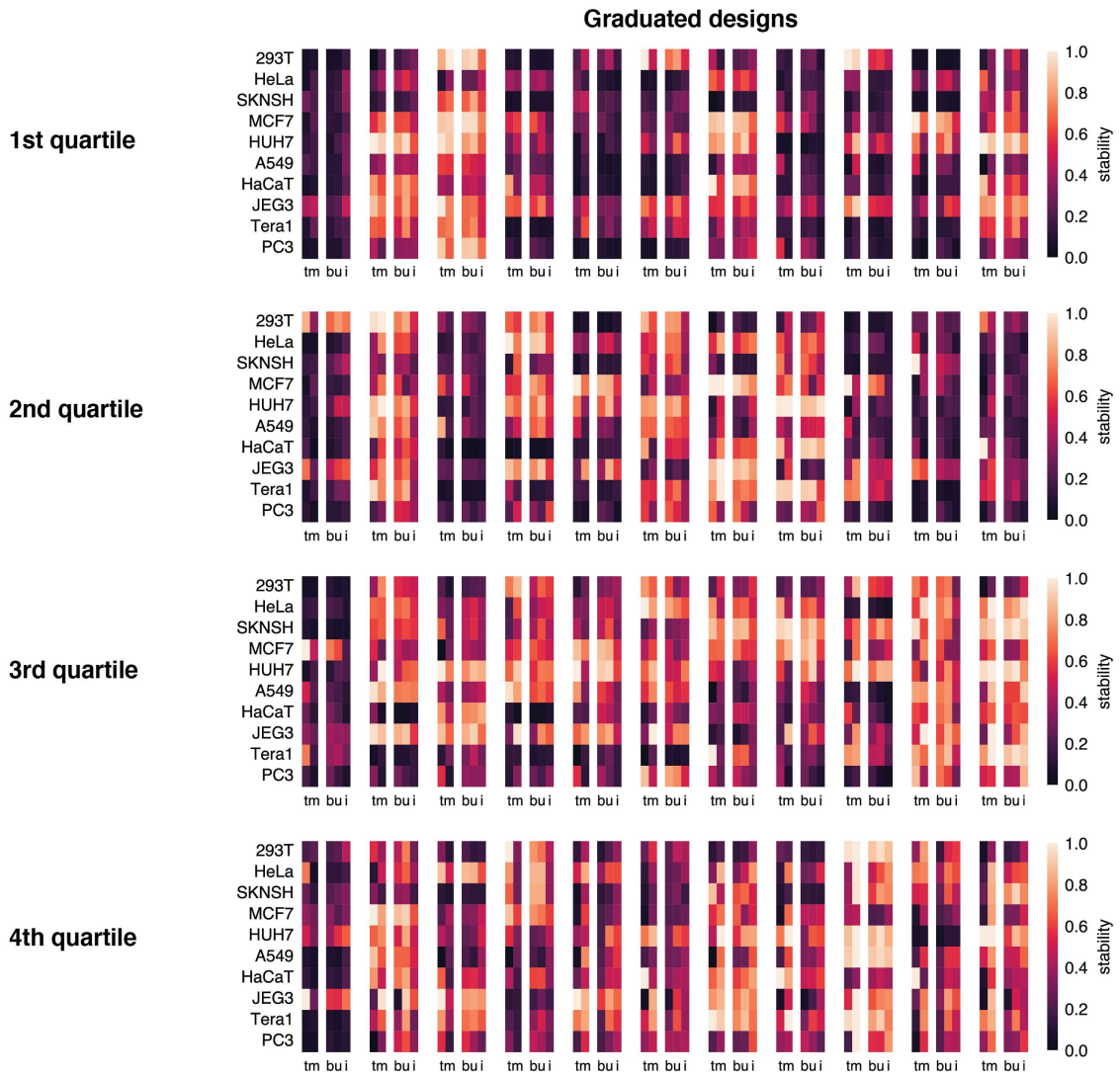


Fig. S31. Measurements and predictions by three different models for graduated designs across design performance quartiles. (**A**). We show ten designs for each quartile of design performance. We show the target and measured stability as well as predictions by the three different prediction models. t: target stability, m: measured stability; b: baseline, u: updated, and i: inverted transfer function model.

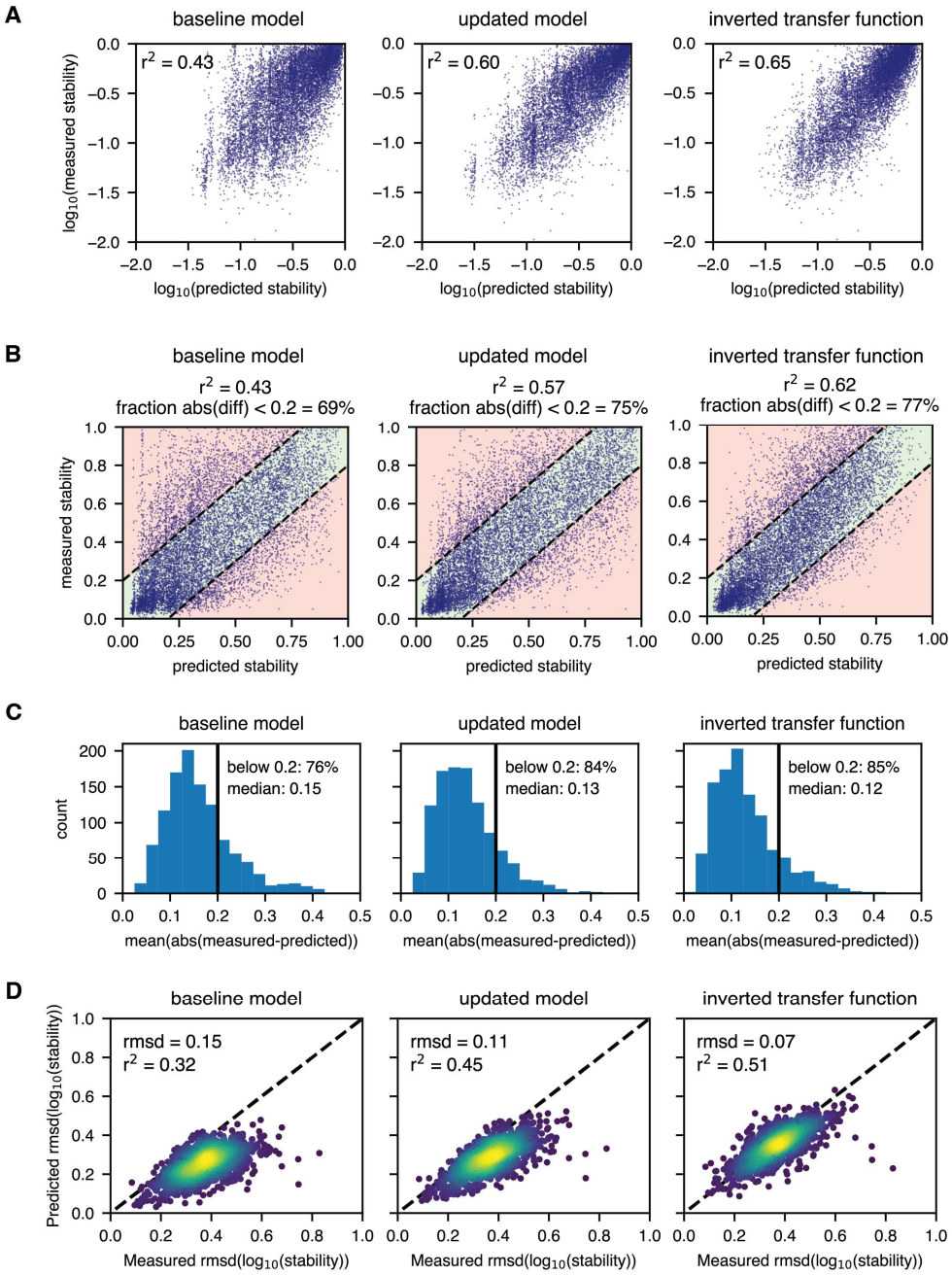


Fig. S32. Prediction performance by different models for graduated designs targeting all cell lines. (**A**) Predicted and measured logarithmic stabilities. (**B**) Predicted and measured linear stabilities. The area where the difference between prediction and measurement is less than 0.2 is shaded green, the rest is shaded red.(**C**) Absolute prediction error per design averaged across cell lines. (**D**) Predicted and measured rmsd values to the target $\log_{10}$ stabilities for all graduated designs.

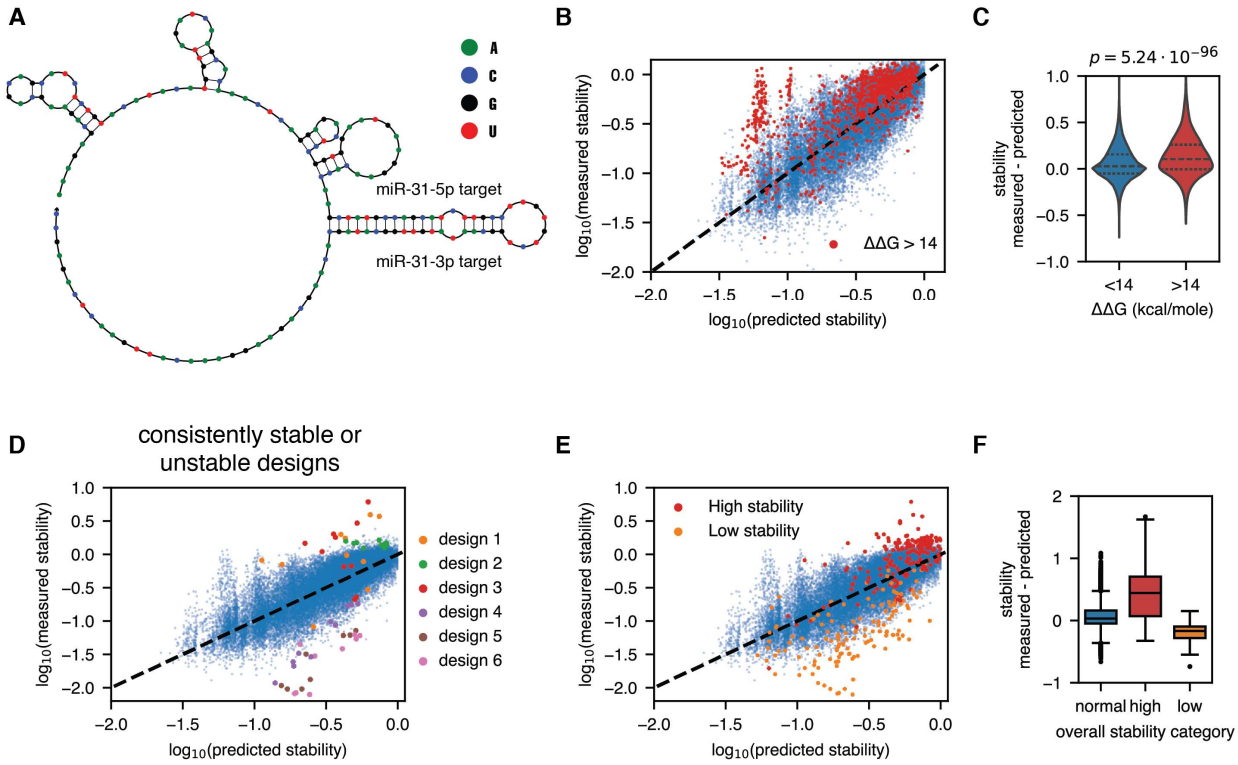


Fig. S33. Secondary structure and global stability explain some prediction failures. (**A**) Example of a design with strong secondary structure due to the use of the 5p and 3p arm of the same miRNA. (**B**) Predicted (inverted transfer function model) and measured stabilities for all designs targeting all cell lines. Designs with strong secondary structure in a dominant miRNA target site are highlighted. (**C**) Difference between measured and predicted stabilities for low and high ΔΔG designs. The p-value was calculated by a Mann-Whitney U test. (**D**) Three particularly stable and three particularly unstable designs in the prediction plot for the inverted transfer function model. The designs tend to be excessively stable or unstable across all measured cell lines. (**E**) All high stability (stability larger than 1.5 in any cell line) and low stability (ratio between predicted and measured stability larger than 6.3 in any cell line) designs. (**F**) Difference between measured and predicted stabilities for normal, high stability, and low stability designs.

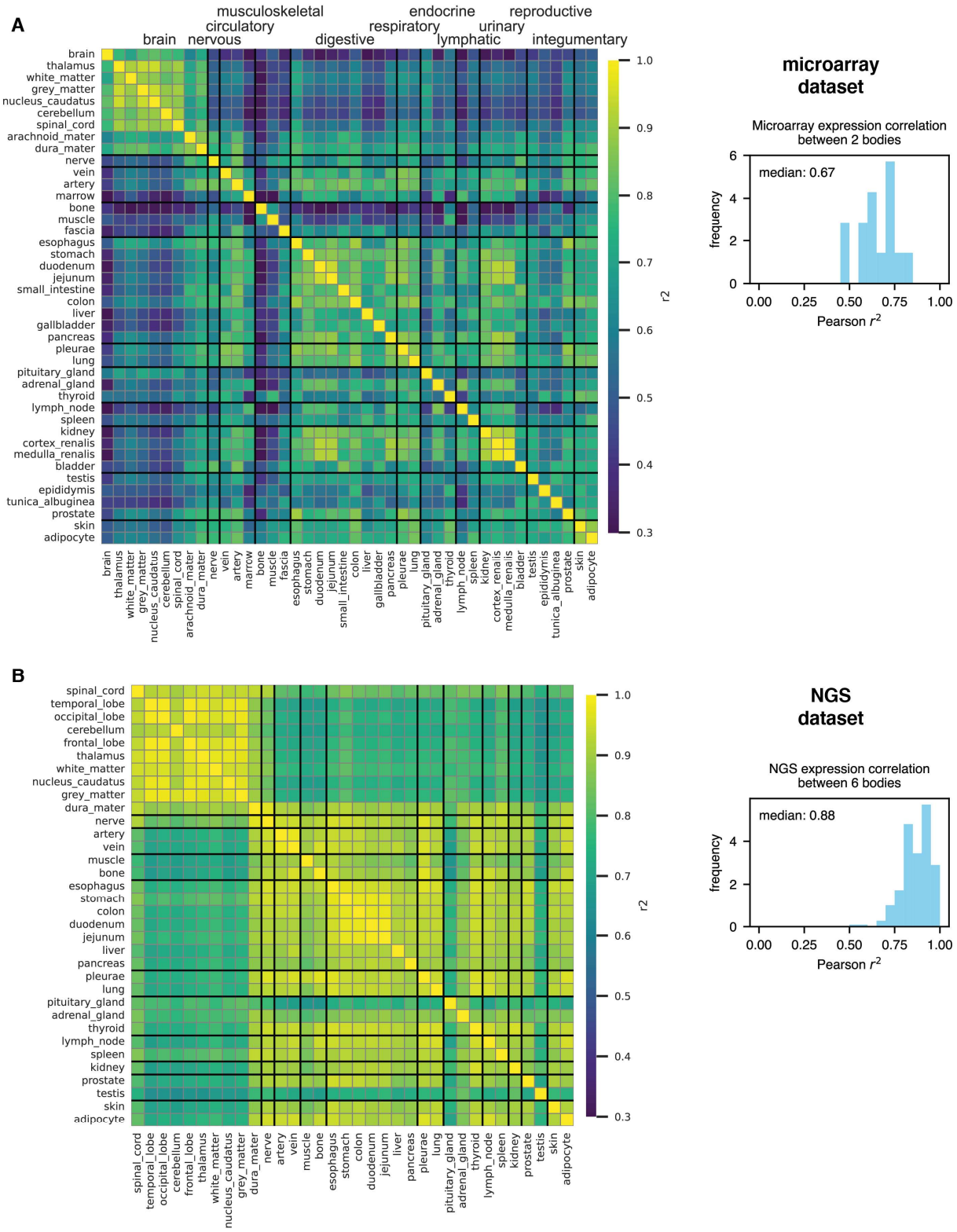


Fig. S34. Correlations between tissues in the human microRNA expression datasets. (**A**) Pearson correlation values between different tissues in the microarray dataset. The microarray tissue dataset was generated by merging expression values for 2 different subjects via a geometric mean. The correlation for the same tissue between the

expression values measured for the two subjects before merging is shown on the right. (**B**) Pearson correlation values between different tissues in the NGS dataset. The dataset was generated by merging expression values for 6 subjects via a geometric mean. The correlation for the same tissue between the expression values measured for the six subjects before merging is shown on the right.

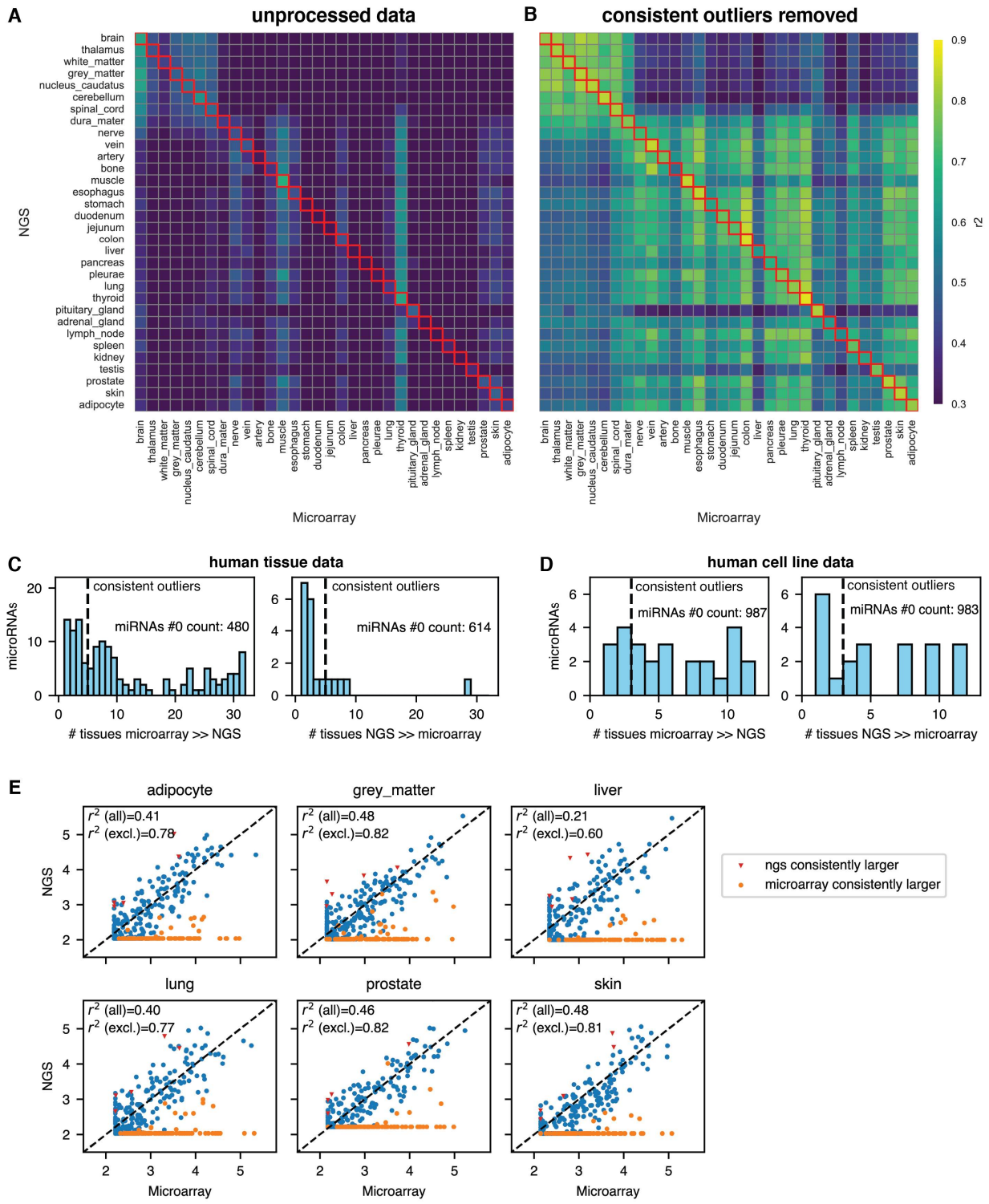


Fig. S35. Consistent outliers reduce the correlation between microarray and NGS-based human tissue datasets. (**A**). Correlation between microarray and NGS tissue datasets for all microRNAs that are in either high confidence in miRbase or listed in MirGeneDB. (**B**) Correlation between tissue datasets after removal of consistent outliers. (**C**) Outliers in each tissue are defined as microRNAs with an expression ratio of 10 or more between the two datasets. Consistent outliers are outliers in five or more tissues. The microarray data contains far more outliers

with a consistently higher expression than what is seen in the NGS dataset. (**D**) Outliers as in (C) but for expression data for our measured cell lines. For the cell line data, the outlier patterns are more symmetric between the two data types. (**E**) Expression levels for microRNAs in the two datasets for six chosen tissues. MicroRNAs that are consistently higher in one of the two datasets are highlighted. The correlation values are given either for all microRNAs or with consistent outliers excluded.

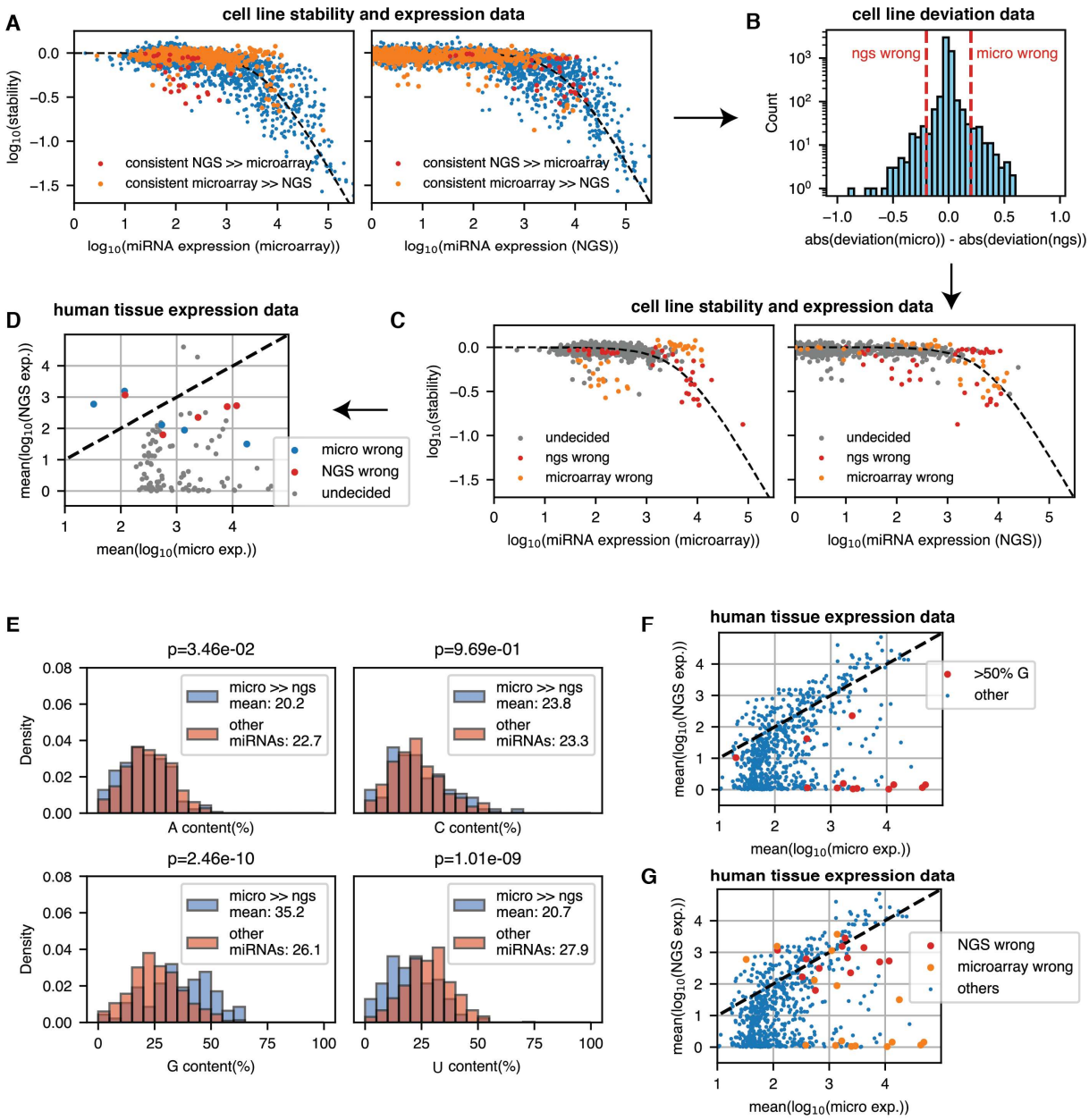


Fig. S36. Merging procedure for the tissue datasets. (A)-(D): Decision procedure for consistent tissue outlier miRNAs identification by comparison with cell line stability data. (**A**) Consistent outlier miRNAs identified in the tissue NGS and microarray data as measured in our cell line expression and stability data. As can be seen in the top left plot, some outliers in which NGS data is consistently much larger in the tissue data are also underpredicted by microarray data for the cell lines. (**B**) Distribution of the difference in absolute deviation between the measured and predicted $\log_{10}$ stabilities in cell lines for NGS and microarray data for all measured miRNAs. A much larger deviation for one or the other in a given cell line is interpreted as a sign of bias. (**C**) Cell line expression and stability of miRNA targets for miRNAs that were called as correct in one of the tissue datasets. The decision was made based on the deviation in (B) (Methods). (**D**) Mean of the log10 expression across all tissues for microRNAs that were either called as correct in one of the two datasets or left undecided. Most microRNAs, especially those with high microRNA expression and no NGS expression, were left undecided because they are not expressed in any of the cell lines according to both cell line expression datasets. (**E**) Composition bias for microRNAs that are much more highly expressed in the microarray data. Most notably, these microRNAs have a higher G and lower U content, in line with an earlier study by Backes *et al.* (*49*). P-values were calculated using a two-sided Mann-Whitney U test.

(**F**) The most biased miRNAs often have a very large G content of over 50%. Backes et al. (*49*) observed the same bias and confirmed that these microRNAs are also measured as low expression by RT-qPCR, which could indicate that miRNAs with very high G content might be systematically wrong in microarray data. (**G**) Mean expression for miRNAs in the two datasets. MicroRNAs that were found to be likely incorrect in one of the two datasets are highlighted.

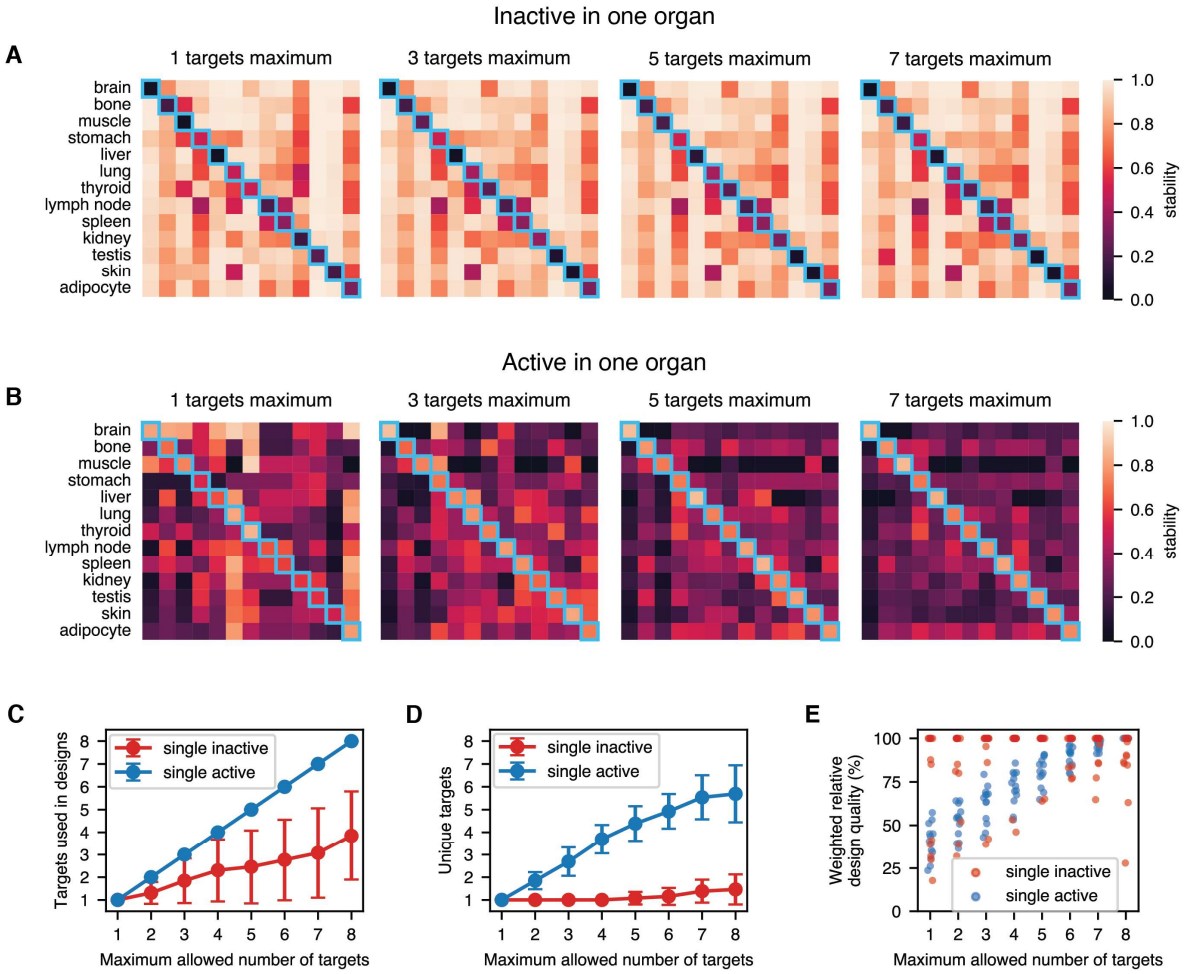


Fig. S37. The impact of the maximum allowed target site number on the predicted performance of designs targeting human tissues. We generated designs with a maximum of between 1 and 8 target sites. The design algorithm was also given the option of using fewer than the maximum number of target sites. The designs either (**A**) eliminate expression in or (**B**) constrain expression to a single organ. In the main text, we show designs with up to six target sites. The behavior of the algorithm differs drastically between the two design objectives. (**C**) Number of targets used by the design algorithm. For inactivity in a single organ, the optimum number of targets is often fewer than are allowed. For activity in a single organ, all available targets are used for every design without exception. (**D**) Unique targets per design. Inactivity in a single organ is overwhelmingly achieved by a single type of miRNA target site. Activity in a single organ is achieved by combining many different target site types. (**E**) Weighted design quality (inverse mse) by the number of allowed target sites. Every dot represents a design targeting a different organ. The design quality is shown relative to the highest achieved quality design generated for a specific target organ across all allowed numbers of target sites. For inactivity in a single organ, a design performance close to the maximum is already achieved by a single target site in most cases. For activity in a single organ, the performance keeps improving with the allowed number of target sites. Note that there are organs for which inactivity in a single organ is also improved by up to six target sites. In those cases, this usually means repeating the same target site six times.