Precision multidimensional assay for high-throughput microRNA drug discovery

Benjamin Haefliger1, Laura Prochazka1, Bartolomeo Angelici1 & Yaakov Benenson1

Development of drug discovery assays that combine high content with throughput is challenging. Information-processing gene networks can address this challenge by integrating multiple potential targets of drug candidates’ activities into a small number of informative readouts, reporting simultaneously on specific and non-specific effects. Here we show a family of networks implementing this concept in a cell-based drug discovery assay for miRNA drug targets. The networks comprise multiple modules reporting on specific effects towards an intended miRNA target, together with non-specific effects on gene expression, off-target miRNAs and RNA interference pathway. We validate the assays using known perturbations of on- and off-target miRNAs, and evaluate an ~700 compound library in an automated screen with a follow-up on specific and non-specific hits. We further customize and validate assays for additional drug targets and non-specific inputs. Our study offers a novel framework for precision drug discovery assays applicable to diverse target families.
Progress in drug discovery is hampered by under-exploration of chemical space and by the difficulty in assessing the full range of drug candidates’ effects on living cells. The former challenge is addressed by extending chemical space coverage, in part using synthetic pathways\(^1\)–\(^3\) engineered using synthetic biology\(^4\)–\(^15\) methods. The latter is partially solved with cell-based assays\(^3\) that allow evaluating drug action in a complex environment. Yet, these assays still generate candidate compounds that perform inadequately in vivo with respect to efficacy and toxicity\(^1\)–\(^4\) in large part because many unwanted interactions\(^1\)–\(^3\) pass undetected in vitro. Multiplex assays and serial testing\(^1\)–\(^3\) have been proposed as a way to gauge off-target effects, yet increasing the number of measured parameters reduces assay throughput and makes it unsuitable for large library screens.

The weaknesses of cell-based assays are amplified with microRNAs (miRNAs) as drug targets. miRNA activity can be enhanced using miRNA mimics\(^1\)–\(^6\) and inhibited with complementary RNA analogs\(^7\) or genetic sponges\(^8\). The search for small molecule miRNA modulators\(^9\)–\(^19\) has relied on qPCR\(^22\) and genetic reporter\(^23\)–\(^24\) assays, to produce a few candidate compounds\(^20\)–\(^25\). However, multiple miRNA molecules can be easily targeted by the same compound due to similarities in nucleotide sequence and phosphoribose backbone, shared maturation pathway, and common pri-miRNA precursors\(^26\); therefore, the likelihood of side effects is high. Given that miRNAs are promising drug targets\(^27\)–\(^28\) playing an important role in a large number of diseases\(^29\), there is a need for miRNA drug discovery tools that adequately address the high risk of non-specific effects with this target family.

We propose to address this challenge via rapid assessment of multiple off-target effects using intracellular genetic information-processing circuits\(^5\). These circuits are complex artificial regulatory networks that convert predefined biomolecular input cues into one or more gene products (outputs), according to a specified relationship. The circuits are implemented with multiple artificial genes that are delivered to and expressed in living cells. In the context of miRNA drug discovery, one can envision a network whose inputs are multiple potential off-target miRNAs and whose output is a single fluorescent reporter. This network would generate high output fluorescence when all its inputs are in their default (ground) state, corresponding to the lack of interference between the drug and the inputs, and low fluorescence otherwise, corresponding to at least one side effect. A network such as this can be formally described as an AND logic circuit, because the output of an AND circuit changes when any one of its inputs changes. Using formal notation and denoting a default state of input $X$ ($X = A, B, \ldots$) as input $X^0$, the following equations hold:

\[
\text{Default output} = (\text{input } A^0) \text{ AND } (\text{input } B^0) \text{ AND } \ldots = \text{no off-target effects} \tag{1}
\]

\[
\text{Change in default output} = \text{NOT}(\text{no off-target effects})
= \text{NOT}((\text{input } A^0) \text{ AND } (\text{input } B^0) \text{ AND } \ldots)
= \text{NOT}(\text{input } A^0) \text{ OR NOT}(\text{input } B^0) \text{ OR } \ldots
= \text{an off-target effect} \tag{2}
\]

Equation (2) shows why a change in the AND gate output results from at least one off-target effect, without specifically identifying it. Adding inputs to the gate will expand the range of sampled off-target effects while keeping a single output without the need for multiplexing. Consequently, the validity and information content of the screen increases dramatically without sacrificing throughput.

Here we describe a novel cell-based assay that utilizes a genetic information-processing circuit to integrate and compress multiple miRNA inputs into a small number of fluorescent reporters to distinguish between off-target and specific effects of candidate compounds. Using an iterative simulation-aided design process, we implement the concept for miR-122, a promising drug target in liver cancer\(^30\) and hepatitis C\(^31\). We validate the assay in HuH-7 cells using miRNA mimics and inhibitors, then further adapt it for automated screening and test a library of $\sim700$ compounds. Finally, we reprogram and revalidate the assay to address additional miRNA drug targets and off-targets. Importantly, we show that compounds that would have been mistaken as specific hits with traditional methods are correctly identified as non-specific modulators. This study presents a precise yet high-throughput approach for miRNA drug discovery. The general concept is applicable to additional target families with appropriate modifications in the sensing and processing components.

### Results

**Basic concepts.** The assay circuit consists of three genetic modules reflecting distinct drug effects, including, respectively, global effects on gene expression (gene expression module), systemic effects on the RNAi pathway or off-target miRNAs (non-specific RNAi module) and the desired effect on the intended miRNA drug target (specific module) (Fig. 1a). The gene expression module reports on cell-wide changes in gene expression and cell viability via a constitutively expressed fluorescent protein. The same protein also serves as a transfection normalization control (see below). The non-specific RNAi module implements a genetic AND gate integrating multiple potential off-target miRNA inputs, into a single output reflected in the intensity of a fluorescent readout. To do so, we first place fully complementary, tandem binding sites for an miRNA input in the 3′-UTR of an mRNA coding for this output, to implement a ‘NOT (input)’ logic\(^3\).

A second miRNA input knocks down a repressor of that same output via fully complementary, tandem binding sites in the repressor’s 3′-UTR, inverting miRNA inhibitory activity. This results in a logic gate that facilitates high output expression when the second miRNA is highly expressed (high input) and the first miRNA is expressed at low levels (low input): ‘high input AND NOT (low input)’ (Fig. 1b, left). Additional high and low miRNA inputs can be added to scale up the gate (Fig. 1b, right). This gate structure suits our purpose because non-specific downregulation of miRNAs will reduce highly expressed miRNA inputs, and non-specific upregulation will elevate low miRNA inputs, with either scenario altering the output. Previous work\(^8\) showed how to improve the inversion of high inputs with additional genetic elements (Fig. 1c, left). We note that the module contains multiple gene expression components and it might also be sensitive to global changes in gene expression.

The output of the AND gate can be a fluorescent reporter; however, following our previously shown strategy for robust integration of multiple modules using a shared transactivator ‘knot’ component\(^32\), we use a transactivator as the immediate AND gate output. This transactivator controls two fluorescent proteins. The first mirrors transactivator expression and serves as the non-specific assay readout; the second protein is furnished in addition with the binding sequence of the intended drug target miRNA in its 3′-UTR. Increased miRNA activity leads to a reduction in this reporter level and vice versa (Fig. 1c, bottom), relative to the non-specific readout. This reporter’s expression (normalized to the non-specific readout) constitutes the specific assay readout.
Validation strategy. We established a set of positive and negative controls to validate the assay modules. Ideally, controls should be chemical counterparts of candidate compounds33. We sought small molecule compounds with proven anti miR-122 activity, as well as those targeting multiple miRNAs or the RNAi pathway. Due to the late emergence of miRNAs as drug targets, controls were difficult to identify (see below), and we sought alternatives as suggested by good practice33. On the basis of prior reports20,23,25, we chose miRNA mimics and locked nucleic acid-based miRNA inhibitors (referred to as LNAs) to respectively increase and decrease miRNA activity in a predictable manner. Perturbing individual miRNA inputs with mimics and LNA emulates individual drug–miRNA interactions, while perturbing multiple inputs simultaneously emulates systemic alteration of miRNA-processing pathways. We designed 15 different assay perturbations comprising subsets of mimics and LNAs that span a range of possible off-target and on-target effects (Fig 1d), and used these perturbations to calculate Z’-factors34 for assay performance evaluation. Each perturbation is associated with its own Z’-factor, providing a comprehensive insight into assay behaviour under different conditions.

Experimental system. As stated above, for the proof-of-concept we used miR-122 as the target. The assay was tested in HuH-7 cells, a liver cancer cell line used as a model for liver tumour35, hepatitis C virus infection56 and miR-122 drug discovery37. Studies showed that miR-122 has an effect on both liver cancer30,38 and hepatitis C virus replication31, making it a promising drug target39.

To determine non-specific assay inputs, we looked for miRNAs expressed at either high or low levels as required by our approach (Fig. 1c). We measured the activities of 21 miRNAs40 with high inputs. Initially we planned to validate the assay with both RNA-based and chemical or genetic agents known to affect miR-122 or the RNAi pathway. We identified a number of relevant bi-directional reporters (Fig. 2a), and retained miR-145, -141, -375 and -146a as non-specific low inputs; and miR-21 and -20a as high inputs. Initially we planned to validate the assay with both RNA-based and chemical or genetic agents known to affect miR-122 or the RNAi pathway. We identified a number of relevant bi-directional reporters (Fig. 2a), and retained miR-145, -141, -375 and -146a as non-specific low inputs; and miR-21 and -20a as high inputs. Initially we planned to validate the assay with both RNA-based and chemical or genetic agents known to affect miR-122 or the RNAi pathway. We identified a number of relevant bi-directional reporters (Fig. 2a), and retained miR-145, -141, -375 and -146a as non-specific low inputs; and miR-21 and -20a as high inputs. Initially we planned to validate the assay with both RNA-based and chemical or genetic agents known to affect miR-122 or the RNAi pathway. We identified a number of relevant bi-directional reporters (Fig. 2a), and retained miR-145, -141, -375 and -146a as non-specific low inputs; and miR-21 and -20a as

Figure 1 | Assay design and validation strategy. (a) High-level representation of the screening assay modules. Module names, on- and off-target miRNA molecules and readouts are indicated. (b) Left: schematic representation of the possible input combinations and outcomes for a two-miRNA input AND gate. The values of 1 and 0 correspond to high and low miRNA expression, respectively. Note that the output is highly expressed only when high input is highly active, and low input is inactive. Squiggly lines are microRNAs, blunt arrows denote repression. Pale hue indicates low expression or activity. Logic gate that determines output activity is shown on top. Right: logic gate and corresponding input combination in the scaled-up four-input AND gate that results in high output expression. (c) Coarse-grained and detailed assay diagrams with four inputs for the non-specific module and one input for the specific module. Genetic implementation for each of the building blocks is shown in respective zoomed-in frames, with pointed arrows indicating activation, blunt arrows indicating repression and component names shown. (d) Validation perturbation table with 15 combinations. The ground state (column 0) describes the miRNA expression levels in the assay cell line furnished with the assay circuit, without perturbation. 1 stands for high and 0 for low expression, respectively. Drug target miRNA expression is shown as 0.5, indicating intermediate activity levels required to detect both up- and downregulation with drug candidates. For each perturbation column, – 1 corresponds to miRNA inhibition, + 1 is miRNA activation/overexpression relative to the ground state and 0 indicates no change. The anticipated expression of non-specific and specific readouts following a perturbation is shown in two bottom rows, with 1, 0 and 2 representing, respectively, the ground state, reduced and elevated readout. CAGop, CAG promoter followed by an intron with two LacO sites; CMV, cytomegalovirus immediate-early promoter; LacI, Lac repressor; rTA, reverse Tet transactivator; TA, transactivator; TRE, TetR responsive element.
Alternative assay design and simulations. The first alternative is the ‘Parallel assay’ where specific and non-specific modules are not connected by the PIT2 ‘knot’ transactivator. Non-specific effects are reflected in a fluorescent protein ZsYellow replacing the PIT2 knot. The specific module is a PIT2-induced bidirecional reporter of miR-122 activity, driving mCherry furnished with miR-122-binding sites and a reference mCerulean. The latter also serves as a global gene expression readout and transfection control, similar to iRFP in the pilot assay. The two other architectures extend the pilot assay circuit with additional feed-forward loops that proved their utility in high input sensing. We implemented the motifs by augmenting the non-specific readout mRNA with binding sites for miRNAs that bind to the PIT2 knot. In an implementation called ‘low inputs feed-forward (LFF) assay’, only the low inputs miR-146a and -141 target the readout; in the ‘complete feed-forward (CFF) assay’, these are miR-146a, -141 and -FF4 (Fig. 3a).

We explored these designs in silico using mechanistic models of the four architectures (Supplementary Note 3; Supplementary Fig. 3). We calculated the dynamic range of the non-specific readout by alternating in silico between high and low non-specific miRNA input concentrations. For high inputs, we concluded that parallel and CFF architectures are superior, and that under wide range of parameter values, the CFF assay improves 2–3-fold over the parallel assay (Fig. 3b; Supplementary Figs 4–6). For low inputs, LFF comes at the top and CFF is close second best. To simulate sensitivity of assays to global changes in the RNAi pathway, we mapped non-specific readout expression as a function of RNA-induced silencing complex concentration and miR-FF4/LacI-miRNA ratio, the latter being a proxy for miRNA-processing efficiency (Fig. 3c,d). Parallel and CFF assays are most sensitive to changes in these parameters. Because the miR-FF4-binding site is embedded in the readout mRNA’s 3’-UTR, CFF is slightly more sensitive than the parallel assay. Thus simulations suggest CFF as the optimal architecture.

Validation of alternative assays. We quantitatively validated and characterized all three variants using a complete set of input perturbations (Fig. 1d), due to uncertainties in simulating complex networks. Z’-factors for specific targeting of miR-122 are >0.5 (excellent) for all assays (Fig. 3e, lower panel). Z’-factors corresponding to all perturbations involving non-specific inputs form distinct distributions in each of the three assays (Fig. 3e, upper panel). As predicted by the model, the CFF assay has the smallest number of Z’-factors <0.5 (2 out of 13), with the rest being ‘excellent’. With other assays, the lack of miR-FF4 feed-forward loop leads to lower performance when high inputs are perturbed. By applying the ‘best-worst-case’ as an indicator of the system’s weakest link, we judge the CFF assay to be superior.

Assembly and testing of pilot assay. We optimized individual circuit building blocks for function and dynamic range under a subset of validation perturbations. We evaluated synthetic transactivators for the central knot and chose the streptogramin-responsive transactivator (Pristinamycin-induced protein (Pip) fused to p65; PIT2)\(^43\). Next, we calibrated DNA composition of the high input sensor to maximize the effect when high inputs are inhibited, and assessed the position effect of low input miRNA-binding sites in the knot’s 3’-UTR to maximize the effect when low inputs are activated (Supplementary Note 2; Supplementary Fig. 2e–i). The optimized components were assembled in a circuit dubbed ‘pilot assay’ (Fig. 2b). The layout follows the structure in Fig. 1c, with the addition of an auxiliary fluorescent reporter mCitrine coupled to the PIT2 transactivator via 2A peptide linker for characterization purposes. The circuit senses five different miRNA inputs, of which four feed into the non-specific RNAi module. Initially we tested a subset of perturbations by targeting each non-specific input individually. We found that the assay performance was satisfactory when judged by the changes in mCitrine/PIT2 (Fig. 2c), but it deteriorated at the non-specific readout level for both high and low input modulation, leading to Z’-factors <0.5, reflecting merely ‘acceptable’ assay performance (Fig. 2d). We concluded that the pilot assay was not robust enough and considered alternative topologies.
In-depth validation of CFF assay. The detailed data of CFF assay (Fig. 4a) response to all perturbations are displayed in Fig. 4 and Supplementary Fig. 7. Zooming into Fig. 4b, we see that specific changes in miR-122 are readily discovered. For the non-specific RNAi module, the induction of miR-20a and miR-21 and the inhibition of miR-FF4 generate the smallest effects, while under other perturbations the assay performs very well. Weak response to increasing high inputs is expected, as they already exert most of their effect in the default state and their main purpose is to serve as detectors of non-specific miRNA inhibition.

We quantified assay’s sensitivity to intermediate perturbation strength by measuring dose response across 50-fold intensity...
Each transfected Mimic/siRNA, (nM)

![Diagram](https://example.com/diagram.png)

Figure 4 | In-depth validation of the CFF assay. (a) Schematic representation of the CFF assay circuit used in these experiments, reproduced here from Fig. 3 for convenience. (b) Performance of the CFF assay when tested with the 15 validation perturbations. The table describes the miRNA/LNA mixtures. Each column shows, which miRNA(s) are changed relative to HuH-7 background (grey columns). ‘–’ indicates inhibition of a miRNA using LNAs and ‘+1’ indicates induction of a miRNA with mimic(s). Grey columns show miRNA expression pattern (0—not expressed, 1—highly expressed, 0.7—intermediate level of the drug target miR-122) in the HuH-7 ground state. Pink columns emphasize specific modulation of miR-122. Pale blue and pink bands indicate the variation (Fig. 5; Supplementary Table 1). The specific module is sensitive to all perturbations in this range; only miR-122 upregulation at the lowest concentration of mimic-122 shows a decreased Z'-factor (Z' = 0.43). The non-specific RNAi module exhibits shallow dose response to increase in high inputs and decrease in miR-FF4, consistent with the end-point data. On the contrary, assay sensitivity to low input activation with respective mimics is very high, resulting in strong readout reduction already.
at the lowest tested concentration. The asymmetry in response to the two types of inputs is advantageous, since it precludes balancing of two opposing unspecific effects of the same strength. This is confirmed on simultaneous activation of all non-specific inputs with respective mimics, whereby the effect of perturbing low inputs overrides that of high inputs. Last, in all cases where the intended drug target miR-122 is affected simultaneously with other miRNAs, we observe changes in non-specific RNAi module readout. These perturbations emulate a scenario when the intended target is modulated together with a number of off-targets. For all such cases the assay correctly reports the side effects. For further details, see section ‘Assay comparison to bidirectional reporters’ below.

Automated screening of small molecule library. Following extensive assay optimization and validation, we performed a pilot small molecule screen. First, we adapted the assay to a high-throughput screening protocol\(^\text{33}\) by automating compound dilution and transfection, and developing an image-processing pipeline for 96-well plates (Fig. 6a, Methods section). The assay was revalidated with the automated protocol using select perturbations and was found to retain its performance (\(Z' > 0.5\)). As the pilot library, we chose the NIH Clinical Collections 1 and 2 with a total of 726 compounds. We tested the library twice using automated liquid handling, transient transfections and triplicate measurements. Assuming that most compounds are inactive\(^\text{34}\), we used all readouts from an entire 96-well compound plate as the reference distribution that is compared, using two-sided t-test, with triplicate measurements of individual compounds from the same plate (Supplementary Figs 8 and 9). A P value cutoff of \(P < 0.1\) (that is, a compound has a ‘true’ effect on a given readout with at least 90% probability) was used to exclude a compound from further analysis based on changes in gene expression readout mCitrine. In screens 1 and 2, 18.3% and 20.2% of compounds, respectively, were excluded. The same criterion was applied to non-specific RNAi module readout (normalized mCerulean), respectively excluding 12.8% and 16.1% of compounds as potential non-specific effectors. For hit identification with the specific module readout (normalized mCherry) we used \(P < 0.01\) cutoff, identifying 39 and 31 hits, respectively, corresponding to an apparent hit rate of 5.1% and 3.3% (Fig. 6b). However, as the relative magnitude of the effects is <30% for all cases, many hits might be false-positives; therefore, secondary validation is needed.

**Figure 6** Screening of NIH clinical collection 1 and 2. (a) Schematic representation of the small-molecule screen and the data-processing pipeline. (b) Volcano plots of screening results with each point representing the mean of a biological triplicate. X values are fold changes of the candidate triplicate compared with the plate average, and Y values represent the P value of a two-sided t-test of each triplicate compared with the plate averages. Green dots are from screen 1, red ones from screen 2. Horizontal dotted lines are drawn at \(P = 0.1\) for the gene expression module (mCitrine) and the non-specific RNAi module (normalized mCerulean), and at \(P = 0.01\) for the specific module (normalized mCherry). (c) Dose response of assay gene expression module readout mCitrine to select compounds originally excluded based on this readout in the automated screen. Compounds found to downregulate gene expression readout mCitrine in the automated screen are in shades of blue, and those upregulating mCitrine are in shades of red. (d) Dose response of gene expression and non-specific readouts for select compounds based on the non-specific RNAi module readout in the automated screen. Black lines show gene expression module readout (mCitrine); blue ones show the non-specific module readout (normalized mCerulean). Values in c and d are mean ± s.d. of biological duplicates. Transfections are described in Supplementary Tables 10–12. a.u., absolute units.
**Validation of screening hits.** We followed up on the miR-122 hits and some of the excluded compounds by measuring dose response of assay readouts. We looked for compounds that significantly affected at least one fluorescent readout in both screen replicas and ranked them according to mean deviation from reference distribution. We measured dose-response of the top 10 modulators of the gene expression module (four up- and six downregulators, Supplementary Fig. 10) and top eight compounds affecting the non-specific module (two up- and six downregulators, Supplementary Fig. 11). We reproduced the effects on the gene expression module with 8 out of 10 compounds, whereas Indinavir and Donepezil (1.6- and 1.5-fold change in the initial screen) failed to elicit the expected response (Fig. 6c; Supplementary Fig. 12). Similarly, compounds that reduced the readout of the non-specific RNAi module in the screen also reduced it in a dose-dependent manner. However, all of them affected the gene expression readout mCitrine at a concentration of 50 μM, five times higher than the one used in the screen, suggesting additional effects on gene expression or cell health. Among compounds that elevated the non-specific module readout, Irenprodil did so in a dose-dependent manner but Oxytetracyclin could not be confirmed (Fig. 6d). For the seven miR-122 hits, including three up- and four downregulators (Supplementary Fig. 13), we confirmed their lack of effect on gene expression and non-specific module readouts apart from an occasional effect on mCitrine at 50 μM. However, modest modulation of miR-122 activity observed in the screen (<1.3-fold) could not be reproduced. On repeated testing with a miR-122 bidirectional reporter, we confirmed the lack of anti-miR-122 effect (Supplementary Fig. 14), indicating that the hits were likely false-positives resulting from ‘multiplicity of testing’ problems.

**Assay customization.** The assay can be customized, or reprogrammed, to address different miRNA drug targets and different sets of non-specific inputs. To show this experimentally, we modified specific and non-specific modules, generating two new assay circuits. The first is a miR-23b drug discovery assay circuit obtained by replacing miR-122-binding sites with those for miR-23b. The second circuit has an augmented non-specific input set obtained through replacement of miR-141-binding sites with the sites for miR-375 and miR-145 (Fig. 7a). We validated the new assay with perturbations modified to match the new miRNA drug target and non-specific inputs. In the miR-23b assay, the Z’-factors for individual perturbations were comparable to those measured with the miR-122 assay, with the exception of miR-23b inhibition that resulted in ‘acceptable’, but not ‘excellent’, performance (Z’ = 0.38, Fig. 7b). This can be explained by the relatively low endogenous activity of miR-23b in HuH-7, and may be rectified with a more sensitive sensor for miR-23b activity. For the miR-122 assay with augmented non-specific inputs, most of the Z’-factors were comparable to the original miR-122 assay, apart from lower sensitivity to LNA-FF4 due to weaker miR-FF4 effect in this circuit (Fig. 7c). These data show that specific and non-specific inputs can be swapped and augmented in a modular fashion requiring only minor optimizations, resulting in drug discovery assays for new on- and off-targets.

**Assay comparison to bidirectional reporters.** A key assay feature is the ‘filtering’ of candidate compounds that would have been otherwise considered specific in simple luciferase or fluorescent reporter assays. To illuminate this advantage, we sought compounds or perturbations that affect RNAi pathway non-specifically. The simplest example is a validating perturbation that affects miR-122 together with the non-specific inputs, for example, a mixture of multiple LNAs with LNA-122 (Fig. 1d, Perturbation 5). When this combination is applied to a bidirectional reporter (Fig. 7d, right chart), it can be mistaken for a specific modulator; analysis of the circuit assay readouts suggests otherwise, because the non-specific RNAi readout mCerulean is significantly reduced (Fig. 7d, charts on the left).

**Discussion**

In this report we describe a large-scale mammalian gene circuit serving as an assay for drug discovery against miRNA targets, enabling highly precise identification of specific target modulators with high throughput. Until now off-target effects have been usually assessed in secondary screens. Although gene circuits have been suggested for use in small-molecule screening before, this is to the best of our knowledge, the first multi-input, customizable assay. It might as well be one of the first large-scale mammalian synthetic circuits that can be directly applied to an unmet technological need. Transient transfection of plasmid sets is sufficient to establish the assay because the readouts are averaged across transfected wells. Thanks to a cell-level computation the circuit performs over a set of potential off-target inputs, assay readouts carry rich information that is difficult to measure otherwise. We used miRNA as a drug target in this study because off-target effects are expected due to shared maturation pathway and common mechanisms of action. To read out the effect of a compound on a certain miRNA, we employ synthetic, fully complementary binding sites to record the miRNA’s activity. They are used as robust and representative reporters of the activity that a miRNA is expected to have on its many endogenous targets.

The same general strategy can be applied in just about any cell-based screening scenario given that appropriate sensing–computing networks are designed and experimentally implemented. For example, very recent report uncovered four consensus genes related to toxicity. Promoters driving these genes can be integrated with an assay such as ours and add to the information content of the non-specific reporter. Our results validate the basic premise behind rationally designed biological information-processing networks, namely that appropriate design frameworks, inspired by and built on solid engineering principles, can give rise to multiple systems with diverse...
properties and intended uses, eliminating the need to construct each new system from scratch. This is not to say that these design frameworks have reached the ‘plug and play’ stage. In this study we describe four different circuits that were tested extensively in silico and in experiments. Eventually, we arrived at a well-performing, customizable architecture and implemented an automated screening protocol, suggesting that these circuits can be used ‘as is’ in exploratory screening campaigns. Our engineering efforts have also augmented the toolkit of synthetic biology with new concepts such as the nested feed-forward motif from CFF assay. Thus, encounters of abstract concepts with real-life applications not only address specific needs, but also provide rich data that are applicable in other contexts of circuit engineering.

Methods

Plasmid construction. Standard cloning techniques were used to construct plasmids. *Escherichia coli* DH5α served as the cloning strain, cultured in LB Broth Miller Difco (BD) supplemented with appropriate antibiotics (Ampicillin, 100 μg ml⁻¹, Chloramphenicol, 25 μg ml⁻¹ and Kanamycin, 30 μg ml⁻¹). Enzymes were purchased from New England Biolabs (NEB). Phusion High-Fidelity DNA Polymerase (NEB) was used for PCR amplification. Oligonucleotides used as primers or for annealing were purchased from Microsynth, IDT or Sigma-Aldrich.
Digestion products or PCR fragments were purified using GenElute Gel Extraction Kit or Gen Elute PCR Clean Up Kit (both Sigma-Aldrich). Ligation was performed using T4 DNA Ligase (NEB) at 16°C for 1 h for sticky ends or at 14°C overnight for blunt-end ligation, followed by transformation into chemically competent cells and plating on LB agar plates with appropriate antibiotics. Clones were screened by colony-PCR using Quick-Load Tag 2X Master Mix (NEB) or by test restriction. Plasmids were sequenced by Microsynth. Detailed cloning procedure for each plasmid can be found in Supplementary Table 2, with primers listed in Supplementary Table 3 and gBlocks in Supplementary Table 4.

Cell culture and transfection. HuH-7 cells were received from the Health Science Research Resources Bank of the Japan Health Sciences Foundation (catalogue no. JCRB0403, Lot # 07152011) and cultured at 37°C, 5% CO₂ in DMEM, low glucose, GlutaMAX (Life Technologies, catalogue no. 21885-025), supplemented with 10% FBS (Sigma-Aldrich, catalogue no. F9665 or Life Technologies, catalogue no. 10279015). 195 ml of LiAc (Na) for transfection was prepared by slow freezing. Specifically, HuH-7 cells were seeded at a density of 30,000 cells per well in 1 ml of medium. The medium was replaced before transfection with medium supplemented with Doxycycline hyclate (Fluka, catalogue no. 44577) at a final concentration of 1 μg/ml - right before transfection. Transfections were performed using Lipofectamine 2000 Transfection Reagent (Life Technologies, catalogue no. P30433). Cells were passaged every 3–4 days using 0.25% Trypsin-EDTA (Life Technologies, catalogue no. 25200-072). Transfection was performed using Lipofectamine 2000 Transfection Reagent (Life Technologies, catalogue no. 11668-019) in uncoated 24-well plates (Thermo Scientific, catalogue no. 142475), 96-well plates (Thermo Scientific, catalogue no. 167008) or black μOpti-MEM for 24/96-well plates, respectively. After 72 h incubation at 37°C in 5% CO₂, the transfections were harvested and the same procedure was repeated for 24-h harvest. All fluorescence microscopy images were obtained with a Plan Apo C270/0.8, M. 50/0.5 compound objective and a Plan Apo 500/24, HC 542/27, BS 520, TxRed HC (HC 624/40, HC 562/40, BS 593), Cy5.5-A (HC 655/40, HC 716/40, BS 685), respectively. For iRFP, we used a 640-nm laser and a 780/60 nm emission filter. Photomultiplier tube voltages were checked and adjusted if needed using standard fluorescent beads (Sphero Rainbow Calibration Particles, RCP-30-5A (eight peaks) and Aliglow Flow Cytometry Alignment Beads, A-16500, Life Technologies) before and after each measurement to ensure constant device performance.

Luciferase assays. Cells were harvested 48 h post transfection. Supernatant was removed, cells were washed with PBS, and 100 μl 1× passive lysis solution (Promega) was added (incubation at 37°C for 15 min). The luciferase reaction was performed using Promega dual luciferase assay kit. We mixed 20 μl of lysed cells with 50 μl of the respective reagent. Measurements were performed with a plate reader and the data were analyzed using softMax Pro 4.5.3 software (both from Molecular Devices). In the case of sensor saturation, measurements were repeated with half of the sample.

Fluorescent microscopy. Cells were measured 48 h after transfection using an inverted fluorescent microscope (Nikon Eclipse Ti) using a Fiber Illuminator (Nikon Intensilight C-HGFI), with optical filter sets (Semrock) and a Digital Camera System (Hamamatsu, ORCA R2). Fluorescent microscopy has the advantage of excitation and emission band pass filters that can be fine-tuned. We measured mCherry, mCitrine, mCherry and mCitrine with the filter set CFP HC (438/24, HC 483/32, BS 458), YFP HC (500/24, HC 542/27, BS 520), TxRed HC (HC 624/40, HC 562/40, BS 593), Cy5.5-A (HC 655/40, HC 716/40, BS 685), respectively. The system was further optimized using a Plan Apo 1.4× objective and 2×1 frames were acquired to cover the majority of a well. For other experiments a Plan Fluo ×10 Ph1 DLL objective was used and 2×2 frames were acquired.

Flow cytometry. Samples were analysed 48 h after transfection using a BD LSRII Fortessa cell analyzer. Medium was removed and cells were incubated with 150/50 μl phenol-red free Trypsin (0.5% Trypsin-EDTA (Gibco, Life Technologies, catalogue #15400-054) 1:2 diluted with PBS (Life Technologies catalogue no. 10010-56) for 5/15 min for 24/96-well plates, respectively. Detached cells were transferred to small FACS tubes (FACS tubes, catalogue no. 02-1412-000) and kept on ice. The fluorophores were measured with a combination of excitation lasers and emission filters. For mCherry, we used a 561-nm excitation laser, 600 nm long-pass filter and 610/20 nm emission filter. For mCitrine, we used a 488-nm laser, 505 nm long-pass filter and 542/27 nm emission filter. For mCerulean we used our 445-nm laser and a 473/10-nm emission filter. For iRFP, we used a 640-nm laser and 780/60 nm emission filter. Photomultiplier tube voltages were checked and adjusted if needed using standard fluorescent beads (Sphero Rainbow Calibration Particles, RCP-30-5A (eight peaks) and Aliglow Flow Cytometry Alignment Beads, A-16500, Life Technologies) before and after each measurement to ensure constant device performance.

Flow cytometry data and image processing. All flow cytometry experiments were analyzed using FlowJo software. Compensation of crosstalk (<1.9%) for mCerulean into the 488–542/27 nm channel of mCitrine was performed if necessary using single-colour controls. The values in the various figures show the mean ± standard deviation. Flow cytometry data of Fig. 6 was processed as follows: (i) Live cells were gated using forward and side scatter. (ii) Within this gate, fluorophore positive gates are constructed using untransfected controls such that 99.9% of cells in this control sample fall outside of the selected gate. (iii) For each positive cell population in a given channel, the mean value of the fluorophore intensity is calculated and multiplied by the frequency of the positive cells to result in absolute intensity (a.u.): absolute intensity (a.u.) of fluorophore X = mean (fluorophore in fluorophore + cells) × frequency (fluorophore + cells)

For the relative intensities, the absolute intensity of the fluorophore of interest was divided by the absolute intensity of a constitutively expressed fluorescent protein that was co-transfected with the other plasmids.
mCitrine: each pixel in mCitrine images with an intensity > 250 a.u. is defined as positive, all others negative. Finally, we calculate the average for these pixels in all three colours (mCitrine, mCerulean, mCherry).

**Statistical analysis**
Ten compound storage plates were used in the screen, with about 80 compounds in each plate. Each plate was 'replicated' three times for the screening assay, and each compound assayed as a triplicate in three different plates. Triplicate assay plates belonging to separate storage plates were analysed separately. For mCitrine absolute readouts were compared, while both mCitrine and mCherry readouts were internally normalized to mCitrine values in each well. The set of measurements made with all compounds in a given storage plate (about 240 values in total for each readout) was used as a reference distribution for individual compounds from this plate for the purpose of exclusion and hit identification.

Specifically, a triplicate measurement of each compound was compared against its respective reference distribution using a two-sided t-test. Compounds that generated non-specific mCitrine and normalized mCerulean readouts that differed from the reference with a P value ≤ 0.1, were considered as potential non-specific modulators and thus were excluded from the analysis. Compounds that generated normalized mCherry readouts that differed from the reference distribution with P value ≤ 0.01, were classified as hits (Fig. 6a). To test whether the data used in the t-test are distributed normally, we built histograms of readouts and fitted them to normal distribution using histfit MATLAB function (Supplementary Figs 8 and 9). Z-factors for different perturbations were calculated with the following formula: $Z = 3 \times (\sigma (\mu) - \sigma (0))/\sigma (\mu)$, where $\sigma (\mu)$ and $\sigma (0)$ are the s.d. and mean of that perturbation respectively, and $\sigma (\mu)$ and $\sigma (0)$ are the s.d. of the control. For perturbations including LNAs, the control is measured with scrambled LNAs, and in perturbations involving miRNA mimic, the control is a scrambled miRNA mimic.

**Modelling**
Model was built and simulations performed using MATLAB and SimBiology toolbox. All simulations were performed with ‘sundials’ solver, with 10^-6 absolute and 0.001 relative tolerance. The end point of the simulation was chosen at 50 h. Parameter scans were performed with a MATLAB code generating the SimBiology model in a loop using bsimomulate command with different parameter values. End-point values were used for analysis. Further details can be found in Supplementary Note 3.

**References**
1. Tempe, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. Proc. Natl Acad. Sci. USA 109, 7085–7090 (2012).
2. Cummings, M., Breitling, R. & Takano, E. Steps towards the synthetic biology of polyketide biosynthesis. FEMS Microbiol. Lett. 351, 116–125 (2014).
3. Rino, K. et al. A universal RNAi-based logic evaluator that operates in mammalian cells. Nat. Biotechnol. 25, 795–801 (2007).
4. Ausländer, S., Ausländer, D., Müller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocircuits. Nature 487, 123–126 (2012).
5. Benenson, Y. Biomolecular computing systems: principles, progress and potential. Nat. Rev. Genet. 13, 455–468 (2012).
6. Ro, D.-K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940–943 (2006).
7. Lan, T. K. & Collins, J. J. Dispersing biologics with engineered enzymatic bacteriophage. Proc. Natl Acad. Sci. USA 104, 11197–11202 (2007).
8. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science 333, 1307–1311 (2011).
9. Weber, W. et al. A synthetic mammalian gene circuit reveals antituberculosis compounds. Proc. Natl Acad. Sci. USA 105, 9994–9998 (2008).
10. Weber, W. & Fussenegger, M. The impact of synthetic biology on drug discovery. Drug Discov. Today 14, 956–963 (2009).
11. Zhao, W., Bonem, M., McWhite, C., Silberg, J. I. & Segatori, L. Sensitive detection of proteasomal activation using the Deg-On mammalian synthetic gene circuit. Nat. Commun. 5, 3612 (2014).
12. Chen, Y. J., Yen, M. C. & Smolke, C. D. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. Proc. Natl Acad. Sci. USA 107, 8531–8536 (2010).
13. Macarron, R. & Hertzberg, R. P. In High Throughput Screening Vol. 565 (eds Janzen, W. P. & Bernasconi, P.) Ch. 1, 1–32 (Humana Press, 2009).
14. Kitano, H. A robustness-based approach to systems-oriented drug design. Nat. Rev. Drug Discov. 2, 202–210 (2007).
15. Muller, P. Y. & Milton, M. N. The determination and interpretation of the therapeutic index in drug development. Nat. Rev. Drug Discov. 11, 751–761 (2012).
16. Feng, Y., Mitchison, T. J., Bender, A., Young, D. W. & Tallarico, J. A. Multi-parameters profiling using cellular effects to characterize small-molecule compounds. Nat. Rev. Drug Discov. 8, 567–576 (2009).
17. Bader, A. G., Brown, D. & Winkler, M. The promise of microRNA replacement therapy. Cancer Res. 70, 7027–7030 (2010).
18. Esau, C. C. Inhibition of microRNA with antisense oligonucleotides. Methods 44, 55–60 (2008).
19. Weibull, W. A statistical distribution function of wide applicability. J. Appl. Mech. Trans. ASME 18, 293–297 (1951).
20. Schmidt, M. F. Drug target miRNAs: chances and challenges. Trends Biotechnol. 32, 578–585 (2014).
21. Kasinski, A. L. & Slack, F. J. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. Nat. Rev. Cancer 11, 849–864 (2011).
22. Ling, H., Fabbri, M. & Calin, G. A. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. Nat. Rev. Drug Discov. 12, 847–865 (2013).
23. Li, Y. et al. HMDD v2.0: a database for experimentally supported human microRNA and disease associations. Nucleic Acids Res. 42, D1070–D1074 (2014).
24. Coulouras, C., Factor, V. M., Andersen, J. B., Durkin, M. E. & Thorgersson, S. N. Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. Oncogene 28, 3526–3536 (2009).
25. Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific microRNA. Science 309, 1311–1315 (2005).
26. Prochazka, L., Angelici, B., Haeffler, B. & Benenson, Y. Highly modular bow-tie gene circuits with programmable dynamic behaviour. Nat. Commun. 5, 4729 (2014).
27. Sittampalam, G. S. et al. (eds) Assay Guidance Manual [Internet]. Available at <http://www.ncbi.nlm.nih.gov/books/NBK53196/ (Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2004).
28. Zhang, J.-H., Chuang, T. D. Y. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J. Biomed. Screen. 4, 67–73 (1999).
29. Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T. & Sato, J. Growth of human hepatoma cells line with differentiated functions in chemically defined media. Cancer Res. 42, 3858–3863 (1982).
30. Lohmann, V. et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285, 110–113 (1999).
31. Kambara, H. et al. Establishment of a novel permissive cell line for the propagation of hepatitis C virus by expression of microRNA miR-122. J. Virol. 86, 1382–1393 (2012).
32. Kutay, H. et al. Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. J. Cell. Biochem. 99, 671–678 (2006).
33. Deal et al. GSK invests in targeting microRNA for the treatment of hepatitis C. Nat. Rev. Drug Discov. 9, 350 (2010).
34. Landgraf, P. et al. A mammalian miRNA expression array based on small RNA library sequencing. Cell 129, 1401–1414 (2007).
35. Ke, M. et al. Complement of Argonaute enhances RNA interference toward perfect match binding sites. Proc. Natl Acad. Sci. USA 105, 9284–9289 (2008).
36. Chien, C.-H. et al. Identifying transcriptional start sites of human microRNAs based on high-throughput sequencing data. Nucleic Acids Res. 39, 9345–9356 (2011).
37. Weber, W., Kramer, B. P., Fux, C., Keller, B. & Fussenegger, M. Novel promoter/transactivator configurations for macrolide- and streptogramin-responsive transgene expression in mammalian cells. J. Gene. Med. 4, 676–686 (2002).
38. Mangan, S. & Alon, U. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. Nat. Rev. Drug Discov. 5, 676–686 (2006).
39. Weibel, W. A statistical distribution function of wide applicability. J. Appl. Mech. Trans. ASME 18, 293–297 (1951).
40. Dickhaus, T. Simultaneous Statistical Inference—With Applications in the Life Sciences (Springer-Verlag, 2014).
47. Mori, M. et al. Hippo signaling regulates microprocessor and links cell-density-dependent miRNA biogenesis to cancer. Cell 156, 893–906 (2014).
48. MacDonald, M. L. et al. Identifying off-target effects and hidden phenotypes of drugs in human cells. Nat. Chem. Biol. 2, 329–337 (2006).
49. Zhang, J. D., Berntenis, N., Roth, A. & Ebeling, M. Data mining reveals a network of early-response genes as a consensus signature of drug-induced in vitro and in vivo toxicity. Pharmacogenomics J. 14, 208–216 (2014).

Acknowledgements
The work was supported by ETH Zurich and by the Swiss National Science Foundation as part of the NCCR Molecular Systems Engineering. We thank M. Fussenegger for providing the plasmids encoding PIT2 and ET activators and their regulated promoter sequences; N. Lapique and I. Müller for plasmid precursors; K. Hoelle for control plasmids; M. Dessing, T. Horn, T. Lopes and V. Jäggin for assistance with flow cytometry and image acquisition, U. Senn for robot programming and handling; and Benenson lab members for discussions.

Author contributions
B.H. conceived the study, performed the great majority of experiments, analysed data and wrote the manuscript. B.A. and L.P. performed some experiments and took part in data analysis and manuscript proofreading. Y.B. conceived the study, performed simulations, analysed data, wrote the manuscript and supervised the project.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: Y.B. and B.H. are authors on a Priority Filing with European Patent Office entitled ‘Screening assay for identifying miRNA-specific drug compounds’, filing number PCT/EP2015/062258. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Haefliger, B. et al. Precision multidimensional assay for high-throughput microRNA drug discovery. Nat. Commun. 7:10709 doi: 10.1038/ncomms10709 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/