Identifying SARS-CoV-2 Variants using Single-Molecule Conductance Measurements

Zahra Aminiranjbar[a,‡], Caglanaz Akin Gultakti[b,c,‡], Mashari Nasser Alangari[d,a], Yiren Wang[e], Busra Demir[b,c], Zeynep Koker[b], Arindam K. Das[e,f], M. P. Anantram[e], Ersin Emre Oren[b,c,*], Joshua Hihath[a,g,*]

[a] Department of Electrical and Computer Engineering, University of California Davis, Davis, California 95616, USA

[b] Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara 06560, Turkey

[c] Department of Materials Science & Nanotechnology Engineering, TOBB University of Economics and Technology, Ankara 06560, Turkey

[d] Department of Electrical Engineering, University of Hail, Hail 2440, Saudi Arabia

[e] Department of Electrical Engineering, University of Washington, Seattle, WA 98115, USA

[f] Department of Computer Science and Electrical Engineering, Eastern Washington University, Cheney, WA 99004,USA

[g] Center for Bioelectronics and Biosensors, School of Electrical, Computer, and Energy Engineering, Arizona State University, Phoenix, Arizona 85281, USA

*Email: eeoren@gmail.com (E. E. Oren); jhihath@asu.edu (J. Hihath)

‡ Z.A and C.A the authors contributed equally to this work.

Keywords: Biosensors, Molecular ElectronicsSARS-CoV-2 Variant Detection, Single-Molecule Break Junction, XGBoost Machine Learning

ABSTRACT: The global COVID-19 pandemic has highlighted the need for rapid, reliable, and efficient detection of biological agents and the necessity of tracking changes in genetic material as new SARS-CoV-2 variants emerge. Here we demonstrate that RNA-based, single-molecule conductance experiments can be used to identify specific variants of SARS-CoV-2. To this end, we i) select target sequences of interest for specific variants, ii) utilize single-molecule break junction measurements to obtain conductance histograms for each sequence and its potential mutations, and iii) employ the XGBoost machine learning classifier to rapidly identify the presence of target molecules in solution with a limited number of conductance traces. This approach allows high-specificity and high-sensitivity detection of RNA target sequences less than 20 base pairs in length by utilizing a complementary DNA probe capable of binding to the specific target. We use this approach to directly detect SARS-CoV-2 variants of concerns B.1.1.7 (Alpha), B.1.351 (Beta), B.1.617.2 (Delta), and B.1.1.529 (Omicron) and further demonstrate that the specific sequence conductance is sensitive to nucleotide mismatches, thus broadening the identification capabilities of the system. Thus, our experimental methodology detects specific SARS-CoV-2 variants, as well as recognizes the emergence of new variants as they arise.

Since the SARS-CoV-2 virus was identified in December 2019 it has spread worldwide, caused nearly seven million fatalities as of October 2023, and has remained a geopolitical, economic, and health issue globally.[1] Much effort during this time has been devoted to developing vaccines to trigger an immune response to the SARS-CoV-2 spike glycoprotein's receptor-binding domain (RBD).[2–8] Unfortunately, as the virus continues to thrive in the human ecosystem new variants of the SARS-CoV-2 virus are continuing to emerge around the globe. Among the emerging variations B.1.1.7 (Alpha), B.1.351 (Beta), B.1.617.2 (Delta), and B.1.1.529 (Omicron) are in the class identified as variants of concern (VOCs).[9–11] Moreover, as modifications to the genome compound in the VOCs, the efficacy of the initial vaccines, which were designed to combat the virus's original genome, has diminished. The quantity, position, and type of mutations on the receptor-binding domain can affect vaccine efficacy over time, and some strains may eventually evade the vaccine induced immunity and trigger another pandemic wave.[12–19] Thus, it is important to monitor circulating variants to track the spread of the disease and implement containment measures in the event of an outbreak. Currently, testing relies heavily on reverse transcription polymerase chain reaction (RT-PCR) and whole-genome sequencing techniques to detect, identify, and track VOCs.[20–23] However, this technique requires significant time and resources for diagnosis and mutation detection. Thus, the COVID-19 pandemic has emphasized the need for fast, sensitive, and cost-effective methods for both diagnostics and variant surveillance, and a new generation of techniques is emerging in an attempt to supplement conventional methods. These include Cas-based assays[24–27], electrochemical biosensors[28–30], FET-based biosensors[31–34], and others.[35–39] Here we present an electrical biosensor based on single-molecule conductance measurements that allow for sensitive,

robust, and cost-effective detection and identification of the VOCs.

The sensing platform is based on the single-molecule break junction (SMBJ) technique (Figure 1g), which has previously been used to demonstrate that: (i) the conductance of double-stranded RNA:DNA hybrids can be measured at the single molecule level[39–44], (ii) the conductance is sensitive to both sequence and base-pair mismatches[41,42], (iii) it is possible to detect and identify pathogenic bacterial strains using RNA sequences down to attomolar concentrations[41], and (iv) detection is possible in complex environments.[42] Thus, in light of the emergence of VOCs, here we examine the utility of single-molecule conductance measurements to identify specific SARS-CoV-2 variants, and to explore how this approach can be generalized to identify multiple targets simultaneously.

In the following section, we first discuss the design considerations involved in selecting DNA probes that are compatible with our measurement system and align with our objectives. Specifically, we examine the conductance of RNA:DNA hybrids designed to detect the wild-type (WT) SARS-CoV-2 virus, and the Alpha, Beta, Delta, and Omicron variants, focusing on spike protein mutations N501Y, E484K, and T478K respectively. We demonstrate that these target variants can be accurately identified even in the presence of interfering sequences. Furthermore, we highlight the potential for improvement in target identification by employing a machine learning model, XGBoost, which effectively identifies the presence of these targets within complex solutions containing off-target sequences. Notably, our approach achieves reliable results with a limited number of samples reads (<100), significantly reducing the overall processing time from sample input to obtaining results to mere seconds.

EXPERIMENTAL SECTION

To begin, we use genomic data from GISAID[45] to identify the mutations present in the VOCs. The genomic data indicates that each VOC has a unique set of mutations within the receptor binding domain (RBD) given in Table S1, and by targeting a subset of these mutations we can unambiguously identify the VOCs of interest. From this analysis we determine that detecting point mutations at residues 501, 484, and 478 within the spike protein allows us to identify wild-type (WT) SARS-CoV-2, and the Alpha, Beta, Delta, and Omicron VOCs (Table S2), provided the conductance values between perfectly matched and mismatched target sequences are distinguishable.

Thus, the next step is to identify target RNA sequences around each mutation (and thereby the corresponding DNA probe sequence) that would maximize our probability of identifying both the perfectly matched and mismatched sequences. An example of this process is shown in Figure 1a for the mutation at the 501st codon of the spike protein, where a Tyrosine (Y, UAU codon) has mutated into an Asparagine (N, AAU codon). As shown in Figure 1b, we focus on short (12 nucleotides) sequences that are well within the capture and measurement capabilities of the SMBJ approach[41–44,46–48], and then aim to (i) maximize the G:C content, as higher GC content in hybrids of a given length correlates with increased conductivity[42], (ii) centralize the mutation point within the sequence to ensure structural stability even when a mismatch is present and (iii) minimize the number of probable confounding sequences within the entire SARS-CoV-2 genome. Based on these criteria, we chose the RNA sequence CCC ACU UAU GGU as our target (Figure 1c) to identify the 501Y mutation, which is present in Alpha, Beta, and Omicron. Next, we chose the complementary DNA sequence as our probe, which we refer to as probe 501. This probe is also able to bind to the RNA sequence CCC ACU **A**AU GGU (501N, Figure 1d), which is present in WT and Delta, as a second target. In addition, we searched for this 12 bp target sequence within the rest of the genome and found no confounding sequences. Finally, since we must also be sensitive to other potential mutations that may occur in this sequence, we chose CCC ACU UA**C** GGU (Figure 1e) as an additional control sequence. In this case, the UAC codon still codes for Tyrosine (Y), so we refer to it as target 501Y'. Thus, the 501 DNA probe is able to bind to RNA targets 501Y, 501Y', and 501N. A similar approach is utilized for mutations at the 484 and 478 positions, and the amino acid and nucleotide sequences for these are given in Figs. S1 and S2, respectively.

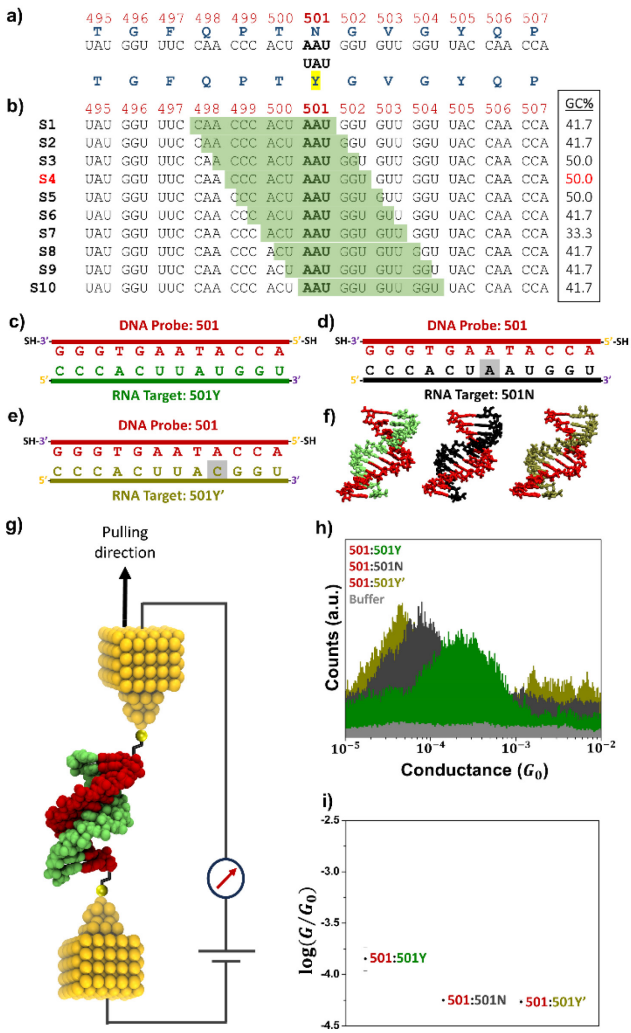

Figure 1. Mutation selection and conductance measurements. a) The section of the genome including the 501 mutation. b) 12 base pair sequence windows around each mutation and the corresponding GC content of each, the selected target is shown in red (S4). c) Designed DNA probe and RNA target for the 501Y mutation, d) the 501N wild-type RNA target, and e) the interfering 501Y' target sequence. f) 3D molecular representations from MD simulations of 501 DNA probe and 501Y, 501N and 501Y' RNA targets g) schematic of 501:501Y DNA:RNA hybrid between Au electrodes, h) Conductance histograms for the 501 DNA probe hybridized to 501Y, 501N and 501Y' RNA targets, i) Conductance values for the individual molecules, with error bars indicating standard error of the mean with N=4.

Previously, we showed that structural instability in RNA:DNA hybrids can result in unmeasurable conductance values[41]. Therefore, the stability of each selected sequence is a crucial factor in obtaining a reliable measurement. Thus, once the probe and targets are identified for a given point mutation, we examine the structural stability of each hybrid using molecular dynamics (MD) simulations. We perform 50 ns MD simulations, analyze the trajectories in terms of root mean square deviation (RMSD), and examine the change in hydrogen bonding over time. MD analyses, whose details are available in section 2 of the supplementary information, show that the selected probe-target hybrids are stable over the course of the simulations. In mismatched sequences, the RNA can still hybridize with its corresponding DNA probe, although forming the structure deviates from the perfectly matched case. Additionally, the melting temperatures of the molecules have been calculated and are presented in SI Table S3, indicating that the RNA:DNA hybrids remain stable at room temperature.

Finally, to ensure that the selected targets can be detected and measured we perform conductance measurements on each using the SMBJ approach (Fig. 1g). Here, 1000s of break-junction cycles are repeated, and then conductance traces with identifiable step regimes (SI: Figure S7) are combined to produce a conductance histogram. As can be seen in this case, each target gives a distinct conductance peak, though the 501:501N and the 501:501Y' hybrids have similar conductance values (Figure 1h and 1i), a point we will return to below. The unique conductance values observed can be attributed to the impact of mismatches on the conformation of RNA:DNA hybrids, as documented in literature.[41] The conductance axes are presented in terms of G_0 where G_0 is the conductance quantum $\frac{2e^2}{h}$ (e is an electron charge and h is the Planck constant).

RNA target sequences and the DNA probes for 484 and 478 mutations are shown in Figure 2a. The conductance results are shown in Figure 2b and 2c along with the conductance histograms resulting from the SMBJ measurements. Details of the SMBJ technique are provided in section 3 of the supplementary information. The 478 probe is designed to detect the 478K mutation, which is present in the Delta and Omicron VOCs, but can also detect the presence of the 478T target (mismatch) which exists in the WT. We also added an additional RNA target to the list which is named 478K' to examine possible interference. Similarly, for the 484 DNA probe, we have 484E (mismatch) present in Alpha, Delta, and WT (Table 1), and 484K (perfectly matched) that persists in Beta. In this case, we include 484K' as a possible mismatched interfering sequence and found a sequence in the SARS-CoV-2 genome that represents a second interference point with a single mismatch to the designed DNA probe (at the end of the sequence) from the Orf1ab protein. As such, we also test for this possibility.

RESULTS AND DISCUSSION

The conductance results from each of these sequences verify that we can distinguish between the perfectly matched complement to our probe sequence, and any of the tested mismatches (Figure 1h and 1i for 501 probe, Figure 2b and 2c for the 484 and 478 probes). Control experiments utilizing single-stranded DNA probe and RNAs indicate the absence of conductance peaks in the histograms within the designated conductance region (refer to Supplementary Information, Figure S10). Figures featuring Gaussian fits have been included in the SI section 3.5 to improve visual clarity. From these sequences, we can readily identify the perfectly matched cases allowing us to unambiguously detect specific VOCs (Figure 3). For example, if we test a sample with the three probes and find the perfectly matched 501Y but not the perfectly matched 484K or 478K we would conclude that the sample is COVID positive, and that it is the Alpha VOC. Similarly, (i) Beta can be identified if 501Y and 484K are present, (ii) for Delta only 478K would be detected, and (iii) for Omicron both the 501Y and 478K peaks would be present (shown as a representative case in Figure 3). Thus, by focusing only on the perfectly matched sequences that are designed to target specific mutations we can positively identify specific VOCs and as new VOCs emerge, new probes can be developed to directly target those (Figure S5).

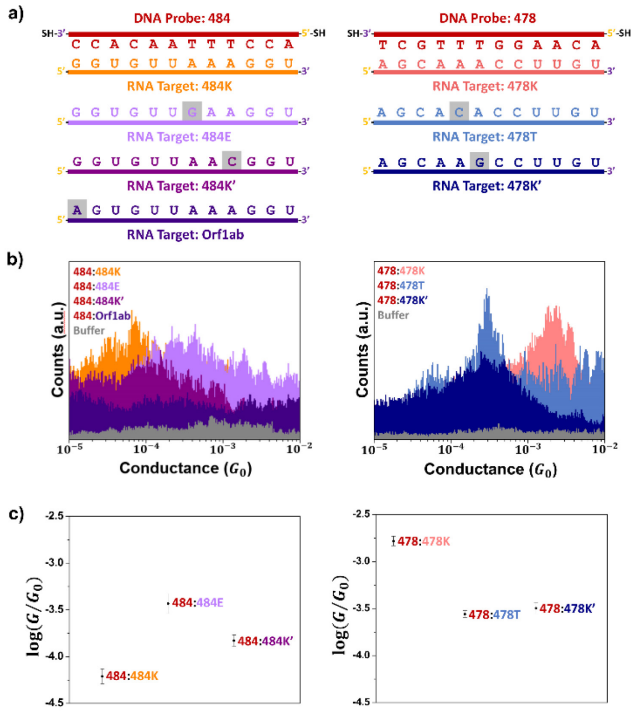

Figure 2. Sequences and conductance measurements for the 484K and 478K mutations. a) The 484 & 478 DNA probes and their complementary and interfering RNA targets, mismatched bases are highlighted. b) Conductance histograms of the DNA:RNA hybrids, color-coded by corresponding hybrids, and c) conductance values from a Gaussian fit of N=3 samples with error bars given by the standard error of the mean.

Although we can positively identify the targeted VOCs based on the perfectly matched sequences, the fact that we get well-defined conductance peaks from the mismatched test sequences (501Y', 484K', and 478K') that represent known variations (501N, 484E, and 478T), and the Orf1ab target raises three important questions. Can we positively identify other variants (e.g. WT) based on these results? Can we ensure a COVID positive result even if the new variants arise? And can we positively identify VOCs even if multiple strains are present in a sample (e.g. in wastewater samples)?

The primary issue in answering these questions is that in some cases the conductance values of the mismatches are not significantly different from one another (e.g., 501N and 501Y', Fig. 1h and 1i). However, if instead of focusing on the dominant peak value alone, we look at the conductance histogram as a "fingerprint" for a specific sequence, we can utilize machine learning techniques to try to distinguish between all possible variations for a given probe sequence. In this approach we note that while conductance values may be similar, the overall distributions of the conductance histograms are different for each case. Thus, we organize our machine learning algorithm to create histograms

from a limited number of traces, and then predict which sample it came from. This is repeated for a number of different traces per histogram. While it is not possible to determine exactly what features of the histogram the algorithm fits to, a generalized way of thinking of this approach is: how many traces does it take to create a histogram that can be identified as being similar to one of the "original" complete histograms?

Thus, to differentiate between all possible targets including the complementary and mismatched ones, and to automate and speed up the diagnosis process, we apply a machine learning algorithm previously developed by our group based on the XGBoost algorithm.[49–51] This approach decreases the number of conductance measurements required for identification while improving the detection and differentiation accuracy. Here, we design our classifier to detect each of the target sequences independently. To implement the algorithm, we first remove the SMBJ current traces that do not represent a molecular junction from the data sets by implementing an exponential curve fitting test. Therefore, any current traces that have an R^2 value larger than a set threshold (0.95) are discarded. This removes a sufficient number of traces without plateaus to enable the machine learning algorithm to accurately identify sequences. Next, we randomly sample H traces and construct 1000 probability histograms for each dataset, with 600 bins over the conductance range from $10^{-7.5}$ G_0 to $10^{-1.5}$ G_0. For each dataset, we use a train/test split ratio of 70/30, which uses 70% of traces to construct histograms and train the XGBoost[50] classifier, while 30% are used for testing. After training, the classifier can distinguish between each of the sequence targets with accuracy > 88.7% (484E vs. Orf1ab) or > 98.5% (the remaining 8 targets) with as only $H = 100$ randomly chosen raw current traces (Fig. 4a). Thus, instead of collecting thousands of traces as is typically done in SMBJ measurements to develop a conductance histogram (see Figs. 1 and 2), we can positively identify the target present with only 100 sample reads (conductance vs. time traces). To determine the minimum number of samples needed we examine the correlation between the accuracy of prediction for every probe:target hybrid and the number of raw current traces (H), as shown in Fig. 4b. Our findings suggest that by independently testing with 3 different probes and their corresponding sequences using 100 individual current traces each ($H = 100$), we can accurately verify not just the presence of the targeted VOCs, but also identify WT and Orf1ab. The results obtained from the three different probes for the ten targets considered are summarized in Table 1. The decision matrix produced by the 0s and 1s in this table uniquely determines the target. Note that the Orf1ab producing gene in the target list (gives a 0 only if COVID negative). However, positively identifying the new variant type would likely require a new probe, with the VOC detection strategy, which is discussed above.

Finally, to answer the third question above and determine whether SMBJ tests can positively identify specific targets even in the presence of multiple different strains in pooled samples (e.g. wastewater testing), we examine whether a target of interest can be identified in a sample containing both the perfectly matched target and other interfering RNA targets. To test this capability, we add each probe to a solution containing all the target RNA sequences relevant to that probe (Fig. 5a). In the 501 and 484 probe cases, because of the dispersion of the conductance histogram for each of the sequences, it is not possible to extract a specific peak for each of the RNA sequences in the sample. However, there is a significant shift in the peak position and overall conductance dispersion when the VOC target is present in the sample, and when it is not (Fig. 5a). This shift is attributed to an increase in counts in the conductance histogram, particularly around the conductance value position of the matching RNA. For instances like 501 and 484, where the conductance values are closely aligned, the shift is discernible primarily through statistical analysis. Examples of the fits are available in the SI section 3.6. A statistical T-test analysis indicates p-values of 0.005 (501) and 0.011 (484) for distinguishing between the cases when the perfectly matched target is present in the mixture and when it is not. In the case of the 478 probe, the peak at 1.65×10^{-3} G_0 appears only in a mixed sample where the perfectly matched sequence is present.

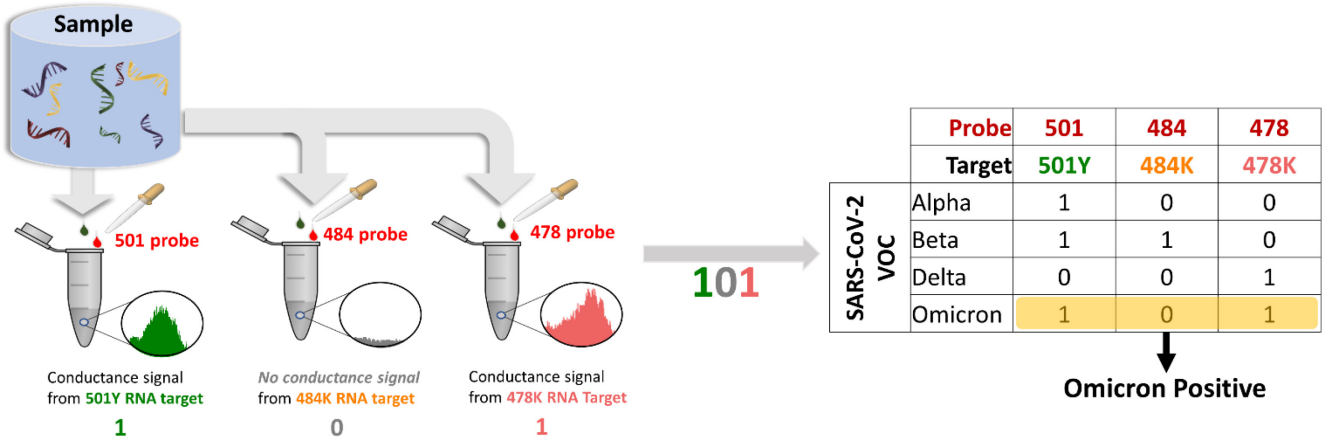

Figure 3. Envisioned process for variant identification using SMBJ conductance values. On the left, a sample is divided into three test samples, one with each of the DNA probes. If the conductance peak corresponding to the perfectly matched sequence for that probe is present, we can use the aggregated data to identify specific VOCs as shown in the table on the right.

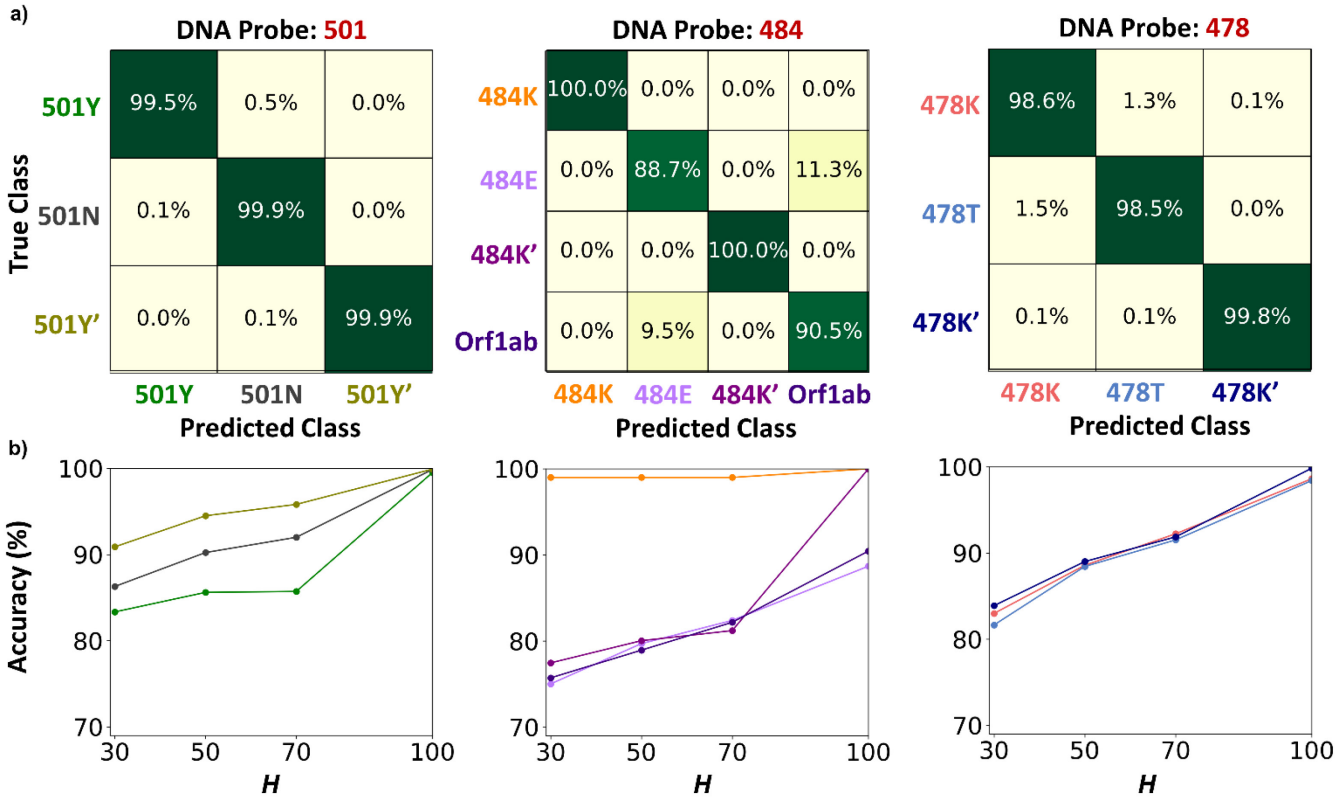

Figure 4. Envisioned process for variant identification using SMBJ conductance values. On the left, a sample is divided into three test samples, one with each of the DNA probes. If the conductance peak corresponding to the perfectly matched sequence for that probe is present, we can use the aggregated data to identify specific VOCs as shown in the table on the right.

While these p-values suggest robust detection of the perfectly matched sequences (and the corresponding VOCs) using conductance measurements with good confidence intervals, it would be preferable to limit the total number of break-junction cycles required to distinguish between cases where the mismatches and targets are both present and cases where only the mismatches are present. Thus, to automate and speed up the diagnostic process and to see if the confidence levels can be improved, we apply the machine learning approach discussed above to distinguish between these cases. Here, we train the classifier using two sets of samples: a mixed sample containing all target sequences for a given probe (both the matched and mismatched sequences) and samples with only the mismatches present. The same 70/30 train/test split ratio is applied to both sets of samples. Using H=100 current traces, the ML system achieves reasonable accuracy in identifying the presence of different targets in mixed samples. For the case of the 501Y target, the system correctly identifies its presence with 97.8% accuracy. Similarly, it achieves 99.7% accuracy for the 484K target of 484 probe and 99.7% for the 478K target (Fig. 5b).

Table 1. Table showing how each of the variants can be identified using results from the XGBoost ML approach.

Probe		501		484			478		All	
	Target	501Y	501N	484K	484E	Orf1ab	478K	478T	501Y', 484K' or 478K'	New Peak
VOC	Wild-Type	0	1	0	1	1	0	1	0	0
	Alpha	1	0	0	1	1	0	1	0	0
	Beta	1	0	1	0	1	0	1	0	0
	Delta	0	1	0	1	1	1	0	0	0
	Omicron	1	0	0	0	1	1	0	0	0
	New Variant	?	?	?	?	1	?	?	?	?
COVID-19 Negative		0	0	0	0	0	0	0	0	0

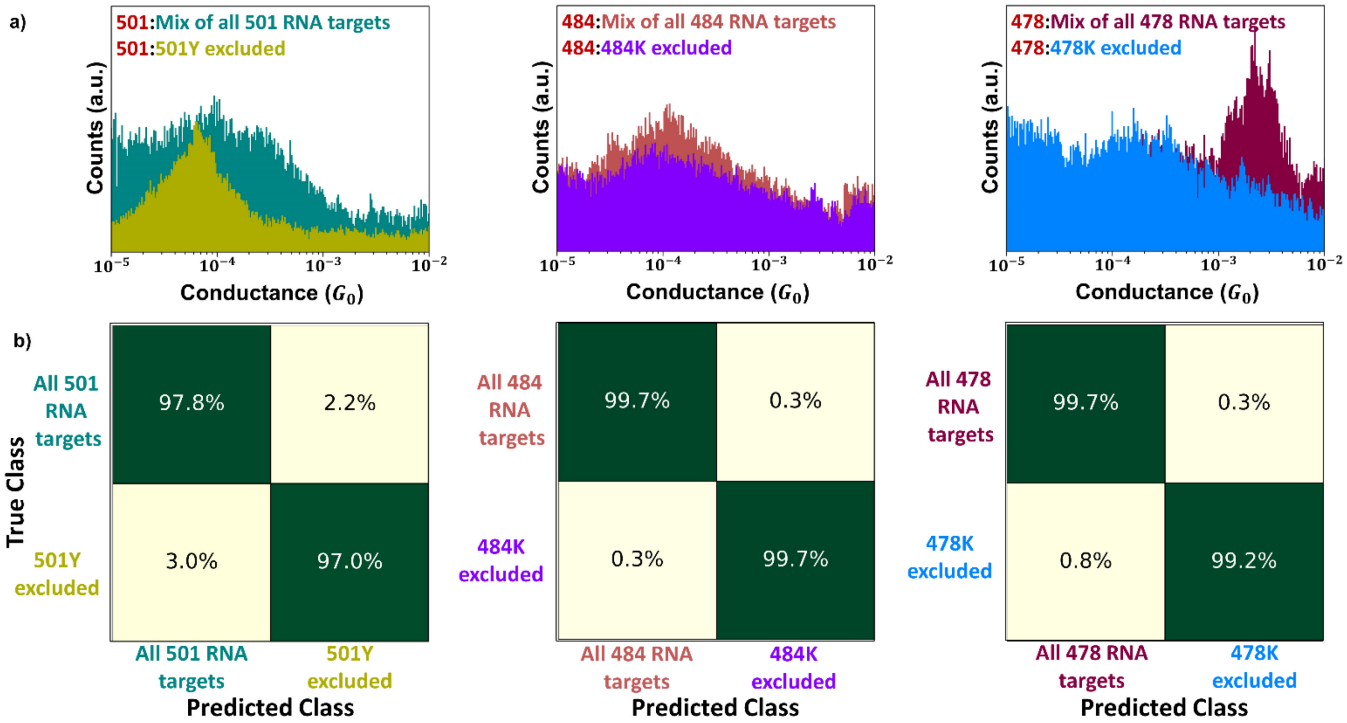

Figure 5. Mutation detection in mixed sample a) distinct conductance peaks in three probe cases, histograms include conductance of the mix of the corresponding RNA targets of each DNA probe, and one excluding the complementary target RNA molecule from sample b) XGBoost classifier confusion matrices based on H=100 raw current traces illustrate the classifier's ability to detect the variant in mixed samples.

To test the robustness of the classifier, we systematically vary the number of samples examined (Fig. S10). This analysis reveals that by increasing the number of samples, the accuracy of the ML system could be improved to the desired threshold. For example, to achieve 98% accuracy, H = 100 histograms are required for the 501Y case, while H = 50 histograms are sufficient for the 484K and 478K cases (Fig. S10). This indicates that even in complex samples, we are still able to positively identify specific VOCs that are perfectly matched with their probes, and to positively identify COVID infections regardless of the variant type. Furthermore, these findings demonstrate the potential of the ML approach to be adapted to emerging cases and to improve the accuracy by adjusting the number of current traces examined. By carefully selecting the appropriate number of current traces, the system can achieve high accuracies even in the presence of other targets in a complex environment. Overall, our ML approach offers fast processing capabilities for sample diagnostics with minimal data collection requirements, once sufficient training data has been generated.

CONCLUSION

In conclusion, this study presents the successful utilization of the single molecule break junction (SMBJ) technique as an electrical biosensor for the detection and identification of genetic material related to specific SARS-CoV-2 VOCs. By leveraging the specificity of the RNA sequences relative to the VOCs, we designed specific probes targeting key regions of the virus's genome, enabling the detection of specific mutants. The inherent resilience of the hybrids' conductance values to mismatches al-

lows for differentiation between multiple targets in mixed samples, thus expanding the capability to detect a wider range of mutations. This finding highlights the potential of the SMBJ technique to effectively identify and distinguish between various VOCs in complex sample environments. Furthermore, the study demonstrates the application of the XGBoost machine learning technique to enhance the diagnostic process. By utilizing XGBoost, we reduced the number of current traces required from the SMBJ system to allow accurate identification of the VOCs and broaden the detection capabilities. This reduction in data collection not only streamlines the diagnosis process but also paves the way for real-time applications for SMBJ-biosensors, enabling faster and more efficient detection of SARS-CoV-2 VOCs or other pathogens.

This combined experimental and data analysis approach greatly expands detection capabilities and reduces the possibility of both false negatives and false positives, which is important when attempting to track variants globally. The strategy is of potential use in scenarios where quick and accurate identification of variants is crucial, such as in clinical settings. The ability to detect specific sequences and identify or exclude specific variants with a small number of samples also reduces the time and resources needed for analysis, enabling prompt decision-making and potentially early intervention. Thus, this detection system enables the identification of specific VOCs, the ability to track their spread against previous variants, and has the potential to identify the emergence of new VOCs.

ASSOCIATED CONTENT

Supporting Information

Method for choosing the target regions of the genome, MD simulation method and simulation analysis, sample preparation and SMBJ technique, extra XGBoost Classification data are provided in the supporting information pdf.

AUTHOR INFORMATION

Corresponding Author

Ersin Emre Oren – *Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara, Turkey; Department of Materials Science and Nanotechnology Engineering, TOBB University of Economics and Technology, Ankara, Turkey;*
Email: eeoren@gmail.com
Josh Hihath – *Center for Bioelectronics and Biosensors, School of Electrical, Computer, and Energy Engineering, Arizona State University, Phoenix, Arizona, USA;*
Department of Electrical and Computer Engineering, University of California Davis, Davis, California, USA;
Email: jhihath@asu.edu

Authors

Zahra Aminiranjbar‡ – *Department of Electrical and Computer Engineering, University of California Davis, Davis, California, USA*
Caglanaz Akin Gultakti‡ – *Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara, Turkey; Department of Materials Science and Nanotechnology Engineering, TOBB University of Economics and Technology, Ankara, Turkey*
Mashari Nasser Alangari –*Department of Electrical Engineering, University of Hail, Hail, Saudi Arabia; Department of Electrical and Computer Engineering, University of California Davis, Davis, California, USA*
Yiren Wang – *Department of Electrical Engineering, University of Washington, Seattle, WA, USA*
Busra Demir – *Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara, Turkey; Department of Materials Science and Nanotechnology Engineering, TOBB University of Economics and Technology, Ankara, Turkey*
Zeynep Koker – *Bionanodesign Laboratory, Department of Biomedical Engineering, TOBB University of Economics and Technology, Ankara, Turkey*
Arindam K. Das – *Department of Computer Science and Electrical Engineering, Eastern Washington University, 99004, Cheney, WA, USA; Department of Electrical Engineering, University of Washington, Seattle, WA, USA*
M. P. Anantram – *Department of Electrical Engineering, University of Washington, Seattle, WA,*

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. Z.A and C.A contributed equally.

Funding Sources

J. Hihath acknowledges funding support from the National Science Foundation Future Manufacturing Program, NSF-2036865/2328217 and the Keck Foundation. *M. P. Anantram* acknowledges NSF Semisynbio grant number 2027165 and Future of Manufacturing grant number 2036865. Authors also acknowledge using TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

Authors acknowledge using TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

REFERENCES

(1) Kissler, S. M.; Tedijanto, C.; Goldstein, E.; Grad, Y. H.; Lipsitch, M. Projecting the Transmission Dynamics of SARS-CoV-2 through the Postpandemic Period. *Science* **2020**, *368* (6493), 860–868. https://doi.org/10.1126/science.abb5793.

(2) Kakavandi, S.; Zare, I.; VaezJalali, M.; Dadashi, M.; Azarian, M.; Akbari, A.; Ramezani Farani, M.; Zalpoor, H.; Hajikhani, B. Structural and Non-Structural Proteins in SARS-CoV-2: Potential Aspects to COVID-19 Treatment or Prevention of Progression of Related Diseases. *Cell Commun Signal* **2023**, *21* (1), 110. https://doi.org/10.1186/s12964-023-01104-5.

(3) Yang, J.; Wang, W.; Chen, Z.; Lu, S.; Yang, F.; Bi, Z.; Bao, L.; Mo, F.; Li, X.; Huang, Y.; Hong, W.; Yang, Y.; Zhao, Y.; Ye, F.; Lin, S.; Deng, W.; Chen, H.; Lei, H.; Zhang, Z.; Luo, M.; Gao, H.; Zheng, Y.; Gong, Y.; Jiang, X.; Xu, Y.; Lv, Q.; Li, D.; Wang, M.; Li, F.; Wang, S.; Wang, G.; Yu, P.; Qu, Y.; Yang, L.; Deng, H.; Tong, A.; Li, J.; Wang, Z.; Yang, J.; Shen, G.; Zhao, Z.; Li, Y.; Luo, J.; Liu, H.; Yu, W.; Yang, M.; Xu, J.; Wang, J.; Li, H.; Wang, H.; Kuang, D.; Lin, P.; Hu, Z.; Guo, W.; Cheng, W.; He, Y.; Song, X.; Chen, C.; Xue, Z.; Yao, S.; Chen, L.; Ma, X.; Chen, S.; Gou, M.; Huang, W.; Wang, Y.; Fan, C.; Tian, Z.; Shi, M.; Wang, F.-S.; Dai, L.; Wu, M.; Li, G.; Wang, G.; Peng, Y.; Qian, Z.; Huang, C.; Lau, J. Y.-N.; Yang, Z.; Wei, Y.; Cen, X.; Peng, X.; Qin, C.; Zhang, K.; Lu, G.; Wei, X. A Vaccine Targeting the RBD of the S Protein of SARS-CoV-2 Induces Protective Immunity. *Nature* **2020**, *586* (7830), 572–577. https://doi.org/10.1038/s41586-020-2599-8.

(4) Niu, L.; Wittrock, K. N.; Clabaugh, G. C.; Srivastava, V.; Cho, M. W. A Structural Landscape of Neutralizing Antibodies Against SARS-CoV-2 Receptor Binding Domain. *Front. Immunol.* **2021**, *12*, 647934. https://doi.org/10.3389/fimmu.2021.647934.

(5) Min, L.; Sun, Q. Antibodies and Vaccines Target RBD of SARS-CoV-2. *Front. Mol. Biosci.* **2021**, *8*, 671633. https://doi.org/10.3389/fmolb.2021.671633.

(6) Moro-Pérez, L.; Boggiano-Ayo, T.; Lozada-Chang, S. L.; Fernández-Saiz, O. L.; De La Luz, K. R.; Gómez-Pérez, J. A. Conformational Characterization of the Mammalian-Expressed SARS-CoV-2 Recombinant Receptor Binding Domain, a COVID-19 Vaccine. *Biol Res* **2023**, *56* (1), 22. https://doi.org/10.1186/s40659-023-00434-5.

(7) Dong, Y.; Dai, T.; Wei, Y.; Zhang, L.; Zheng, M.; Zhou, F. A Systematic Review of SARS-CoV-2 Vaccine Candidates. *Sig Transduct Target Ther* **2020**, *5* (1), 237. https://doi.org/10.1038/s41392-020-00352-y.

(8) Dickey, T. H.; Tang, W. K.; Butler, B.; Ouahes, T.; Orr-Gonzalez, S.; Salinas, N. D.; Lambert, L. E.; Tolia, N. H. Design of the SARS-CoV-2 RBD Vaccine Antigen Improves Neutralizing Antibody Response. *Sci. Adv.* **2022**, *8* (37), eabq8276. https://doi.org/10.1126/sciadv.abq8276.

(9) Akkız, H. The Biological Functions and Clinical Significance of SARS-CoV-2 Variants of Corcern. *Front. Med.* **2022**, *9*, 849217. https://doi.org/10.3389/fmed.2022.849217.

(10) Aleem, A.; Akbar Samad, A. B.; Vaqar, S. Emerging Variants of SARS-CoV-2 and Novel Therapeutics Against Coronavirus (COVID-19). In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2023.

(11) A Comparative Overview of SARS-CoV-2 and Its Variants of Concern. *Infez Med* **2022**, *30* (3). https://doi.org/10.53854/liim-3003-2.

(12) Abu-Raddad, L. J.; Chemaitelly, H.; Butt, A. A. Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351 Variants. *N Engl J Med* **2021**, *385* (2), 187–189. https://doi.org/10.1056/NEJMc2104974.

(13) Bates, T. A.; Leier, H. C.; Lyski, Z. L.; McBride, S. K.; Coulter, F. J.; Weinstein, J. B.; Goodman, J. R.; Lu, Z.; Siegel, S. A. R.; Sullivan, P.; Strnad, M.; Brunton, A. E.; Lee, D. X.; Adey, A. C.; Bimber, B. N.; O'Roak, B. J.; Curlin, M. E.; Messer, W. B.; Tafesse, F. G. Neutralization of SARS-CoV-2 Variants by Convalescent and BNT162b2 Vaccinated Serum. *Nat Commun* **2021**, *12* (1), 5135. https://doi.org/10.1038/s41467-021-25479-6.

(14) Becker, M.; Dulovic, A.; Junker, D.; Ruetalo, N.; Kaiser, P. D.; Pinilla, Y. T.; Heinzel, C.; Haering, J.; Traenkle, B.; Wagner, T. R.; Layer, M.; Mehrlaender, M.; Mirakaj, V.; Held, J.; Planatscher, H.; Schenke-Layland, K.; Krause, G.; Strengert, M.; Bakchoul, T.; Althaus, K.; Fendel, R.; Kreidenweiss, A.; Koeppen, M.; Rothbauer, U.; Schindler, M.; Schneiderhan-Marra, N. Immune Response to SARS-CoV-2 Variants of Concern in Vaccinated Individuals. *Nat Commun* **2021**, *12* (1), 3109. https://doi.org/10.1038/s41467-021-23473-6.

(15) Garcia-Beltran, W. F.; Lam, E. C.; St. Denis, K.; Nitido, A. D.; Garcia, Z. H.; Hauser, B. M.; Feldman, J.; Pavlovic, M. N.; Gregory, D. J.; Poznansky, M. C.; Sigal, A.; Schmidt, A. G.; Iafrate, A. J.; Naranbhai, V.; Balazs, A. B. Multiple SARS-CoV-2 Variants Escape Neutralization by Vaccine-Induced Humoral Immunity. *Cell* **2021**, *184* (9), 2372-2383.e9. https://doi.org/10.1016/j.cell.2021.03.013.

(16) Hoffmann, M.; Arora, P.; Groß, R.; Seidel, A.; Hörnich, B. F.; Hahn, A. S.; Krüger, N.; Graichen, L.; Hofmann-Winkler, H.; Kempf, A.; Winkler, M. S.; Schulz, S.; Jäck, H.-M.; Jahrsdörfer, B.; Schrezenmeier, H.; Müller, M.; Kleger, A.; Münch, J.; Pöhlmann, S. SARS-CoV-2 Variants B.1.351 and P.1 Escape from Neutralizing Antibodies. *Cell* **2021**, *184* (9), 2384-2393.e12. https://doi.org/10.1016/j.cell.2021.03.036.

(17) Singh, D. D.; Parveen, A.; Yadav, D. K. SARS-CoV-2: Emergence of New Variants and Effectiveness of Vaccines. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 777212. https://doi.org/10.3389/fcimb.2021.777212.

(18) Muik, A.; Wallisch, A.-K.; Sänger, B.; Swanson, K. A.; Mühl, J.; Chen, W.; Cai, H.; Maurus, D.; Sarkar, R.; Türeci, Ö.; Dormitzer, P. R.; Şahin, U. Neutralization of SARS-CoV-2 Lineage B.1.1.7 Pseudovirus by BNT162b2 Vaccine–Elicited Human Sera. *Science* **2021**, *371* (6534), 1152–1153. https://doi.org/10.1126/science.abg6105.

(19) Kuzmina, A.; Khalaila, Y.; Voloshin, O.; Keren-Naus, A.; Boehm-Cohen, L.; Raviv, Y.; Shemer-Avni, Y.; Rosenberg, E.; Taube, R. SARS-CoV-2 Spike Variants Exhibit Differential Infectivity and Neutralization Resistance to Convalescent or Post-Vaccination Sera. *Cell Host & Microbe* **2021**, *29* (4), 522-528.e2. https://doi.org/10.1016/j.chom.2021.03.008.

(20) Udugama, B.; Kadhiresan, P.; Kozlowski, H. N.; Malekjahani, A.; Osborne, M.; Li, V. Y. C.; Chen, H.; Mubareka, S.; Gubbay, J. B.; Chan, W. C. W. Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano* **2020**, *14* (4), 3822–3835. https://doi.org/10.1021/acsnano.0c02624.

(21) Maryam, S.; Ul Haq, I.; Yahya, G.; Ul Haq, M.; Algammal, A. M.; Saber, S.; Cavalu, S. COVID-19 Surveillance in Wastewater: An Epidemiological Tool for the Monitoring of SARS-CoV-2. *Front. Cell. Infect. Microbiol.* **2023**, *12*, 978643. https://doi.org/10.3389/fcimb.2022.978643.

(22) Sharma, S.; Shrivastava, S.; Kausley, S. B.; Rai, B.; Pandit, A. B. Coronavirus: A Comparative Analysis of Detection Technologies in the Wake of Emerging Variants. *Infection* **2023**, *51* (1), 1–19. https://doi.org/10.1007/s15010-022-01819-6.

(23) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G.; Haagmans, B. L.; Van Der Veer, B.; Van Den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J.-L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P.; Drosten, C. Detection of 2019 Novel Coronavirus (2019-nCoV) by Real-Time RT-PCR. *Eurosurveillance* **2020**, *25* (3). https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045.

(24) Broughton, J. P.; Deng, X.; Yu, G.; Fasching, C. L.; Servellita, V.; Singh, J.; Miao, X.; Streithorst, J. A.; Granados, A.; Sotomayor-Gonzalez, A.; Zorn, K.; Gopez, A.; Hsu, E.; Gu, W.; Miller, S.; Pan, C.-Y.; Guevara, H.; Wadford, D. A.; Chen, J. S.; Chiu, C. Y. CRISPR–Cas12-Based Detection of SARS-CoV-2. *Nat Biotechnol* **2020**, *38* (7), 870–874. https://doi.org/10.1038/s41587-020-0513-4.

(25) Park, J. S.; Hsieh, K.; Chen, L.; Kaushik, A.; Trick, A. Y.; Wang, T. Digital CRISPR/Cas-Assisted Assay for Rapid and Sensitive Detection of SARS-CoV-2. *Advanced Science* **2021**, *8* (5), 2003564. https://doi.org/10.1002/advs.202003564.

(26) Liu, H.; Chang, S.; Chen, S.; Du, Y.; Wang, H.; Wang, C.; Xiang, Y.; Wang, Q.; Li, Z.; Wang, S.; Qiu, S.; Song, H. Highly Sensitive and Rapid Detection of SARS-CoV-2 via a Portable CRISPR-Cas13a-based Lateral Flow Assay. *Journal of Medical Virology* **2022**, *94* (12), 5858–5866. https://doi.org/10.1002/jmv.28096.

(27) Figueiredo, D.; Cascalheira, A.; Goncalves, J. Rapid, Multiplex Detection of SARS-CoV-2 Using Isothermal Amplification Coupled with CRISPR-Cas12a. *Sci Rep* **2023**, *13* (1), 849. https://doi.org/10.1038/s41598-022-27133-7.

(28) Najjar, D.; Rainbow, J.; Sharma Timilsina, S.; Jolly, P.; De Puig, H.; Yafia, M.; Durr, N.; Sallum, H.; Alter, G.; Li, J. Z.; Yu, X. G.; Walt, D. R.; Paradiso, J. A.; Estrela, P.; Collins, J. J.; Ingber, D. E. A Lab-on-a-Chip for the Concurrent Electrochemical Detection of SARS-CoV-2 RNA and Anti-SARS-CoV-2 Antibodies in Saliva and Plasma. *Nat. Biomed. Eng* **2022**, *6* (8), 968–978. https://doi.org/10.1038/s41551-022-00919-w.

(29) Hussein, H. A.; Hanora, A.; Solyman, S. M.; Hassan, R. Y. A. Designing and Fabrication of Electrochemical Nano-Biosensor for the Fast Detection of SARS-CoV-2-RNA. *Sci Rep* **2023**, *13* (1), 5139. https://doi.org/10.1038/s41598-023-32168-5.

(30) Ji, D.; Guo, M.; Wu, Y.; Liu, W.; Luo, S.; Wang, X.; Kang, H.; Chen, Y.; Dai, C.; Kong, D.; Ma, H.; Liu, Y.; Wei, D. Electrochemical Detection of a Few Copies of Unamplified SARS-CoV-2 Nucleic Acids by a Self-Actuated Molecular System. *J. Am. Chem. Soc.* **2022**, *144* (30), 13526–13537. https://doi.org/10.1021/jacs.2c02884.

(31) Wang, L.; Wang, X.; Wu, Y.; Guo, M.; Gu, C.; Dai, C.; Kong, D.; Wang, Y.; Zhang, C.; Qu, D.; Fan, C.; Xie, Y.; Zhu, Z.; Liu, Y.; Wei, D. Rapid and Ultrasensitive Electromechanical Detection of Ions, Biomolecules and SARS-CoV-2 RNA in Unamplified Samples. *Nat. Biomed. Eng* **2022**, *6* (3), 276–285. https://doi.org/10.1038/s41551-021-00833-7.

(32) Wang, X.; Kong, D.; Guo, M.; Wang, L.; Gu, C.; Dai, C.; Wang, Y.; Jiang, Q.; Ai, Z.; Zhang, C.; Qu, D.; Xie, Y.; Zhu, Z.; Liu, Y.; Wei, D. Rapid SARS-CoV-2 Nucleic Acid Testing and Pooled Assay by Tetrahedral DNA Nanostructure Transistor. *Nano Lett.* **2021**, *21* (22), 9450–9457. https://doi.org/10.1021/acs.nanolett.1c02748.

(33) Wu, Y.; Ji, D.; Dai, C.; Kong, D.; Chen, Y.; Wang, L.; Guo, M.; Liu, Y.; Wei, D. Triple-Probe DNA Framework-Based Transistor for SARS-CoV-2 10-in-1 Pooled Testing. *Nano Lett.* **2022**, *22* (8), 3307–3316. https://doi.org/10.1021/acs.nanolett.2c00415.

(34) Wang, H.; Liu, J.; Wei, J.; Xiao, K.; Chen, Y.; Jiang, Y.-L.; Wan, J. Au Nanoparticles/HfO_2/Fully Depleted Silicon-on-Insulator MOSFET Enabled Rapid Detection of Zeptomole COVID-19 Gene With Electrostatic Enrichment Process. *IEEE Trans. Electron Devices* **2023**, *70* (3), 1236–1242. https://doi.org/10.1109/TED.2022.3233544.

(35) Zhang, T.; Deng, R.; Wang, Y.; Wu, C.; Zhang, K.; Wang, C.; Gong, N.; Ledesma-Amaro, R.; Teng, X.; Yang, C.; Xue, T.; Zhang, Y.; Hu, Y.; He, Q.; Li, W.; Li, J. A Paper-Based Assay for the Colorimetric Detection of SARS-CoV-2 Variants at Single-Nucleotide Resolution. *Nat. Biomed. Eng* **2022**, *6* (8), 957–967. https://doi.org/10.1038/s41551-022-00907-0.

(36) AbdElFatah, T.; Jalali, M.; Yedire, S. G.; I. Hosseini, I.; Del Real Mata, C.; Khan, H.; Hamidi, S. V.; Jeanne, O.; Siavash Moakhar, R.; McLean, M.; Patel, D.; Wang, Z.; McKay, G.; Yousefi, M.; Nguyen, D.; Vidal, S. M.; Liang, C.; Mahshid, S. Nanoplasmonic Amplification in Microfluidics Enables Accelerated Colorimetric Quantification of Nucleic Acid Biomarkers from Pathogens. *Nat. Nanotechnol.* **2023**. https://doi.org/10.1038/s41565-023-01384-5.

(37) Arizti-Sanz, J.; Bradley, A.; Zhang, Y. B.; Boehm, C. K.; Freije, C. A.; Grunberg, M. E.; Kosoko-Thoroddsen, T.-S. F.; Welch, N. L.; Pillai, P. P.; Mantena, S.; Kim, G.; Uwanibe, J. N.; John, O. G.; Eromon, P. E.; Kocher, G.; Gross, R.; Lee, J. S.; Hensley, L. E.; MacInnis, B. L.; Johnson, J.; Springer, M.; Happi, C. T.; Sabeti, P. C.; Myhrvold, C. Simplified Cas13-Based Assays for the Fast Identification of SARS-CoV-2 and Its Variants. *Nat. Biomed. Eng* **2022**, *6* (8), 932–943. https://doi.org/10.1038/s41551-022-00889-z.

(38) Bošković, F.; Zhu, J.; Tivony, R.; Ohmann, A.; Chen, K.; Alawami, M. F.; Đorđević, M.; Ermann, N.; Pereira-Dias, J.; Fairhead, M.; Howarth, M.; Baker, S.; Keyser, U. F. Simultaneous Identification of Viruses and Viral Variants with Programmable DNA Nanobait. *Nat. Nanotechnol.* **2023**, *18* (3), 290–298. https://doi.org/10.1038/s41565-022-01287-x.

(39) Pattiya Arachchillage, K.; Chandra, S.; Williams, A.; Rangan, S.; Piscitelli, P.; Florence, L.; Ghosal Gupta, S.; Artes Vivancos, J. *A Single-Molecule RNA Electrical Biosensor for COVID-19*; preprint; Chemistry, 2023. https://doi.org/10.26434/chemrxiv-2023-77mrp.

(40) Pattiya Arachchillage, K. G. G.; Chandra, S.; Williams, A.; Piscitelli, P.; Pham, J.; Castillo, A.; Florence, L.; Rangan, S.; Artes Vivancos, J. M. Electrical Detection of RNA Cancer Biomarkers at the Single-Molecule Level. *Sci Rep* **2023**, *13* (1), 12428. https://doi.org/10.1038/s41598-023-39450-6.

(41) Li, Y.; Artés, J. M.; Demir, B.; Gokce, S.; Mohammad, H. M.; Alangari, M.; Anantram, M. P.; Oren, E. E.; Hihath, J. Detection and Identification of Genetic Material via Single-Molecule Conductance. *Nature Nanotech* **2018**, *13* (12), 1167–1173. https://doi.org/10.1038/s41565-018-0285-x.

(42) Veselinovic, J.; Alangari, M.; Li, Y.; Matharu, Z.; Artés, J. M.; Seker, E.; Hihath, J. Two-Tiered Electrical Detection, Purification, and Identification of Nucleic Acids in Complex Media. *Electrochimica Acta* **2019**, *313*, 116–121. https://doi.org/10.1016/j.electacta.2019.05.036.

(43) Li, Y.; Artés, J. M.; Qi, J.; Morelan, I. A.; Feldstein, P.; Anantram, M. P.; Hihath, J. Comparing Charge Transport in Oligonucleotides: RNA:DNA Hybrids and DNA Duplexes. *J. Phys. Chem. Lett.* **2016**, *7* (10), 1888–1894. https://doi.org/10.1021/acs.jpclett.6b00749.

(44) Li, Y.; Artés, J. M.; Hihath, J. Long-Range Charge Transport in Adenine-Stacked RNA:DNA Hybrids. *Small* **2016**, *12* (4), 432–437. https://doi.org/10.1002/smll.201502399.

(45) Khare, S.; Gurry, C.; Freitas, L.; B Schultz, M.; Bach, G.; Diallo, A.; Akite, N.; Ho, J.; Tc Lee, R.; Yeo, W.; Core Curation Team, G.; Maurer-Stroh, S.; GISAID Global Data Science Initiative (GISAID), Munich, Germany; Bioinformatics Institute, Agency for Science Technology and Research, Singapore; Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, Brazil; Institut Pasteur de Dakar, Dakar, Senegal; National Institutes of Biotechnology Malaysia, Selangor, Malaysia; Smorodintsev Research Institute of Influenza, St. Petersburg, Russia; Genome Institute of Singapore, Agency for Science Technology and Research, Singapore; China National GeneBank, Shenzhen, China; A*STAR Infectious Disease Labs (ID Labs), Singapore; National Public Health Laboratory, National Centre for Infectious Diseases, Ministry of Health, Singapore; Department of Biological Sciences, National University of Singapore, Singapore. GISAID's Role in Pandemic Response. *China CDC Weekly* **2021**, *3* (49), 1049–1051. https://doi.org/10.46234/ccdcw2021.255.

(46) Venkatramani, R.; Keinan, S.; Balaeff, A.; Beratan, D. N. Nucleic Acid Charge Transfer: Black, White and Gray. *Coordination Chemistry Reviews* **2011**, *255* (7–8), 635–648. https://doi.org/10.1016/j.ccr.2010.12.010.

(47) Hihath, J.; Xu, B.; Zhang, P.; Tao, N. Study of Single-Nucleotide Polymorphisms by Means of Electrical Conductance Measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (47), 16979–16983. https://doi.org/10.1073/pnas.0505175102.

(48) Bruot, C.; Xiang, L.; Palma, J. L.; Li, Y.; Tao, N. Tuning the Electromechanical Properties of Single DNA Molecular Junctions. *J. Am. Chem. Soc.* **2015**, *137* (43), 13933–13937. https://doi.org/10.1021/jacs.5b08668.

(49) Wang, Y.; Alangari, M.; Hihath, J.; Das, A. K.; Anantram, M. P. A Machine Learning Approach for Accurate and Real-Time DNA Sequence Identification. *BMC Genomics* **2021**, *22* (1), 525. https://doi.org/10.1186/s12864-021-07841-6.

(50) Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System; 2016; pp 785–794. https://doi.org/10.1145/2939672.2939785.

(51) Wang, Y.; Khandelwal, V.; Das, A. K.; Anantram, M. P. Classification of DNA Sequences: Performance Evaluation of Multiple Machine Learning Methods. In *2022 IEEE 22nd International Conference on Nanotechnology (NANO)*; IEEE: Palma de Mallorca, Spain, 2022; pp 333–336. https://doi.org/10.1109/NANO54668.2022.9928773.

Table of Contents