Article

Nucleic Acid Quantitation with Log-Linear Response Hybridization Probe Sets

Lucia Ruojia Wu, John Z. Fang, Dmitriy Khodakov, and David Yu Zhang

ACS Sens., **Just Accepted Manuscript** • Publication Date (Web): 01 Jun 2020

Nucleic Acid Quantitation with Log-Linear Response Hybridization Probe Sets

Lucia R. Wu[♯,1] John Z. Fang[♯,1] Dmitriy Khodakov,[1] and David Yu Zhang[*1,2]

[1]*Department of Bioengineering, Rice University, Houston, TX*
[2]*Systems, Synthetic, and Physical Biology, Rice University, Houston, TX*

(Dated: May 25, 2020)

ABSTRACT: Concentrations of different nucleic acid species in biological samples span many orders of magnitude. Real-time PCR maps the concentration of a target nucleic acid sequence log-linearly into cycle threshold to enable quantitation with a wide dynamic range, but suffers from enzymatic biases. Here, we present a general design for constructing hybridization probe sets with highly log-linear response curves to enable accurate enzyme-free quantitation across large ranges (more than 6 logs) of target DNA concentrations. The sensitivity of each component probe is accurately adjusted via formulation stoichiometry to reduce the standard error of target quantitation down to 7%. As a proof of concept, we show multiplexed quantitation of three microRNA species in total RNA of human brain and liver.

Keywords: Nucleic acids thermodynamics; toehold probe; nucleic acid hybridization; dynamic range; RNA quantitation; enzyme-free biosensor

Nucleic acids are critical biological molecules, and precise quantitative understanding of different nucleic acid species holds the key to elucidating biological pathways [1, 2] and guiding clinical treatment of diseases [3]. Technology platforms of varying complexity, prices, multiplexing, and accuracy have been developed over the past 30 years for sequence-specific nucleic acid quantitation (Table 1); these technologies can be divided into 2 categories: enzymatic and non-enzymatic. Enzymatic technologies include qPCR and reverse-transcription PCR (RT-PCR), Next-generation Sequencing (NGS), digital PCR [4], various isothermal amplification methods [5] and most microarray-based assays. Non-enzymatic or direct DNA / RNA detection technologies, including electrochemical [6, 7] and microresonator-based methods [8, 9], have gained scientific interest in the past 10 years because they are tolerant to chemical and biological impurities in the DNA sample, and because they might be used as affordable solutions in point-of-care settings.

Most nucleic acid sensors and assays exhibit linear input-output response curves in their quantitative dynamic range. The linear response allows more accurate quantitation than assays with sublinear saturation responses (e.g. ELISA assays for proteins). One notable exception is quantitative PCR (qPCR), in which the amount of target nucleic acid in a sample is mapped in a log-linear fashion to cycle threshold Ct. This log-linear response curve allows accurate quantitation across many orders of magnitude of target concentrations, and relaxes the dynamic range requirements for sensors (e.g. photodiodes) [10]. However, qPCR is inherently based on enzymatic amplification, and thus requires purified DNA / RNA samples for ultrasensitive detection or sample dilution to mitigate PCR inhibitor issue. Furthermore, buffer incompatibilities often render multi-enzyme reactions unstable. Finally, qPCR also frequently fails to accurately quantitate in multiplexed settings due to issues such as primer dimerization, sequence bias, and dNTP exhaustion.

Recent advances in electrochemical [6, 7, 11, 12], microresonator-based [8, 9] sensors are now enabling attomolar levels of molecular sensitivity, and are potentially obviating the need for enzymatic amplification of nucleic acids. Unamplified detection and quantitation of nucleic acids are desirable not only for their ease of use, but also because the absence of amplicons would substantially mitigate the risk of amplicon contamination that result in false positives when applied to ultrasensitive detection. Accurate quantitation across many orders of magnitude of potential target concentrations would greatly increase the values of these technologies, because expression levels and copy numbers of biological nucleic acids can vary greatly [1, 13, 14].

Here, we present a generalized method for designing a set of enzyme-free hybridization probes that collectively exhibit a log-linear binding response to a target DNA or RNA sequence. Components of the probe set can be individually tuned to improve the log-linearity of the probe set, and enable accurate quantitation. Furthermore, being enzyme-free and highly specific to even single nucleotide variations, these high dynamic range probe sets can enable multiplexed quantitation of DNA and RNA species. We have demonstrated preliminary proof-of-concept experiments in bulk solution and on surfaces (microarray), culminating in a demonstration of 3-plex microRNA quantitation from total RNA. Though the log-linear response hybridization probe set is not yet optimized for sensitivity, mutiplexing and reaction time, we envision that our technology may serve as an alternative solution for broad dynamic range quantitation of nucleic acids.

Material and Methods

Oligonucleotides. All DNA and RNA oligonucleotides were purchased from Integrated DNA Technologies. Fluorophore- or quencher- modified DNA oligonucleotides were purified by HPLC; unmodified DNA or RNA oligonucleotides were non-purified.

Standard Buffer Conditions. DNA oligonucleotides used in non-RNA-related experiments were resuspended in 1×PBS buffer (0.15 M Na+, purchased as 10× stock from Sigma Aldrich) with 0.1% TWEEN 20 (Sigma Aldrich) and

	LLR	qPCR/RT-qPCR	Microarray	NGS
Dynamic range	**10^6**	10^8	$10^3 - 10^4$	$\approx 10^4$
Enzymatic amplification	**No**	Yes	Yes / No*	Yes
Single nucleotide specificity	**Yes**	Yes / No*	No	**Yes**
Response type	**Log-linear**	**Log-linear**	Linear	Linear
Detection limit	10^{-11} M**	$< 10^{-15}$ **M**	10^{-12} M	10^{-15} **to** 10^{-12} **M**
Multiplexing	≤ 4	≤ 4	> 100	> 100

TABLE 1: Comparison of nucleic acid quantitation methods. * Can be yes or no depending on the specific method used. ** Can be improved by using readout equipment of higher sensitivity or using signal amplification methods.

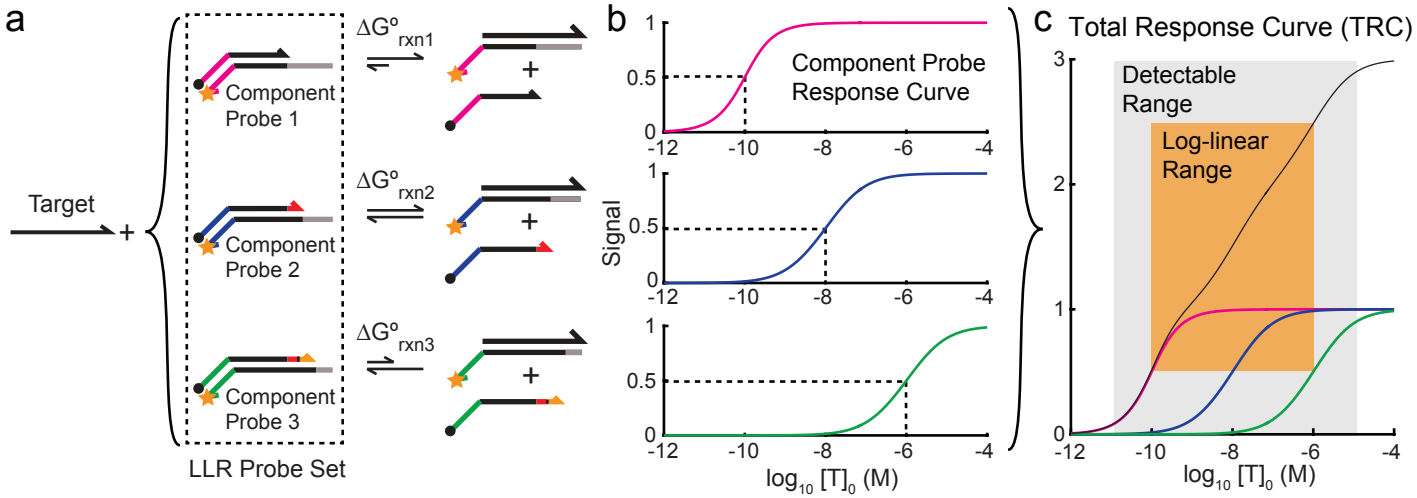

FIG. 1: Hybridization probe sets exhibiting log-linear response. **(a)** Several toehold probes (component probes) are designed to react with the same nucleic acid target, but with different sensitivities. **(b)** The response curve of each component probe is sigmoidal, but different probes' curves are horizontally shifted depending on their $\Delta G^\circ_{\text{rxn}}$ and formulation stoichiometry. **(c)** Designed properly, the aggregate signal response of several different toehold probes will exhibit a large log-linear range, in which a multiplicative increase in target concentration is reflected as an additive increase in fluorescence signal. The detectable range extends approximately 1 order of magnitude above and below the limits of the log-linear range; target concentration inference accuracy drops with the distance from boundary of the log-linear range.

stored at 4 °C. RNA oligonucleotides and DNA oligonucleotides used in RNA-related experiments were resuspended in above mentioned buffer with additional 0.1% RNaseZap (Life Technologies) and stored at 4 °C. Human liver or brain total RNA (Life Technologies) were stored in THE RNA Storage Solution (Life Technologies) at -20 °C. Reaction buffer was 1×PBS with 0.1% TWEEN 20 for non-RNA-related experiments, and 1×PBS with 0.1% TWEEN 20 and 0.1% RNaseZap for RNA-related experiments.

Response Curve Characterization and Target Concentration Quantitation. Data shown in manuscript Figs. 3-5 were obtained using similar protocols. First, stock concentration of each oligonucleotide was calibrated with NanoDrop 2000 (Thermo Scientific). Component probes were prepared by mixing protector and complement oligonucleotide solutions in reaction buffer for at least 1 hour at room temperature. In each preparation, there was at least 10 nM of both complement and protector at desired stoichiometry, so the reaction should reach equilibrium in 1 hour, based on our previous experience [1]. LLR probe sets were prepared by mixing all the pre-hybridized component probes; in multiplexed experiments, the 3 LLR sets (15 probes) were combined into the same solution (see Table S3-1 for detailed stoichiometry of each component probe).

Next, the pre-hybridized probes or LLR probe sets were mixed with different concentrations of target at the desired final concentration; the reaction was incubated for 12 to 24 hours at 25 °C in a dark box. Fluorescence was measured with a Horiba FluoroMax-4 spectrofluorimeter and Hellma Semi-Micro 114F spectrofluorimeter cuvettes. Excitation and emission wavelengths were 558/580 nm for TAMRA, 648/663 nm for Alexa647, and 491/514 nm for Alexa488; these are characterized to be the wavelengths that generate maximum fluorescence for each fluorophore in the buffers we used. Slit widths were 10 nm for TAMRA, 8 nm for Alexa647 and Alexa488, and integration time was 10 s. An external temperature controller maintained the reaction temperature to 25 ± 0.2 °C. Raw fluorescence signal of each sample was calculated as the average of last 3 reads.

MicroRNA Quantitation in Human Total RNA Samples. LLR probe sets were prepared using the protocol described above. 3 μL of 1 mg/mL human liver or brain total RNA sample was mixed with 4 μL of formamide (Sigma Aldrich) and 13 μL of the LLR probe sets solution, result-

ing in a 20 μL solution containing desired concentration of each component probe (Table S3-1), 0.15 mg/mL total RNA, and 20% formamide. The background fluorescence was measured using a similar solution containing 3 μL of THE RNA Storage Solution instead of total RNA samples. Reactions were incubated for 12 to 24 hours at 25 °C in a dark box; fluorescence was measured using a similar protocol to that described above, except for the cuvette was Starna 16.10F4-Q-10/Z15 Sub-Micro Fluorometer Cell.

Microarray Experiments. All chemicals were purchased from Sigma-Aldrich and used as received with no additional purification. HPLC-purified alkyne-modified oligonucleotides were purchased from Integrated DNA Technologies (IDT). Water clear glass slides (Corning) were washed and oxidized by consecutive immersing in 10 M sodium hydroxide and Piranha solutions (70% sulfuric acid, 30% hydrogen peroxide) for 30 and 60 min respectively. Then, the amino-silane layer were deposited on the glass surface by incubation of the slides in the solution containing 2% (3-aminopropyl)triethoxysilane (APTES), 3% water and 95% ethanol (by volume) for 20 min followed by ethanol washing and baking at 120 °C for 1 hour. Subsequently, the APTES modified surface was reacted with 10 mM p-phenylene diisothiocyanate (PDITC) dissolved in DMF/pyridine mixture (9/1 volume ratio) for 1 hour, followed by ethanol washing and drying at 120 °C for 20 min.

The PDITC functionalized surface was then converted into azide-functionalized surface by reacting with 20 mM solution of the 11-azido-3,6,9-trioxaundecan-1-amine in 0.1 M sodium phosphate buffer pH 8.0 for 1 hour. Attachment of the 3′-alkyne modified oligonucleotide capture probes to the azide-functionalized surface was performed via copper catalyzed 1,3-dipolar cycloaddition. The final water based reaction mixture contained 50% DMSO (by volume), 0.2 M triethylammonium acetate buffer pH 7.0, 0.5 mM ascorbic acid, 0.5 mM Cu-TBTA complex and 500 nM of the alkyne modified oligonucleotide capture probe. The immobilization reaction was let to proceed for 1 hour at room temperature and relative humidity of 50%. In total, six oligonucleotides were manually spotted on each microarray surface using positive displacement pipet. After immobilization the microarrays was washed with hot 0.1x SSC buffer (70 °C) containing 0.1% SDS.

MicroRNA labeling. Both synthetic microRNAs and total RNA specimens from brain and liver tissues were fluorescently labeled using Hi-Power Labeling kit (Exiqon) strictly following manufacturer's recommendation. For each labeling reaction 1 μg of synthetic RNA oligonucleotide (150 pmole) or 3 μg of total RNA specimen was used. Labeled samples were stored at -80 °C until further use.

Microarray hybridization and imaging. Microarray hybridization was performed in 2x PBS buffer with 0.1% SDS and 0.5 nM control hybridization oligonucleotide. Hybridization was performed at 40 °C. The hybridization microarrays were washed twice with 2x PBS supplied with 0.1% SDS, twice with 1x PBS, and twice with 0.2x PBS. The arrays were then dried with stream of nitrogen and immediately imaged. Imaging of the microarrays was performed in Cy3 fluorescent channel using Zeiss Axio Observer A1m microscope equipped with 10x EC Plan-Neofular objective. All images were taken with 1s exposure time.

Surface-based LLR. LLR probe sets targeting the miR-122, miR-9, and miR-10b miRNAs were redesigned and resynthesized with 3′ biotin functionalizations. The sequences of the probes were modified in the non-homologous region to compensate the loss of thermodynamic favorability contributed by the fluorophore-quencher interactions. Each probe set was reacted with samples of end-labeled synthetic microRNA.

After reactions reached equilibrium (at least 5 hours), 0.1 μL of the reaction mixture was spotted on Super-Streptavidin microarray substrate slides (Arrayit) in reaction/deposition buffer (1x PBS, 0.02% Tween-20). The deposition process was carried out for 1 hour at room temperature and relative humidity of 50%. After the deposition the slides were washed twice with the reaction/deposition buffer followed by one washing with 0.2x PBS, then dried with nitrogen and imaged.

Safety Comments. Sodium hydroxide, Piranha solutions (70% sulfuric acid, 30% hydrogen peroxide),(3-aminopropyl)triethoxysilane (APTES) and p-phenylene diisothiocyanate (PDITC) are chemical hazards. These chemicals were handled only in designated area with precaution and personal protection equipment.

Design Theory

A log-linear response (LLR) probe set is constructed from a number of *component probes* (CP), each capable of hybridizing to the same target sequence, but with different sensitivities (Fig. 1). Each component probe has a characteristic *response curve*, indicating signals elicited at different target concentrations. Under certain combinations of component probe sensitivities and concentrations, the *total response curve* (TRC) of the probe set mixture will possess a log-linear dynamic range, in which a multiplicative increase in target concentration results in a constant increase in signal.

The detectable range of a LLR probe set extends beyond its log-linear range, as concentration changes are reflected monotonically in TRC signals (Fig. 1c). Quantitation accuracy worsens as target concentration exits the log-linear range, to a limit of about 1 order of magnitude above and below the maximum and minimum of the range. For comparison, traditional single-probe system essentially has no log-linear range, and has about 2 orders of magnitude of detectable range.

Here, we implement LLR probe sets using toehold probes [15] as component probes, because probe sensitivities can be precisely and independently tuned via formulation stoichiometry. As will be explained later in this section, errors in probe sensitivity result in poor log-linearity of the overall probe set, and reduce the accuracy of quantitation from the observed signal. Conversely, accurately tuned component probes are the foundation of probe sets with log-linear

responses that enable scalable bijective mapping between signals and concentrations, with little to no information loss when inferring concentration.

A toehold probe is a partially double-stranded DNA molecule comprising a protector (P) and a complementary strand (C). When a target (T) is present in solution, it may initially hybridize to CP via a single-stranded region, and subsequently displace P (see Supplementary Section S1 for mechanistic details). Toehold probes are designed such that at operational temperature and buffer conditions, the binding of P or T to C is mutually exclusive and trimolecular species CPT are high-energy intermediates rarely observed in the solution. In this manuscript, we use a fluorophore-functionalized C and quencher-functionalized P such that CP is originally dark; through the course of target hybridization, the fluorophore is unquenched and provides increased solution fluorescence. In principle, the signal may be any output such as chemiluminescence or electrochemistry readout dependent on probe-target binding.

Sensitivity spacing. Here, we define probe sensitivity $s = -\log_{10}([T_{1/2}])$ to be the negative logarithm of the target concentration that elicits a half-maximum binding response; e.g. $s = 10$ means the component probe is half-saturated at 100 pM of target. Probe sensitivity s can be analytically solved for a toehold probe given the $\Delta G^\circ_{\text{rxn}}$ of its reaction with the target, its concentration, and its formulation stoichiometry:

$$\begin{aligned} \text{CP} + \text{T} &\rightleftharpoons \text{TC} + \text{P} \\ e^{-\Delta G^\circ_{\text{rxn}}/R\tau} = K_{\text{eq}} &= \frac{[\text{TC}]\ [\text{P}]}{[\text{CP}]\ [\text{T}]} \end{aligned}$$

By definition, when initial T concentration is $[\text{T}_{1/2}]$, then at equilibrium $[\text{TC}] = [\text{CP}] = [\text{CP}]_0\ /\ 2$:

$$\begin{aligned} e^{-\Delta G^\circ_{\text{rxn}}/R\tau} &= \frac{[\text{P}]_0 + [\text{CP}]_0/2}{[\text{T}_{1/2}] - [\text{CP}]_0/2} \\ [\text{T}_{1/2}] &= e^{\Delta G^\circ_{\text{rxn}}/R\tau} \cdot ([\text{P}]_0 + [\text{CP}]_0/2) + [\text{CP}]_0/2 \end{aligned}$$

More generally, the yield response of a component probe to different target concentrations can be solved and plotted as a sigmoid in semilog axes (Fig. 1).

We first performed simulations to study the combinations of component probe concentrations and sensitivities s that would result in LLR (Fig. 2). First, we considered the intuitive probe set design that contained component probes with equal concentration and uniformly spaced s (e.g. $s = 5, 9, 13$). The initial results showed that for a given desired log-linear dynamic range of the probe set (8 logs in Fig. 2a), a probe set consisting of more component probes, and thus smaller Δs, provided better log-linearity (quantitated as the linear correlation coefficient r^2 of the normalized signal against the logarithm of the target concentration).

Given the above simulation results, we would expect that log-linearity of the probe sets improves monotonically with decreasing Δs: the higher the number of component probes and the more tightly spaced, the better the results. However, this is actually incorrect: Fig. 2b shows that there is an optimal Δs spacing of between 1 and 1.5 for probe sets with 3 or more component probes. The fact that the optimal $\Delta s \geq 1$ is fortuitous is because it reduces the complexity and price of probe sets: a probe set with desired log-linearity between 10^{-x} and 10^{-y} M requires $(\frac{x-y}{\Delta s} + 1)$ component probes.

Fig. 2c shows the effects of normally distributed errors in $\Delta G^\circ_{\text{rxn}}$ for each component probe (mean of N = 1000 Monte Carlo simulations) on r^2. In our experience the typical error in $\Delta G^\circ_{\text{rxn}}$ has a standard deviation of roughly 2 kcal/mol, consistent with literature [16], which results in roughly 0.05 reduction in r^2, and corresponds to a standard error of target concentration quantitation greater than 150%. Probe $\Delta G^\circ_{\text{rxn}}$ can be adjusted through sequence redesign, but it is frequently impractical to consistently achieve $\Delta G^\circ_{\text{rxn}}$ within 0.5 kcal/mol of the designed number even with iterative optimization.

Compared to errors in probe $\Delta G^\circ_{\text{rxn}}$, errors in probe concentration have a relatively smaller impact on the overall log-linearity of the probe set (Fig. 2d). The typical experimental standard deviation of concentration error should be roughly 5%, which results in less than 0.001 reduction of r^2. Based on our analyses, accurate target concentration quantitation to less than 20% error requires $r^2 > 0.997$, so concentration errors will usually not significantly reduce quantitation accuracy from theoretical optimum (Fig. 2d). Additionally, even in the case of gross concentration error, probe concentrations can be adjusted based on initial experimental results.

Thus far, we have been using the r^2 of signal against log([Target]) as a metric of goodness. For quantitation purposes, what we care about is the deviation between the true target concentration and the inferred target concentration based on a linear interpolation. The two metrics are different because quantitation error is not calculated with the logarithm of the concentration, and the latter examines horizontal (rather than vertical) deviation on the TRC. Fig. 2e shows the scatter plot summary of the mean quantitation error for 10,000 simulations, plotted against the r^2 of the TRC (see Supplementary Section S2 for details). There is a strong but not perfect correlation between r^2 and quantitation error. Generally though, it appears that $r^2 \geq 0.998$ is sufficient to ensure quantitation errors of below 20%.

Although the uniformly spaced component probes of equal concentration represents an easy way of achieving high log-linearity, this approach does not produce the LLR probe set with the highest r^2 value. For a 4-probe LLR set with equal concentrations and sensitivities spaced 1 unit apart, we achieve $r^2 = 0.9998$ ($1\text{-}r^2 = 2{\cdot}10^{-3}$) over a 3 log target concentration range. By contrast, when we performed numerical optimizations of 6,000 randomly generated sets of components probes (4 in each set), over 200 of the optimized sets yielded $r^2 > 0.99999$ ($1 - r^2 < 10^{-5}$), and 2 of them yielded $r^2 > 0.999999$ ($1 - r^2 < 10^{-6}$). Fig. 2f shows an example of optimized probe sets. However, we were unable to generalize design guidelines for these optimized sets (see

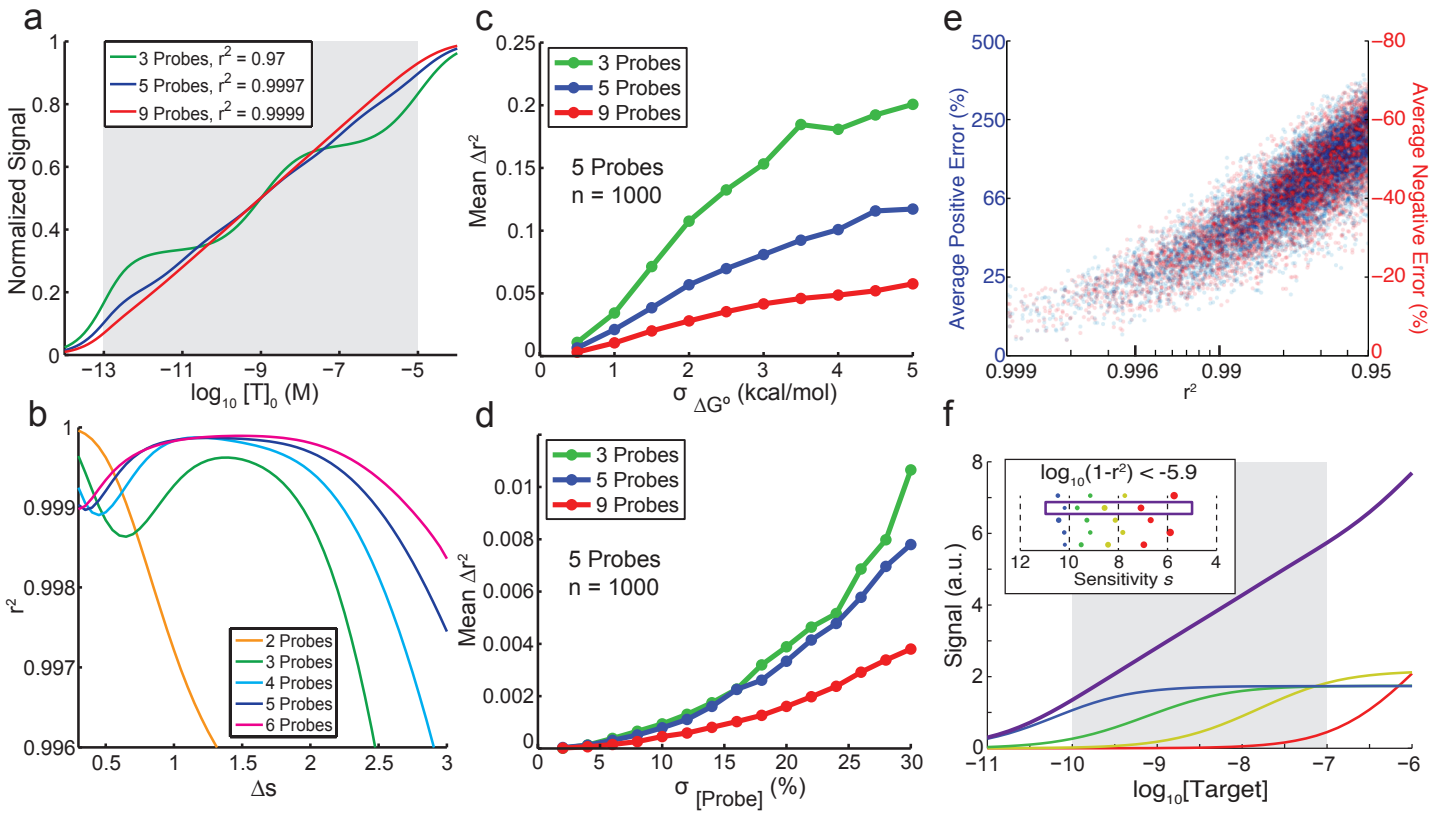

FIG. 2: Theoretical analysis of LLR probe set design. **(a)** For a fixed dynamic range (here 8 logs, shown in gray), increasing the number of component probes improves the log-linearity of the TRC. The component probes considered for this analysis have sensitivities (s) uniformly spaced between 13 and 5 (inclusive), and each has a concentration of 10^{-13} M and a formulation stoichiometry of $[\mathrm{P}]_0/[\mathrm{CP}]_0 = 1$. **(b)** Optimal component probe sensitivity spacing for uniformly spaced and equal concentration component probes. For probe sets with more than 2 component probes, sensitivity spacing (Δs) between 1 and 1.5 provide the best r^2 between the total binding signal and the logarithm of the target concentration. **(c)** Effects of deviations in component probe $\Delta G^\circ_{\mathrm{rxn}}$ on log-linearity of the total response curve. Each data point represents the mean reduction of r^2 for 1000 simulations, in which each component probe has its $\Delta G^\circ_{\mathrm{rxn}}$ normally distributed with standard deviation shown. **(d)** Effects of deviations in component probe concentration have smaller effects on log-linearity compared to errors in $\Delta G^\circ_{\mathrm{rxn}}$. **(e)** Quantitation accuracy improves with r^2, based on simulations of 10,000 probe sets. Target concentrations inferred from a log-linear fit to the signal were sometimes higher than their actual values, and sometimes lower. Higher inferred concentrations could have unlimited upwards error ($+\infty\%$), whereas lower inferred concentrations are limited to -100%. For this reason, positive errors are plotted in blue and negative errors are plotted in red, to two different scale bars. **(f)** LLR probe sets with non-uniformly spaced component probe sensitivities. The inset shows 5 representative LLR probe sets, each consisting of 4 probes, exhibiting $r^2 > 0.99999$ over 3 orders of magnitude of target concentration. Circles denote component probes; center positions represent sensitivity, and areas represent relative concentrations.

Supplementary Section S2 for discussion).

Sensitivity tuning. Given the importance of component probe sensitivity on the log-linearity and quantitation accuracy of LLR probe set , it is critical to use a probe architecture that facilitates precise and decoupled sensitivity tuning. Recently, we showed that toehold probes can be accurately and finely tuned for yield and selectivity by adjusting the protector concentration [17] (Supplementary Section S3). The sensitivity of a toehold component probe can be similarly tuned by changing formulation stoichiometry $V \equiv [\mathrm{P}]_0/[\mathrm{CP}]_0$.

Assuming an initial observed sensitivity s_0, initial probe stoichiometry V_0, and initial probe concentration $[\mathrm{CP}]_0$, our analysis indicates that a desired sensitivity s' can be achieved by reformulating the probe with stoichiometry V':

$$\frac{V' + 1/2}{10^{-s'}/[\mathrm{CP}]_0 - 1/2} = \frac{V_0 + 1/2}{10^{-s_0}/[\mathrm{CP}]_0 - 1/2} = e^{-\Delta G^\circ_{\mathrm{rxn}}/R\tau}$$

$$V' = \frac{2 \cdot 10^{-s'} - [\mathrm{CP}]_0}{2 \cdot 10^{-s_0} - [\mathrm{CP}]_0}(V_0 + \frac{1}{2}) - \frac{1}{2}$$

If $V \gg \frac{1}{2}$ and $\frac{10^{-s}}{[\mathrm{CP}]_0} \gg \frac{1}{2}$, the above equations can be approximated as $V' \approx 10^{s_0 - s'} \cdot V_0$. In these conditions, increasing $[\mathrm{P}]_0$ 10-fold will lower the sensitivity by 1 unit. In this article, we used the exact calculation formula for V'.

To ensure that each component probe can be independently stoichiometrically tuned, we designed the nonhomologous regions of different component probes to be pairwise orthogonal. Such decoupling of component probe systems,

preventing the protector of one component probe to serve as the protector of another, is necessary to ensure reliable adjustment of component probes and consequent accurate target concentration quantitation.

In principle, probe sets could be constructed by using any probe architecture (e.g. molecular beacons [18, 19]), and high log-linearity may be achievable through brute-force sequence optimization of probes for any particular target sequence and set of experimental conditions. In practice, however, other molecular designs are unable to be generalized to arbitrary biological sequences and biologically relevant concentrations. Additionally, the accurate and decoupled sensitivity adjustment provided by toehold probes shortens the design-test-optimize cycle. This is particularly true when performing multiplexed analysis, wherein a large number of component probes may in aggregate show significant nonspecific interactions. Finally, analysis of biological samples such as total RNA may necessitate the use of chemical additives such as formamide or DTT, or higher temperatures to denature RNA. On-the-fly stoichiometric tuning allows probe sets to be adapted to be log-linear in virtually any condition, whereas typical probes "hard-coded" to provide high log-linearity in one set of temperatures/buffers will fail in another.

Experimental Results

Probe and Probe Set Characterization. To validate our design theory, we first constructed an LLR probe set targeting a subsequence of the human *htt* gene, and initially used a synthetic DNA oligonucleotide bearing the target sequence as our sample (see Supplementary Section S1 for sequence design details). The probe set consists of 5 component probes (Fig. 3a), designed to possess $\Delta G^\circ_{\text{rxn}}$ of 1.18, -0.46, 2.69, 4.11, and 5.25 kcal/mol (Supplementary Section S3), so that at $[\text{P}]_0/[\text{CP}]_0 = 1$ and $[\text{CP}]_0 = 100$ pM, the theoretical probe sensitivities are $s = 8.94$, 9.92, 7.85, 6.81, and 5.98, respectively.

Pre-hybridized component probes were mixed with target and incubated at room temperature (25 °C) for 12-24 hours with no exposure to light; all reactions were performed in 1×PBS. The individual response curve of each component probe is shown in Fig. 3b; from this data, we fitted the true $\Delta G^\circ_{\text{rxn}}$ and maximum signal (corresponding to 100% yield) of each component probe. Note that the first and second component probes are listed in unusual order, because experimentally the second component probe was observed to be less sensitivity than the first, even though by design criteria it should have been the most sensitive.

We next adjusted stoichiometric ratio $V \equiv [\text{P}]_0/[\text{CP}]_0$ to modulate component probe sensitivity. In Fig. 3b, probe CP3 originally possessing a sensitivity of 8.50 when $[\text{P}]_0/[\text{CP}]_0 = 1$; this was successfully tuned to $s = 8.80$, 8.00, and 7.00 via $V = 0.25$, 4.35, and 48.2, respectively. Because stoichiometric tuning is highly predictable, we did not need to characterize the adjusted individual response curve of each component probe. Rather, we determined the tuned stoichiometry via

$$V' = \frac{2 \cdot 10^{-s'} - [\text{CP}]_0}{2 \cdot 10^{-s_0} - [\text{CP}]_0}(V_0 + \frac{1}{2}) - \frac{1}{2}$$

Each component probe was individually tuned to yield $s =$ 10, 9, 8, 7, and 6 used the observed sensitivities s at $V =$ 1. Consistent with our theoretical analysis, stoichiometric tuning significantly improved the log-linearity of the probe set.

To assess the quantitative accuracy of the LLR probe set, we compared the true target concentration $[\text{T}]_{\text{true}}$ to $[\text{T}]_{\text{fitted}}$, to the target concentration interpolated from a best-fit log-linear regression. Due to the log-linear nature of the TRC, we assume that the inferred concentration error on the log scale ($\log_{10}[\text{T}]_{\text{fitted}}$ - $\log_{10}[\text{T}]_{\text{true}}$) follows a normal distribution, and the standard deviation of this value is $\sigma_{\log}$. Therefore, the quantitation error, calculated as

$$\begin{aligned}\text{Error} &= ([\text{T}]_{\text{fitted}} - [\text{T}]_{\text{true}})/[\text{T}]_{\text{true}} \\ &= 10^{(\log_{10}[\text{T}]_{\text{fitted}} - \log_{10}[\text{T}]_{\text{true}})} - 1\end{aligned}$$

obeys an asymmetrical log-normal distribution, instead of normal distribution for standard linear quantitation methods. In this article, we report standard deviation of inferred concentration error as $\sigma = 10^{\pm\sigma_{\log}}$, and the relative quantitation error as $10^{\pm\sigma_{\log}} - 1$.

Fig. 3d shows the untuned LLR probe set ($V = 1$ for all probes) results in $r^2 = 0.96$ and quantitation error $\sigma = 10^{\pm 0.24}$, corresponding to an average of 74% positive quantitation error and 42% negative quantitation error. Stoichiometric tuning of each component probe improved r^2 to 0.997 and reduced concentration inference standard deviation to $10^{\pm 0.07}$, corresponding to 17% positive error and 15% negative error (Fig. 3e).

The necessity of stoichiometric tuning to achieve desired probe sensitivities is vividly demonstrated in the comparison of Component Probes 1 and 2 (Fig. 3ab). The protector for CP1 is one nucleotide longer than that of CP2; consequently mFold and NUPACK predicted that CP1 would have more positive ΔG° and lower sensitivity than CP2. Experimental characterization shows the opposite, with CP2 exhibiting marginally lower sensitivity. Without stoichiometric tuning, CP1 and CP2 would not have been able to function together as part of an LLR probe set, and at least one of the component probes would have needed to be redesigned and resynthesized.

To show sequence generality of our tuning approach, we constructed and tuned LLR probe sets to 3 additional DNA targets (Fig. 4abc). Observed r^2 ranged between 0.994 and 0.9993, corresponding to concentration inference standard deviations between $10^{\pm 0.10}$ (21% error) and $10^{\pm 0.03}$ (7% error). The total dynamic range of the probe sets is over 6 orders of magnitude (10 pM to 10 μM), with a log-linear region of 4 logs (100 pM to 1 μM). We were able to quantitate shorter DNA targets (24 nt, minimal to be unique in a genome) as well as longer targets (dT2, 41nt), while maintaining specificity to every nucleotide of longer targets. In

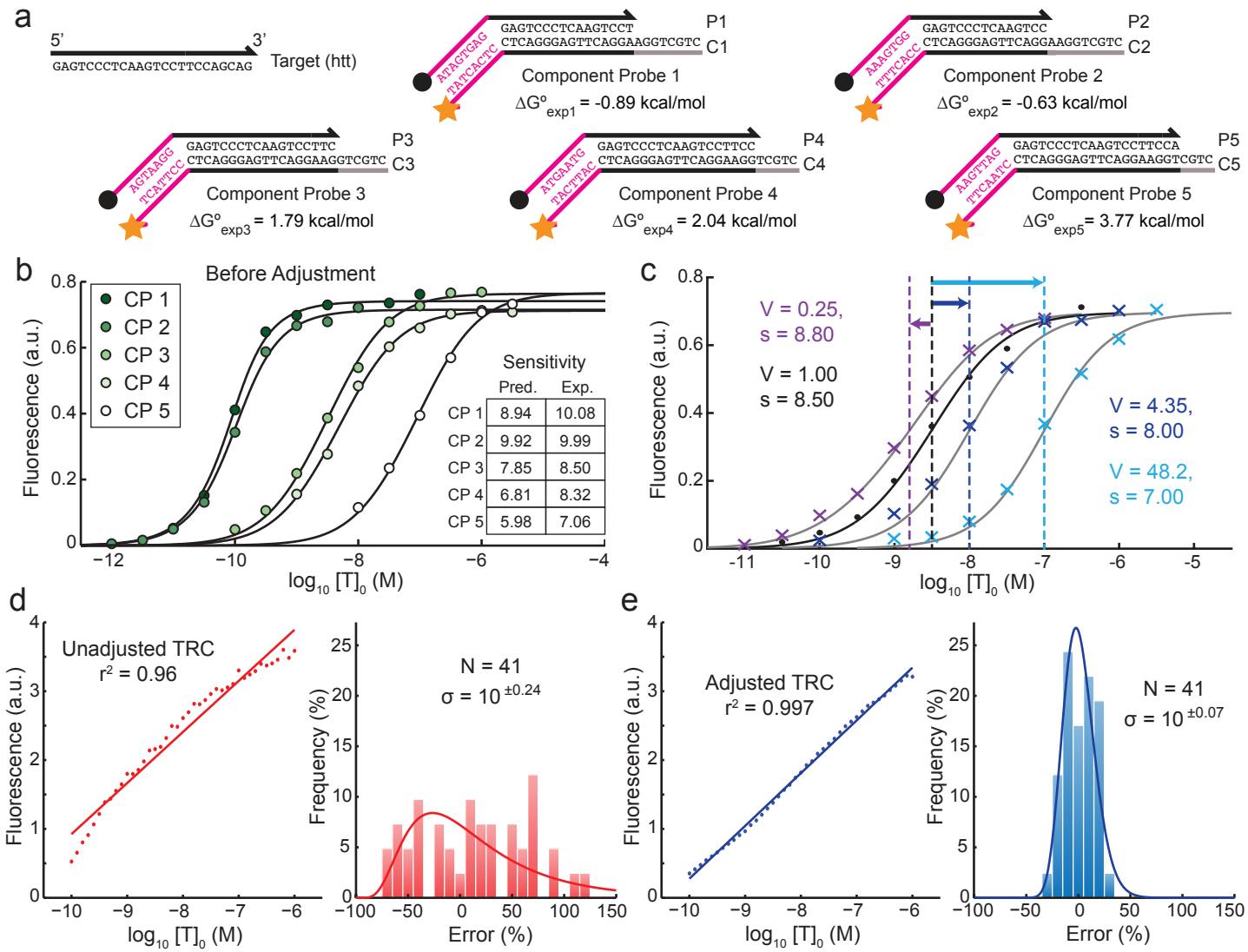

FIG. 3: Stoichiometric tuning of component probes to improve log-linearity of the LLR probe set. **(a)** Target and component probe sequences. Here, the probes are functionalized with fluorophores, and the protectors are functionalized with quenchers, so that binding to target results in an increase in observed fluorescence. Experimental best-fit $\Delta G^{\circ}_{\mathrm{rxn}}$ values are shown. **(b)** Experimental response curves for each component probe. Best-fits (black sigmoids) to experimental data (dots) show up to -2.07 kcal/mol of deviation from designed. All experiments were performed at 25 °C in 1×PBS; $[\mathrm{CP}]_0$ = 100 pM and $[\mathrm{P}]_0/[\mathrm{CP}]_0$ = 1 for all component probes. **(c)** Stoichiometric adjustment of sensitivity s for probe CP3. Consistent with theory, a roughly 10-fold increase of $[\mathrm{P}]_0$ corresponds to a roughly 1 unit decrease of sensitivity. **(d)** TRC of the unadjusted probe set, using $[\mathrm{P}]_0/[\mathrm{CP}]_0$ = 1 for all component probes. From observed fluorescence values, target concentration can be back-calculated given the log-linear best-fit trendline. The histogram to the right shows the concentration error between the back-calculated target concentration and the true target concentration. **(e)** Stoichiometric tuning improves log-linearity and quantitation accuracy of the LLR probe set.

probe sets for longer targets, functionalized probe and protector sequences are correspondingly longer.

Like the component toehold probes [15], LLR probe set maintains high sequence selectivity. Sequences that differ by even a single nucleotide from the target sequence will elicit a much weaker response (Fig. 4d), and variant sequences of equal to or more than 100-fold the amount of target sequences are needed to produce the same signal as the target. Variants different from the target by 2 or more mismatches are suppressed within the whole dynamic range. Such high selectivity of probe sets indicates that the background oligonucleotides in a biological sample are unlikely to cause signal increase. Moreover, in a multiplexed setting, different probe sets are unlikely to interfere with one another. Finally, the performance of the LLR probe set is robust to probe formulation and reaction conditions (Supplementary Section S4).

Surface-based Quantitation. Because the limit of detection of our LLR probe sets are bounded by the sensitivity of readout instrumentation, which in turn is limited by the background autofluorescence of water and the cuvette, we expanded LLR to microarray setting by applying LLR

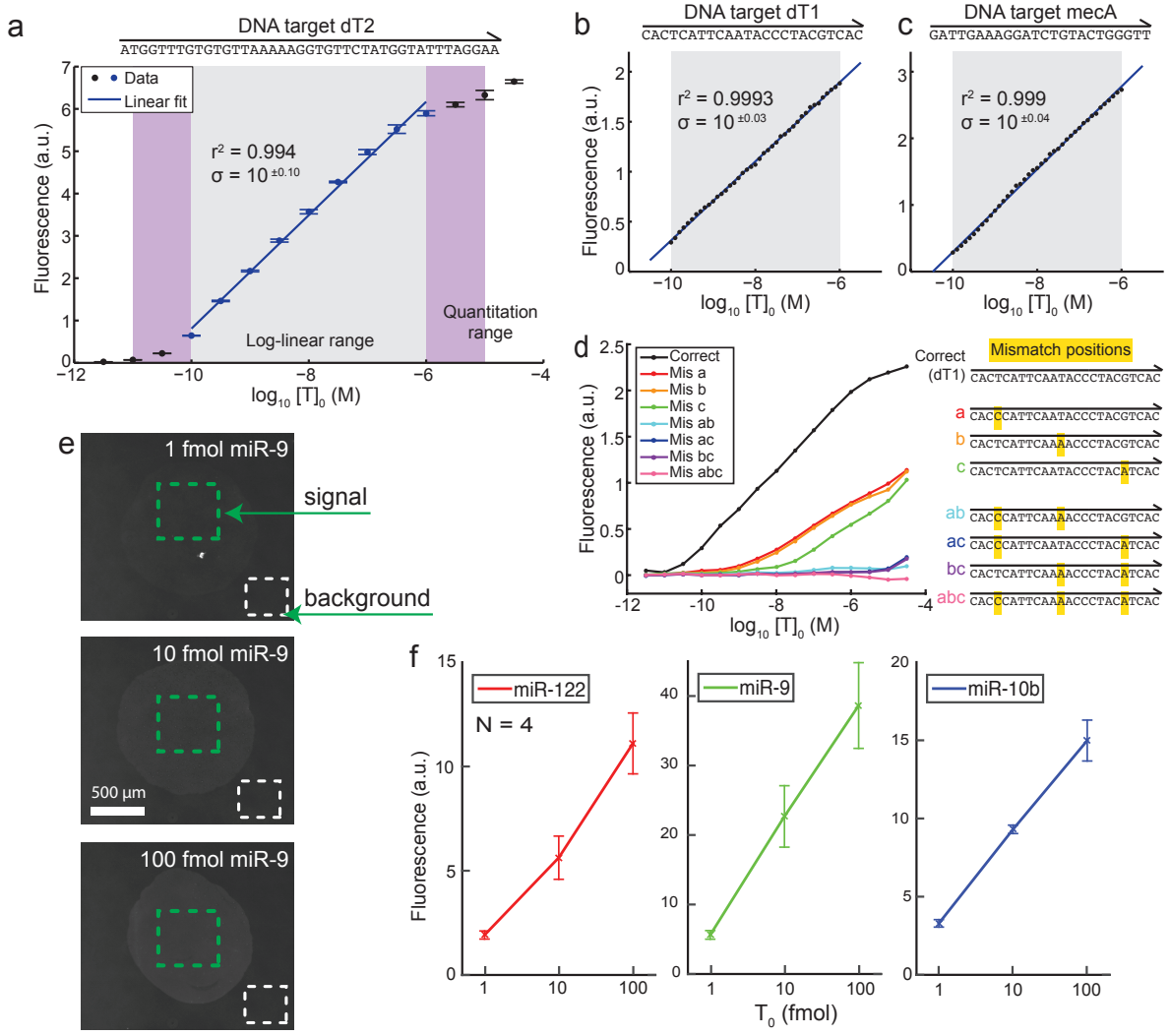

FIG. 4: Observed TRCs for LLR probe sets. **(a, b, c)** Probe sets for each target were designed to possess a log-linear range over 4 orders of magnitude, from 100 pM to 1 μM (gray shaded region). The total quantifiable dynamic range is larger (more than 6 orders of magnitude), from approximately 10 pM to 10 μM. Each error bar represents 1 standard deviation of N = 3 experimental runs. **(d)** Probe sets exhibit high sequence selectivity. Single nucleotide variants of the target (at position a, b, or c) produce total response curves shifted by roughly 2 units of sensitivity. Two- and three-nucleotide variants produce nearly unobservable signal. **(e, f)** Surface-based detection of LLR probe set response. End-labeled target miRNA are hybridized to biotinylated probes, which are then captured to a streptavidin-coated surface and washed to remove unbound miRNA. Observed fluorescence is log-linear in target concentration.

probe sets to surfaces of glass slides. Fig. 4ef shows the results of biotin-functionalized LLR probe sets reacting with fluorescently labeled miRNA targets, deposited onto glass slide surfaces; observed fluorescence is remains proportional to the logarithm of the target concentrations. Our results show a limit of detection of roughly 1 fmol, corresponding to 10 pM analyte in 100 μL hybridization volume. This detection limit in principle can be further improved through better microscopy equipment and contact printing equipment to generate smaller spots.

Multiplexed RNA Quantitation. To demonstrate multiplexed quantitation with LLR probe sets, we designed and tuned 3 LLR probe sets to 3 distinct target sequences. Within each probe set, each component probe bears an identical fluorophore (Alexa488, TAMRA, or Alexa647), but different probe sets produce spectrally distinct signals. Fig. 5b shows the TRC of each LLR probe set against its intended target (in single-plex setting). The miR-9 probe set experimentally exhibited significantly worse log-linearity than we expected; we believed that the secondary structure inherent in miR-9 complicated the kinetics of the reaction so that equilibrium was not rapidly achieved. In general, the sensitivity and $\Delta G^{\circ}_{\mathrm{rxn}}$ of toehold probes to RNA targets deviated from predicted values more than to DNA targets, possibly due to the less complete RNA-DNA hybridization parameters (Supplementary Section S3).

We next applied all three probe sets simultaneously in a multiplexed setting. Three target mixture samples were prepared by mixing known quantities of synthetic RNA molecules bearing the miR-122, miR-9, and miR-10b se-

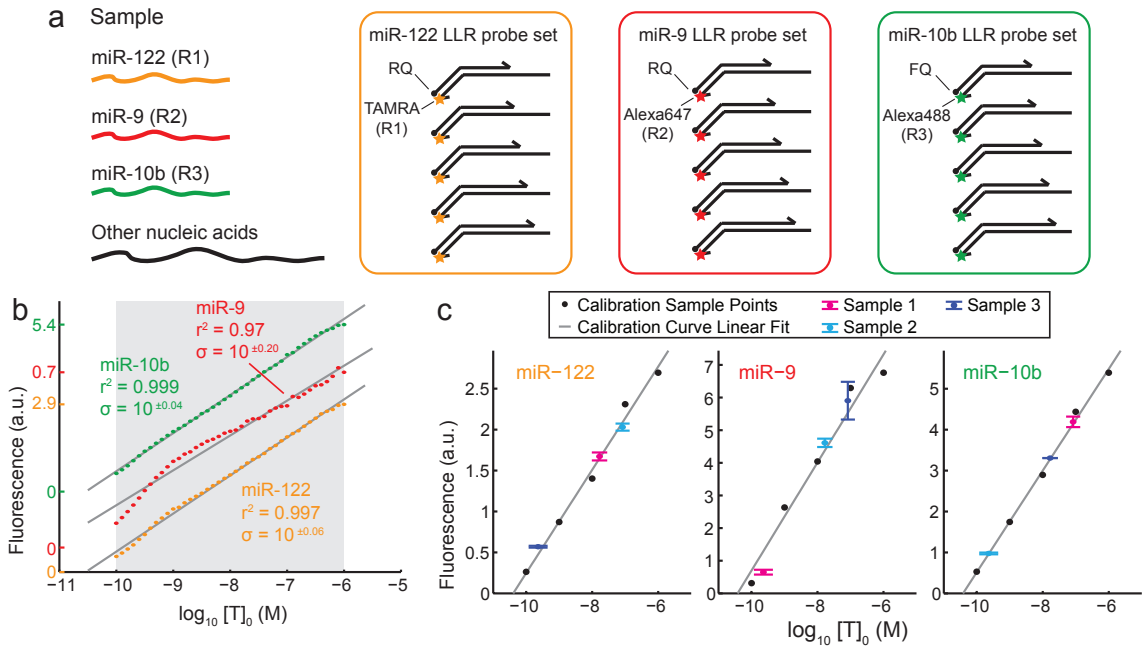

FIG. 5: Multiplexed quantitation of three synthetic RNA targets. **(a)** Three probe sets were designed to provide log-linear responses to three distinct RNA targets, each bearing the sequence of an important microRNA. Each probe set uses a spectrally distinct fluorophore to enable simultaneous quantitation from the same sample. **(b)** TRCs of the three probe sets in single-plex setting. The probe set targeting miR-9 exhibited imperfect log-linearity even after stoichiometric tuning, possibly due to secondary structure in the target. **(c)** Multiplexed quantitation results. Black dots represent calibration data points, and three different samples with distinct combinations of miR-122, miR-09, and miR-10b are displayed in different colors. All reactions were performed in 1x PBS at 25 °C in triplicates. The average positive quantitation error was 25.1% for miR-122, 121.9% for miR-9, and 5.7% for miR-10b. The average negative quantitation error was 19.2% for miR-122, and 59.4% for miR-9.

quences. The observed fluorescence results for each sample are shown in separate colors in Fig. 5c, and plotted against the known input concentrations. At all concentrations tested, presence of one target did not significantly affect the signals in the other two fluorescence channels (Supplementary Section S5). The same experiment was performed in 1×PBS with 20% formamide (Supplementary Section S5).

MicroRNAs miR-122, miR-9, and miR-10b (our targets in Fig. 5) are present biologically at different concentrations in different tissue types [20], and different expression patterns in microRNAs can be used to classify cancer subtypes [21]. To demonstrate our probe sets for multiplexed quantitation in biological samples, we applied our probe sets to human brain and liver tissue total RNA samples (Life Technologies). Fig. 6 shows the inferred target concentrations, using TRCs calibrated based on synthetic RNA of known concentrations (see Supplementary Section S5 for detailed procedures). Our results (Fig. 6) were qualitatively consistent with literature microarray data [13].Using both LLR and microarray methods, miR-122 expression level was observed to be the highest out of the 3 miRNA tested in liver tissue, and miR-9 was the highest in brain tissue; miR-10b has the lowest expression level in both liver and brain (Table 2).

To better understand microarrays as a comparison technology, we constructed and imaged microarrays to quantitate the same three microRNA targets: miR-122, miR-9, and miR-10b (Supplementary Section S6). Calibration results on synthetic miRNA species showed that fluorescence signal saturated over time, but at different rates for differ-

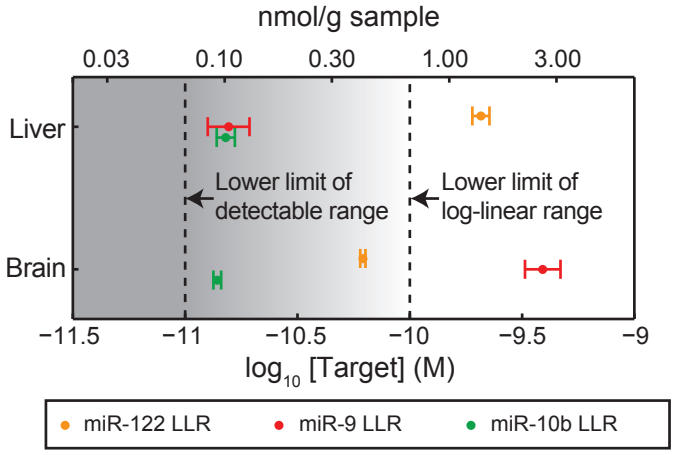

FIG. 6: Multiplexed quantitation of microRNA concentrations in total RNA of human liver and brain, using the 3-plex probe set from Fig. 5. The fluorimeter we used for readout possessed limit of detection of roughly 10^{-11}M (10 pM), so we designed the lower end of our probe sets' log-linear ranges to be roughly 100 pM. Using more sensitive detection modalities will allow higher dynamic range and more accurate quantitation of microRNAs. All data points and error bars represent the means and standard deviations of n = 3 experimental runs.

	Tissue	miR-122	miR-9	miR-10b
LLR (nmol/g sample)	Liver	**1.382**	0.104	0.101
	Brain	0.412	**2.604**	0.093
Microarray (counts)	Liver	**65518**	3504	477
	Brain	447	**42659**	433

TABLE 2: Comparison of miRNA quantitation between LLR method and microarray method. For LLR method, data was shown as nmol miRNA per gram of total RNA sample. Literature values obtained by microarray are shown as raw counts [20]. Highest expression level out of the 3 miRNA tested in the tissue is highlighted in green.

ent miRNA targets. Because surface hybridization was significantly slower than solution hybridization, overnight microarray hybridization did not necessarily reach equilibrium for all targets, reducing quantitation accuracy. Microarray analysis of brain and liver total RNA samples showed clear over-expression of miR-122 in liver and miR-9 in brain, but did not provide accurate quantitation of the low-expressed microRNAs, due to the high-observed background signal. Interestingly, a negative control spot with a probe targeting a sequence not present in the transcriptome showed low fluorescence, indicating that nonspecific fluorescence was sequence dependent.

As another comparison, we performed single-plex RT-qPCR for microRNA quantitation, using the Thermo Fisher Taqman Advanced miRNA Assay (Supplementary Section S7). Although qPCR is typically considered the gold standard for DNA quantitation, the short lengths of microRNAs necessitate extra polymerase- and ligase- based steps to achieve PCR specificity. RT-qPCR results on brain and liver samples again demonstrated overexpression of miR-122 in liver and miR-9 in brain, but reported significantly lower results for the low expression level miRNAs than either LLR probe sets or microarray results (both ours and literature results).

In comparing our LLR probe set results with microarray results and RT-qPCR results, there was relatively higher agreement between LLR probe sets and microarrays. One possible explanation for this is the existence of precursor microRNAs (pre-miRNA) in total RNA samples, which bear the mature miRNA sequence but also contain 5′ and 3′ dangles. These pre-miRNA cannot be ligated to RT-qPCR adaptors, and are thus not reflected in RT-qPCR results. Because all current methods for miRNA quantitation may suffer from different sources of bias, there is currently no definitive method for absolute quantitation of miRNA. Continued technology development may result in additional independent measurements that can obtain consensus values for the scientific community to use.

Discussion

In this work, we introduced, analyzed, and validated hybridization probe sets with log-linear response (LLR), and showed that they can be used for accurate quantitation of nucleic acid concentrations across many orders of magnitude. Extremely high log-linearity is necessary for accurate quantitation, which in turn requires component probes to be adjusted to precise sensitivities. Because component probes used the toehold probe architecture, we could use formulation stoichiometry for decoupled modulation of probe sensitivity. In the absence of stoichiometric tuning, and based on our observed deviations between designed and experimentally observed probe $\Delta G^{\circ}_{\mathrm{rxn}}$ (up to 2 kcal/mol), we expect that empirical sequence optimization to adjust probe sensitivity would be quite tedious, especially in multiplexed settings. Finally, due to the high sequence selectivity of the LLR probe sets, multiplexed quantitation of nucleic acid sequences is possible, including in biological (total RNA) samples.

The current log-linear probe sets based on fluorophore readout have limitations in sensitivity, long incubation time, limited multiplexing, calibration curve-dependent and sensitive to RNA secondary structure. As the goal of this manuscript was to introduce and demonstrate quantitation and tuning capabilities of LLR probe sets, we did not optimize for molecular sensitivity, and our fluorimeter photomultiplier tube constrained our limit of detection to roughly 10 pM. Using readout equipment of higher sensitivity (such as fluorescence microscopes [22], electrochemical sensors [11, 12]) could result in a significant improvement in the limit of detection as well as the log-linear dynamic range. Alternatively, using signal amplification methods such as coupled horseradish peroxidase [23] or beta galactosidase [24] moieties in place of fluorophores for immobilized targets can also improve sensitivity. Our quantitation results in biological total RNA samples are qualitatively consistent with literature microarray results; quantitatively, there is some difference. We believe this is due to the change in secondary structure and nonspecific RNA-RNA interactions in different hybridization conditions.

Quantitation of nucleic acids, especially RNAs, may provide important scientific understanding of diseases and guide clinical diagnoses; for example, several panels [25] rely on the relative expression levels of tens of RNA for cancer recurrence prediction. The range of concentrations that biological RNAs exhibit spans 5 or more orders of magnitude due to tissue- or disease-specific expression differences, cell counts, and sample volumes. Current clinical diagnostics relying on RT-PCR require polymerase-based nucleic acid amplification. Amplification-free quantitation of RNA, in principle, offer at least 3 distinct advantages over RT-PCR: (1) improved performance in multiplexed settings, due to reduced primer dimers and amplification bias, (2) lower risk of amplicon contamination, compared to nested PCR which requires open tube steps, and (3) higher suitability for point of care applications because precise temperature cycling equipment is not needed.

RNA sequencing (RNAseq) is another popular method for massive multiplexed profiling of RNA concentrations [26]. When sequencing a raw sample, many reads are wasted during redundant sequencing of high concentration species (such as ribosomal RNA and globin mRNA) that dominate the sample (typically, 90+%). RNAseq users typically deplete

their total RNA samples of these high concentration species via kits such as RiboZero [27], but this depletion step introduces biases due to nonspecific depletion of other species, and does not allow efficient sequencing of rare RNA species either. The authors envision that adapting LLR probe sets to RNAseq could allow efficient coverage of both rare and abundant RNA species for comprehensive gene expression profiling.

In this work, we did not apply LLR probe sets to quantitating mRNA or dsDNA, because the relatively low abundance of a specific mRNA or dsDNA sequence in biological samples necessitates further improvement in readout instrumentation for improved limits of detection. Additionally, mRNA sequences can be thousands of nucleotides long, and exhibit significant secondary structures that interfere with our hybridization-based LLR approach; dsDNA cannot directly hybridize to probes. The authors envision that RNA and dsDNA samples could be fragmented before use, either through incubation in Magnesium buffer or through sonication. dsDNA samples require an additional denature step; due to the low concentration and high sequence complexity of biological samples, the two complementary DNA strands will not hybridize back after denaturation. For historical clinical samples that are formalin-fixed paraffin-embedded (FFPE), extracted DNA and RNA would have already been fragmented. In addition to fragmentation, we may use higher temperature (40-65 °C) and/or chemicals to reduce secondary structure (e.g. betaine and DMSO), paired with hybridization probes with stronger binding energy.

In addition to quantitation of nucleic acid sequences, our log-linear probe sets could also be potentially applied to quantitating proteins or metabolites, by using conjugated antibodies or aptamers on nucleic acid oligos. For example, aptamer- and antibody-coupled primers have been used in a PCR setting to quantitate protein expression [28–30]; toehold probes could be similarly adapted. Combination assays involving simultaneous quantitation of nucleic acids, proteins, and metabolites may emerge as a powerful new class of diagnostics [31] to inform clinical decision.

Acknowledgements. The authors thank J. Sherry Wang and Jin H. Bae for discussions. The authors thank Jianyi Nie and Peng Dai for proofreading assistance. This work was funded by RP180147 (CPRIT) and R01HG008752 (NHGRI) to DYZ.

Author contributions. LRW, JZF, and DYZ conceived the project. LRW, JZF, and DYZ performed theoretical analysis. LRW, JZF, and DK performed experiments and data analysis. LRW, JZF, and DYZ wrote the manuscript. LRW and JZF contributed equally to this work.

Supporting Information. Supporting Information Available: The following files are available free of charge. PDF file 'LLR_SI.pdf', containing all supporting text and figures.

Additional information. LRW consults for Nuprobe Global. DYZ consults for and holds significant equity in Nuprobe Global, hold significant equity in Torus Biosystems, and consults for Avenge Bio. Correspondence may be addressed to DYZ (dyz1@rice.edu).

References

[1] Lockhart, D. J., & Winzeler, E. A. Genomics, gene expression and DNA arrays. *Nature* **405,** 827-836 (2000).

[2] Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G.R., Wu, G.R., Matthews, L. and Lewis, S., Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* **33,** D428-D432 (2005).

[3] Frampton, G.M., Fichtenholtz, A., Otto, G.A., Wang, K., Downing, S.R., He, J., Schnall-Levin, M., White, J., Sanford, E.M., An, P. and Sun, J., Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* **31,** 1023-1031 (2013).

[4] Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C. and Kitano, T.K. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Analytical chemistry,* **83,**8604-8610 (2011).

[5] Craw, Pascal, and Wamadeva Balachandran. Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review. *Lab on a Chip,* **12** 2469-2486 (2012).

[6] Huang, H., Bai, W., Dong, C., Guo, R. and Liu, Z. An ultrasensitive electrochemical DNA biosensor based on graphene/Au nanorod/polythionine for human papillomavirus DNA detection. *Biosensors and Bioelectronics*, **68,** 442-446 (2015).

[7] Liu, A., Wang, K., Weng, S., Lei, Y., Lin, L., Chen, W., Lin, X. and Chen, Y., Development of electrochemical DNA biosensors. *TrAC Trends in Analytical Chemistry*, **37,** 101-111 (2012).

[8] Toren, P., Ozgur, E., & Bayindir, M. Oligonucleotide-based label-free detection with optical microresonators: strategies and challenges. *Lab on a Chip*, **16,** 2572-2595 (2016).

[9] Toren, P., Ozgur, E., & Bayindir, M. Real-time and selective detection of single nucleotide DNA mutations using surface engineered microtoroids. *Analytical chemistry*, **87,** 10920-10926 (2015).

[10] Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3,** 1101-1108 (2008).

[11] Soleymani, L., Fang, Z., Sun, X., Yang, H., Taft, B. J., Sargent, E. H., & Kelley, S. O. Nanostructuring of Patterned Microelectrodes To Enhance the Sensitivity of Electrochemical Nucleic Acids Detection. *Angew. Chem.*, **121(45)**, 8609-8612 (2009).

[12] Di Li, Song, S., & Fan, C. Target-Responsive Structural Switching for Nucleic Acid-Based Sensors. *Accounts of Chemical Research*, **43(5)**, 631-641 (2010).

[13] Barad, O., Meiri, E., Avniel, A., Aharonov, R., Barzilai, A., Bentwich, I., Einav, U., Gilad, S., Hurban, P., Karov, Y. and Lobenhofer, E.K., MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.* **14,** 2486-2494 (2004).

[14] Nelson, P. T., Baldwin, D. A., Scearce, L. M., Oberholtzer, J. C., Tobias, J. W., & Mourelatos, Z. Microarray-based, high-throughput gene expression profiling of microRNAs. *Nature Methods*, **1(2)**, 155-161 (2004).

[15] Zhang, D.Y., Chen, S. X. & Yin, P. Thermodynamic optimization of nucleic acid hybridization specificity. *Nat. Chem.* **4,** 208-214 (2012).

[16] SantaLucia, J. & Hicks, D. The Thermodynamics of DNA Structural Motifs. *Ann. Rev. Biochem.* **33,** 415-440 (2004).

[17] Wu, L.R., Wang, J.S., Fang, J.Z., Evans, E.R., Pinto, A., Pekker, I., Boykin, R., Ngouenet, C., Webster, P.J., Beechem, J. and Zhang, D.Y. Continuously tunable nucleic acid hybridization probes. *Nat. Meth.* **12,** 1191-1996 (2015).

[18] Tyagi, S. & Kramer, F.R. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14,** 303-308 (1996).

[19] Vallee-Belisle, A., Ricci, F., & Plaxco, K. W. Engineering biosensors with extended, narrowed, or arbitrarily edited dynamic range. *J. Am. Chem. Soc.* **134,** 2876-2879 (2012).

[20] Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136,** 215-233 (2009).

[21] Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A. and Downing, J.R., MicroRNA expression profiles classify human cancers. *Nature* **435,** 834-838 (2005).

[22] Levsky, J. M., & Singer, R. H. Fluorescence in situ hybridization: past, present and future. *J. Cell Sci.* **116,** 2833-2838 (2003).

[23] Schonhuber, W., Fuchs, B., Juretschko, S., & Amann, R. Improved sensitivity of whole-cell hybridization by the combination of horseradish peroxidase-labeled oligonucleotides and tyramide signal amplification. *Appl. Environ. Microbiol.* **63,** 3268-3273 (1997).

[24] Warren, A. D., Gaylord, S. T., Ngan, K. C., Dumont Milutinovic, M., Kwong, G. A., Bhatia, S. N., & Walt, D. R. Disease detection by ultrasensitive quantification of microdosed synthetic urinary biomarkers. *Journal of the American Chemical Society* **136,** 13709-13714 (2014).

[25] Cronin, M., Sangli, C., Liu, M.L., Pho, M., Dutta, D., Nguyen, A.,

Jeong, J., Wu, J., Langone, K.C. and Watson, D., Analytical validation of the Oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer. *Clinical chemistry* **53,** 1084-1091 (2007).

[26] Ozsolak, F., & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics,* **12,** 87-98 (2011).

[27] Haas, B. J., Chin, M., Nusbaum, C., Birren, B. W., & Livny, J. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC genomics,* **13,** 734 (2012).

[28] Sano, T., Smith, C. L., & Cantor, C. R. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science,* **258,** 120-122 (1992).

[29] Schweitzer, B., Wiltshire, S., Lambert, J., O'Malley, S., Kukanskis, K., Zhu, Z., Kingsmore, S.F., Lizardi, P.M. and Ward, D.C., Immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proceedings of the National Academy of Sciences,* **97,** 10113-10119 (2000).

[30] Pinto, A., Redondo, M. C. B., Ozalp, V. C., & O'Sullivan, C. K. Real-time apta-PCR for 20000-fold improvement in detection limit. *Molecular bioSystems,* **5,** 548-553 (2009).

[31] Ullal, A. V., Peterson, V., Agasti, S. S., Tuang, S., Juric, D., Castro, C. M., & Weissleder, R. Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. *Science translational medicine,* **6,** 219ra9 (2014).

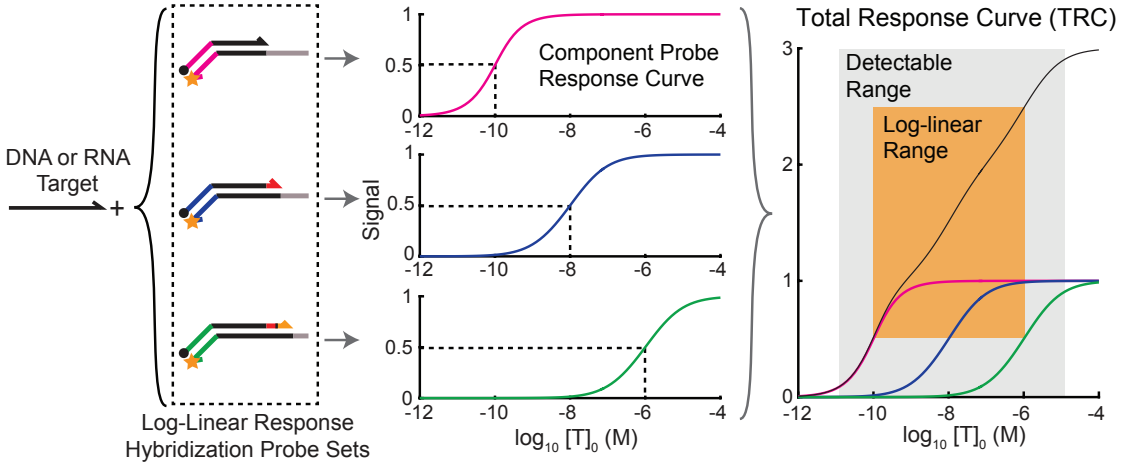

FIG. 0: For TOC only