Optimization of logic gates for one-step detection of microRNAs via Split loop-mediated isothermal amplification (Split-LAMP)

CB018

Medha Shridharan

INTRODUCTION

microRNAs (miRNAs) are an abundant class of small endogenous, nonprotein-coding RNAs which regulate gene expression with critical functions in many pathologic and physiologic processes such as carcinogenesis and thus, their expression can be of high diagnostic value.

To develop point-of-care diagnostics for miRNA detection, loop-mediated isothermal amplification (LAMP) is ideal. Traditional LAMP uses strand displacement to generate single stranded loops for Forward Internal Primer (FIP) and Backward Internal Primer (BIP) binding and subsequent exponential amplification (Fig. 1 (appendix)), comprising an F2/B2 sequence concatemerized with the reverse complement of the F1/B1 sequence. A loop is formed when F1/B1 binds to the F1c/B1c sequence. F3/B3 primers then displace the FIP/BIP amplified strand from the double-stranded DNA template to create the loop structure. FIP/BIP binding to the loop structure results in branching of the DNA strand to form further loops and allow exponential amplification to occur. LF and LB primers are sometimes added to increase branching from the reverse complement loop, accelerating the LAMP reaction. A key limitation of current miRNA LAMP designs is that only one miRNA is used as input, but single miRNAs are seldom diagnostically predictive. However, if the LAMP reaction could depend on the input of a miRNA panel using an ANDgate approach, the miRNA signals can be integrated into a single output for miRNA panel point-of-care testing (POCT).

Our lab is developing a novel split-LAMP technique. BIP and FIP are split into their component primers to detect B2 and F2 concentrations, as opposed to the linear template analyte in traditional LAMP. The B2/F2 primers strand invade to generate single-stranded regions for subsequent elongation by the B1c/F1c primers (Fig. 2 (appendix)). As this system is dependent on the presence of both B2 and F2, substituting B2 and F2 with miRNAs functionally creates an ANDgate. As B2 and F2 generate the displaced strand that is then discarded, it does not matter whether these are DNA or RNA as both will result in dead-end amplification rather than a loop.

In preliminary experiments shown in Fig. 3 (appendix), our lab showed that B1c and B2 can replace BIP in a LAMP reaction and that exponential increases in [F2] and [B2] affect amplification time. However, as primer interactions in the LAMP system are complex and non-

intuitive, the primer concentrations that allow good discrimination of F2 and B2 concentrations in combination must be determined experimentally. Herein, we identify ideal primer concentrations in the Split-LAMP, demonstrating proof-of-concept for a miRNA Split-LAMP system for POCT and elucidating the nature of primer interactions of Split-LAMP.

METHODS

Materials and apparatus. Bst DNA polymerase large fragment and dNTPs were purchased from New England Biolabs (USA). EvaGreen dye is purchased from Biotium. All nucleic acids, including ultramer (template) and primers, were synthesised by Integrated DNA Technologies, Inc. The sequences of all nucleic acids are listed in Table 1 (Appendix). Buffer replacement solution (BRS) was made inhouse.

Experimental procedures. Primers were added to reaction mixtures with BRS and EvaGreen dye. The LAMP reaction mixture was then added into a 96-well plate which was maintained at 65°C for 1h, heated at a rate of 1.6°C/s. Fluorescence signals were collected every 1min with a QuantStudio™ 5 real-time quantitative PCR (qPCR) instrument. The resulting data was analysed to generate Ct (quantification cycle) values for each amplification.

Risk assessment. Experimental procedures were minimal risk and carried out in a BSL-0 laboratory. Standard laboratory personal protective equipment was required. No hazardous biological agents or hazardous chemicals were used.

Data analysis. Well-defined real-time amplification plots are obtained from different primer concentrations. The POI (point of inflection) value, defined as the time corresponding to the maximum slope of the fluorescence curve, is used to quantitatively determine [B2] and [F2]. With increasing [B2] and [F2], the corresponding POI is gradually shortened, indicating a stronger signal from the system. Standard curves are obtained by plotting POI values against the logarithm (lg) of [B2] and [F2] and the linear relationship is obtained.

RESULTS

[Ultramer] = 53pg/ml.

As there is greater resolution between curves at 53pg/ml, it is concluded that this is the ideal concentration of ultramer for use in split-LAMP (Fig. 4 (appendix)).

[LB] = 100nM or [LB] = 10nM.

The POI values in Fig 5 (appendix). exhibit an excellent linear relationship with the lg of [B2] ranging from 160nM to 160pM. Although amplification time resolution between 160nM and 1.6nM curves increases at low LB concentrations, [B2] = 1.6nM and [B2] = 0.16nM cannot be differentiated. At [LB] = 10nM or 100nM, smaller [B2] such as 1.6nM and 160pM can effectively be differentiated if the reaction is run for 60 minutes (Fig. 5 (appendix)). We conclude that both [LB] = 10nM and [LB] = 100nM are feasible for use in this system. [LB] = 100nM is chosen for further experimentation.

[LF] = 0nM.

From Fig. 6 (appendix), the POI values exhibit an excellent linear relationship with the lg of F2 concentrations ranging from 160nM to 160pM. As all standard curves have similar gradients and correlation coefficients, it is concluded that the presence of LF does not make a large differenece to the Split-LAMP system and [LF] should be maintained at 0nM.

[LF] = 0nM does not significantly affect the dynamic range of B2.

From the standard plot in Fig. 7 (appendix), we observe that the POI values maintain an excellent linear relationship in both the presence and absence of LF. The fluorescence curves also demonstrate that the POI and dynamic range for B2 with and without LF are very similar.
The triplicate of experiments with [LF] = 100nM (M = 28.2, SD = 2.76) compared to the triplicate with [LF] = 0nM (M = 30.7, SD = 6.73) demonstrated no significant change in dynamic range between the curves of [B2]=160nM to [B2]=0.16nM, $t(4) = 0.5919$, $p = 0.5857$.
Hence, we conclude that removing LF from the Split-LAMP system makes no significant change to the dynamic range of B2.

[B1C] = 100nM.

From Fig. 8 (appendix), [B1c]=160nM, which can be approximated to 100nM, has the highest correlation coefficient of 0.9353 and hence the POI values are most consistent on the logarithmic scale. It has a large dynamic range of 19.21 cycles and large gradient of 7.017 as compared to [B1c] = 1600nM. Hence, [B1c]=100nM is ideal

[F1C] = 100nM.

From Fig. 9 (appendix), [F1c]=160nM, which can be approximated to 100nM, has the highest correlation coefficient of 0.9658. Additionally, it has a large dynamic range of 17.44 cycles and hence, [F1c]=100nM is ideal.

ANDgate is dependent on [F2] and [B2].

From Fig. 10 (appendix), we observe that the Split-LAMP reactions are generally dependent on [B2] and [F2] on in the range of 160nM to 1.6nM, which can be approximated as an effective dynamic range of 1-200nM for B2 and F2. If a cutoff Ct value is set at an appropriate time to amplification (e.g. Ct = 8 for 16nM B2/F2), amplification before the cutoff marks an ANDgate readout in which there be a positive signal only if both [B2] and [F2] are above 80nM.

Fig. 11 (appendix) illustrates proof-of-concept that F2 and B2 can be used in combination to affect amplification time and hence the ANDgate is dependent on [F2] and [B2]. It is worth noting that the shift in POI only appears significantly on a logarithmic scale, making it difficult to distinguish 2-fold or 3-fold increases in the concentration of B2 or F2. This is primarily due to the fact the LAMP detection method is based on exponential amplification.

Specificity of Split-LAMP system.

In order to determine the specificity of Split-LAMP amplification, the system was tested with nonspecific background RNAs. We observe from Fig. 12 (appendix) that amplification of B2 still occurs in the presence of background RNAs. Although there is some background signal, samples containing RNA can be clearly visually distinguished from those containing the B2 target. The reaction mixture with only B2 primer had the fastest amplification with POI = 7.205. Hence, the

split-LAMP system is highly specific and no false positives occur. It is worth noting that the background signal can be attributed to F2, which is added in excess in the reactions.

Bayesian Optimisation model introduces new points in the solution space.

Bayesian optimization is a powerful methodology which minimizes unknown objective functions which are expensive to evaluate. It builds a Gaussian process regression model for the objective and quantifies its uncertainty. After each experiment, the model is updated, and an acquisition function which incorporates both data from prior experiments and model uncertainty is used to decide which point in the solution space to evaluate next. In this work, to accelerate the process of identifying ideal primer concentrations the to maximise the dynamic range of F2 serial dilutions, we implemented a Bayesian Optimization model which received primer concentration data (i.e. the R^2*gradient of standard plots) from prior experiments as input. The model then selected points to evaluate which we would not have explored using domain knowledge alone, increasing the likelihood of finding points which outperform expectations. The details of this model are discussed further in the appendix. The sets of primer concentrations are listed in Table 2 (appendix) and the results of the first iteration are illustrated in Fig. 13 (appendix).

We conclude that the set of primer concentrations in Test 5 where [B2] = 75nM, [B1c] = 100nM, [LB] = 70nM, [F1c] = 25nM and [LF]=125nM outperforms expectations. The POI values in Fig. 13 maintain an excellent linear relationship with the highest correlation coefficient of 0.9546 and the highest gradient and dynamic range of 6.6392 and 18.565 cycles respectively.

DISCUSSION

LF in excess impedes strand invasion by F2.

Loop primers (LF and LB) act as a source of loop opening independent of F2/B2, allowing for amplification by F1C/B1C respectively. Although the competitive effect of LB and B2 is not obvious in the split-LAMP system, the addition of LF in F2 serial dilutions appeared to make little to no difference. Hence, controlling the concentration of LB/LF is key in the Split-LAMP system.

Resolution of F2 curves decreases with decreasing [B2] and vice versa.

At low [B2], it is observed that the resolution of F2 serial dilution curves largely decreases. A similar effect on B2 serial dilution curves at low [F2]. This is likely due to the fact that at low [B2], amplification becomes more dependent on stochastic strand invasion and loop opening by B1c.

[B2] is inconsequential at high [B1c].

In excess B1c, serial dilution curves of B2 in the range of 160nM to 160pM exhibit low resolution. This is likely due to the stochastic nature of B1C strand invasion. At low [B1C], strand invasion of the B1C primer is more dependent on the prior invasion of B2 primer. However, at high [B1C], B1c strand invasion and elongation occurs more often as the loop is more likely to open sporadically. Hence, B1c strand invasion becomes independent of B2 primer.

Nonspecific DNA synthesis may occur in Split-LAMP.

It is worth noting that theoretically the fluorescence curve of the blank (160pM) will be a straight line. The blank signal arises gradually, which may originate from the unprimed but templated nonspecific DNA synthesis.

Limitations of Bayesian Optimization model.

Given sufficient iterations, the model will very likely introduce many new points successfully. However, the model is also limited by assumptions. Firstly, Bayesian Optimization assumes that points close together on the Gaussian Process will have similar output values as it assumes a continuous function. It is impossible to verify this assumption as the actual value of the function is unknown. Additionally, the current metric used as input in the model is R^2*gradient of the standard curve generated for each set of primer concentrations. This metric currently does not take the mean POI of each set of primer concentrations into account. If the mean POI is too high or too low, this could hinder the effectiveness of the model for practical applications.

CONCLUSION

In this work, we demonstrated a proof-of-concept for diagnostic miRNA quantification via Split-LAMP. The optimal primer concentrations to produce a signal (measured in amplification time) responsive to a 100-fold change in concentrations of the two miRNAs were empirically determined. From these concentration parameters, we approximated the effective dynamic range of Split-LAMP and identified a practical cutoff value such that a positive output is observed only if both [B2] and [F2] are above a threshold concentration. We found that split-LAMP is highly specific with low risk of false positives. Additionally, we developed a Bayesian Optimization model to introduce new points in the solution space to optimize the F2 dynamic range, with one primer concentration ratio suggested by the model outperforming expectations. The next step for this work would be to test RNA B2 and F2 primers to demonstrate the effectiveness of the miRNA quantification system. In the future, the optimized split-LAMP technique may be significant in POCTs for cancer diagnosis.

BIBLIOGRAPHY

[1] Abdullah AL-maskri, A. A., Ye, J., Talap, J., Hu, H., Sun, L., Yu, L., Cai, S., & Zeng, S. (2020). Reverse transcription-based loop-mediated isothermal amplification strategy for real-time miRNA detection with phosphorothioated probes. *Analytica Chimica Acta*, *1126*, 1–6. https://doi.org/10.1016/j.aca.2020.06.007

[2] Biolabs, N. E. (n.d.). *Loop-mediated isothermal amplification*. NEB. Retrieved November 17, 2022, from https://www.neb.sg/applications/dna-amplification-pcr-and-qpcr/isothermal-amplification/loop-mediated-isothermal-amplification-lamp

[3] Deng, R., Zhang, K., & Li, J. (2017). Isothermal amplification for microRNA detection: From the test tube to the cell. *Accounts of Chemical Research*, *50*(4), 1059–1068. https://doi.org/10.1021/acs.accounts.7b00040

[4] Li, C., Li, Z., Jia, H., & Yan, J. (2011). One-step ultrasensitive detection of microRNAs with loop-mediated isothermal amplification (LAMP). *Chem. Commun.*, *47*(9), 2595–2597. https://doi.org/10.1039/c0cc03957h

[5] Sun, Y., Tian, H., Liu, C., Sun, Y., & Li, Z. (2017). One-step detection of microRNA with high sensitivity and specificity via target-triggered loop-mediated isothermal amplification (TT-Lamp). *Chemical Communications, 53*(80), 11040–11043. https://doi.org/10.1039/c7cc06140d

[6] Rasmussen, C. E. Gaussian processes for machine learning. MIT Press, 2006.

APPENDIX

Figures referenced in report.

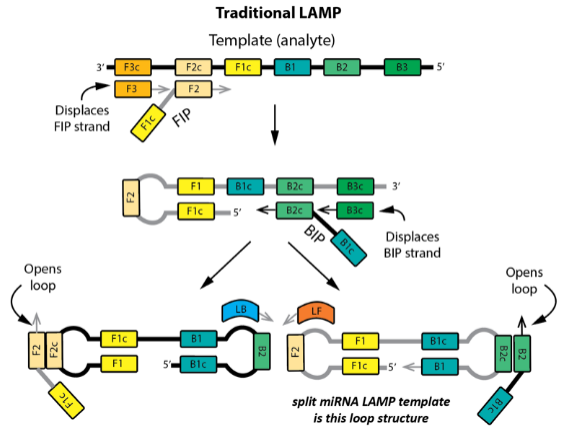

Fig. 1 Strand displacement by primers in traditional LAMP.

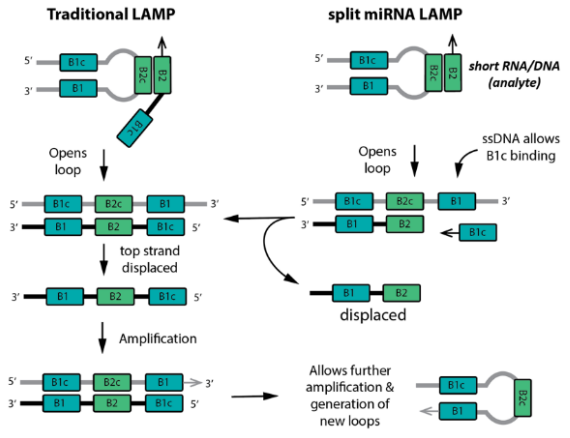

Fig. 2 In split-LAMP, BIP and FIP are aplit into their component primers. This results in more strand invasion and amplification.

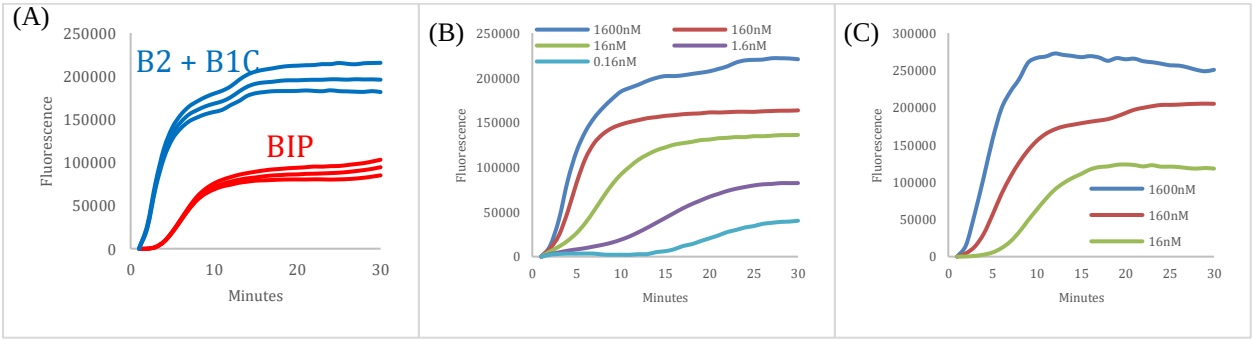

Fig. 3 (A) A standard 20uL LAMP reaction was conducted with [BIP] = 1.6uM (red). BIP was replaced by components where [B2]=[B1C]=1.6uM (blue). [FIP]=[BIP] =1.6uM, [LF]=[LB]=0.4uM, [Ultramer] = 5.3ng/ml. Adding B1c and

B2 as separate oligonucleotides rather than a single BIP primer results in faster amplification. Each experiment was conducted in triplicate. (B) [FIP] = 1.6uM, [LF] = 0.4uM, LB = 0.1uM and [B1c] = 0.1uM, and the stated concentrations of B2 were used. Exponential increase in [B2] affects amplification time. (C) [BIP] = 1.6uM, [LB] = 0.4uM, [LF] = 0.1uM and [F1c] = 0.1uM, and the stated concentrations of F2 were used. Exponential increase in [F2] had similar effect.

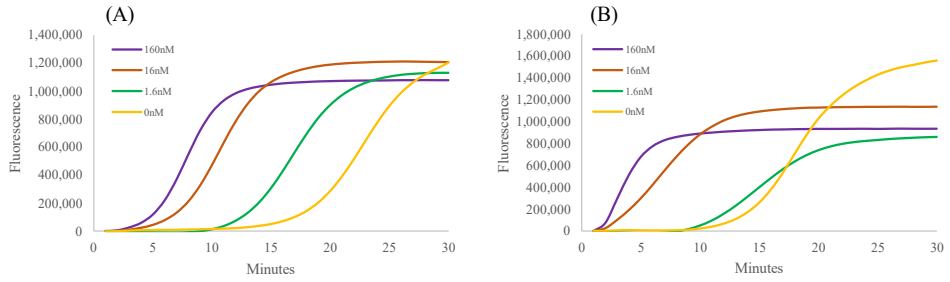

Fig. 4 Real-time flourescence from varying B2 concentration in the range of (from left to right) 160nM to 0nM with (A) 53pg/ml and (B) 530pg/ml respectively. [F1C] = 0.1uM, [F2] = 1.6uM, [B1C] = 0.1uM, [LF] = 0.1uM [LB] 0.1uM.

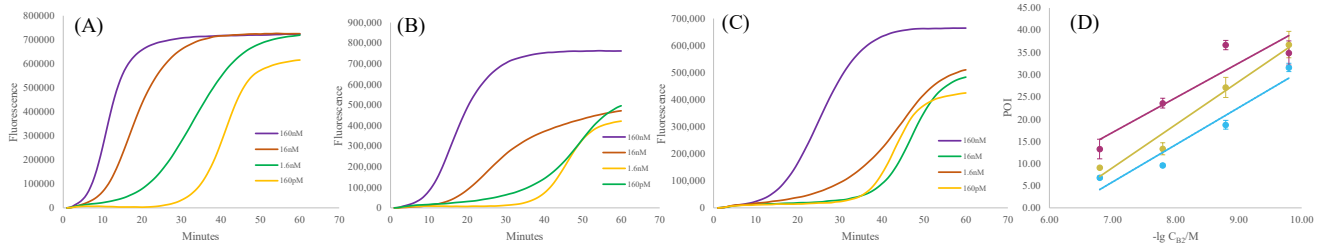

Fig. 5 (A) to (C) represent real-time flourescence curves aroused by varying [B2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [LB] was (A) 100nM, (B) 10nM, (C) 0nM respectively. [Ultramer] = 53pg, [B1C]=[F2]=[F1C]=[LF]=0.1uM. Each experiment was conducted in triplicate. (D) Standard curves for B2 serial dilutions at various [LB]. (100nM) POI = -52.689 – 8.3627 lg [B2](M), R^2 = 0.9314**,** (10nM) POI = -58.717 – 9.6806lg [B2](M), R^2 = 0.9656, (0nM) POI = -37.554 – 7.7939 lg [B2](M), R^2 = 0.8531.

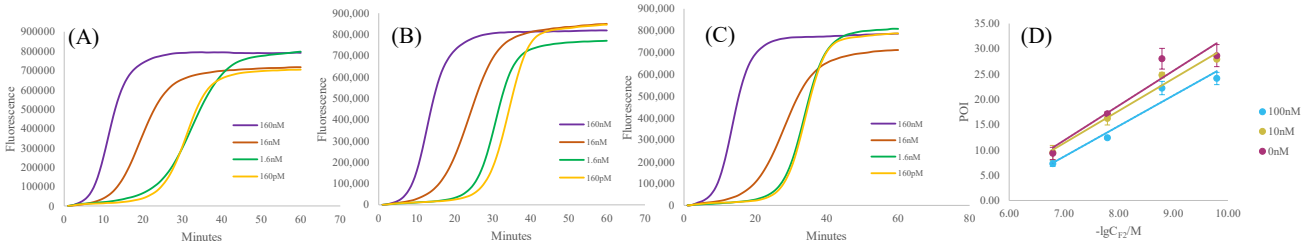

Fig. 6 (A) to (C) represent real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [LF] was (A) 100nM, (B) 10nM, (C) 0nM respectively. [Ultramer] = 53pg, [B1C]=[B2]=[F1C]=[LB]=0.1uM. Each experiment was conducted in triplicate. (D) Standard curves for B2 serial dilutions at various [LB]. (100nM) POI = -33.539 - 6.0344 lg [F2](M), R^2 = 0.9452, (10nM) POI = -37.524 - 6.9435 lg [F2](M), R^2 = 0.9908, (0nM) POI = -36.34 – 6.8859 lg [F2](M), R^2 = 0.9158.

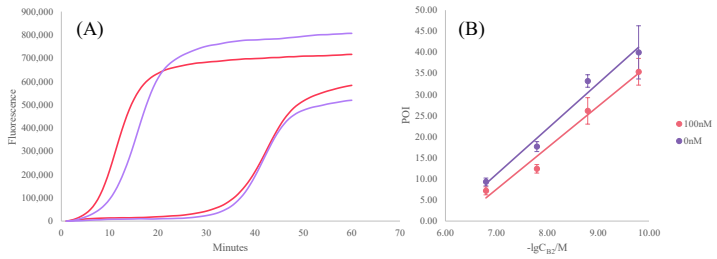

Fig 7. (A) Real-time fluorescence curves aroused by varying [B2] to 160nM (left) and 0.16nM (right). LF concentrations are 100nM (red) and 0nM (purple). [F1C] = [B1C] = [F2] = [LB] = 0.1uM, [Ultramer] = 5.3ng/ml. Each experiment was conducted in triplicate. (B) Standard curves for B2 serial dilutions at stated [LF]. (100nM)=POI = -61.314 + 9.8342 lg C_{B2} (M), R^2 = 0.9748, (0nM) POI = POI = -64.249 + 10.761 lg C_{B2} (M), R^2 = 0.9776.

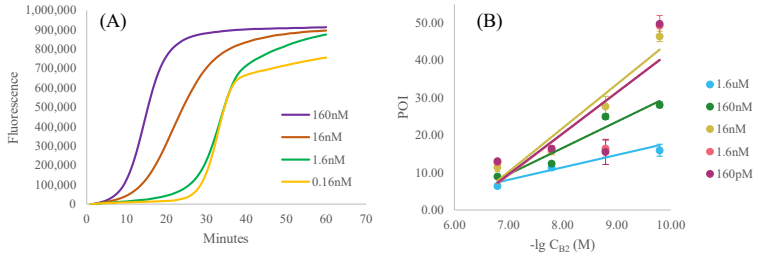

Fig. 8 (A) represents real-time flourescence curves aroused by varying [B2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [B1c] = 160nM, [Ultramer] = 53pg, [B1C]=[B2]=[LB]=0.1uM, [LF]=0nM. Each experiment was conducted in triplicate. (B) Standard curves for B2 serial dilutions at various [B1c]. (1.6uM) POI = -15.105 – 3.3138 lg [B2](M), R^2 = 0.8933, (160nM) POI = -39.603 – 7.0174lg [B2](M), R^2 = 0.9353, (16nM) POI = -71.896 – 11.719 lg [B2](M), R^2 = 0.9295, (1.6nM) POI = -67.881 – 11.03 lg [B2](M), R^2 = 0.6828, (160pM) POI = -67.881 – 11.03 lg [B2](M), R^2 = 0.6828.

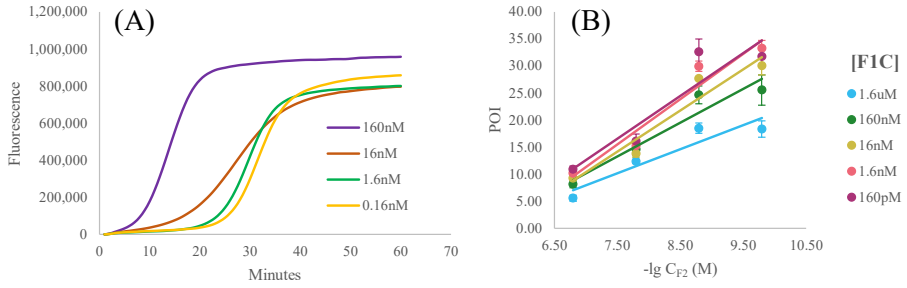

Fig. 9 (A) represents real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [F1c] = 160nM, [Ultramer] = 53pg, [B1C]=[B2]=[LB]=0.1uM, [LF]=0nM. Each experiment was conducted in triplicate. (B) Standard curves for F2 serial dilutions at various [F1c]. (1.6uM) POI = -18.806 – 3.827 lg [F2](M), R^2 = 0.8555, (160nM) POI = -32.493 – 6.0615lg [F2](M), R^2 = 0.9658, (16nM) POI = -43.066 – 7.6255 lg [F2](M), R^2 = 0.9265, (1.6nM) POI = -47.513 – 8.3977 lg [F2](M), R^2 = 0.9393, (160pM) POI = -48.924 – 8.7311 lg [F2](M), R^2 = 0.9119.

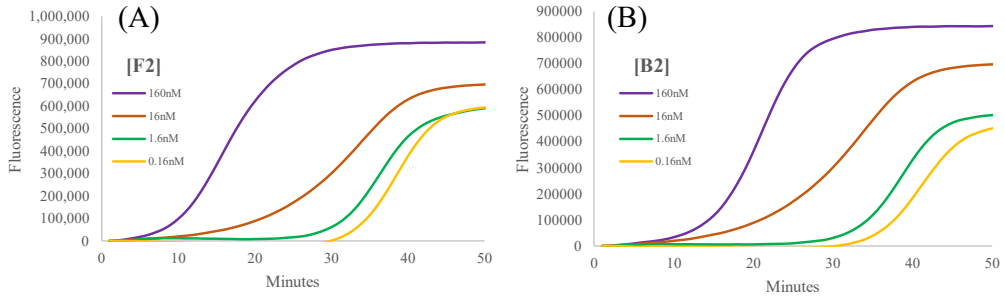

Fig. 10 (A) represents real-time fluorescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow), where [B2] = 16nM. (B) represents similar curves aroused by varying [B2] from from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow), where [F2] = 16nM. [Ultramer] = 53pg, [B1C]=[F1C]=[LB]=0.1uM. Each experiment was conducted in triplicate.

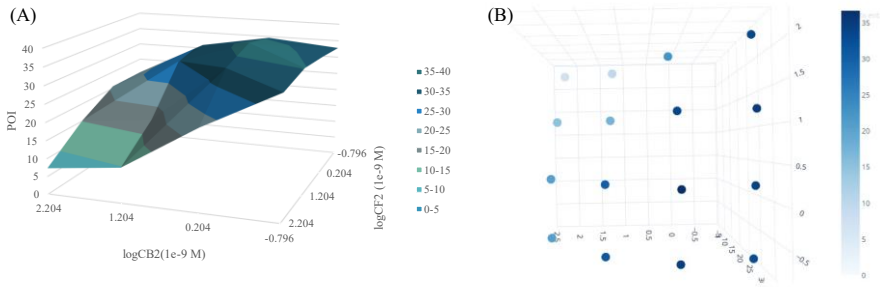

Fig. 11 (A) Surface plot of POI values in the F2/B2 reaction space. (B) Colour gradient of points in the F2/B2 reaction space. The maximum POI of 36.7 is achieved at [F2] = 1.6nM and [B2] = 1.6nM.

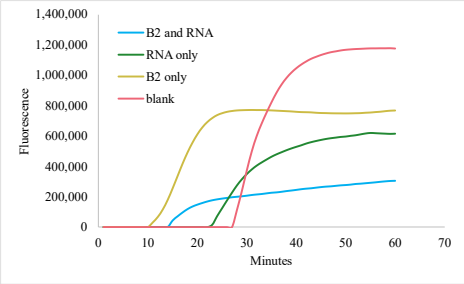

Fig. 12 Real-time flourescence curves aroused by B2 in split-LAMP. The curves show reaction mixtures with both B2 and background RNA (blue), background RNA only (green), B2 only (yellow) and blank (pink). The mean POI values for each of the reaction mixtures are 10.916, 17.202, 7.205 and 22.737 respectively. 0.5uL of 1.995 ug/ul HEK239 total cell RNA was added to a standard 20uL LAMP reaction. [B2] = 16nM, [F2] = 1.6uM, [F1C] = [B1C] = [LB] = 0.1uM.

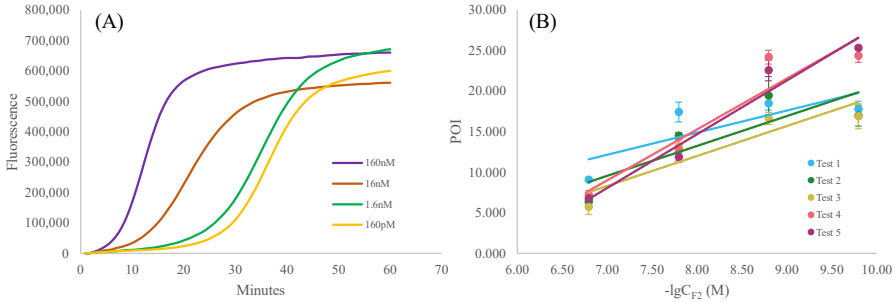

Fig. 13 (A) represents real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [F1c] = 160nM, [Ultramer] = 53pg, [B1C]=[B2]=[LB]=0.1uM, [LF]=0nM. Each experiment was conducted in triplicate. (B) Standard curves for F2 serial dilutions at different sets of primer concentrations recommended by model. (Test 1) POI = -6.90 – 2.7243 lg [F2](M), R^2 = 0.6271, (Test 2) POI = -16.244 – 3.6863lg [F2](M), R^2 = 0.7001, (Test 3) POI = -17.574 – 3.6987 lg [F2](M), R^2 = 0.8422, (Test 4) POI = -34.958 – 6.2802 lg [F2](M), R^2 = 0.8997, (Test 5) POI = -38.436 – 6.6392 lg [F2](M), R^2 = 0.9546.

Sequences of nucleic acids used in this work.

Nucleic acid	Sequence
B1c	GATGACAGTGACATCCTGCCT
B2	**TGGCAGTGTCTTAGCTGGTTGT**
LB	TGACAGGACATCGGTGACAGT
BIP	GATGACAGTGACATCCTGCCTA**GGCAGTGTCTTAGCTGGTTGT**
F1c	CGGAGAGGTCGCGATAGTCA
F2	**GCTTATCAGACTGATTTGA**
LF	TCACTGATCTGGCCGTAGACCA
FIP	CGGAGAGGTCGCGATAGTCAT**GCTTATCAGACTGATGTTGA**
Ultramer (double loop structure)	CGGAGAGGTCGCGATAGTCAT**GCTTATCAGACTGATGTTGA**TGGTCTACGGC CAGATCAGTGACTGACTATCGCGACCTCTCCGGTGATGACAGTGACATCCTGC CTGTGACAGGACATCGGTGACAGTT**ACAACCAGCTAAGACACTGCC**TAGGC AGGATGTCACTGTCAT GTCTACGGCCAGATCAGTGACT GTGACAGGACATCGGTGACAGT G GACTATCGCGACCTCTCCGGTGATGACAGTGACATCCTGCCT T

T CTGATAGCGCTGGAGAGGC CTACTGTCACTGTAGGACGGA

agttgtagtcagactattcgTA 5’ 3’

Tccgtcacagaatcgaccaaca

Table 1. Sequences of nucleic acids used in work. The analytical performance of the Split-LAMP assay was evaluated by challenging the assay system with hsa-miR-21 and hsa-miR-34 sequences respectively (shown in bold).

Real-time fluorescence curves at different concentrations of F1c.

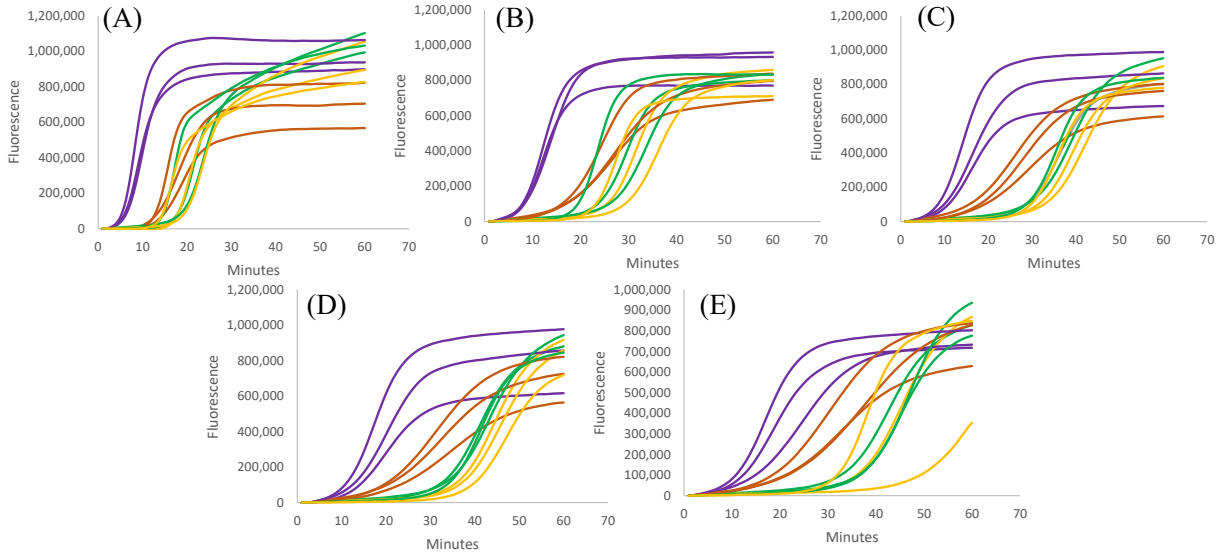

Fig. 14 (A) to (E) represent real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [F1c] in each is as follows: (A) 1.6uM, (B) 160nM, (C) 16nM, (D) 1.6nM and (E) 0.16nM. [Ultramer] = 53pg, [B1C]=[B2]=[LB]=0.1uM, [LF]=0nM. Each experiment was conducted in triplicate.

Real-time fluorescence curves at different concentrations of B1c.

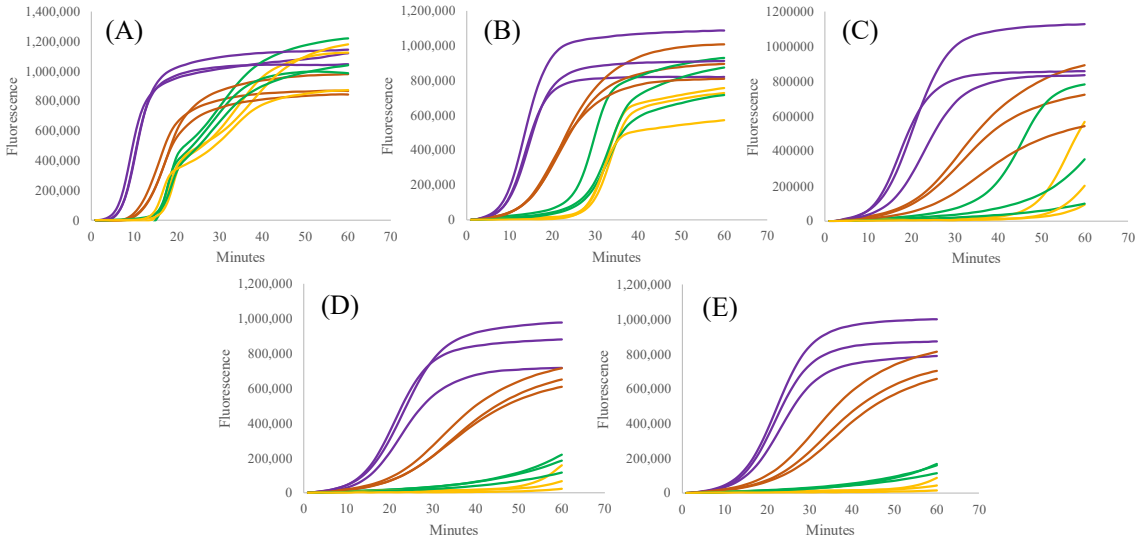

Fig. 15 (A) to (E) represent real-time flourescence curves aroused by varying [B2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [B1c] in each is as follows: (A) 1.6uM, (B) 160nM, (C) 16nM, (D) 1.6nM and (E) 0.16nM. [Ultramer] = 53pg, [F1C]=[B2]=[LB]=0.1uM, [LF]=0nM. Each experiment was conducted in triplicate.

Real-time fluorescence curves at different concentrations of F2 and B2.

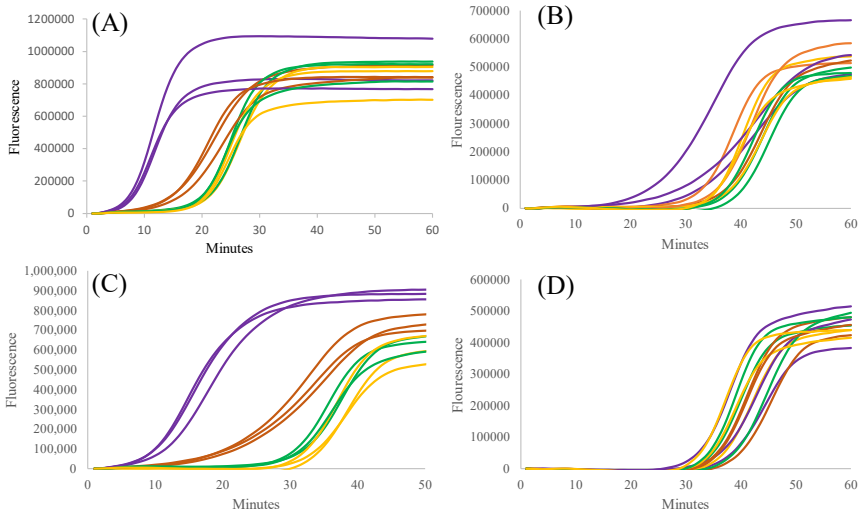

Fig. 16 Real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [B2] was (A) 160nM, (B) 16nM, (C) 1.6nM, (D) 160pM. [Ultramer] = 53pg, [B1C]=[F1C]=[LB]=0.1uM. Each experiment was conducted in triplicate.

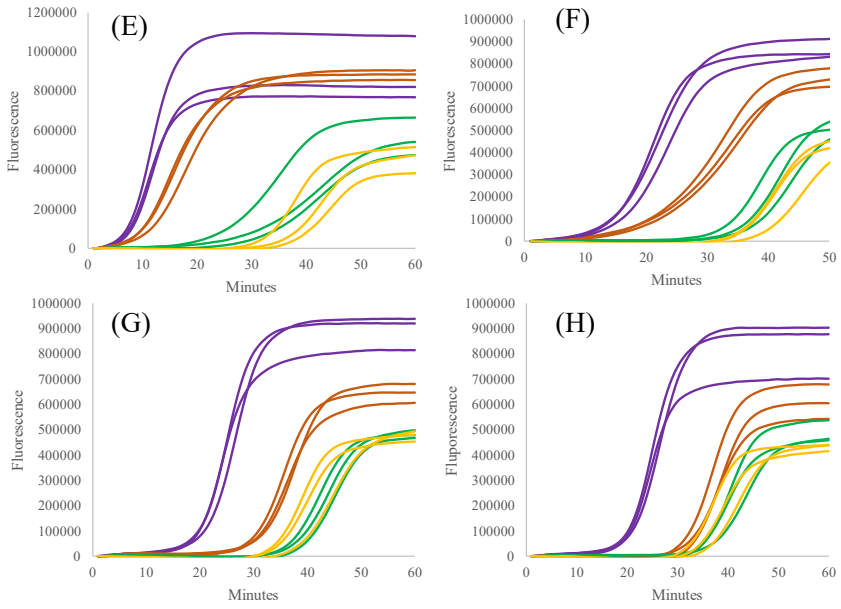

Fig. 17 Real-time flourescence curves aroused by varying [B2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow). [F2] was (E) 160nM, (F) 16nM, (G) 1.6nM, (H) 160pM. [Ultramer] = 53pg, [B1C]=[F1C]=[LB]=0.1uM. Each experiment was conducted in triplicate.

Details of Bayesian Optimization Framework

Our Bayesian Optimization framework relies on a Gaussian Process Regression to build a surrogate model and quantify its uncertainty. In this section, we first introduce Gaussian Process Regression, and then describe our Bayesian Optimization framework.

A Gaussian process provides a statistical distribution over functions where every point x in the input space is normally distributed with mean $\mu(x)$ and standard deviation $\sigma(x)$. After each experiment is conducted, Bayes Theorem is applied to update the Gaussian Process model to its posterior using information from the new data point obtained. In the context of this work, the critical takeaway is that the posterior model can be efficiently computed using simple linear algebra [6].

Bayesian optimization is a framework to solve a minimization problem with an unknown objective function $\ell(x)$ which is expensive to evaluate. It utilizes a Gaussian process to model

$\ell(x)$ and its associated uncertainty. At each iteration, the Gaussian process model is used to select the most promising candidate $\hat{x}$ for evaluation. An experiment is then conducted to evaluate the costly function at $\hat{x}$ in this iteration. Subsequently, the Gaussian process is updated to its posterior distribution $\tilde{\ell}(x)$ with the new data pair $(\hat{x}, \ell(\hat{x}))$. These iterations can be repeated until an optimum is found. The critical step is the selection of the candidate point, which is done via an acquisition function which incorporates both data from prior experiments and model uncertainty.

The maximiser $\hat{x}$ of the acquisition function is selected as the next candidate point. Intuitively, the acquisition function balances between exploiting points which we know perform well (using a mean term), and learning more about points we are uncertain about (using a standard deviation term). The balance between these two competing objectives is controlled using the parameter β. In this work, we use the Upper Confidence Bound (UCB) acquisition function, which comprises of the mean of the Gaussian process plus its standard deviation weighted by the trade-off parameter, β. That is,

$$UCB(x) = \mu(x) + \sqrt{\beta}\sigma(x)$$

Results of the Bayesian Optimization model

Test Number	**[B2] (nM)**	**[B1c] (nM)**	**[LB] (nM)**	**[F1c] (nM)**	**[LF] (nM)**	**R^2*gradient** (3dp)
1	15	165	180	175	30	1.708
2	75	120	165	155	35	3.436
3	90	195	195	65	10	3.115
4	90	80	40	120	60	5.652
5	75	100	70	25	125	6.340

Table 2. Sets of primer concentrations suggested by model on first iteration and results obtained. Test 5 obtained the best results with the highest R^2*gradient value.

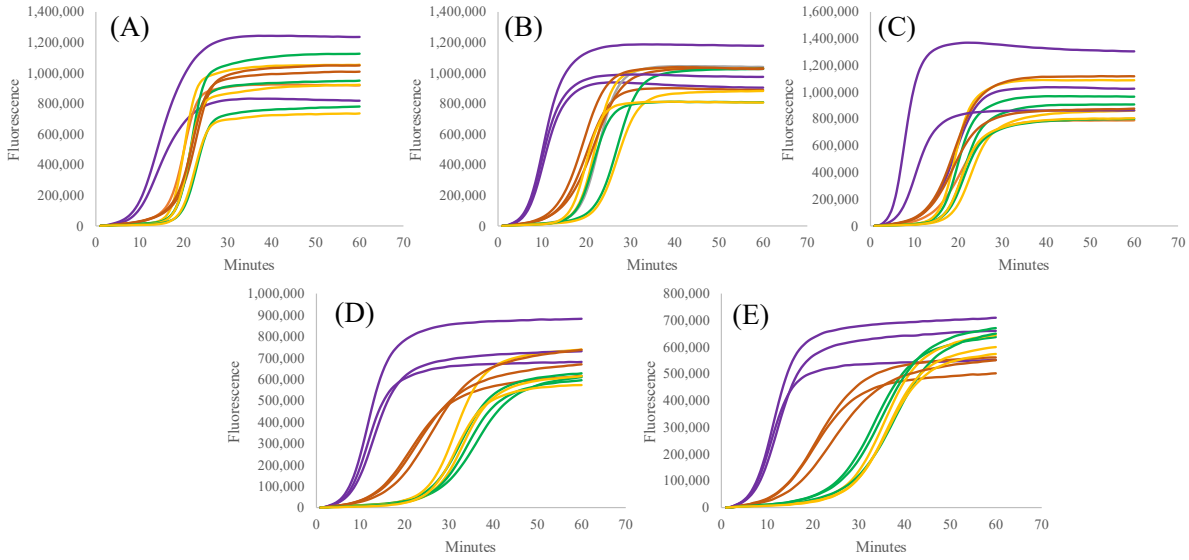

Fig. 18 (A) to (E) illustrates real-time flourescence curves aroused by varying [F2] from 160nM (purple), 16nM (orange), 1.6nM (green) to 160pM (yellow), with each representing a different primer concentration ratio suggested by the model on its first iteration. [Ultramer] = 53pg, [B1C]=[F1C]=[LB]=[B2]=0.1uM. Each experiment was conducted in triplicate. Test 5 yielded the best results with the largest dynamic range and strongest correlation coefficient in its standard plot.