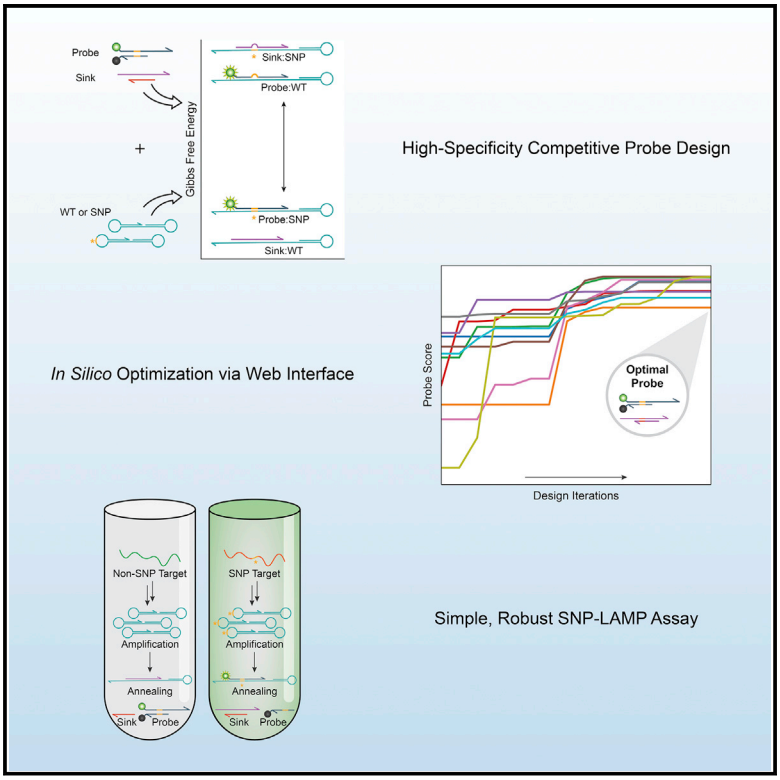
Cell Reports
Methods

Report

Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection

Graphical abstract



Highlights

- A simple and rapid assay to detect specific SNPs in complex nucleic acid samples
- Approach leverages competitive probes to specifically detect SNP over non-SNP sequences
- Demonstrated utility across diverse targets and sample types
- Web application for researchers to design SNP-LAMP probes for any target of interest

Authors

Leland B. Hyman, Clare R. Christopher, Philip A. Romero

Correspondence

promero2@wisc.edu

In brief

Detecting single-nucleotide polymorphisms (SNPs) is important for understanding human disease, the emergence of pathogen variants, and the application of genetic agricultural breeding programs. Hyman et al. develop a simple and rapid assay to detect SNPs in complex nucleic acid samples. A web-based tool allows researchers to design custom SNP probes.

Hyman et al., 2022, Cell Reports Methods 2, 100242
July 18, 2022
https://doi.org/10.1016/j.crmeth.2022.100242

CellPress

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

Cell Reports Methods

CellPress
OPEN ACCESS

Report

Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection

Leland B. Hyman,[1] Clare R. Christopher,[1] and Philip A. Romero[1,2,3,4,*]
[1]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA
[2]Department of Chemical & Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA
[3]The University of Wisconsin Carbone Cancer Center, Madison, WI, USA
[4]Lead contact
*Correspondence: promero2@wisc.edu
https://doi.org/10.1016/j.crmeth.2022.100242

MOTIVATION Single-nucleotide polymorphisms (SNPs) are the most common source of genetic variation between individuals and have implications in human disease, pathogen drug resistance, and agriculture. SNPs are often detected using DNA sequencing or PCR, both of which can be difficult to deploy for on-site testing or in low-resource settings. There is a need for new SNP detection methods that can be performed rapidly without advanced laboratory equipment or a cold supply chain.

SUMMARY

In this work, we developed a simple and robust assay to rapidly detect SNPs in nucleic acid samples. Our approach combines loop-mediated isothermal amplification (LAMP)-based target amplification with fluorescent probes to detect SNPs with high specificity. A competitive "sink" strand preferentially binds to non-SNP amplicons and shifts the free energy landscape to favor specific activation by SNP products. We demonstrated the broad utility and reliability of our SNP-LAMP method by detecting three distinct SNPs across the human genome. We also designed an assay to rapidly detect highly transmissible severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants from crude biological samples. This work demonstrates that competitive SNP-LAMP is a powerful and universal method that could be applied in point-of-care settings to detect any target SNP with high specificity and sensitivity. We additionally developed a publicly available web application for researchers to design SNP-LAMP probes for any target sequence of interest.

INTRODUCTION

Single-nucleotide polymorphisms (SNPs) are the most widespread source of genetic variation between individuals (Gao et al., 2019). A single base substitution can induce profound changes in the structure of a protein, altering its enzymatic function (Ueki et al., 2019), cellular trafficking (Megaraj et al., 2011), or solubility (Chasman and Adams, 2001). As a result, SNPs play crucial roles in many different biological phenomena, including human (Bakir-Gungor and Sezerman, 2011) and animal disease (Koopaee and Koshkoiyeh, 2014), pathogen drug resistance (Ramanathan et al., 2017), and agricultural blight (Rahimi et al., 2019). SNPs are often detected using DNA sequencing (Tahir and Sardaraz, 2020), which requires a laboratory setting for sample preparation, in addition to large, expensive, and slow DNA sequencing instruments. PCR methods are a simpler alternative, but the requirement for thermocycling is still a disadvantage in the field. Thus, neither of these methods are ideal for low-resource settings such as rural areas or developing countries (Duffy et al., 2017). There is a substantial need for rapid and point-of-care SNP detection assays that can be performed on-site without advanced laboratory equipment or a cold supply chain.

Loop-mediated isothermal amplification (LAMP) is a simple and robust method for sequence-specific detection of nucleic acids (Notomi et al., 2000; Nagamine et al., 2002; Wong et al., 2018). Unlike PCR, the amplification process occurs continuously at an isothermal temperature, facilitating fast amplification times and use of a simple heated block rather than a thermocycler (Velders et al., 2018). Previous work has developed LAMP-based assays to detect SNPs. Mismatch SNP-LAMP incorporates the SNP base into the 3′ terminus of a LAMP primer, causing a mismatch and preventing polymerase extension when the non-SNP sequence is present (Badolo et al., 2012; Duan et al., 2016; Liu et al., 2018). However, heterogeneity in primer synthesis and promiscuous mismatch extension by the LAMP polymerase (Chen et al., 2015; Wang et al., 2017) can lead to unpredictable amplification times and false positive events. Other SNP-LAMP strategies use fluorescent DNA probes to detect SNPs in LAMP products (Jiang et al., 2015).

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

The probe consists of a DNA duplex with a quenched fluorophore that is complementary to the SNP sequence and thus preferentially binds the SNP over the non-SNP. The difference in binding energies from a single mismatched base can be quite small, leading to a substantial signal from non-SNP sequences that is difficult to distinguish SNP sequences (Wang and Zhang, 2015).

We have developed a SNP-LAMP method that can rapidly and robustly detect SNPs with a simple workflow. Our approach leverages LAMP-based target amplification and competitive fluorescent probes (Tyagi and Kramer, 1996; Li et al., 2002; Morlan et al., 2009; Wang and Zhang, 2015) to specifically detect SNP over non-SNP sequences. Competitive "sink" strands preferentially bind non-SNP sequences and help to widen the free energy gap between highly similar SNP and non-SNP sequences. We devised a thermodynamics-based computational optimization algorithm to design probe sets with high sensitivity and specificity for a target SNP. We developed a web application to make these optimal probe designs accessible to any researchers interested in our assay, regardless of their computational knowledge. We demonstrated the ability of competitive SNP-LAMP to detect specific SNPs from highly complex total RNA samples in a simple one-pot reaction. Finally, we developed a simple and streamlined assay to detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains that may be usable for monitoring emerging variant outbreaks. Competitive SNP-LAMP is a powerful and robust solution for detecting SNPs that is simple and inexpensive enough to be deployed on a large scale and in low-resource settings.

RESULTS

Competitive probes for highly specific LAMP-based detection of SNPs

We sought to identify a robust isothermal approach to detect SNPs in diverse nucleic acid samples. Previous work has found the signal of strand displacement probes can be enhanced by including a "sink" complex that competes for binding with the non-SNP sequence (Tyagi and Kramer, 1996; Li et al., 2002; Morlan et al., 2009). This competitive probe alters the free energy landscape and has achieved remarkable specificity for PCR-based SNP detection (Wang and Zhang, 2015). We adapted this approach to detect SNPs in LAMP products.

Our system includes a strand displacement probe duplex that is complementary to the SNP sequence and a competitive sink duplex that is complementary to the non-SNP sequence (Figures 1A–1D). Non-SNP LAMP products will preferentially bind to the sink strand over the fluorescent probe, thereby reducing this undesired signal (Figures 1C and 1D). As a proof of concept, we designed a set of probe and sink duplexes to detect the c.776A > C mutation in the human *ACTB* gene. Our competitive probe system could clearly distinguish DNA oligos representing the non-SNP and SNP sequences ($p = 1.6 \times 10^{-7}$) (Figure 1B). The fluorescence signal of the SNP sample was 2.7 times higher than the non-SNP sample. In addition, the non-SNP signal was only 15% higher than the non-template control, indicating that the sink complex successfully reduced probe binding to the non-SNP oligo. We also performed the reaction without the sink complex and found that both the SNP and non-SNP targets activated, with no significant difference in signal between the two ($p = 0.46$). These results suggested that competitive fluorescent probes are suitable for SNP detection within single-stranded LAMP amplicon regions.

Computational design of competitive SNP-LAMP probes

Our competitive probe system consists of four DNA strands that can be of different lengths and complementary to different regions surrounding the target SNP. Our initial probe design involved manually tweaking the oligonucleotide sequences to achieve the desired behavior. However, the full design space is massive and shifting a strand by even a single base can drastically alter a design's specificity. These factors make it challenging to manually design competitive SNP probes with optimal signal and specificity.

We developed a computational framework to design competitive probe combinations that maximally discriminate between SNP and non-SNP targets. Probe design is a multi-objective optimization problem that must balance two potentially conflicting goals: high fold activation in the presence of the target SNP and high SNP specificity over the non-SNP. We used thermodynamic modeling (Dirks et al., 2007) to evaluate how a given design will behave in the absence of any target, in the presence of the SNP target, and in the presence of the non-SNP target. These simulations provide an estimate of the fluorescence signal (concentration of unquenched probe) produced under these three conditions. We define an aggregate objective function that captures the two design objectives:

$$f = pSNP \cdot \log\left(\frac{pSNP}{\max\left[p_{non}SNP, p_{bg}\right]}\right),$$

where p_{SNP}, p_{nonSNP}, and p_{bg} are the proportion of unquenched probe in SNP, non-SNP, and background (no template) thermodynamic simulations, respectively. p_{SNP} represents the fluorescence signal produced by the SNP, while the second term captures the difference between SNP target and non-SNP target signals. We seek to maximize this aggregate objective over all of the possible probe designs.

Our system consists of a probe duplex that contains a fluorophore-quencher pair and a sink duplex. These four DNA strands are each specified by their number of complementary bases before and after the SNP base position (Figure 2A). To better understand the probe design space, we simulated 10,000 random probe sets, each specific to a randomly generated SNP target sequence between 25 and 40 bases long. We found that optimized probe sets are incredibly rare, with only 0.45% of these random designs displaying a signal greater than 70% and specificity more than 100-fold (Figure 2B). We generated a set of 12 biophysical features describing each probe (Table S1) and performed principal component analysis (PCA) to visualize the design space (Figure S1). The optimized probe designs fall within a specific region that occupies roughly one-fourth of the total design space area.

We used a hybrid genetic algorithm (GA)-hill climbing algorithm to optimize our objective function over probe design space

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

CellPress
OPEN ACCESS

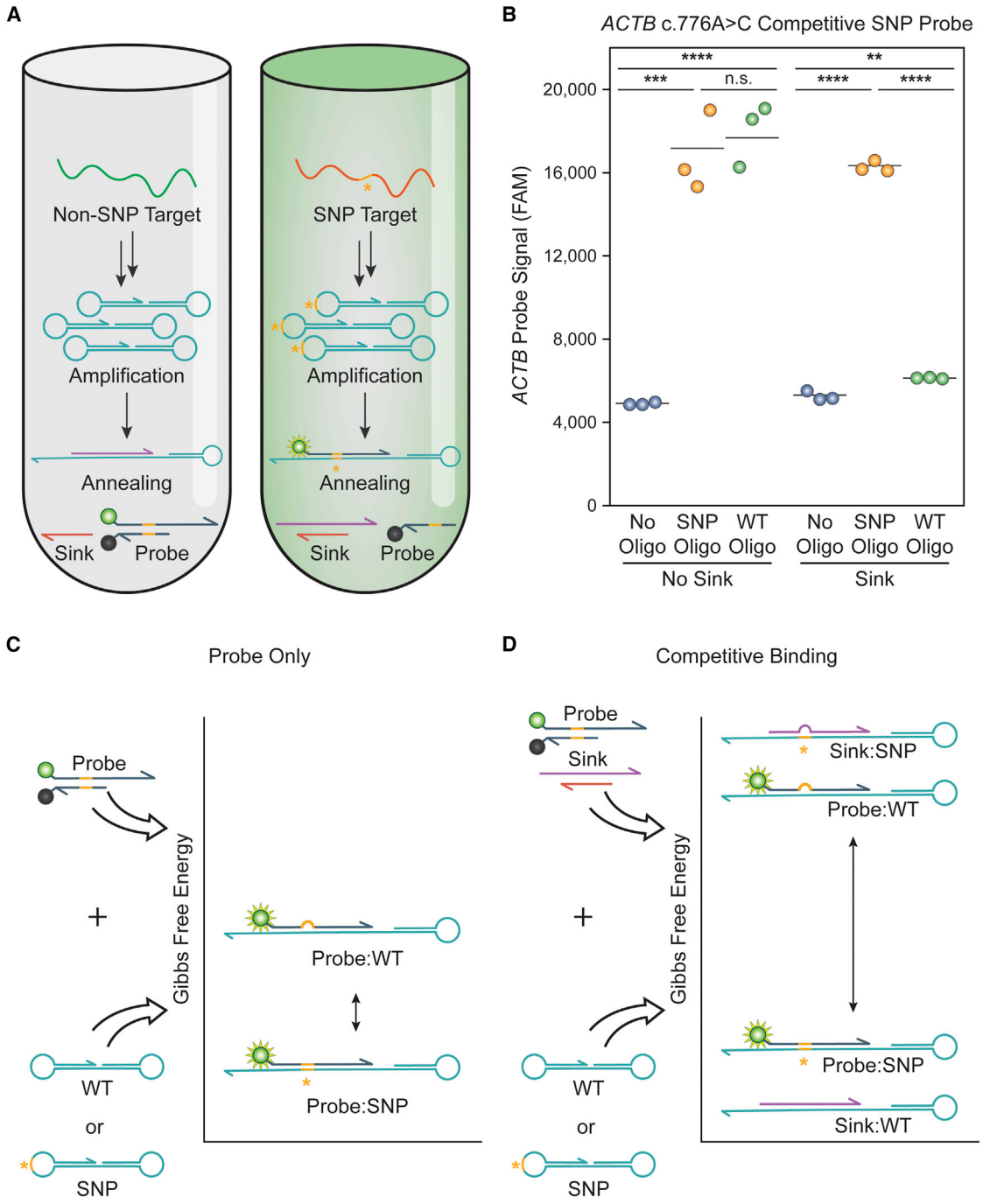


Figure 1. A competitive SNP-LAMP detection strategy

(A) Schematic of the competitive SNP-LAMP assay. SNP and non-SNP targets are amplified in a standard LAMP reaction, after which the amplicons are melted and annealed to a SNP-specific probe and a non-SNP-specific sink complex. The probe strand predominantly binds to the SNP sequence, producing a fluorescent signal.

(B) Detection of a c.776A > C mutation in the *ACTB* gene using competitive SNP-LAMP. Adding a sink complex greatly reduced the signal from a representative WT oligo target ($p = 8.3 \times 10^{-5}$), while producing no significant change in signal from a SNP oligo target ($p = 0.65$).

(C) SNP-LAMP detection using a normal fluorescent DNA probe. Due to the small difference in binding energy between a perfect match and a single mismatch, probe specificity may be poor.

(D) SNP-LAMP competitive probe strategy and thermodynamics. In addition to the probe complex, a WT-specific sink complex is added, which competes for WT binding and greatly increases specificity. This is reflected in a large ΔG difference between the probe:SNP and probe:WT duplexes at equilibrium.

n.s., $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$, respectively.

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

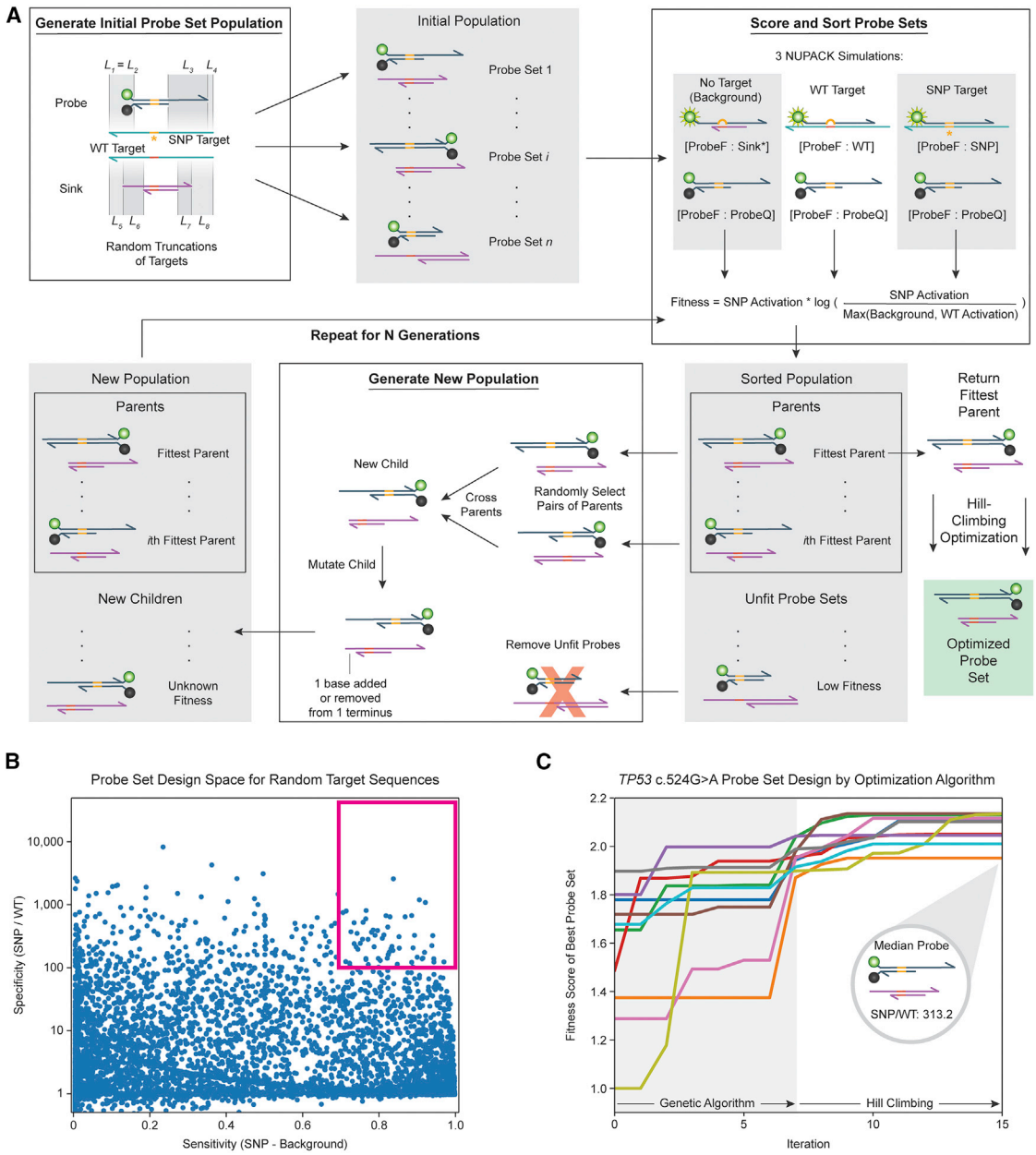


Figure 2. A hybrid GA and hill-climbing optimization strategy for SNP-LAMP probe design

(A) During GA optimization, probes are randomly generated, mixed, and selected by fitness in a series of generations, mimicking natural evolution. After the designated number of generations, the fittest probe is further optimized by a hill-climbing algorithm to reach the nearest fitness maximum.

(B) Design space for 10,000 random probes, each designed for a unique SNP target. High-fitness probes are rare: only 0.45% of designs fall within the optimal design space (magenta).

(C) Convergence of the algorithm over 7 GA generations followed by hill climbing. Across all 10 runs, the median optimal probes' background-subtracted SNP/WT specificity was consistently high, with a median of 313.2 and a standard deviation (SD) of 47.6.

(Figure 2A). The GA searches the space broadly by iteratively breeding, mutating, and selecting top design candidates. The top sequence from the GA is then optimized via hill climbing to exhaustively search the local design space and ensure that we have reached a local maximum. The resulting probe design should balance signal and specificity for highly optimized SNP detection. We tested our algorithm by designing a series of probe sets toward a c.524G > A mutation in the *TP53* tumor suppressor gene. We ran the algorithm 10 times from random starting points and observed its convergence to an optimal probe sequence (Figure 2C). With only 7 GA generations and an initial population size of 128, the algorithm reliably converged to high specificity and high signal probe sets. Across all 10 runs, the median final probe design had a SNP/WT specificity of 313.2 with a

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

SD of 47.6, and produced 85.5% of its maximum possible fluorescent signal when detecting the SNP sequence. Furthermore, the algorithm showed steady fitness improvement in nearly every case, suggesting that it can generate enough diversity to avoid becoming trapped in local minima. This algorithm should therefore serve as a reliable and effective means for competitive composition-based SNP probe design.

One-pot competitive SNP-LAMP robustly detects SNPs in total RNA samples

We tested the generality and reliability of our computational design method by designing SNP-LAMP probes to detect three distinct SNPs across the human genome. The three targets are the c.186A > G mutation in the *MT-C O 2* mitochondrial housekeeping gene, the c.524G > A mutation in the *TP53* tumor suppressor gene, and the c.4799T > C mutation in the *NOTCH1* oncogene. We verified that the SK-BR-3 and MOLT-4 human cell lines differ at these three targeted sites (Figure S2), and we used total RNA samples from each to test the performance of our designs.

We developed a one-pot SNP-LAMP assay in which a nucleic acid sample is added to a tube along with the LAMP primers, the designed probe and sink duplexes, and a standard LAMP master mix (Figure 3A). The sample is first incubated at 65°C to allow LAMP-based target amplification, then heated to melt the LAMP products and slowly annealed to allow probe/sink hybridization to reach equilibrium. We performed our one-pot SNP-LAMP assay on each of the three targeted mutations and found it could reliably distinguish SNP template versus no template and SNP versus non-SNP template in all of the cases (Figures 3B–3D). We observed a small but statistically significant difference ($p < 0.05$) in fluorescence between the WT and background samples for two of the three targets (MT-CO2 and TP-53). This suggests that despite the improvements in specificity from competitive probes, it is still necessary to include both SNP and non-SNP controls in these assays.

To assess how our designed probe sets may perform with rare variant allelic frequencies, we performed experiments with varying proportions of the SNP and non-SNP DNA (Figure S3A). Both *P53* and *NOTCH1* probes showed a high sensitivity for the SNP oligo even at a prevalence of 1 in 100 WT copies, showing a significant fluorescence increase over the wild-type (WT) only condition with p values of 2.3×10^{-5} and 3.1×10^{-3}, respectively. We next tested whether SNP-LAMP reactions would perform similarly using *in vitro* transcribed RNAs at varying ratios (Figures S3B and S3C). Surprisingly, we found that SNP-LAMP reactions were not as tolerant of low target proportions, showing high variance in fluorescence and failing to differentiate between 0% target RNA and 50% target RNA in some cases. This may be due to the exponential nature of LAMP, in which the frequency of the SNP sequence in the final reaction is not proportional to its frequency in the initial RNA population. Therefore, the SNP-LAMP assays in this work are best suited in cases in which the SNP allele is highly represented in the sample.

A rapid test to distinguish SARS-CoV-2 variants

The COVID-19 pandemic has illustrated the importance of rapid point-of-care testing in disease mitigation and tracking (Hogan et al., 2020). This need for low-cost, rapid, and point-of-care strain tracking inspired us to develop a competitive SNP-LAMP test for specific SARS-CoV-2 variants.

We designed a probe set to target the D614G mutation in the viral spike protein, which is thought to increase the viral load in infected patients and rapidly became dominant in the SARS-CoV-2 population (Korber et al., 2020). We tested our designs in a one-pot SNP-LAMP assay using *in vitro* transcribed RNA fragments for the WT and mutant spike protein variants (Figure 3E). Our SNP-LAMP assay readily distinguished the D614G spike RNA from WT ($p = 5.8 \times 10^{-5}$), with a small but significant difference between the 614D variant and background ($p = 0.0041$).

Encouraged by these results, we further optimized the D614G SNP-LAMP test to develop it into a more useful diagnostic assay. Our main objectives were to achieve a low detection limit and to reduce the reaction time. To improve the sensitivity, we switched to a two-step reaction format (Figure 3A), since we have previously observed slight LAMP inhibition from the probe in one-pot reactions. To reduce the reaction time, we screened additional primer sets and identified one that amplified faster than our first design. This reduced the LAMP reaction time required from 60–75 min to 30 min. We also shortened the annealing cycle on the thermocycler to reduce the time required from 70 min to 38 min, thus reducing the total thermocycler time to 68 min. We also designed an additional competitive probe specific to the 614D variant to better differentiate it from the 614G variant.

For point-of-care SARS-CoV-2 tests, it can also be advantageous to directly add a crude biological sample to the reaction. We therefore quantified the limit of detection using target RNAs spiked into viral transport media, saliva, and blood samples. Following previously described LAMP workflows (Kobayashi et al., 2021), we diluted each crude sample type in a buffer containing guanidinium and DTT to stabilize the input RNAs, heat inactivated it, and directly added the mixture to each LAMP reaction. With a general LAMP indicator, we found the limit of detection to be between 6 and 60 copies per reaction in water and viral transport media (Figure S4). We observed only slightly worse performance in blood and saliva, with 60 copies detected in at least 2 out of 3 reactions. This limit of detection is similar to others reported for direct SARS-CoV-2 LAMP detection from saliva (Newman et al., 2021), suggesting that our assay could reasonably be used for these applications. Finally, we tested whether adding these crude biological samples would adversely affect the performance of our SNP-LAMP probes. The 614D- and 614G-specific probes correctly distinguished their respective variants ($p < 0.05$) and showed significantly higher fluorescence than background ($p < 0.05$) when 1 fM RNA (600 copies per reaction) was spiked into water, viral transport media, saliva, and blood (Figure 3F). These results suggest that our SNP-LAMP assay is tolerant of crude samples with low input RNA concentrations, making it well suited for development into diagnostic assays.

A web application for SNP-LAMP probe design

To make our work more accessible to a general audience who may not wish to implement our probe design strategy from scratch, we have developed a publicly available web application

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

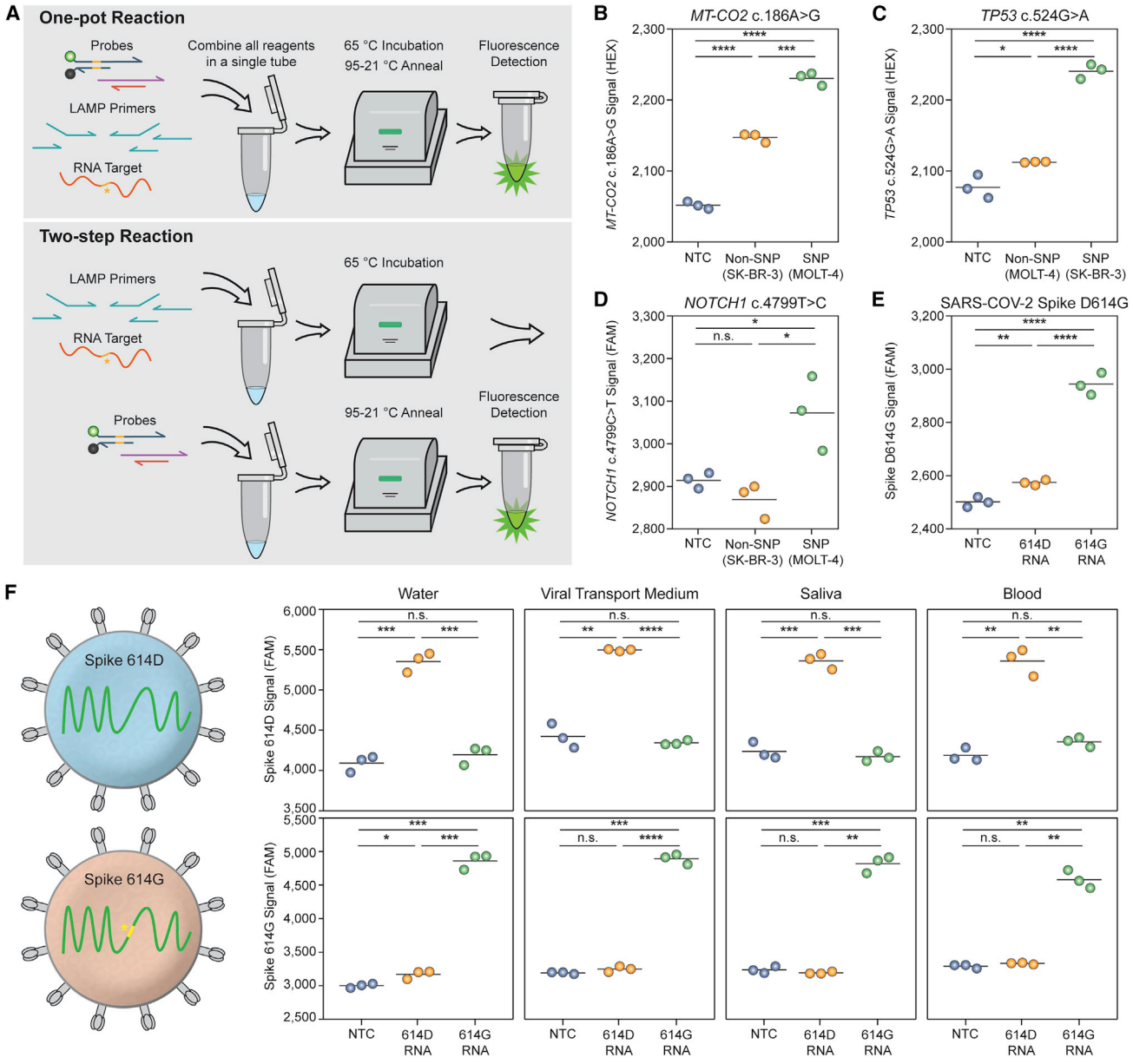


Figure 3. One-pot and two-step SNP-LAMP assays toward four unique targets

Endpoint fluorescence values are shown after LAMP and annealing.

(A) Schematic of the one-pot and two-step SNP-LAMP assay workflows. While the one-pot format is simpler, the two-step format provides improved signal and sensitivity.

(B) A one-pot SNP-LAMP reaction targeting the c.186A > G mutation in the *MT-C O 2* gene. MOLT-4 total RNA produced a significantly larger signal than SK-BR-3 RNA ($p = 1.05 \times 10^{-4}$) and background ($p = 3.96 \times 10^{-6}$).

(C) A one-pot SNP-LAMP reaction targeting the c.524G > A mutation in the *TP53* tumor suppressor gene. SK-BR-3 total RNA produced a significantly larger signal than MOLT-4 RNA ($p = 1.5 \times 10^{-5}$) and background ($p = 6.4 \times 10^{-5}$).

(D) A one-pot SNP-LAMP reaction targeting the c.4799T > C mutation in the *NOTCH1* oncogene. MOLT-4 total RNA produced a significantly larger signal than SK-BR-3 RNA ($p = 0.011$) and background ($p = 0.019$).

(E) A one-pot SNP-LAMP assay targeting the SARS-CoV-2 spike D614G mutation. The probe reliably distinguished a 614G spike RNA fragment from a 614D fragment ($p = 5.8 \times 10^{-5}$) and a non-template control ($p = 3.7 \times 10^{-5}$).

(F) Two-pot SNP-LAMP reactions targeting SARS-CoV-2 spike 614D and 614G variants. The assay showed high sensitivity in a variety of biological samples, distinguishing 1 fM (600 copies per reaction) of the target RNA from background in all cases ($p < 0.05$). Furthermore, the assay successfully distinguished the SNP RNA from the non-SNP RNA in all of the cases ($p < 0.05$).

n.s., $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$, respectively.

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

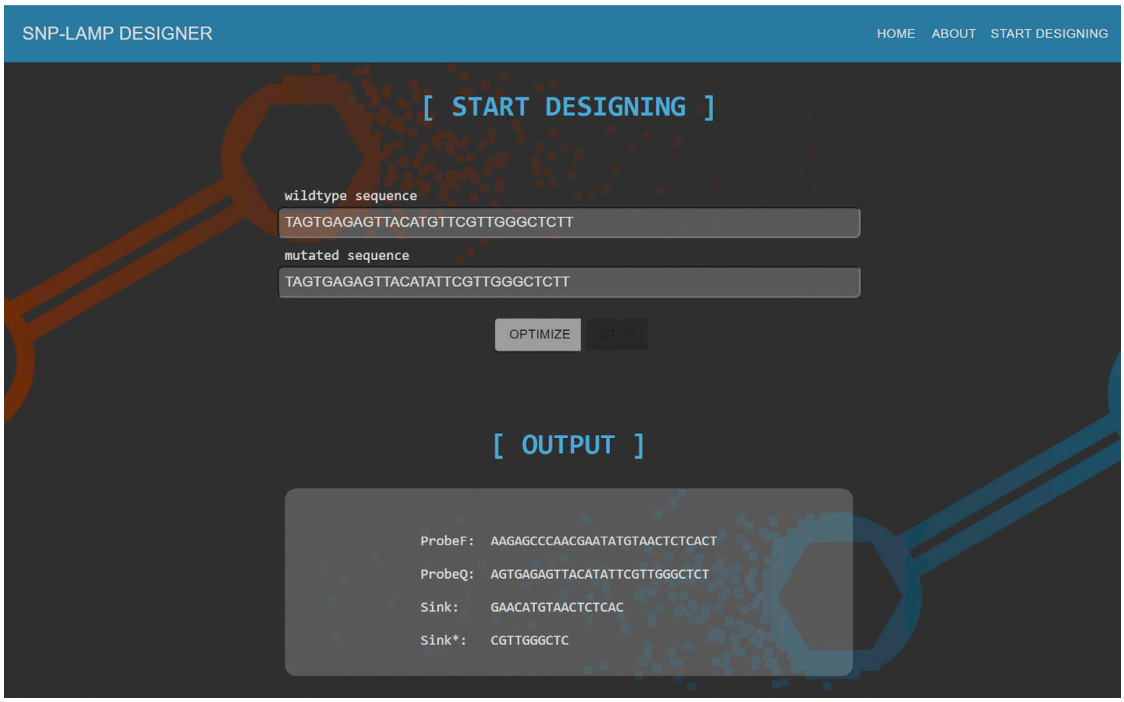


Figure 4. A web application to design optimal probes for SNP-LAMP

Researchers can enter virtually any set of wild-type and mutated sequences through a simple interface and receive an optimal competitive probe design within seconds.

and host it at https://snp-lamp-design.netlify.app/(Figure 4). This application provides a simple interface for probe design: a user enters a WT and mutated target sequence and can receive an optimal probe set within seconds. We achieved this by generating a database of 250,000 high-specificity probe designs that can be quickly searched and adapted to the target sequences. This allows researchers to design SNP-LAMP probes for virtually any target sequence, regardless of their computational background.

DISCUSSION

SNPs make up the majority of genetic variations between individuals (Gao et al., 2019) and have implications in human disease (Gao et al., 2019), pathogen antibiotic resistance (Ramanathan et al., 2017), and agricultural production (Koopaee and Koshkoiyeh, 2014; Rahimi et al., 2019). Detecting SNPs is challenging due to their high similarity with non-SNP sequences, and DNA sequencing is the most reliable and commonly used method to identify SNPs. DNA sequencing methods require sample preparation and bulky instrumentation, and thus cannot be easily deployed for onsite or low-resource testing (Tahir and Sardaraz, 2020). In this work, we have developed a simple method to distinguish complex nucleic acid samples that differ by only a single base. Our approach leverages LAMP-based target amplification and competitive sink DNA strands to favor specific activation by SNP products. We demonstrated the robustness of our SNP-LAMP assay by detecting three distinct SNPs in highly complex total RNA samples. We also developed a simple and streamlined assay to detect SARS-CoV-2 strains that may be usable at a large scale and in low-resource settings. Furthermore, we developed a publicly available tool that allows researchers to easily design these SNP probes for their own purposes.

Previous LAMP-based SNP detection methods are based on either primer mismatches or SNP-specific fluorescent probes. These strategies can be unreliable due to promiscuous polymerase mismatch extension activity (Chen et al., 2015), which results in false positives, limited ability to resolve signal differences between highly similar SNP and non-SNP targets, and constraints in designing primer/probe sets that restrict the SNP loci that can be targeted. We were not able to get either of these approaches to work in our hands despite trying a total of three different designs. Our competitive SNP-LAMP approach overcomes these limitations by using competing sink strands to drastically enhance the specificity for the target SNP. This resulted in a highly reliable SNP detection method that worked on the first attempt for all of the targets tested.

We developed a computational pipeline to design competitive SNP-LAMP probes with high sensitivity and specificity. The possible probe space is massive and very few meet our SNP detection criteria. To traverse this design space, we used a hybrid GA and hill-climbing approach to optimize the thermodynamic properties of a probe set. When initialized from 10 random starting states, this algorithm consistently converged to highly specific and sensitive designs. Although we did not perform a rigorous head-to-head comparison, we did observe that our computationally designed probe sets had superior SNP:non-SNP specificity relative to manually designed probes. This result

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

is expected since computational optimization can search a much larger design space than human rational design. We believe that our computational probe design method can be readily generalized to target any possible SNP of interest.

Our SNP-LAMP method can rapidly detect SNPs with a simple protocol that could be performed by individuals with minimal laboratory training and equipment. The nucleic acid sample is added to a tube containing LAMP primers and a LAMP master mix. This sample is then heated at 65°C for 30 min and the probes are added and hybridized by annealing from 95C to 21C over 38 min. The fluorescence of the sample is measured to provide an assay result in approximately 1 h of thermocycler time. The inexpensive cost of our assay is also a major advantage for large-scale testing. While Sanger sequencing assays typically cost $4–$7 USD per reaction at core facilities and private biotechnology companies, the LAMP protocols used here consume approximately $1 USD in reagents.

Our SNP-LAMP method is simple, rapid, and low cost, and can be performed on basic laboratory equipment. There are several additional modifications that would enable a true point-of-care assay that could be deployed in the field or other low-resource settings. Our current SNP-LAMP assay involves multiple incubation temperatures that require a temperature-adjustable heating device. The temperature-annealing step of our method is necessary to bring the reaction to thermodynamic equilibrium because the probes were designed based on thermodynamics. Isothermal detection schemes could be devised by considering hybridization kinetics and incorporating single-stranded "toehold" sequences to direct probe binding via DNA branch migration (Srinivas et al., 2013). Our assay also relies on a fluorescent readout that can be challenging to perform onsite. There are other label-free detection methods that rely on simple DNA hybridization and strand displacement to generate electrical readouts (Ahmed et al., 2007; Zhang et al., 2010). In theory, our SNP-LAMP strategy could be adapted to a have a simple electrical readout.

Field-testing applications, as well as point-of-care assays in developing regions, could also benefit from eliminating cold supply chain requirements. LAMP reagents can be lyophilized and deployed at room temperature (Carter et al., 2017), and even packaged in pre-made reaction tubes that require only a liquid sample to be added. Since our method does not rely on pre-annealing the probe and sink complexes before the reaction, they could simply be lyophilized with the other LAMP reagents. We also believe our method may work well on crude nucleic acid samples that have not been processed or purified. Both LAMP and DNA hybridization processes are extremely tolerant to cellular debris, additives, and inhibitors (Alhassan et al., 2007; Chen et al., 2018). In this case, a crude sample could be added to lyophilized assay reagents to provide a streamlined workflow that requires minimal hands-on processing and laboratory equipment.

SNPs are crucial drivers of many biological processes, with important implications in human diseases ranging from cystic fibrosis (Gisler et al., 2013) to cancer (Gao et al., 2019). SNPs can also contribute to other undesirable phenomena such as antibiotic resistance in microbes (Ramanathan et al., 2017) and agricultural breeding issues (Koopaee and Koshkoiyeh, 2014; Rahimi et al., 2019). Competitive SNP-LAMP provides a reliable, simple, rapid, and low-cost SNP detection method that could be deployed for on-site testing in the field or in developing areas of the world (Liu et al., 2018). This method will empower researchers across the life sciences by providing a universal solution for point-of-care SNP detection.

Limitations of the study

Our competitive SNP-LAMP approach has several limitations that may restrict its application in some contexts. Some applications require detecting SNPs at low variant allele frequencies (VAFs). These include rare dominant active mutations found in genes with high copy numbers or also mixed samples from multiple individuals. We demonstrated that the *TP53* and *NOTCH1* probes could detect their target sequence at SNP frequencies below 1%. However, these experiments were performed DNA oligos and the results did not generalize to actual SNP-LAMP assays. LAMP is a stochastic process with exponential kinetics, and the final LAMP products may not reflect the original variant frequencies within the sample. In cases in which SNPs must be detected at low VAFs, a linear amplification method such as rolling circle amplification (RCA) (Ali et al., 2014) may provide more reliable frequency estimates.

Related to the above point about low VAFs, in some instances we found that competitive SNP-LAMP could not detect SNPs at 1:1 SNP-to-non-SNP ratios. When analyzing varying proportions of SNP versus non-SNP SARS-CoV-2 spike variants (Figures S3B and S3C), we found that the 614G probe could detect as low as 10% 614G (i.e., distinguish between 0% 614G + 100% 614D versus 10% 614G + 90% 614D). However, when detecting the opposite mutation, the 614D probe could only distinguish between 100% 614G and 100% 614D. The 0% 614D + 100% 614G sample was not statistically different from the 50% 614D + 50% 614G sample. The inability to detect SNPs at 50% prevents the analysis of homozygous loci in diploid organisms. This does not appear to be a general limitation because the 614G probe could detect down to 10% 614G. It is likely that further primer and probe optimization would allow robust detection of 614D below 50%.

Another confounding factor in our assay is the small but often statistically significant fluorescence signal produced by non-SNP targets. We expect that controls representing both SNP and non-SNP targets will be necessary in any developed SNP-LAMP tests to account for this background signal.

As we were developing our competitive SNP-LAMP method, we found that the competitive probes could cause inhibition of LAMP-based target amplification. We found that one-pot SNP-LAMP reactions could only include up to 100 nM of each probe and sink strand before inhibition caused an issue. This constraint limits the total fluorophore in the system and therefore maximum fluorescence output that can be achieved. The fluorescence signal is still easily detectable on a standard laboratory qPCR instrument; however, the signal:background ratio is only ~1.1. We overcame this issue by increasing the LAMP primer concentrations in one-pot assays. Ultimately, we found two-step SNP-LAMP assays to be more powerful, since 10-fold higher probe concentrations can be used. For any assay that requires higher sensitivity, as with our SARS-CoV-2 demonstrations, the two-step assay described here is superior.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
- METHOD DETAILS
  - LAMP primer design
  - Thermodynamic simulations of probe set binding
  - Probe design space simulations
  - Approximate Tm calculations
  - Principal component analysis
  - Probe set design by genetic algorithm
  - Probe set optimization by hill-climbing
  - DNA complexes and primer mixes
  - Production of mRNAs using in *vitro* transcription
  - Total RNA extraction from cell lines
  - SNP-LAMP and annealing assays
  - Crude biological sample preparation
  - Sanger sequencing of LAMP amplicons
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical testing
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2022.100242.

ACKNOWLEDGMENTS

We would like to acknowledge funding from the Damon Runyon Cancer Research Foundation (DRR-40–16 to P.A.R.).

AUTHOR CONTRIBUTIONS

L.B.H. and P.A.R. conceived the project. L.B.H. and C.R.C. performed the experiments. L.B.H. analyzed the data, with feedback from P.A.R. L.B.H. and P.A.R. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 16, 2021
Revised: February 28, 2022
Accepted: June 2, 2022
Published: June 13, 2022

REFERENCES

Ahmed, M.U., Idegami, K., Chikae, M., Kerman, K., Chaumpluk, P., Yamamura, S., and Tamiya, E. (2007). Electrochemical DNA biosensor using a disposable electrochemical printed (DEP) chip for the detection of SNPs from unpurified PCR amplicons. Analyst *132*, 431–438. https://doi.org/10.1039/b615242b.

Alhassan, A., Thekisoe, O.M., Yokoyama, N., Inoue, N., Motloang, M.Y., Mbati, P.A., Yin, H., Katayama, Y., Anzai, T., Sugimoto, C., and Igarashi, I. (2007). Development of loop-mediated isothermal amplification (LAMP) method for diagnosis of equine piroplasmosis. Vet. Parasitol. *143*, 155–160. https://doi.org/10.1016/j.vetpar.2006.08.014.

Ali, M.M., Li, F., Zhang, Z., Zhang, K., Kang, D.K., Ankrum, J.A., Le, X.C., and Zhao, W. (2014). Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. Chem. Soc. Rev. *43*, 3324–3341. https://doi.org/10.1039/c3cs60439j.

Badolo, A., Okado, K., Guelbeogo, W.M., Aonuma, H., Bando, H., Fukumoto, S., Sagnon, N., and Kanuka, H. (2012). Development of an allele-specific, loop-mediated, isothermal amplification method (AS-LAMP) to detect the L1014F kdr-w mutation in Anopheles gambiae s. l. Malar. J. *11*, 227. https://doi.org/10.1186/1475-2875-11-227.

Bakir-Gungor, B., and Sezerman, O.U. (2011). A new methodology to associate snps with human diseases according to their pathway related context. PLoS One *6*, e26277. https://doi.org/10.1371/journal.pone.0026277.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607. https://doi.org/10.1038/nature11003.

Bunney, P., Zink, A., Holm, A., Billington, C., and Kotz, C. (2017). Orexin activation counteracts decreases in nonexercise activity thermogenesis (NEAT) caused by high-fat diet. J. Virol. Methods *176*, 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040.

Chasman, D., and Adams, R.M. (2001). Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation11Edited by F. Cohen. J. Mol. Biol. *307*, 683–706. https://doi.org/10.1006/jmbi.2001.4510.

Chen, F., Zhao, Y., Fan, C., and Zhao, Y. (2015). Mismatch extension of DNA polymerases and high-accuracy single nucleotide polymorphism diagnostics by gold nanoparticle-improved isothermal amplification. Anal. Chem. *87*, 8718–8723. https://doi.org/10.1021/acs.analchem.5b01545.

Chen, R.P., Blackstock, D., Sun, Q., and Chen, W. (2018). Dynamic protein assembly by programmable DNA strand displacement. Nat. Chem. *10*, 474–481. https://doi.org/10.1038/s41557-018-0016-9.

Dirks, R.M., Bois, J.S., Schaeffer, J.M., Winfree, E., and Pierce, N.A. (2007). Thermodynamic analysis of interacting nucleic acid strands. SIAM Rev. *49*, 65–88. https://doi.org/10.1137/060651100.

Duan, Y., Yang, Y., Wang, Y., Pan, X., Wu, J., Cai, Y., Li, T., Zhao, D., Wang, J., and Zhou, M. (2016). Loop-mediated isothermal amplification for the rapid detection of the F200Y mutant genotype of carbendazim-resistant isolates of Sclerotinia sclerotiorum. Plant Dis. *100*, 976–983. https://doi.org/10.1094/PDIS-10-15-1185-RE.

Duffy, D., Mottez, E., Ainsworth, S., Buivan, T.P., Baudin, A., Vray, M., Reed, B., Fontanet, A., Rohel, A., Petrov-Sanchez, V., et al. (2017). An in vitro diagnostic certified point of care single nucleotide test for IL28B polymorphisms. PLoS One *12*, e0183084. https://doi.org/10.1371/journal.pone.0183084.

Gao, C., Zhuang, J., Zhou, C., Li, H., Liu, C., Liu, L., Feng, F., Liu, R., and Sun, C. (2019). SNP mutation-related genes in breast cancer for monitoring and prognosis of patients: a study based on the TCGA database. Cancer Med. *8*, 2303–2312. https://doi.org/10.1002/cam4.2065.

Gisler, F.M., von Kanel, T., Kraemer, R., Schaller, A., and Gallati, S. (2013). Identification of SNPs in the cystic fibrosis interactome influencing pulmonary progression in cystic fibrosis. Eur. J. Hum. Genet. *21*, 397–403. https://doi.org/10.1038/ejhg.2012.181.

Hogan, C.A., Sahoo, M.K., Huang, C., Garamani, N., Stevens, B., Zehnder, J., and Pinsky, B.A. (2020). Five-minute point-of-care testing for SARS-CoV-2: not there yet. J. Clin. Virol. *128*, 104410. https://doi.org/10.1016/j.jcv.2020.104410.

Jiang, Y.S., Bhadra, S., Li, B., Wu, Y.R., Milligan, J.N., and Ellington, A.D. (2015). Robust strand exchange reactions for the sequence-specific, real-time

detection of nucleic acid amplicons. Anal. Chem. *87*, 3314–3320. https://doi.org/10.1021/ac504387c.

Kobayashi, G.S., Brito, L.A., Moreira, DdP., Suzuki, A.M., Hsia, G.S.P., Pimentel, L.F., de Paiva, A.P.B., Dias, C.R., Lourenço, N.C.V., Oliveira, B.A., et al. (2021). A novel saliva RT-LAMP workflow for rapid identification of COVID-19 cases and restraining viral spread. Diagnostics *11*, 1400. https://doi.org/10.3390/diagnostics11081400.

Koopaee, H.K., and Koshkoiyeh, A.E. (2014). SNPs genotyping technologies and their applications in farm animals breeding Programs: Review. Braz. Arch. Biol. Technol. *57*, 87–95. https://doi.org/10.1590/S1516-89132014000100013.

Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., et al. (2020). Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell *182*, 812–827.e19. https://doi.org/10.1016/j.cell.2020.06.043.

Centers for Disease Control and Prevention. Preparation of Viral Transport Medium. (2020). SOP#: DSR-052-05. https://www.cdc.gov/coronavirus/2019-ncov/downloads/Viral-Transport-Medium.pdf.

LAMP primer designing software: PrimerExplorer (n.d.). Available at: http://primerexplorer.jp/lampv5e (Accessed: July 27, 2020)

Li, Q., Luan, G., Guo, Q., and Liang, J. (2002). A new class of homogeneous nucleic acid probes based on specific displacement hybridization. Nucleic Acids Res. *30*, E5. https://doi.org/10.1093/nar/30.2.e5.

Liu, X., Zhang, C., Zhao, M., Liu, K., Li, H., Li, N., Gao, L., Yang, X., Ma, T., Zhu, J., et al. (2018). A direct isothermal amplification system adapted for rapid SNP genotyping of multifarious sample types. Biosens. Bioelectron. *115*, 70–76. https://doi.org/10.1016/j.bios.2018.05.021.

Marmur, J., and Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. *5*, 109–118. https://doi.org/10.1016/S0022-2836(62)80066-7.

Megaraj, V., Zhao, T., Paumi, C.M., Gerk, P.M., Kim, R.B., and Vore, M. (2011). Functional analysis of nonsynonymous single nucleotide polymorphisms of multidrug resistance-associated protein 2 (ABCC2). Pharmacogenetics Genom. *21*, 506–515. https://doi.org/10.1097/FPC.0b013e328348c786.

Morlan, J., Baker, J., and Sinicropi, D. (2009). Mutation detection by real-time PCR: a simple, robust and highly selective method. PLoS One *4*, e4584. https://doi.org/10.1371/journal.pone.0004584.

Nagamine, K., Hase, T., and Notomi, T. (2002). Accelerated reaction by loop-mediated isothermal amplification using loop primers. Mol. Cell. Probes *16*, 223–229. https://doi.org/10.1006/mcpr.2002.0415.

Newman, C.M., Ramuta, M.D., McLaughlin, M.T., Wiseman, R.W., Karl, J.A., Dudley, D.M., Stauss, M.R., Maddox, R.J., Weiler, A.M., Bliss, M.I., et al. (2021). Initial evaluation of a mobile SARS-CoV-2 RT-LAMP testing strategy. J. Biomol. Tech. *32*, 137–147. https://doi.org/10.7171/jbt.21-32-03-009.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., and Hase, T. (2000). Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. *28*, E63. https://doi.org/10.1093/nar/28.12.e63.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. *44*, D733–D745. https://doi.org/10.1093/nar/gkv1189.

Rahimi, Y., Bihamta, M.R., Taleei, A., Alipour, H., and Ingvarsson, P.K. (2019). Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. BMC Plant Biol. *19*, 541. https://doi.org/10.1186/s12870-019-2165-4.

Ramanathan, B., Jindal, H.M., Le, C.F., Gudimella, R., Anwar, A., Razali, R., Poole-Johnson, J., Manikam, R., and Sekaran, S.D. (2017). Next generation sequencing reveals the antibiotic resistant variants in the genome of Pseudomonas aeruginosa. PLoS One *12*, e0182524. https://doi.org/10.1371/journal.pone.0182524.

Srinivas, N., Ouldridge, T.E., Šulc, P., Schaeffer, J.M., Yurke, B., Louis, A.A., Doye, J.P.K., and Winfree, E. (2013). On the biophysics and kinetics of toehold-mediated DNA strand displacement. Nucleic Acids Res. *41*, 10641–10658. https://doi.org/10.1093/nar/gkt801.

Tahir, M., and Sardaraz, M. (2020). A fast and ScalableWorkflow for SNPs detection in genome sequences using Hadoop map-reduce. Genes *11*.

Tyagi, S., and Kramer, F.R. (1996). Molecular beacon probes that fluoresce on hybridiztion. Nature publishing Group *14*, 303–308.

Ueki, M., Kimura-Kataoka, K., Fujihara, J., Iida, R., Kawai, Y., Kusaka, A., Sasaki, T., Takeshita, H., and Yasuda, T. (2019). Evaluation of the functional effects of genetic variants–missense and nonsense SNPs, indels and copy number variations–in the gene encoding human deoxyribonuclease I potentially implicated in autoimmunity. Sci. Rep. *9*, 13660. https://doi.org/10.1038/s41598-019-49935-y.

Varoquaux, G., Buitinck, L., Louppe, G., Grisel, O., Pedregosa, F., and Mueller, A. (2015). Scikit-learn: machine learning in Python. GetMobile: Mobile Comput. Commun. *19*, 29–33. https://doi.org/10.1145/2786984.2786995.

Velders, A.H., Schoen, C., and Saggiomo, V. (2018). Loop-mediated isothermal amplification (LAMP) shield for Arduino DNA detection. BMC Res. Notes *11*, 93. https://doi.org/10.1186/s13104-018-3197-9.

Wallace, R.B., Shaffer, J., Murphy, R., Bonner, J., Hirose, T., and Itakura, K. (1979). Hybridization of synthetic oligodeoxyribonucleotides to Φ_X174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. *6*, 3543–3558. https://doi.org/10.1093/nar/6.11.3543.

Wang, G., Ding, X., Hu, J., Wu, W., Sun, J., and Mu, Y. (2017). Unusual isothermal multimerization and amplification by the strand-displacing DNA polymerases with reverse transcription activities. Sci. Rep. *7*, 13928. https://doi.org/10.1038/s41598-017-13324-0.

Wang, J.S., and Zhang, D.Y. (2015). Simulation-guided DNA probe design for consistently ultraspecific hybridization. Nat. Chem. *7*, 545–553. https://doi.org/10.1038/nchem.2266.

Wong, Y.P., Othman, S., Lau, Y.L., Radu, S., and Chee, H.Y. (2018). Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. J. Appl. Microbiol. *124*, 626–643. https://doi.org/10.1111/jam.13647.

Yuan, X.Y., Wang, Y.L., Meng, K., Zhang, Y.X., Xu, H.Y., and Ai, W. (2019). LAMP real-time turbidity detection for fowl adenovirus. BMC Vet. Res. *15*, 256. https://doi.org/10.1186/s12917-019-2015-5.

Zadeh, J.N., Steenberg, C.D., Bois, J.S., Wolfe, B.R., Pierce, M.B., Khan, A.R., Dirks, R.M., and Pierce, N.A. (2011). NUPACK: Analysis and design of nucleic acid systems. J. Comput. Chem. *32*, 170–173.

Zhang, Z., Zeng, D., Ma, H., Feng, G., Hu, J., He, L., Li, C., and Fan, C. (2010). A DNA-origami chip platform for label-free SNP genotyping using toehold-mediated strand displacement. Small *6*, 1854–1858. https://doi.org/10.1002/smll.201000908.

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
human pooled saliva	Innovative Research	Cat# IRHUSL50ML
single-donor human whole blood	Innovative Research	Cat# IWB1NAH10ML
Critical commercial assays		
WarmStart® LAMP Kit (DNA & RNA)	New England Biolabs	Cat# E1700L
SUPERaseIn™ RNase Inhibitor	Invitrogen	Cat# AM2696
HiScribe™ T7 High Yield RNA Synthesis Kit	New England Biolabs	Cat# E2040S
Experimental models: Cell lines		
MOLT-4	American Type Culture Collection	Cat# CRL-1582
SK-BR3	American Type Culture Collection	Cat# HTB-30
Oligonucleotides		
Oligonucleotides used in this study	This paper	See Table S2
Software and algorithms		
NUPACK	Zadeh et al., 2011	http://nupack.org/
Other		
Code for probe design and data analysis	This paper	10.5281/zenodo.6575233
Web application for probe design	This paper	https://snp-lamp-design.netlify.app

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Philip Romero (promero2@wisc.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited at Zenodo (10.5281/zenodo.6575233) and is publicly available as of the date of publication. The DOI is also listed in the key resources table.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

We subcultured MOLT-4 cells (American Type Culture Collection) in a 1:8 ratio every two days in RPMI-1640 Medium (Gibco) supplemented with 10% FBS (Gibco) and 1X Antibiotic-Antimycotic (Gibco) at 37°C and 5% CO_2. We subcultured SK-BR3 cells (American Type Culture Collection) in a 1:4 ratio every two days in DMEM, high glucose (Gibco) supplemented with 10% FBS and 1X Antibiotic-Antimycotic 37°C and 5% CO_2. MOLT-4 cells were derived from a 19-year-old male and SK-BR3 cells were derived 43-year-old female. Our cell lines were not authenticated.

METHOD DETAILS

LAMP primer design

We first identified SNP targets in target cell lines based on data from the Broad Institute Cancer Cell Line Encyclopedia (Barretina et al., 2012). We then used the genomic location to retrieve target WT and SNP sequences from the NCBI genome browser (O'Leary

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

et al., 2016). For each gene target, we designed LAMP primers to target these sequence regions using PrimerExplorer V5 software (LAMP primer designing software: PrimerExplorer, n.d.) (Eiken Chemical Co.), placing the SNP base within the LAMP dumbbell loop structure.

Thermodynamic simulations of probe set binding

We predicted the equilibrium concentration of each DNA complex using the *complexes* and *concentrations* functions from the NUPACK software package (Dirks et al., 2007). For each probe set, we simulated three conditions: one with 1μM of a linear DNA strand representing the SNP target sequence, one with 1μM of the WT sequence, and a background condition with no WT or SNP target present. For simplicity, we assumed that 1μM of each probe and sink strand is present in each simulation, as well as 65mM NaCl and 8mM $MgCl_2$, matching our SNP-LAMP conditions. We then calculated the percentage of free probeF strand in each simulation to predict the fitness of each probe using the following equation:

$$Fitness = SNP * \log\left(\frac{SNP}{Max(WT, Background)}\right) \quad \text{(Equation 1)}$$

Probe design space simulations

We simulated 10,000 probe sets by first randomly generating 10,000 pairs of SNP and target sequence pairs. We then created a probe set for each target sequence by truncating a random number of bases from each terminus of the SNP sequence and its complement, and the WT sequence and its complement. We ensured that every complex in each probe set had a duplex at least 6 bases long, and that the probe duplex contained a blunt end for fluorophore and quencher placement. We then performed the thermodynamic simulations described above to predict the fitness of each probe using Equation 1, listed above.

Approximate Tm calculations

We performed approximate Tm calculations for each DNA duplex in order to screen out highly unfit probes before more computationally intensive thermodynamic simulations. For duplexes greater than 13 bases in length, we used a formula derived by Wallace et al. (1979):

$$Tm = 64.9 + 41.0 * \left(\frac{countGC - 16.4}{countGC + countAT}\right) \quad \text{(Equation 2)}$$

For duplexes less than or equal to 13 bases in length, we used the Marmur rule (Marmur and Doty, 1962):

$$Tm = (4 * countGC) + (2 * countAT) \quad \text{(Equation 3)}$$

Principal component analysis

We performed principal component analysis using the sci-kit-learn package for Python 3.3 (Varoquaux et al., 2015). We calculated a set of 12 features for each probe representing the binding affinity of the probe and sink complexes using the approximate Tm calculations described above, along with several other relevant design aspects. Features are listed in Table S1. We fit our PCA to the 10,000 random simulated probe sets we generated, and reduced these 12 features to 2 principal components for plotting and for probe screening in our algorithm.

Probe set design by genetic algorithm

The genetic algorithm starts by generating an initial population of 128 probe sets that contain a random number of complementary bases on either side of the SNP base position. We include additional constraints that the probe duplex must be greater than 6 bases in length, containing a blunt end to accommodate the fluorophore:quencher pair, and that the bottom strands of both the probe and sink duplexes are completely covered by their complement. This initial population is then evolved by (1) filtering candidate designs to remove designs that do not occupy the high-fitness region of the PCA space (Figure S1), (2) evaluating each member's fitness using NUPACK (Dirks et al., 2007) and our design objective function, (3) selecting the top ranked probes as 'parents' for the next generation, (4) randomly crossing parents by selecting a probe complex from one and a sink complex from the other to generate 'children', (5) mutating the children by randomly adding/removing a single terminal base to generate new population, and (6) repeating steps 1-5 over 7 generations, halving the population size at each generation. After the genetic algorithm optimization, we identify the top design and perform hill climbing to exhaustively search the local design space and ensure we have reached a maxima. This final probe design should be highly optimized for SNP detection, with a high specificity and signal.

Probe set optimization by hill-climbing

In each hill-climbing iteration, we generated 16 possible mutant probe sets by adding or removing 1 base from each terminus of each probe and sink strand. We performed NUPACK thermodynamic simulations on each mutant probe set and calculated its fitness as in

Please cite this article in press as: Hyman et al., Competitive SNP-LAMP probes for rapid and robust single-nucleotide polymorphism detection, Cell Reports Methods (2022), https://doi.org/10.1016/j.crmeth.2022.100242

our genetic algorithm. We then moved to the mutant probe set with the greatest fitness gain over the current probe, and repeated the hill-climbing algorithm. When none of the mutant probe sets had a greater fitness than the current one, we returned the current probe set as the final probe set design.

Upon completion of our algorithm, we retrieved the fittest probe in the final generation, and added a fluorophore and quencher to a blunt end of the probe duplex. We then added poly-T tails to all unmodified 3′ termini in the probe and sink complexes in order to prevent polymerase extension on the LAMP product.

DNA complexes and primer mixes

We ordered all DNA oligos from Integrated DNA Technologies (Coralville, Iowa), and dissolved each into nuclease-free water (Thermo Fisher) prior to storage at −20°C. We prepared LAMP primer mix and probe set mix stocks in nuclease-free water (Thermo Fisher) and stored at −20°C. On the day of experiment, we thawed each mixture on ice while protecting from light, and subsequently added to LAMP reactions.

Production of mRNAs using in *vitro* transcription

DNA templates for SARS-CoV-2 614D and 614G fragments were synthesized by Integrated DNA Technologies (Coralville, Iowa), each containing an upstream T7 promoter. We PCR amplified each fragment using Phusion® High-Fidelity DNA Polymerase (New England Biolabs) and purified the resulting amplicons with the DNA Clean & Concentrator-5 kit (Zymo Research). We then performed *in vitro* transcription from these amplicons using the HiScribe™ T7 High Yield RNA Synthesis Kit (New England Biolabs), and purified the resulting RNA using a GeneJET RNA Purification Kit (Thermo Scientific). We quantified each RNA sample's concentration using a NanoDrop™ Spectrophotometer (Thermo Scientific) and stored RNA stocks at −80°C in RNase-free water. Primer and DNA fragment sequences are given in Table S2.

Total RNA extraction from cell lines

We subcultured MOLT-4 cells and SK-BR3 cells as described above. We collected approximately 2.5 million cells of each type and purified their total RNA using a GeneJET RNA Purification Kit (Thermo Scientific). We quantified the concentration of each RNA sample using a NanoDrop™ Spectrophotometer (Thermo Scientific) and stored RNA stocks at.

−80°C in RNase-free water.

SNP-LAMP and annealing assays

We performed SNP-LAMP and annealing assays in triplicate using a Bio Rad CFX Connect qPCR machine. Except in anneal-only experiments, we incubated the reactions at 65°C to allow LAMP amplification while monitoring FAM, HEX, and/or SYBR fluorescence channels. We subsequently heated the reaction to 95°C for two minutes. We then annealed the probes by lowering the temperature by 1°C every 30 s and monitored FAM or HEX fluorescence channels at each step. For SARS-CoV-2 variant mixture experiments, limit of detection experiments, and crude sample assays, we reduced this annealing cycle to 10 s per degree to speed up the workflow. For these experiments, we used a two-step reaction workflow where we ran LAMP in 10 μL at 65ºC for 30 min, added probes to a final volume of 20μL and a concentration of 1μM per probe strand, and then performed the anneal cycle. Each reaction comprised a total volume of 10μL, consisting of 1X WarmStart LAMP Master Mix (New England Biolabs), and 0.5 U/μL SUPERase•In™ RNase Inhibitor (Invitrogen), with primer concentrations given in Table S3 and probe concentrations given in Table S4. Primer, target oligo, and probe set sequences are given in Table S2. LAMP durations and input RNA amounts are listed in Table S3, while probe and target oligo concentrations are listed in Table S4. In cases where no target oligo or RNA was present, we added water as a non-template control.

Crude biological sample preparation

We performed limit of detection and crude sample experiments by spiking in known quantities of *in vitro* transcribed SARS-CoV-2 spike 614D or 614G RNA fragments into water, viral transport media (VTM) prepared following CDC guidelines (Centers for Disease Control and Prevention, 2020), human pooled saliva (Innovative Research), and lysed single-donor human whole blood (Innovative Research). To prevent RNA degradation, 20mM Tris at pH 8.0, 133μM DTT and 400μM Guanidinium were present in the VTM, saliva, and blood mixtures prior to RNA addition. We diluted saliva 1:3 before spiking in RNA at the concentrations listed, heat-inactivating for 3 min at 95ºC, and adding to the LAMP reactions. We diluted VTM 1:3, heat inactivated before spiking in RNA at the concentrations listed, and added to the LAMP reactions. For blood, we diluted 1:5 with 1% Triton X-100, added RNA at the concentrations listed, heat inactivated, spun down the samples in a mini benchtop centrifuge for 2 min, and added the supernatant to the LAMP reactions. In all cases, we added 1μL of each crude sample directly to each 10μL LAMP reaction.

Sanger sequencing of LAMP amplicons

Using the *MT-CO2*, *ACTB*, *TP53*, and *NOTCH1* LAMP primers listed in Table S2, we performed LAMP amplification from 10ng of MOLT-4 or SK-BR-3 total RNA in 25μL reactions for 90 min, as described above. We verified reaction completion by eye using turbidity (Yuan et al., 2019). Upon reaction completion, we added 10μg of RNase A (Thermo Fisher) and incubated at 37°C for 15 min to destroy cellular RNA. We then purified products from each LAMP reaction using a Zymo DNA Clean & Concentrator-5 kit, eluting in nuclease-free water. We quantified each product using a NanoDrop™ Spectrophotometer (Thermo Scientific) and

submitted for sanger sequencing using the product's corresponding FIP and BIP primers listed in Table S2. In the case of *MT-CO2,* the primers used for sequencing differed from those used in SNP-LAMP experiments, as we later found a LAMP primer set with superior reaction speed. However, both primer sets targeted the same SNP mutation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical testing

To obtain the p values reported in this work, we first performed a two-tailed t-test to verify that a significant effect is present. When the p value for this test was above 0.05, we reported it directly. Otherwise, we performed a one-sided t-test and reported the p value. We performed all tests with n = 3, assuming homoscedasticity.

ADDITIONAL RESOURCES

The web application for SNP-LAMP probe design is available at https://snp-lamp-design.netlify.app/.