



Cite this: *Sens. Diagn.*, 2026, 5, 242

Artificial intelligence-powered signal analysis of loop-mediated isothermal amplification (LAMP) for the screening of Kaposi sarcoma at the point of care

Darke Hull, ^a Juan Boza, ^a Jason Manning, ^a Xinying Chu, ^b Ethel Cesarman, ^c Aggrey Semeere, ^d Jeffrey Martin ^e and David Erickson ^{*bf}

Unlike the polymerase chain reaction (PCR), loop-mediated isothermal amplification (LAMP) lacks a consistent thermal cycle, making quantification particularly challenging. Previously, we demonstrated that LAMP can accurately diagnose Kaposi sarcoma (KS) from skin lesion biopsies at the point of care (receiver operating characteristic area under the curve (AUC) = 0.967). A common approach in LAMP analysis involves setting a minimum absorbance threshold and time cutoff for positivity, which can introduce bias. We present a less biased, automated signal processing approach involving the fitting of a signal curve to five, two-parameter algebraic function fits, and the training of an artificial intelligence (AI) model on those parameters and their variances. An extreme gradient boosting (XGB) model was trained and tested on a primary dataset consisting of 1317 LAMP curves (from 451 unique patient samples with replicates). Five-fold *k*-validation on the train/test set yielded an receiver operating curve (ROC) area under the curve (AUC) of 0.952 ± 0.029 . Each of the five-fold models were then validated on a separate secondary dataset of 966 LAMP curves (from 414 unique patient samples with replicates) and achieved an AUC of 0.950 ± 0.005 . While the traditional methodology (which did not implement *k*-validation or a test/train split) outperformed the AI model's train/test set performance, the AI model generalized better and achieved a higher accuracy on the validation set (0.950 ± 0.005 vs. 0.9347). It performed even better when the analysis was applied directly to the raw signal data without additional pre-processing steps such as artifact filtering. This suggests that the AI model is more generalizable to new data and is able to discriminate KS-present and KS-absent samples better than traditional methods.

Received 14th May 2025,
Accepted 12th October 2025

DOI: 10.1039/d5sd00068h

rsc.li/sensors

Introduction

Real-time loop mediated isothermal amplification (LAMP) faces challenges in unbiased, reliable, and quantitative clinical implementation. This is particularly important in instances for which quantitative thresholds such as viral load are required. This is the case for Kaposi sarcoma (KS), a cancer identified by characteristic skin lesions and for which

the Kaposi sarcoma-associated herpesvirus (KSHV) is a necessary but insufficient cause.¹ In 2020, the majority of KS cases (73.0%) and fatalities (86.6%) occurred in African countries.² In sub-Saharan Africa, the paucity of pathology services makes traditional diagnosis by histopathologic interpretation inaccessible. Previous work has leveraged the viral origin of KS to provide highly specific and sensitive screening (AUC = 0.967) based on KSHV viral load testing *via* LAMP on DNA extracts from skin lesion biopsies.³ As with all cancers, accuracy in new diagnostic methods is critical to avoid false positives (which could lead to needless use of chemotherapy) and false negatives (which could delay diagnosis or lead to failure in diagnosis).

Similar to the polymerase chain reaction (PCR), LAMP is a nucleic acid amplification method for DNA detection and quantitation, though standardized methods for signal processing have not been established. In real-time LAMP, the reaction is monitored to collect the entire amplification curve, which can then be used to determine the amplification

^a Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY 14850, USA

^b Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, 14850, USA. E-mail: de54@cornell.edu

^c Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY, 10021, USA

^d Infectious Diseases Institute, Makerere University College of Health Sciences, Kampala, Uganda

^e Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA 94143, USA

^f Division of Nutritional Science, Cornell University, Ithaca, NY 14850, USA



time and thereby quantitate the analyte. Compared to the polymerase chain reaction (PCR), which discretely replicates DNA through thermocycling, LAMP works isothermally, continuously and semi-unpredictably amplifying DNA. Researchers leverage the PCR's cycle-dependency to determine amplification at a particular cycle threshold (CT) value. The parallel for LAMP curves is the time to threshold (T_t). Positive CT and T_t values both correspond to the cycle/time where the amplification signal, which is proportional to the initial DNA copies, surpasses a threshold.⁴ There is not, however, a standardized way of assigning T_t or its corresponding threshold. The most basic method is choosing a threshold which gives a desirable sensitivity/specificity when applied to the data to find T_t . Systematic thresholds have been employed; one LAMP study defined a positive result as the assay color exceeding the baseline plus three standard deviations of the baseline noise.⁵ Another systematic method assigned the threshold as the average fluorescence of the background and positive control.⁶ A review on LAMP technologies, Moehling 2021, concluded that LAMP is best suited for qualitative analysis rather than quantitation because of its narrow quantitative range and low linearity at lower concentrations.⁷ Despite this, less biased methods of quantification have been developed for both the PCR and LAMP. For the scope of this paper, elements which introduce bias refer to analysis and modelling decisions made by a researcher, which ultimately affect the final reported accuracy of a method. While bias is inevitable, avoidable forms of bias include choosing a threshold post-hoc, choosing models which overfit data, and not employing a test/train split.

Several studies have leveraged mathematical transformations and fits to further reduce bias. One study found the inflection point using the second derivative of the LAMP signal and used that point as the T_t .⁸ Similarly, another study found the maximum of the second derivative of a 7-parameter sigmoid function fit.⁹ Also using a sigmoid function, another study fit the LAMP signal to a four-parameter sigmoid function, and used the corresponding time point to the average of the maximum and minimum signal as the T_t .¹⁰ Algebraic function fitting has also been used with the PCR. One study fit an exponential function to the exponential portion of PCR curves and used it to calculate amplification efficiency.¹¹ Another compared fitting logistic, sigmoid, Gompertz, and Chapman functions to qRT-PCR signals for quantification.¹² Similarly, another study used a logistic function, Richard's curve, on LAMP signals, defining TTP as the intersection of the linear extrapolation of the maximum slope and the starting signal value.¹³ Researchers have leveraged artificial intelligence (AI) methods for the quantification of LAMP curves. Convolutional neural networks (CNNs) have been used to determine the signal of colorimetric LAMP assays from images of test tubes.¹⁴ Recently, a paper and preprint describing a custom LAMP platform employing YOLOv8 (an CNN-based object detection model) was effective for the detection of antimicrobial

resistance genes in UTI-causing gram-negative bacteria.^{15,16} Decision tree algorithms with binary outputs (pregnant/not pregnant) have been trained on LAMP features consisting of the rise up times (defined as the first time the absorbance exceeded 0.05) of three genes.¹⁷ Amplification curve analysis has leveraged k -nearest neighbour algorithms for a novel, data-driven multiplexing technique.¹⁸ The authors compared their method to the melting curve analysis (MCA), which performed slightly (3.41%) better. The authors concluded that while MCA is more accurate, it is not suitable for point of care (POC) scenarios where precise temperature control is difficult. In addition to AI LAMP methods, AI has been used in other DNA based diagnostics. AI has been applied to a wide range of cell-free-DNA diagnostics as a way to bolster these less invasive methods. One review of cell-free-DNA diagnostics, including prenatal and cancer screening, concludes that recent advances in gene sequencing and biomarker detection present immense promise for precision medicine.¹⁹ Interestingly, AI powered DNA detection is not limited to medical applications. One review on forensic DNA profiling discussed AI powered human identification methods and suggested that this technology for applications like the prediction of visual appearance, ancestry, and age present an objective and permanent progression of the field.²⁰ While applications of AI DNA detection are widespread, to our knowledge no other AI-based LAMP signal analysis methods have been described in the literature.

LAMP is more demanding of robust data analysis methodologies than the PCR because of its basis in time rather than cycles and its tendency of non-template amplification. To address this, several publications have developed less biased methods of LAMP data analysis, often employing algebraic fit-curves. Choosing which fit-curves to include and not employing a test/train data split are ways in which bias persists in these less biased methods. A method is presented for an unbiased classification of KS-present (positive) and KS-absent (negative) samples using an AI model trained on fitted algebraic curves. The model is compared to the method presented in the original publication of the data.³ The dataset used to train all models was gathered point-of-care (POC) from medical facilities in Africa, where artifacts such as voltage drops make quantification especially difficult.

Experimental

All experiments were performed in accordance with the Guidelines for ethical conduct of research, and the protocol was approved by the School of Biomedical Sciences Research Ethics Committee at Makerere University. Written informed consent was obtained from all participants in this study.

All code for this paper was written in python in Google Colaboratory. Data smoothing was done with the hampel package or a moving average. For curve fitting, cipy.optimize was used. The xgboost package was used for the gradient boost method. Several supporting methods for k -fold



validation and ROC analysis were used from sklearn. Google Gemini, Microsoft Copilot, and ChatGPT were used as coding and syntax tools. All generative AI tools were used with staunch skepticism and human verification.

The first dataset used herein (the test/train set) consisted of real-time LAMP curves of an assay targeting a sequence of KSHV DNA in skin lesion biopsies collected in Uganda.³ Following collection, skin punch biopsies were longitudinally bisected for histopathological diagnosis and nucleic acid testing on DNA extracted using the DNeasy blood and tissue kit (QIAGEN 69504). Briefly, tissue samples were enzymatically digested and DNA purification was accomplished *via* spin column membranes. For the LAMP assays, a 5 μ L aliquot of DNA extract was added to the LAMP mastermix targeting KSHV. LAMP reactions were performed at 68 $^{\circ}$ C for 50 minutes in the TINY device which tracked real-time fluorescence (Evagreen fluorescent dye, Biotium).²² The dataset consisted of 1317 LAMP runs, representing 451 unique patient samples with replicates. The replicate method measured each tissue biopsy in duplicate and assigned the T_t as the time at which a threshold was reached. For duplicates with T_t disparities in the top 5% of samples, two more samples were retested. Herein, for samples with more than two replicates in the dataset, the first two replicates were excluded. For each of the remaining replicates, the fitting parameter values were averaged. A test/train split was then applied on the patient level, after this replicate-removing method was finished to mitigate data leakage downstream.

Five-fold k -validation was performed on the test/train dataset. This is accomplished using a built in function in the sklearn package along with custom code. k -validation entails dividing the dataset k times (in this case 5 times) and training k separate models, each using $(k - 1)/k$ of the data and testing on the remaining $1/k$ of the data, so that each point is used as a testing point 1 time and a training point $k - 1$ times. The testing performance of each of the k models are then averaged. In this paper, deviation will be reported and calculated in the same manner as standard deviation, though it should be noted that since the distribution of the test/train sets are linearly dependent and semi-nonrandom, it cannot be considered a true standard deviation. The same is true for the reporting of validation AUC, which will be reported as the average performance of each of the models on the validation set, with a deviation calculated in the same manner as standard deviation.

The second dataset used herein (the validation set) consists of the subsequent 966 LAMP runs, representing 414 unique patients.²¹ This dataset was not used in the training or testing of the model, but was used to validate each of the five models generated by the 5-fold validation. The test/train set and validation set do not contain any shared patient ID's to mitigate data leakage. Both the train/test and validation sets are the TINY device's blue signal divided by its yellow signal.

The original method assigned T_t when the smoothed curve (moving average window = 10) increased by more

than the chosen slope threshold. Each of five, 2-parameter fitting functions were employed for the training of the AI model. Several models were applied to the LAMP dataset to evaluate prediction accuracy. Parameters for maximum depth, estimators and learning rate were varied between (1, 64) (1, 2000), and (0.01, 1), respectively to optimize AUC. An 80%/20% train/test split and 5-fold k -validation were employed. All signal processing steps (*e.g.* smoothing, baseline normalization, and artifact filtering) were applied systematically to all samples to address device-related signal artifacts and did not alter underlying biological signals.

Results and discussion

The curve was smoothed (moving average window = 10). Then, the minimum value from each curve was subtracted from the curve to normalize the minima to zero (Fig. 1). Qualitatively, after this processing the KS-present-positive (red) and KS-absent-negative (blue) samples appear better grouped, though still with much overlap between the two groups.

Each amplification curve was fit to a 2-parameter algebraic function: linear, quadratic, sigmoid, exponential (fit to 2.25

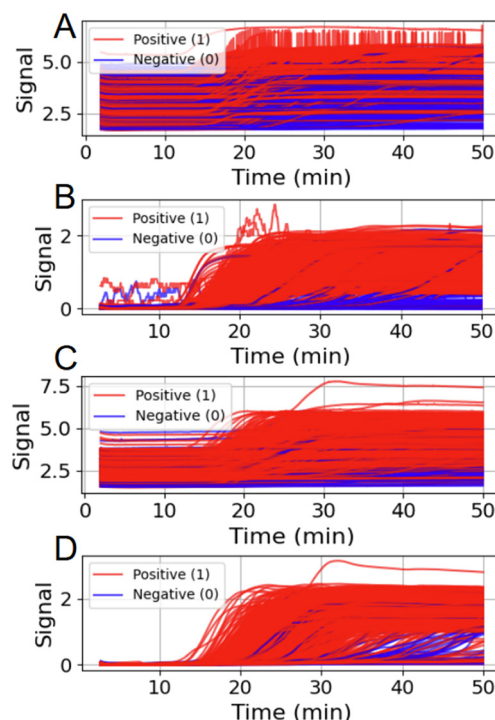


Fig. 1 The data cleaning method employed for curve cleaning. Fig. 1A shows the raw test/train set curves. Fig. 1B shows the smoothed curves (moving average window = 10), which have had the minimum values of each curve subtracted from the curve to normalize the minima to 0. Fig. 1C shows the validation set curves. Fig. 1D shows the smoothed curves (moving average window = 10) which have had the minimum values of each curve subtracted from the curve to normalize the minima to 0. For all plots, red samples are KS-present (positive) and blue samples are KS-absent (negative). This figure includes all LAMP runs, including replicates.



min), and Gompertz (SI1). Potential cutoffs were assigned as the average of each adjacent, ordered point. Each cutoff was evaluated and plotted, and the ROC curves and AUC values are shown in the SI figures (SI2).

In exploratory testing, other AI methods were applied to less success than XGB to include random forest, feed forward neural network, and perceptron algorithms. These exploratory models were evaluated based on classification accuracy¹⁸ (number correct/total) and employed a test/train split but not *k*-validation. The most successful model was extreme Gradient Boost (XGB), which was then optimized.

An extreme gradient boost model was trained on the fitting parameters from the 2-parameter mathematical models and the covariance of the fit generated from the `scipy.optimize` package's `curve_fit`. The model feature importances are disclosed in the SI (SI3). The featured importances are dominated by the steepness of the sigmoid fit (parameter *n*), suggesting that the model mainly predicts

based on how steeply sigmoidal the curve is. The optimized model achieved a 5-fold *k*-validation testing score AUC of 0.952, with a fold deviation of 0.029 (Fig. 2). The ROC curve was generated by assigning a likelihood of being positive to each sample, then plotting each point as a different minimum likelihood to be considered positive.

The XGB testing ROC curve falls below the originally published model when trained on the train/test set by less than one standard deviation of the folds at those false positive rates. However, when both models (the original method compared to the average of the 5-fold models) are validated on new, unseen data, the XGB validation ROC performs over a full standard deviation above the original method (Fig. 3). Moreover, the original method filtered for signal artifacts such as voltage drops, a systematic adjustment that may introduce bias similarly to how threshold fitting and determination does. The 5-fold XGB model performs nearly five (4.96) standard deviations above the validation set without artifact filtering (Fig. 3). Moreover, the 5-fold XGB model performs better on the test set than the original method without artifact filtering.

The gap in performance between the train/test set and validation set of the original model suggests that the method overfits to the original dataset. The shortcomings of the original model include the requirement of an artifact filtering step. This aspect introduces bias into the model that can be mitigated by proper employment of AI methods and demonstrably increase validated performance. The similarity between the average AI performance of the 5 folds between the train/test and validation sets suggests that it generalizes better to new, unseen data. The deviation of the performance of the folds of the AI model on the train/test set is likely due to the nature of 5-fold validation changing a large portion of the dataset each fold.

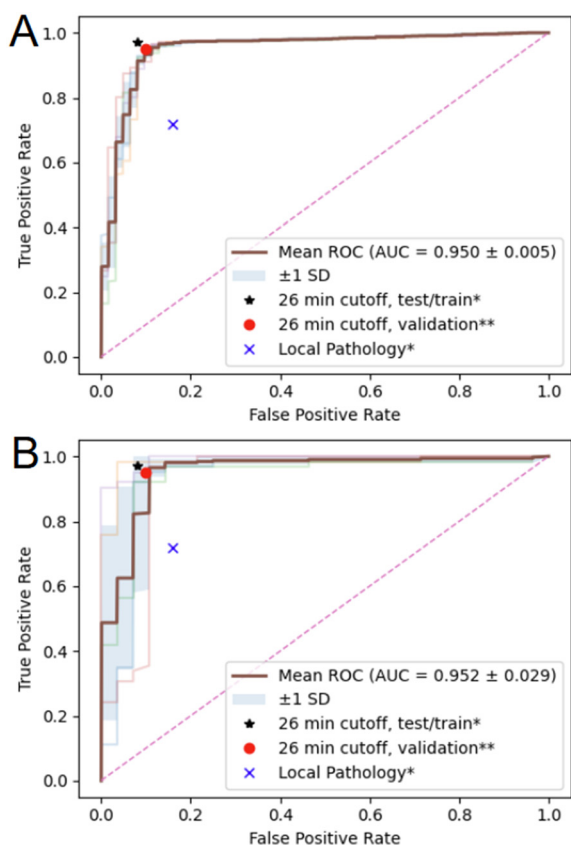


Fig. 2 A ROC curve for the 5-fold AI model trained on each 2-parameter algebraic fitting function and their fitting covariances of the validation set (A) and the test/train set (B). Each fold is denoted by a thin solid line, the average curve is denoted by a thicker brown line, and the shaded region represents one standard deviation of the values at a given false positive rate. False positive rate and true positive rate values from the original publication of the test/train data*, validation data**, and local pathology* are included for reference and labelled ref. 3* and 21**.

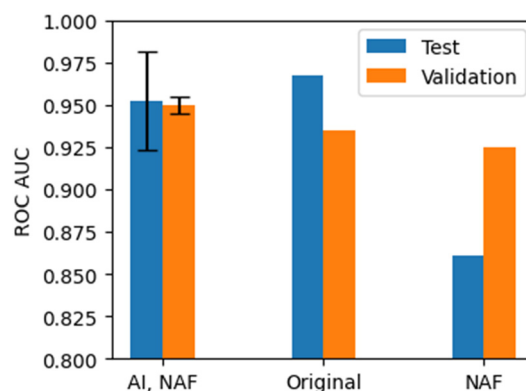


Fig. 3 ROC AUC values for the 5-fold XGB AI model which did not involve an artifact filtering step (AI, NAF), the original model with artifact filtration (original), and the original model with no artifact filtration step (NAF). Error bars on the AI model represent the calculated standard deviations of the AUCs of the 5 folds. It should be noted that since 5-fold validation set distributions are linearly dependent, it is not a true standard deviation but similarly represents the spread of the model performance. It should also be noted that the y axis spans the range of 0.8 to 1.



Conclusions

The proposed AI model presents an unbiased, highly generalizable method for LAMP signal analysis for KS diagnostic classification. The model folds perform 2.79 standard deviations above (94.9 vs. 93.5) the previously published model that involves an artifact filtration step and 4.96 standard deviations above a version of the previously published model that did not apply the artifact filtration step. This is important as false negatives are extremely detrimental to the patient because of low follow-up rates. The proposed model also takes measures to reduce bias by not determining a threshold, neither arbitrarily nor systematically, and fully automating the process. It should be noted, though, that the originally published method does not employ a test/train split or a validation dataset. The AI model was trained on 80% of the model and tested on 20% of the data, simulating how well the system represents new, unseen data. Through the process of *k*-validation, this was repeated a total of 5 times on 5 distinct models, and the results were averaged. Each of the 5 models were then applied to the unseen validation dataset, which performed on average higher than the original model on the same validation set. This suggests the new model may represent new, unseen data better than the original. While the AI model is useful diagnostically, it lacks the interpretability of the traditional method which provides a viral load. The original model, however, is theoretically able to screen more rapidly while the AI model requires the full 50-minute period. Further model optimization could reduce the required runtime, such as training the model on curtailed curves. Unlike several other methods, the AI method described here is mostly unbiased. There is, however, some bias in choosing which algebraic curves to train the AI model on. While this paper presents a novel method of data analysis, more development is required before it is ready for widespread use in TINY devices for KS diagnosis. Before the model could be implemented in the data analysis workflow for TINY devices, most practically as post-processing, the dataset should additionally be validated on a dataset that differs from the test/train set more than the validation set does. The most robust additional validation set would be at a new medical facility location, with additional, newly trained healthcare providers. Additional future work could include training a similar regression model to estimate viral load in the sample, since each sample is paired with PCR data. Such a model could reduce the bias of determining the T_t value on a dataset without a test/train split. The TINY is also currently exploratorily being used with colorimetric indicators, which may result in differing types of signal curves. Future work may include evaluating if the methods described in this paper extend to colorimetric LAMP indicators.

Exploration of novel methods of data analysis bolsters the effectiveness of POC devices without additional hardware complexity or cost, making their development critical alongside development of physical devices. While these

results are tailored for KS and are not directly transferable to other LAMP studies, similar methods could be employed for curve analysis on future projects.

Author contributions

Darke Hull – conceptualization, formal analysis, investigation, methodology, project administration, software, validation, visualization, and writing. Juan Boza – conceptualization, project administration, and supervision. Jason Manning – data curation, reviewing & editing. Xinying Chu – data curation and formal analysis. Ethel Cesarman – supervision, validation, writing, reviewing & editing. Jeffrey Martin – supervision, validation, writing, reviewing & editing. Aggrey Semeere – investigation, supervision, validation, writing, reviewing & editing. David Erickson – funding acquisition, validation, writing, reviewing & editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

All code for this project is available at <https://github.com/d-ericksonlab/KS-AI-Analysis>. Patient data will not be made publicly available.

Supplementary information (SI): the SI includes fitting functions, ROC curves for the fitting variables, and model feature importances. See DOI: <https://doi.org/10.1039/d5sd00068h>.

Acknowledgements

This work was supported by the U.S. National Institutes of Health/National Cancer Institute grant U01 CA269199, U.S. National Institutes of Health/National Institute of Biomedical Imaging and Bioengineering grant U54EB034654, and a training grant to D. Hull T32HD113301.

Notes and references

- 1 E. Cesarman, B. Damania, S. E. Krown, J. Martin, M. Bower and D. Whitby, *Nat. Rev. Dis. Primers*, 2019, **5**, 9.
- 2 L. Fu, T. Tian, B. Wang, Z. Lu, Y. Gao and Y. Sun, *Lancet Glob. Health*, 2023, **11**, e1566–e1575.
- 3 D. McCloskey, A. Semeere, R. Ayanga, M. Laker-Oketta, R. Lukande, M. Semakadde, M. Kanyesigye, M. Wenger, P. LeBoit, T. McCalmont, T. Maurer, A. Gardner, J. Boza, E. Cesarman, J. Martin and D. Erickson, *Sci. Adv.*, 2023, **9**, eadc8913.
- 4 I. P. Oscorbin, G. Y. Shevelev, K. A. Pronyaeva, A. A. Stepanov, D. V. Shamovskaya, O. V. Mishukova, D. V. Pyshnyi and M. L. Filipenko, *Int. J. Mol. Sci.*, 2021, **22**, 5743.
- 5 G. Papadakis, A. K. Pantazis, N. Fikas, S. Chatziioannidou, V. Tsiakalou, K. Michaelidou, V. Pogka, M. Megariti, M. Vardaki, K. Giarentis, J. Heaney, E. Nastouli, T. Karamitros, A. Mentis, A. Zafropoulos, G. Sourvinos, S. Agelaki and E. Gizeli, *Sci. Rep.*, 2022, **12**, 3775.



- 6 P. Hardinge and J. A. H. Murray, *Sci. Rep.*, 2019, **9**, 7400.
- 7 T. J. Moehling, G. Choi, L. C. Dugan, M. Salit and R. J. Meagher, *Expert Rev. Mol. Diagn.*, 2021, **21**, 43–61.
- 8 A. H. Buultjens, K. Vandellannoote, L. K. Sharkey, B. P. Howden, I. R. Monk, J. Y. H. Lee and T. P. Stinear, *ACS Biomater. Sci. Eng.*, 2021, **7**, 4982–4990.
- 9 H. Q. Nguyen, V. D. Nguyen, H. V. Nguyen and T. S. Seo, *Sci. Rep.*, 2020, **10**, 15123.
- 10 J. A. Hambalek, J. E. Kong, C. Brown, H. E. Munoz, T. Horn, M. Bogumil, E. Quick, A. Ozcan and D. Di Carlo, *ACS Sens.*, 2021, **6**, 3201–3210.
- 11 W. Liu and D. A. Saint, *Anal. Biochem.*, 2002, **302**, 52–59.
- 12 S. Zhao and R. D. Fernald, *J. Comput. Biol.*, 2005, **12**, 1047–1064.
- 13 S. Subramanian and R. D. Gomez, *PLoS One*, 2014, **9**, e100596.
- 14 M. A. Rohaim, E. Clayton, I. Sahin, J. Vilela, M. E. Khalifa, M. Q. Al-Natour, M. Bayoumi, A. C. Poirier, M. Branavan, M. Tharmakulasingam, N. S. Chaudhry, R. Sodi, A. Brown, P. Burkhart, W. Hacking, J. Botham, J. Boyce, H. Wilkinson, C. Williams and M. Munir, *Viruses*, 2020, **12**, 972.
- 15 B. Mezgebo, R. Chaffee, L. R. Castellanos, S. Ashraf, J. Burke-Gaffney, J. D. D. Pitout, B. I. Iorga, M. E. MacDonald and D. R. Pillai, *bioRxiv*, 2025, preprint, DOI: [10.1101/2025.03.21.644632](https://doi.org/10.1101/2025.03.21.644632).
- 16 L. R. Castellanos, R. Chaffee, H. Kumar, B. K. Mezgebo, P. Kassau, G. Peirano, J. D. D. Pitout, K. Kim and D. R. Pillai, *J. Clin. Microbiol.*, 2024, **62**, 11.
- 17 H. Kunii, T. Kubo, N. Asaoka, A. Z. Balboula, Y. Hamaguchi, T. Shimasaki, H. Bai, M. Kawahara, H. Kobayashi, H. Ogawa and M. Takahashi, *Biochem. Biophys. Res. Commun.*, 2021, **569**, 179–186.
- 18 K. Malpartida-Cardenas, L. Miglietta, T. Peng, A. Moniri, A. Holmes, P. Georgiou and J. Rodriguez-Manzano, *Sens. Diagn.*, 2022, **1**, 465–468.
- 19 W. H. A. Tsui, S. C. Ding, P. Jiang and Y. M. D. Lo, *Genome Res.*, 2025, **35**, 1–19.
- 20 M. Barash, D. McNevin, V. Fedorenko and P. Giverts, *Forensic Sci. Int.: Genet.*, 2024, **69**, 102994.
- 21 R. Snodgrass, A. Gardner, L. Jiang, C. Fu, E. Cesarman and D. Erickson, *PLoS One*, 2016, **11**, e0147636.
- 22 X. Chu, J. C. Manning, J. M. Boza, M. McGaskey, R. Ayanga, H. Muwando, M. Wenger, R. Lukande, T. Maurer, E. Cesarman, J. Martin, A. Semeere and D. Erickson, *Authorea*, 2025, preprint, DOI: [10.22541/au.175508845.57943650/v1](https://doi.org/10.22541/au.175508845.57943650/v1).

