communications medicine

Article

https://doi.org/10.1038/s43856-024-00582-z

A micro-disc-based multiplex method for monitoring emerging SARS-CoV-2 variants using the molecular diagnostic tool Intelli-OVI

Md Belal Hossain[1,2], Yoshikazu Uchiyama[3], Samiul Alam Rajib[1], Akhinur Rahman[1], Mitsuyoshi Takatori[1], Benjy Jek Yang Tan[1], Kenji Sugata[1], Mami Nagashima[4], Mamiyo Kawakami[4], Hitoshi Ito[4], Ryota Kumagai[4], Kenji Sadamasu[4], Yasuhiro Ogi[5], Tatsuya Kawaguchi[5,6], Tomokazu Tamura[7,8,9], Takasuke Fukuhara[7,8,9,10,11], Masahiro Ono[12,13], Kazuhisa Yoshimura[4] & Yorifumi Satou[1]

Abstract

Background Highly transmissible viruses including SARS-CoV-2 frequently accumulate novel mutations that are detected via high-throughput sequencing. However, there is a need to develop an alternative rapid and non-expensive approach. Here we developed a novel multiplex DNA detection method Intelli-OVI for analysing existing and novel mutations of SARS-CoV-2.

Methods We have developed Intelli-OVI that includes the micro-disc-based method IntelliPlex and computational algorithms of objective variant identification (OVI). More than 250 SARS-CoV-2 positive samples including wastewater ones were analysed to verify the efficiency of the method.

Results IntelliPlex uses micro-discs printed with a unique pictorial pattern as a labelling conjugate for DNA probes, and OVI allows simultaneous identification of several variants using multidimensional data obtained by the IntelliPlex method. Importantly, de novo mutations can be identified by decreased signals, which indicates that there is an emergence of de novo variant virus as well as prompts the need to design additional primers and probes. We have upgraded probe panel according to the emergence of new variants and demonstrated that Intelli-OVI efficiently identified more than 20 different SARS-CoV-2 variants by using 35 different probes simultaneously.

Conclusions Intelli-OVI can be upgraded to keep up with rapidly evolving viruses as we showed in this study using SARS-CoV-2 as an example and may be suitable for other viruses but would need to be validated.

Plain language summary

As the COVID-19 pandemic progresses, it is increasingly becoming important to be able to detect emerging new variants of concerns of SARS-CoV-2, the virus that causes COVID-19, for accurate surveillance and timely interventions. We developed a rapid diagnostic method for detecting multiple SARS-CoV-2 variants and tested it using various starting materials such as sputum, nasopharyngeal swabs and wastewater. The method could accurately detect multiple subvariants of Omicron and showed potential for rapid adaptability to detect the virus as it evolves. This technology could enable continuous monitoring of emerging SARS-CoV-2 variants and the opportunity to intercept transmission with timely interventions to prevent viral spread.

Simultaneous detection or quantification of multiple targets, such as proinflammatory cytokines in serum and cancer-related transcriptome using microarray has accelerated progress in biomedical research. Furthermore, recent revolutionary advances in single-cell technology have enabled us to obtain information on more than 20,000 gene transcripts at single-cell resolution[1], and to address unanswered biological questions that arise with conventional approaches. Bioinformatic approaches are useful to interpret such multidimensional information and extract valuable information for practical use. For example, dimension reduction methods assist the visualisation of multidimensional information and the attainment of transcriptomic proximity among tens of thousands of cells[2].

A full list of affiliations appears at the end of the paper. e-mail: y-satou@kumamoto-u.ac.jp

The emergence of various SARS-CoV-2 variants has caused multiple infection "waves" since the outbreak began in 2019[3,4]. These emerging variants contain mutations in the spike protein that increase the viral infectivity or immune escape ability, with or without increasing pathogenicity[5–8]. Thus, efficient surveillance of SARS-CoV-2 variants is important for monitoring the emergence of new variants of concern (VOCs) to regulate and minimize viral outbreaks (https://www.who.int/publications-detail-redirect/WHO_2019-nCoV_surveillance_variants)[9]. The rapid detection of SARS-CoV-2 variants has been established using RT-qPCR with TaqMan DNA probes specific for a characteristic mutation[10,11]. However, the emergence of variants accumulating multiple mutations has made the use of an RT-qPCR approach less valuable. There is an increasing need for simultaneous multiple mutation detection to distinguish recent Omicron subvariants. SARS-CoV-2 variant mutations are most comprehensively analysed via whole viral genome sequencing (WGS)[12,13]. However, WGS is time-consuming and expensive; sequence results are obtained after a few days and the procedure requires expertise and substantial financial costs. To establish a sustainable surveillance method in a world living with COVID-19, the urgent need to develop a new monitoring method for SARS-CoV-2 mutation analysis that is more efficient and rapid than WGS has become increasingly apparent.

The IntelliPlex assay is a nucleic acid detection method with DNA probes conjugated on a microdisc printed with a unique pictorial pattern (πCode microdisc). The method is useful in characterising viral infections, such as human immunodeficiency virus (HIV) and hepatitis C virus[14,15].

The IntelliPlex SARS-CoV-2 variant detection kit version 2 (v2) is currently available for research use only. To validate the method, we analysed SARS-CoV-2-positive clinical specimens that had previously been analysed by WGS and examined whether the method can be used and upgradable to identify SARS-CoV-2 variants. Further, we developed an IntelliPlex package for objective variant identification (Intelli-OVI) that could identify known and new variants by combining multiple mutation/deletion panel data with informatics interpretation.

Methods

Regulatory approvals

Clinical samples were collected from the Kumamoto Medical Association Inspection Center in Japan. RT-qPCR analyses of the samples were performed at the same centre. cDNA synthesis of the extracted viral RNA from the samples, as well as subsequent sequencing and downstream analyses, were performed at the Division of Genomics and Transcriptomics in the Joint Research Center for Human Retrovirus Infection (JRCHRI), Kumamoto University. Written informed consent was obtained from some participants, while others were automatically included in the study but were provided with clear instructions on how to opt-out if they did not wish to participate.

Additional clinical samples were collected from the Tokyo Metropolitan Institute of Public Health. For these additional samples, WGS was performed at the Tokyo Metropolitan Institute of Public Health. Patients at the Tokyo Metropolitan Institute of Public Health consented to participate via an opt-out procedure, which included clear instructions on how to opt-out if they did not wish to participate. The study was approved by the Ethics Review Committee of the Faculty of Life Sciences, Kumamoto University (approval no. 2223).

Specimen storage and collection

As a part of the regular COVID-19 diagnostic service, sputum or nasopharyngeal swabs were collected from the Kumamoto City Area from September 30, 2020, to January 29, 2022. Collected samples and extracted RNA were preserved at −80 °C, according to the "New coronavirus infection suspected patient specimen collection and transportation manual," for further analysis (https://www.niid.go.jp/niid/ja/2019-ncov/2518-lab/9325-manual.html).

Extraction of RNA and performance of RT-qPCR

Sputum and/or nasopharyngeal swab samples (140 μL) were collected from potential patients for diagnostic purposes. RNA was extracted from the collected specimens using the QIAamp viral RNA Mini Kit (QIAgen, Hilden, Germany) and EZ1 Virus Mini Kit version 2.0 (QIAgen), following the manufacturer's protocols. Buffer AVE (RNase-free water with 0.04% sodium azide) was used to elute 60 μL of the final RNA product, of which 5 μL was used for RT-qPCR; the remaining RNA product was stored at −80 °C for further downstream analysis. A LightCylcer96 system (Roche, Basel, Switzerland) was used to perform RT-qPCR as per the protocol developed by NIID, Japan[16,17] to detect the N gene using the One-Step PrimerScript™ RT-PCR Kit (Perfect Real Time, Takara).

Collection and processing of wastewater samples

The wastewater samples were collected from the wastewater treatment plants in Tokyo. 400 mL of influent sewage was collected by the grab sampling method and immediately stored at −18 °C until analysis. 4 mL of sewage sediment was collected by centrifuging 400 mL of each sample at 15,700×*g* for 30 min (CR21N, Himac). RNA was extracted by applying 1 mL of sewage sediment to the Aptima Specimen Transport Tube (Hologic), mixing by inversion, and centrifuging at 1000×*g* for 5 min (CF9RX, Himac). Upon centrifugation, 140 μL of supernatant was used to extract 60 μL of RNA solution using the QIAamp Viral RNA Mini Kit (QIAGEN). RT-qPCR was performed as described above to detect the N gene of SARS-CoV-2 and 10 μL of extracted RNA was used in the subsequent IntelliPlex analysis.

Characterisation of SARS-CoV-2 using WGS

A QIAseq DIRECT SARS-CoV-2 kit (QIAgen) was used to synthesize cDNA, enrich viral sequences, and amplify and index the library following the manufacturer's recommendations[17]. QIASeq DIRECT UDI Set-A and Set-B (QIAgen) were used to multiplex the samples. At the end of the process, SARS-CoV-2 libraries were prepared for each sample (25 μL), and the quality of the enriched SARS-CoV-2 libraries was assessed by electrophoresis using a TapeStation 4150 system (Agilent Technologies). Finally, after denaturation, the prepared libraries were sequenced using MiSeq Micro and Nano reagent Kits (version 2; 300 cycles) and a MiSeq desktop sequencing system (Illumina).

A FASTQ file (single read, R1) was generated for each sample following sequencing. Adaptor sequences were trimmed from the reads using the 'cutadapt' tool[18], and the trimmed sequences were then cleaned using the PRINSEQ tool[19]. R1 reads with a Phred score of more than 20 were retained at this step for downstream analysis. Adaptor trimmed and cleaned reads were aligned against the SARS-CoV-2 reference genome (NC_045512.2; isolate Wuhan-Hu-1) using the BWA-MEM algorithm[20]. The SAMTools program[21], was used to remove PCR duplicates, and Freebayes (v1.2.0) command-line tools (https://arxiv.org/abs/1207.3907v2) were used for variant calling from the duplicate removed BAM files. The resultant variant call format files were visualised using integrative genomics viewer[22]. Consensus viral genome sequences were also obtained from the duplicate removed BAM files and the resultant consensus sequences were deposited to the global initiative for sharing all influenza data (GISAID; Supplementary Data 4)[23]. Consensus sequences were used to assign the lineage using the Pangolin Lineage Assigner Tools[24], as well as to detect any novel mutation within the viral genome sequences using the CoV-GLUE web application tool (https://www.preprints.org/manuscript/202006.0225/v1).

IntelliPlex πCode assay and data retrieval

SARS-CoV-2-positive clinical samples and wastewater samples were analysed in the IntelliPlex πCode platform using the IntelliPlex SARS-CoV-2 Variant Analysis commercial Kit (v2) and upgraded kit, following the manufacturer's instructions. The kit combines the RT-PCR master mix with a set of primers (reverse primers were biotinylated at the 5′ end) to amplify the characteristic mutation positions of the spike protein and two relatively conserved regions of the ORF1ab and N proteins. The primer, probe, and PCR information are provided in Supplementary Data 1 and 2.

Briefly, 5 μL of RNA sample (10 μL in case of wastewater samples) was used for RT-PCR amplification using the thermal cycler "Mini Amp: ABI" (Thermo Fisher Scientific) with the recommended PCR conditions,

followed by a denaturation step at 98 °C for 7 min and a 4 °C incubation for 5 min. For target-probe hybridisation, 100 µL of hybridisation buffer, 20 µL of microdisk-conjugated probes and 10 µL of denatured PCR amplicon were added to a well of a flat-bottom 96-well plate. The IntelliPlex 1000 πCode processor housed the hybridisation, washing, and streptavidin-phycoerythrin (SAPE) labelling of the hybridised targets. Finally, the PlexBio 100 fluorescent analyser was used to quantify the mean fluorescent intensity (MFI) for each of the target-probe conjugates. To run the PlexBio 100 fluorescent analyser, the necessary commands were provided using the built-in analysis software 'DexiPher', and after analysis completion, the data were retrieved as an Excel file with raw MFI values for each of the probe targets.

A specific πCode microdisc without a coupled probe was used as a blank to quantify the background MFI. Absolute MFI for each of the probe targets was calculated using the following equation and used for result interpretation.

$$\text{Absolute MFI} = \text{Raw MFI of a target} - \text{MFI of the blank}$$

Upgradation of the IntelliPlex πCode assay

The assay upgrade process involves several steps, including designing primers and probes to detect novel mutations, coupling and validating the coupling efficiency of the probes to πCode microdisks, and finally analysing clinical samples using a custom IntelliPlex assay. Among the novel variants, when certain mutations arose in the primer positions of target regions, an assessment of PCR efficiency was performed via RT-qPCR using original primers and several new primer candidates. Primer concentration and PCR conditions were set to the same as those used in the default IntelliPlex assay.

To design probes for capturing novel mutations, we obtained information about how the IntelliPlex company optimized the probes while developing a commercial kit. Then, we independently designed and optimized probes used for novel mutations in this study by considering the information. To design a new probe, several points need to be considered. For a certain target, several probe candidates were designed with a T_m value ranging between 37 °C and 40 °C. The T_m value was calculated using the net primer (http://www.premierbiosoft.com/netprimer/index.html) with default settings. Usually, a probe sequence between 12 and 18 nucleotides in length has a T_m value of 37–40 °C, but the overall length of a probe used in the IntelliPlex assay should be approximately 25 nucleotides to obtain a better resolution, according to the IntelliPlex guideline. To achieve this, the addition of a poly-T sequence at the 5′ end of the original probe sequence, up to 25 nucleotides, is recommended. To avoid non-specificity associated with this random poly-T sequence at the 5′ end of the probe, it is also recommended to replace all Ts with As when there is a T in the original template sequence. Probes were designed to avoid hairpin and self-dimer formation and to exclude primer-binding sites of target regions. C6 amine modification was applied at the 5′ end of all probes to facilitate coupling with πCode microdiscs.

Probes were coupled to πCode microdiscs following the provided user manuals. Unique πCode microdiscs were assigned for each new probe to differentiate them from existing panel probe-microdisc conjugates. Bulk and small-scale coupling methods were employed, and after washing, probe-coupled microdiscs were eluted and stored. Briefly, for bulk coupling, 200 µL of 3 µM probe solution was added into preactivated πCode microdisc, and for small-scale coupling, 100 µL of 3 µM probe solution was added into 10 µL of methanol resuspended preactivated πCode microdisc. The probe and preactivated πCode microdisc were mixed using an Intelli mixer for 2 h at 37 °C setting program at "MODE-U2 and RPM-99" for bulk coupling and "MODE-uu and RPM-99" for small-scale coupling. After consecutive washes with rinse buffer, blocking buffer, and wash buffer, the probe-coupled microdisc was eluted in storage buffer at a concentration of 100 microdiscs/µL and stored at 4 °C. The πCode ID for the microdisc coupled to different probes is provided in Supplementary Data 3.

Next, the performance of each of the nucleic acid probes was validated individually before proceeding to the multiplex assay. To confirm the coupling of a nucleic acid probe to πCode microdisc and its fluorescence specificity, RNA samples of a SARS-CoV-2 variant/recombinant virus/complementary oligo having probe-specific target sequence (positive control, signal fluorescence), and other variants having mismatch sequences at the probe binding position (negative control, noise) were RT-PCR amplified using a potential primer set (reverse primer 5′ biotinylated). After RT-PCR amplification, the amplicon is denatured at 98 °C for 7 min and followed by incubation for 5 min at 4 °C. A working solution of probe-microdisc conjugates was prepared by diluting the coupled πCode microdisc stocks to a final concentration of 5 discs/µL in a storage buffer. The coupled πCode microdisc solution, hybridisation buffer, and denatured PCR product were added into the well of a 96-well plate at a volume of 20 µL, 100 µL, and 10 µL, respectively. The IntelliPlex 1000 πCode processor was used for the automated hybridisation, washing, and SAPE labelling of the captured target probes. The custom assay of the PlexBio 100 fluorescent analyser was used to quantify the MFI for the target-probe conjugates. The output Excel file contains the MFI and the number of the πCode microdisc. As a stringency control, for an optimal probe, a signal/noise ratio ≥ 10 was tried to obtain. However, depending on the sequence of some probe-target regions, the ratio was compromised a little after carefully evaluating the clear separation between signal and noise in clinical samples.

Finally, in a similar manner to testing the coupling efficiency, the multiplex assay using updated primers and probes was performed. First, 5 µL of sample RNA was RT-PCR amplified using a primer mix for different target regions. The forward primer for amplicon D was replaced by a potential primer candidate for all SARS-CoV-2 variants, and the working concentration for each of the primer sets remained the same as in the previous assay, except for amplicon D. The working concentration for primers of amplicon D was increased by 2.5× considering the RT-qPCR analysis and the number of targets to be captured in amplicon D (Supplementary Data 2). Second, for the hybridisation reaction, new custom microdisc-conjugated probes were mixed with the existing probe panel of the IntelliPlex assay. The working concentration of the custom microdisc-conjugated probe was set to 100 microdiscs/µL. For the hybridisation assay, 0.5 µL each of the new probes was mixed with 20 µL of IntelliPlex default probe panel and then were added to a 96-well plate along with 100 µL of hybridisation buffer, and 10 µL of denatured PCR product. The IntelliPlex 1000 πCode processor was used for the automated hybridisation, washing, and SAPE labelling of the captured target probes. After fluorescent detection using the custom assay of PlexBio 100 fluorescent analyser, data were retrieved against each of the unique ID πCode microdiscs.

The built-in analysis software 'DexiPher' of the PlexBio 100 fluorescent analyser has several analysis tabs; each of the tabs' functions is mainly automatic requiring minimal command to run which is described in the user manuals. One of the tabs is "Custom Assay" for custom IntelliPlex analysis. Upon selecting the wells to be analysed of the 96-well plate, the custom assay run can be started. The time required to finish the fluorescent analysis depends on the number of wells to be analysed; approximately 30 s per well is required. The output data is exported as an Excel file having four tabs. The "Summary result" tab contains the micro-discs count and MFI value against each of the πCode microdiscs for each well (sample).

Usually, before the probe-target hybridization reaction, an equal volume of multiple probe-conjugated πCode discs is mixed and a working concentration of probe mixtures is prepared by adding the required volume of storage buffer. From the working probe mixture, a certain volume is pipette out and added to the well of 96 plates. Therefore, the output πCode discs count is variable and for some probes, it could be higher than the working concentration.

Generation of the recombinant SARS-CoV-2 variant

We generated the recombinant BA.2.75 chimeric viruses by the CPER method as followed by our previous study [25]. In brief, we inserted the PCR product of BA.2.75 encoding the spike protein gene into the pCSII-CoV-2-G8 by XhoI and XbaI site with an in-fusion cloning enzyme (Takara). Then, the F8 fragment was amplified by the CoV-2-F8-Fw and CoV-2-F8-Rv with

a PrimeSTAR GXL polymerase (Takara). CPER assembly together with the other fragments was conducted and the reaction was transfected into HEK293-C34 cells. To observe the cytopathic effect of virus replication, the culture supernatants were collected and inoculated into the naïve VeroE6/TMPRSS2 cells. The supernatants were collected at 2 days post-infection (dpi) and stored at −80 °C. The viral sequence was confirmed via sequencing with SeqStudio Genetic Analyser (Thermo Fisher Scientific) and outsourced service (Fasmac).

MDS, kNN, and LOF analysis

Panel data with n items can be represented as an n-dimensional vector. However, we normalised the panel data by dividing each element by the magnitude of the vector, since the element values of the vector change with the amount of virus. This allows each case to be arranged on an n-dimensional unit circle, and each case can be classified according to which direction it faces from the origin. Even if the same panel barcode pattern has different output values, if the panel barcode pattern is the same, it is possible to classify the viruses as the same variant because they face the same direction.

Since panel values are n-dimensional data, each case can be represented as one point in n-dimensional space. If the n-dimensional data can be reduced to 2-dimensional data, they can be represented by a scatter plot, which makes it possible to intuitively determine which variant the test data indicate or whether the test data indicate an unknown variant. MDS was used for dimensional reduction[26]. MDS is a method that generates a transformation matrix using a distance matrix consisting of the Euclidean distances between all data and then determines new coordinate points given by the eigenvectors of the transformation matrix. Since MDS is a linear transformation that preserves the Euclidean distance between data, it can be interpreted as a reproduction of the relative position of each case in n-dimensional space in 2-dimensional space.

To distinguish between types of variants, we used the kNN[27], with panel data as input data. In kNN, the Euclidian distance between all the training and test data is calculated, and the class of training data with the closest distance is discriminated as the type of variant of the test data. The leave-one-out method[27], was used to evaluate the proposed method. This method extracts one case as test data from m training data, and the process of discriminating the test data with $m-1$ training data is repeated m times. Therefore, the training and test data are treated independently. When kNN was applied to the cases of Intelli-OVI analysis, the distinction performance was 100%.

Because a test sample is classified as one of the known variants when we used kNN, the unknown variants cannot be detected. Therefore, we employed LOF for novelty detection[28]. LOF is a method to quantify the density around the test sample in multidimensional space. If a test sample is a known variant, the density around the test sample is high because known variants are distributed around the test sample. However, when the test sample is an unknown variant, the density is sparse because there are no known variants in the neighbourhood. When the density of a test sample is low compared to the neighbouring data, the LOF outputs a large value. Therefore, unknown variants can be detected by thresholding the LOF values. The problem with LOF is that the calculation accuracy is low when the number of samples of known variants is small. Therefore, we generated bootstrap samples of known variants and estimated the multidimensional normal distributions. The LOF value was calculated after increasing the number of samples of known variants by random sampling from each multidimensional normal distribution. The scripts for all analyses were written in Python programming language. The source code for each data analysis algorithm has been deposited to GitHub (https://github.com/satoulab/SARS-CoV-2_Intelli-OVI).

Statistics

All data analyses were performed using GraphPad Prism 7 software (GraphPad Software, San Diego, CA, USA), including the student t-test. Correlation and regression analysis between the MFI data and sample Ct values was performed with R (v4.0.3). ROC analysis was conducted on the LOF values to evaluate the diagnostic potential of the Intelli-OVI approach in identifying the emergence of new variants. Discrimination was assessed using the area under the curve (AUC) as discussed in our previous study[29]. Optimal cutoff values maximizing the sensitivity and specificity of Cv were calculated using R (v4.0.3). Moreover, the source code and relevant data used in this study for the ROC/AUC analysis have been deposited to GitHub, respectively as an R script and.txt file (https://github.com/satoulab/SARS-CoV-2_Intelli-OVI).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Evaluation of the IntelliPlex SARS-CoV-2 variant detection kit v2 with clinical samples

We tested the commercially available IntelliPlex SARS-CoV-2 variant detection kit v2 on various clinical samples that had previously been characterized using WGS. The principle of the IntelliPlex assay is shown in Supplementary Fig. 1. There are various πCode microdiscs (PlexBio) for DNA probe labelling, each printed with a unique pictorial pattern (Supplementary Fig. 1a). Tens of thousands of different πCode microdiscs are available facilitating identification of multiple targets in a single reaction. DNA probes are conjugated to microdiscs prior to hybridization with amplicons generated using biotinylated primers (Supplementary Fig. 1b). After the washing step, streptavidin conjugated with a fluorescent compound, phycoerythrin (PE), is added to quantify target DNA on each πCode microdisc on the bottom of the 96-well plate. There are two steps for signal detection in each well: first, bright-field imaging identifies the position of each πCode microdisc. Second, dark-field imaging identifies the fluorescent emission from each πCode microdisc which is measured to quantify the amount of PCR amplicon containing the same sequence as each probe (Supplementary Fig. 1c). The assay probe panel covered the major mutations that distinguish SARS-CoV-2 variants. Four regions of interest containing the probe-binding region in the *S* gene were amplified via RT-PCR (Fig. 1a). The panel also contains PCR primers and probes of the *ORF1ab* and *N* genes, as highly conserved regions, to quantify the amount of virus (Fig. 1a). Each reverse PCR primer was conjugated with biotin, and each DNA probe was conjugated with a unique πCode microdisk (Fig. 1b), which was used to identify and quantify target DNA in a subsequent detection step. We used multiple probes to distinguish each variant as a different pattern in the mutation panel (Fig. 1c). We examined clinical samples ($n = 62$) collected from Kumamoto, Japan, between February 2021 to September 2021, when the conventional strains (B.1.1.284 and B.1.1.214) as well as the R.1, alpha (B.1.1.7), and delta (B.1.617.2.29.1) variants were epidemic in Japan (Supplementary Fig. 2a). Even in samples with a low viral load, the IntelliPlex assay efficiently provided mutation patterns distinguishing each variant (Fig. 1d). Although the fluorescence signals were occasionally positive in both wild-type and mutation probes, such as the 484, 501, or 452 probes, the fluorescence signal of probes that completely matched the targets was significantly higher than that for probes with a single base mismatch (Supplementary Fig. 2b). Such an overlap of signal between the wild-type and mutation probe was observed for only these three probes out of 20 probes for the spike region of v2 kit. Importantly, a significant linear correlation was found between the RT-qPCR cycle threshold (CT) value and MFI of the *ORF1ab, N* gene, and L452R probe (Supplementary Fig. 2c), with a similar trend observed for most of the other probes except for P681R (Supplementary Fig. 2c: regression analysis). The reason behind this unusual trend of MFI for the P681R probe is unknown but could be the presence of a higher GC clamp at the 3' end of the probe (Supplementary Data 1), which may induce the formation of secondary structures by the probe or strong self-dimerization of the PCR amplicon[30,31].

Further, we analysed more clinical samples ($n = 31$) with emerging variants collected from a megacity like Tokyo, Japan, between December 2021 and February 2022 (Supplementary Fig. 2d). The samples included Omicron variants BA.1 and BA.2 that had just emerged at the time of

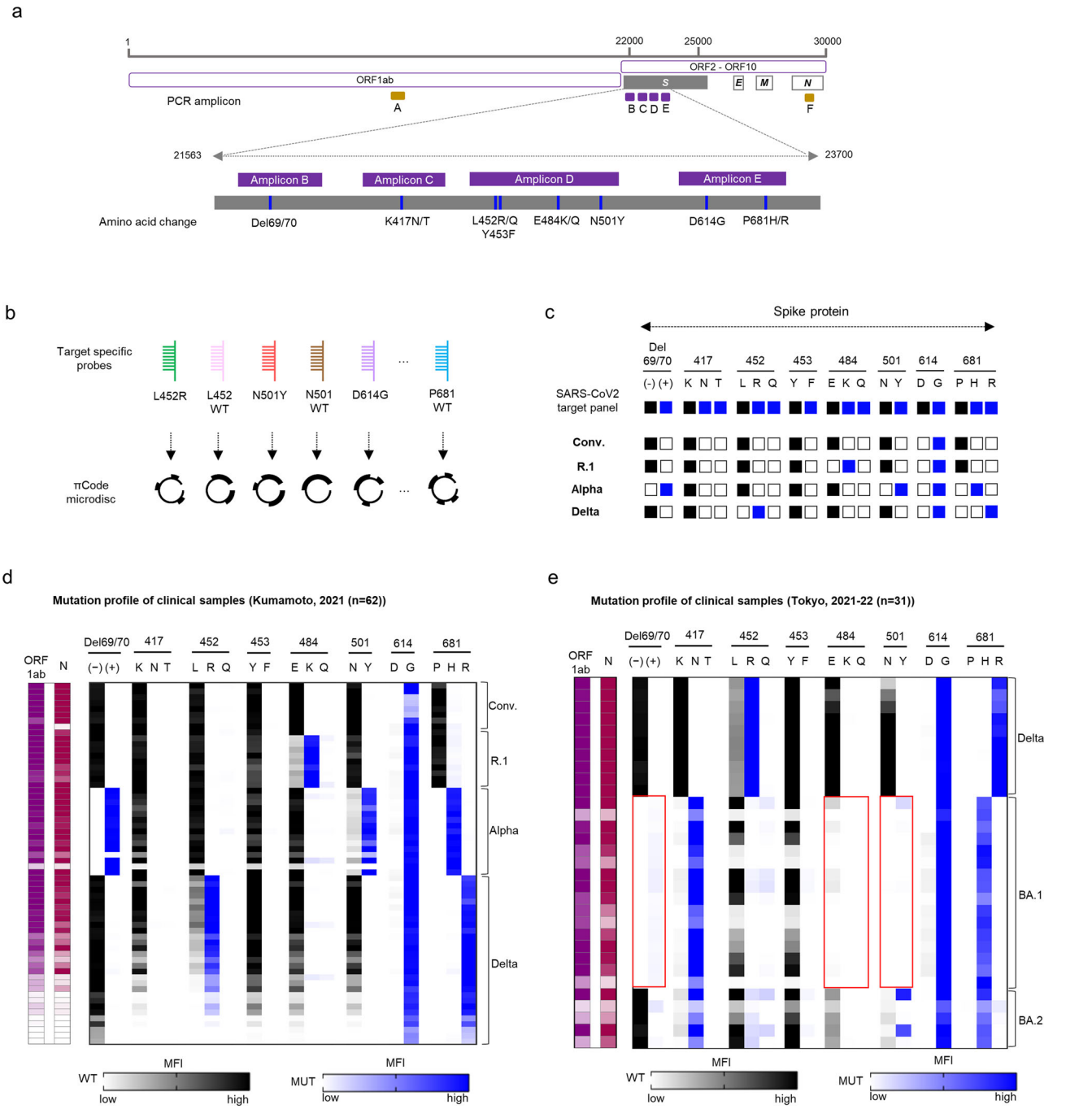

Fig. 1 | Evaluation of the IntelliPlex system for SARS-CoV-2 variant detection manipulating the multiple probe information. a Schematic illustration of the SARS-CoV-2 genome showing the six target regions for PCR amplification. PCR amplicons contained the mutations of interest to be captured by the specific probes. The highlights indicate four PCR amplification regions of the spike (*S*) gene harbouring some key mutation positions (blue bar on the *S* gene). **b** Schematic of example probes used to capture specific target mutations or wild-type sequences in the IntelliPlex system for SARS-CoV-2 variants detection. Each of the target-specific probes is coupled to a distinct pictorial pattern πCode microdisc to facilitate multiplexing of the system. **c** SARS-CoV-2 spike protein target positions for which wild type or mutations can be detected with the probes in the IntelliPlex variant detection panel. The default panel covers major mutations of the R.1, alpha, and delta variants separating each other with a unique mutation pattern. The black and blue rectangles in the target panel represent wild and mutant sequences, respectively, for the amino acids at the specified positions. **d** Clinical samples ($n = 62$) collected from Kumamoto, Japan during the early half of 2021 were analysed using the IntelliPlex assay with the default probe panel. A mutation pattern heatmap for each clinical sample was generated using the absolute MFI obtained for each target. The corresponding SARS-CoV-2 target panel is shown above the heatmap. The MFI depends on the presence of different targets in an individual sample. A coloured scale is shown at the bottom of the heatmap with increasing shades indicating higher MFI values, suggesting a higher amount of PCR amplicon. The control region (highly conserved regions of ORF1ab and N) heatmap at the left compares the PCR efficiency of other target regions. MUT: Mutation, WT: Wild type. **e** Clinical samples ($n = 31$) from Tokyo, collected between the end of 2021 and the start of 2022, were analysed using the IntelliPlex assay. An atypical mutation pattern was observed for BA.1 and BA.2 (the novel variant of that time). For some positions, faint or no signals were observed (marked as a red box on the heatmap), indicating the emergence of a novel mutation at that position or mutations at the primer/probe binding region. This atypical pattern indicates the emergence of a new variant but cannot be characterized completely. SARS-CoV-2 severe acute respiratory syndrome coronavirus 2, PCR polymerase chain reaction.

analysis. Although samples from delta variants were efficiently identified by the predicted heatmap in the mutation panel, there was little or no signal in the Del69/70, 484, and 501 probe regions of the Omicron BA.1 sample (Fig. 1e).

Probe signal impairment mechanism and probe panel upgrade to cover new virus variants

Additional mutations in Omicron variants could be responsible for the impaired signal in some probe areas. There are two steps in the IntelliPlex assay, PCR amplification and probe hybridisation (Supplementary Fig. 1b, c). To determine the emergence of mutations in the primer and probe regions, the sequences of analysed samples and sequences available in public databases for other Omicron subvariants that evolved at that time were aligned with the SARS-CoV-2 Wuhan-Hu strain. As anticipated, novel mutations in the probe regions and the primer-binding region of amplicon D in Omicron variants were observed (Fig. 2a). Omicron BA.1 contained two mutations in the forward primer region (Fig. 2a), which was consistent with a significant decrease in signal from amplicon D, including 452, 453, 484, and 501 probe regions (Fig. 1e). We analysed the amplification efficiency of PCR using a qPCR method and found that there was little amplification in amplicon-D PCR when we used cDNA from the BA.1 sample (Supplementary Fig. 3a). To overcome this drawback, we designed alternative primers avoiding the 3' end mutation of the original primer observed in BA.1 while accepting the mutation common to all Omicron subvariants (Supplementary Fig. 3b). We analysed the efficiency of PCR amplification (Supplementary Fig. 3c) and discovered that the P1_MT primer showed the most efficient PCR amplification with Omicron BA.1, without reducing amplification efficiency with the delta variant (Supplementary Fig. 3c). Based on these results, we used the P1_MT primer in subsequent experiments. Next, we designed additional probes that cover new mutations in Omicron variants and created several candidates with similar melting temperatures as other probes in the SARS-CoV-2 probe panel v2 (Supplementary Fig. 3d–i). Further, we analysed the efficiency of each candidate probe by evaluating both sequence-matched and -mismatched variants. Based on these results, we selected the optimal probes with a high signal/noise ratio and incorporated them in the probe panel, called SARS-CoV-2 kit v3 (Fig. 2b), to analyse variants in the subsequent experiments.

Upgraded SARS-CoV-2 kit v3 precisely identified a wide range of variations in SARS-CoV-2 variants

To evaluate the detection efficiency of the upgraded kit, we analysed a wide range of SARS-CoV-2 variants collected from 2020 to 2022 in Japan. The SARS-CoV-2 kit v3 efficiently provided expected heatmap patterns for Conventional, R.1, alpha, delta, BA.1, BA.2, BA.2.12.1, and BA4/5 variants (Fig. 2c). We observed impaired signal in only the 452-probe region in variant BA.2.56, indicating possible novel mutation in that position. Upon checking the sequences, the presence of a novel mutation L452M (S: C22916A) was confirmed for BA.2.56.

The multiple mutation panel information would be more accurate in identifying variants than the conventional qRT-PCR method. Next, we aimed to visualise the multiple panel data as a two-dimensional graph using a dimension reduction algorithm. Dimension reduction resulted in the clustering of data with a similar character, thus, enabling the examiner to visually inspect the distribution of each variant cluster. We compared two major dimension reduction methods, multidimensional scaling (MDS) and principal component analysis (PCA). We discovered that discrimination of the SARS-CoV-2 variants using PCA was not as clear as that observed in the MDS analysis (Fig. 2d), because MDS is a method of forming new axes while preserving the Euclidian distance of each data point. The BA.2.56 variant with a novel mutation at position 452 (L452M), and therefore having a missing signal, clustered in a position distinct from all other variants in the MDS results (Fig. 2d). These results demonstrated that the IntelliPlex assay combined with MDS might be helpful to identify the emergence of new SARS-CoV-2 variants.

SARS-CoV-2 kit v4 upgraded for various Omicron subvariants

In late 2022, several sub-lineages of Omicron variants emerged globally. We investigated any additional mutation positions in the receptor binding domain (RBD) of the spike protein (Fig. 3a). Since PCR amplicons in SARS-CoV-2 kit v3 did not cover the R346T mutation position; therefore, we designed a new PCR targeting that position (Supplementary Fig. 4a). We designed probes for both wild type and mutant sequences and selected the optimal probes after checking their target hybridization specificity (Supplementary Fig. 4b). For some mutations, the specificity of the probes was validated using either RNA from recombinant virus or the complementary oligo because the clinical samples for the variants were not available at that time. We added ten more probes to distinguish each Omicron subvariant, ultimately resulting in a total of 35 probes in the upgraded kit (Fig. 3b). To validate the efficiency of discriminating sub-lineage level of BA.2, we analysed clinical samples collected in early 2023 in Tokyo. The findings demonstrated that each sub-variant of Omicron showed a unique pattern in the heatmap (Fig. 3c). Furthermore, we checked the distribution of different virus variants through MDS analysis and found a unique cluster from each SARS-CoV-2 variant (Supplementary Fig. 4c). Collectively, these findings demonstrated that the new multiple-dimension nucleic-acid detection system can be upgraded by adding new probes and primers following the emergence of SARS-CoV-2 variants even before the variants become endemic in a country (Fig. 3d).

Development of IntelliPlex package for OVI to monitor known and new SARS-CoV-2 variants

As mentioned above, multiple target information would be useful in identifying not only known but also unknown variants. To apply this method for practical use, it would be valuable to provide the result with quantitative value in an objective manner. To achieve that, we developed a data analytics workflow with three bioinformatic similarities and/or anomaly detection algorithms: MDS, local outlier factor (LOF) and k-nearest neighbour (kNN), to discriminate between known and new variants and to identify each of the known variants in an objective manner (Fig. 4a). As a proof of concept, we analysed the clinical sample data shown in Fig. 1d, e. We used the sample data collected from Kumamoto as reference data and the sample data collected from Tokyo as test data. We performed MDS clustering of Kumamoto samples to show the distribution of different known variants (Fig. 4b). We used those clinical samples from Kumamoto (CT value ≤ 25; $n = 39$) as a reference to visualise the distribution of new samples (test samples; $n = 31$) collected from Tokyo (Fig. 4c). Test samples 1–10 clustered in the same area as other delta variants in the reference, in contrast, test samples 11–31 distributed outlier of the known variants cluster. To make a rational judgement, we used the LOF algorithm to obtain the outlier value for each of the test samples with respect to the reference known variants. A representative example of either a known or unknown variant is shown in Fig. 4d. The LOF value of the test sample with known variants was as low as that of the reference known sample when we analysed them as a test sample. In contrast, a higher LOF value was obtained from the other test samples with Omicron BA.1 or BA.2. There was a statistically significant difference between known and unknown (novel) test samples (Fig. 4e). When we used 4.5 as a threshold value, the sensitivity and specificity for isolating the test samples belonging to known existing variants were remarkably high (Supplementary Fig. 5a, b). Further, we utilised the kNN algorithm to determine which known variants are most similar to the test samples with LOF values below the threshold (Fig. 4f). We successfully identified all known variants as delta variants. We named the clinical test package combining the IntelliPlex SARS-CoV-2 variant detection kit and data analytics workflow as the Intelli-OVI. This could aid examiners in unbiasedly and objectively identifying SARS-CoV-2 variants (See Methods for more details).

To further validate the efficiency of the Intelli-OVI method for diverse Omicron sub-variants, we applied the Intelli-OVI to analyse the multiple targets fluorescent (MFI) data obtained from SARS-CoV-2 kit v3 analysis

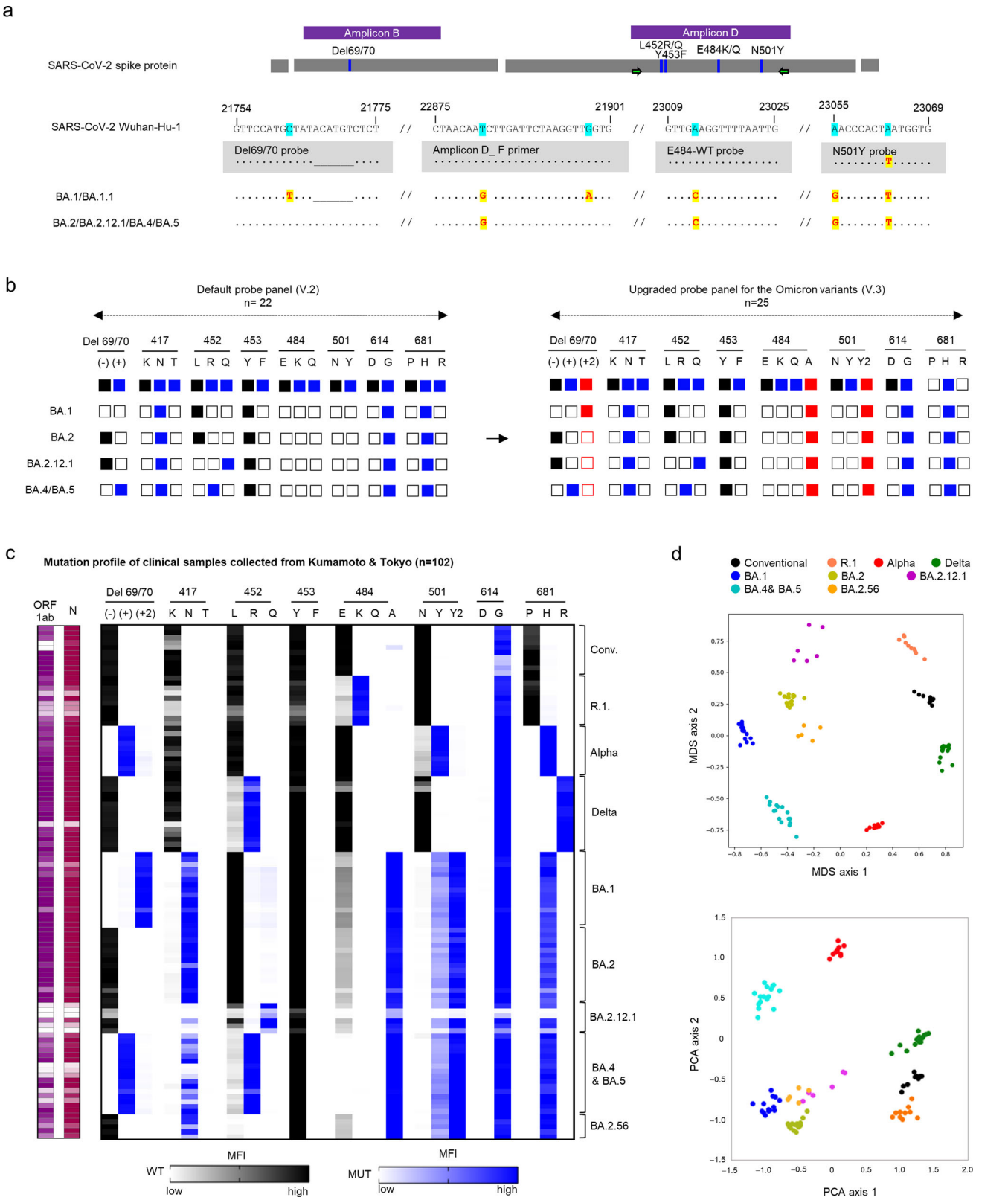

Fig. 2 | Upgradability of the probe panel and its evaluation using clinical samples of diverse variants. a The underlined reasons for the impaired signal obtained for the Omicron BA.1 and BA.2. Whole genome sequences of different subvariants of Omicron revealed novel mutation at the 5' end of forward primer of amplicon-D, E484, and upstream of N501Y positions hampering the PCR amplification and the target-probe hybridization, respectively. Besides in BA.1, mutations were detected at the 3' end of the forward primer of amplicon-D and N501Y probe binding region. **b** Addition of three new probes to characterize different subvariants of Omicron. **c** Mutation pattern heatmap for clinical samples ($n = 102$) from diverse variants, based on the absolute MFI values of each of the targets. The coloured scale shown at the right indicates the relative MFI, which is related to the amount of PCR amplicon. A novel variant, BA.2.56, generated an atypical heatmap pattern with faint or no signals for position 452. **d** Two major dimension reduction methods, MDS and principal component analysis (PCA) were compared to inspect the pattern of clustering of different variants and their distribution. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

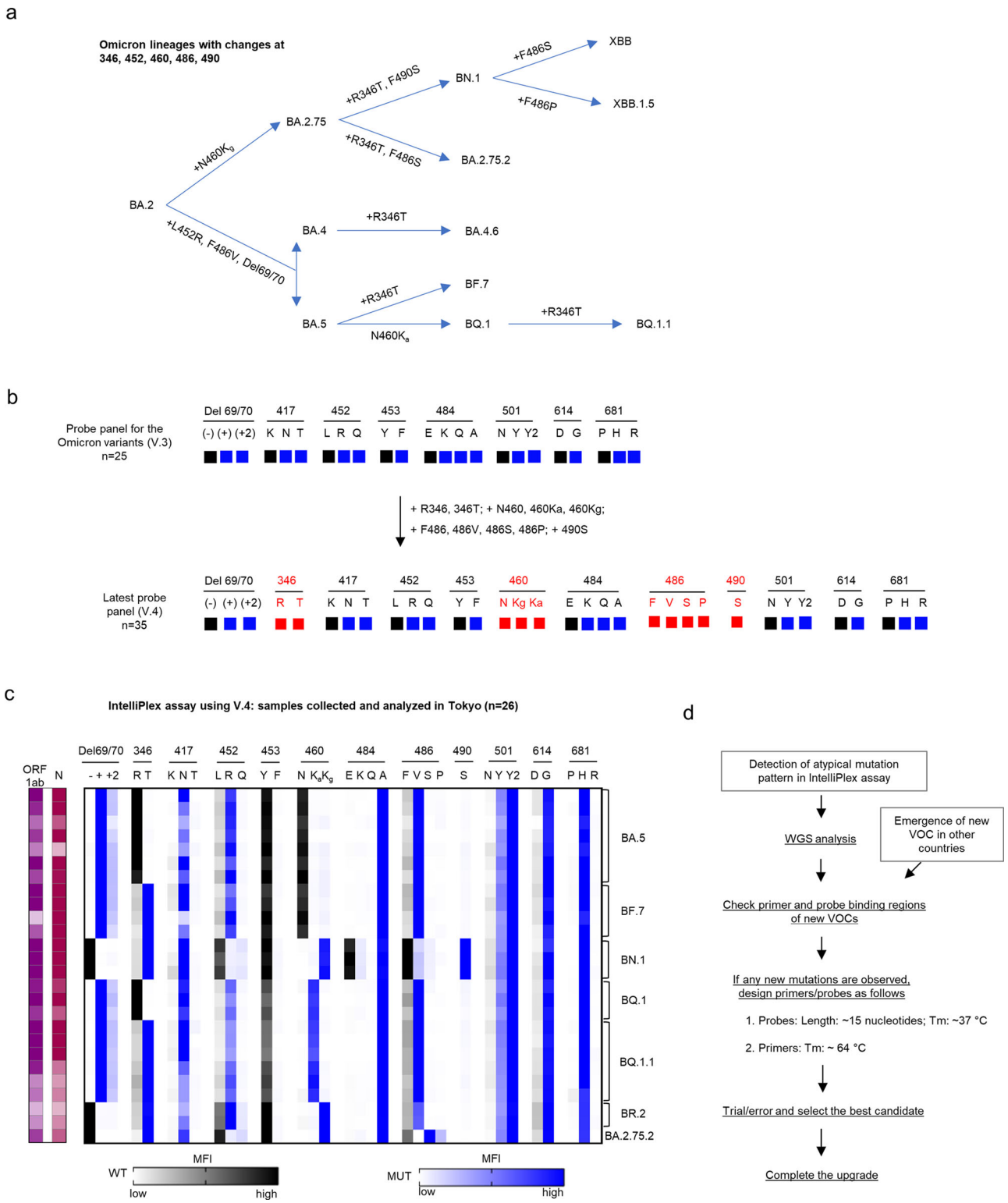

Fig. 3 | Far-reaching upgradability of the probe panel to capture emerging sub-lineages of Omicron and its evaluation with clinical samples. a Emerging sub-lineages of Omicron BA.2 with additional novel mutations at different positions in the receptor binding domain of the spike protein. N460Ka/Kg: variants evolved with the same amino acid mutation but nucleotide mutation is different at this position (T22942A/G). **b** Addition of 10 new probes to increase the window of the probe panel distinguishing the existing and emerging VOC. **c** The most recent clinical samples ($n = 26$) were collected and analysed in Tokyo, Japan to evaluate the distinguishability of the latest probe panel (35 probes) among different sub-lineages of the emerging variants. A mutation pattern heatmap was generated for the different clinical samples based on the absolute MFI values of each of the targets. MUT mutation, WT wild type. **d** Step-by-step workflow for updating and upgrading the variant detection probe panel and the overall system. VOCs variants of concern, WGS whole genome sequencing.

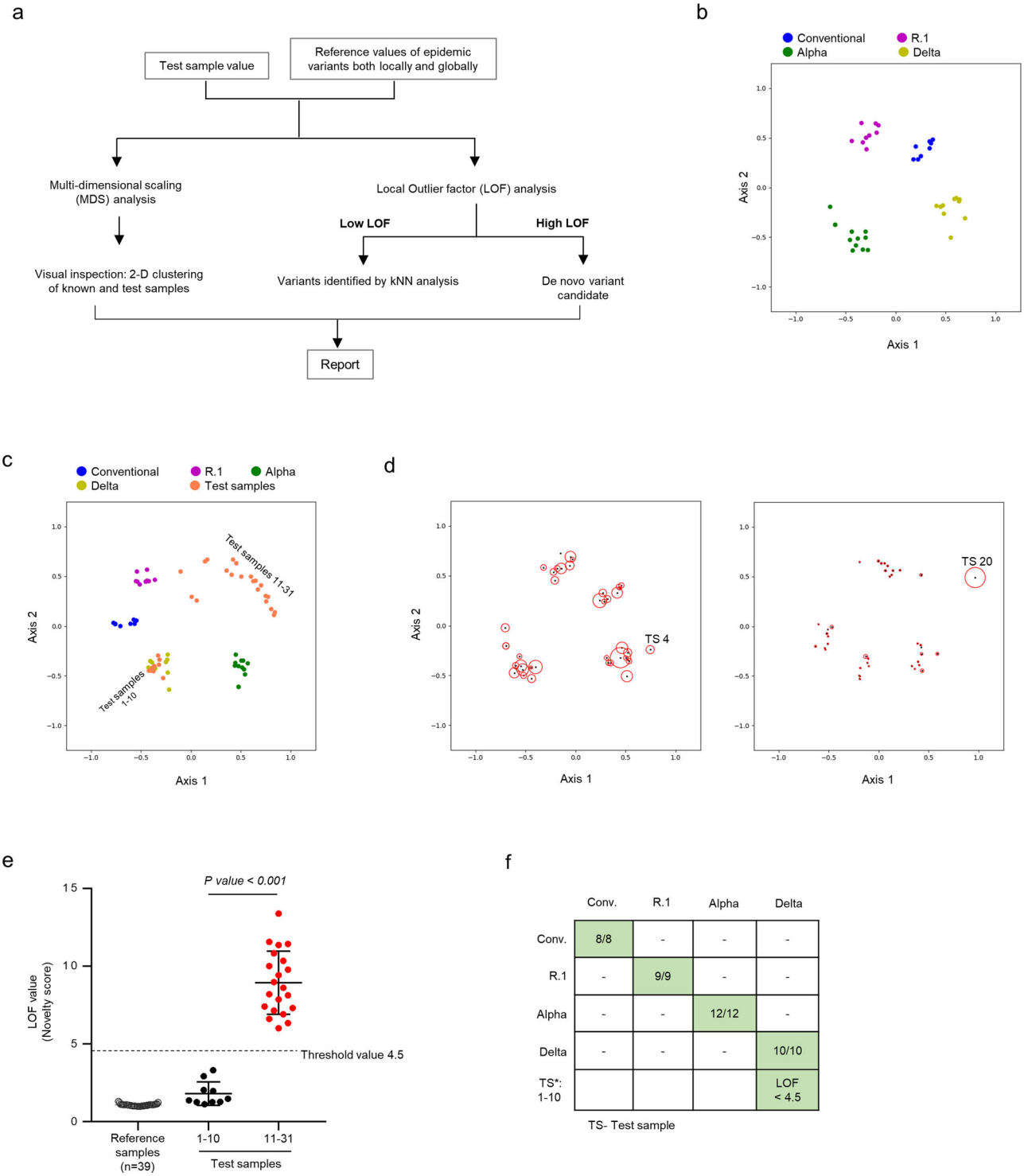

Fig. 4 | Development of Intelli-OVI for the easy screening of known and novel variants. a A workflow of Intelli-OVI for the quantitative detection of different SARS-CoV-2 variants as well as identifying the presence of possible novel (unknown) cases. Three statistical algorithms (MDS, LOF, and kNN) were used to interpret the detection of different variants. MFI mean fluorescent intensity, 2-D two-dimensional. **b** Multi-panel MFI data of different known variants with higher virus titre (CT value ≤ 25) was used to analyse the MDS clustering for the Kumamoto, Japan clinical samples (Fig. 1d). **c** MDS analysis using clinical samples data of Kumamoto as a reference to visualize the distribution of new samples (test samples) collected from Tokyo (Fig. 1e). Test samples 1–10 (delta variants) clustered in the same area as other delta variants in the reference, other test samples (BA.1, BA.2) distributed outlier of the known variants cluster. **d**, **e** The LOF algorithm was used to obtain the outlier value for each of the test samples concerning the reference. Example of outlier detection using LOF algorithm for the test sample 4 (representative of the delta cluster) and Test sample 20 (representative of the unknown cluster) (**d**). Comparison between the LOF value of two clusters of test samples (**e**). Statistical significance was obtained using a Student *t*-test (unpaired, mean ± SD). The threshold was calculated by analysing each of many clinical samples as unknown concerning the other as reference (Supplementary Fig. 5a, b). **f** The kNN algorithm identified each of the test samples as a neighbour to the nearest reference cluster. Variant detection by the kNN algorithm was made when LOF was less than the threshold of 4.5. MDS multidimensional scaling, LOF local outlier factor, kNN k-nearest neighbour.

shown in Fig. 2c, and demonstrated that new variants exhibited higher LOF values (Supplementary Fig. 6).

Application of Intelli-OVI for variant monitoring of a SARS-CoV-2 cluster

Once we developed the Intelli-OVI system, we employed it for variant monitoring of a SARS-CoV-2 cluster at an office in Tokyo, Japan. First, the IntelliPlex assay was performed using the SARS-CoV-2 kit v4 to analyse RNA samples ($n = 16$). The output fluorescent value (MFI) was obtained against each of the targets used in the subsequent analysis for OVI. The OVI output was a distinct cluster and a higher LOF value in MDS and LOF analysis, respectively (Fig. 5a, b), for the test samples when compared to previously analysed samples used as a reference. Therefore, the Intelli-OVI system identified the cluster as caused by a new variant. We checked the MFI heatmap pattern (mutation profile) of the test samples and compared it with that of the globally dominant variants at that time. From mutation profile analysis, we suspected the cluster variant to be XBB.1.5 or XBB.1.16. We later confirmed the variant as XBB.1.16 via WGS analysis of six samples (Fig. 5c). These results provided evidence that Intelli-OVI can efficiently monitor and detect the variant of SARS-CoV-2.

Analysis of more sub-lineages of Omicron and wastewater samples

The Omicron variant continues to evolve, with certain sub-lineages having become a global concern. Since the SARS-CoV-2 upgraded kit v4 has 35 probes, it is expected to obtain the characteristic mutation profile for diverse sub-lineages. Though it is impractical to characterize all variants, the probe panel should include the characteristic mutation profile for the globally dominant variants and VOC. We analysed more clinical samples collected in mid-to-late 2023 in Tokyo. We could identify all the potential sub-lineages of concern even though certain minor sub-lineages shared the mutation profile of some dominant sub-lineages (Fig. 5d). Further, we analysed five wastewater samples that were collected during different SARS-CoV-2 infection waves. Because the viral load of wastewater samples was very low, we could not obtain enough MFI signals to have the characteristic mutation profile for three of the five cases, having Ct values higher than 35. The other two samples collected during the 7th infection wave in Japan had a comparatively higher viral load (Ct value 33 and 33.3), and we could identify the virus variant (Fig. 5d). Because the wastewater sample contained a mixture of different viruses, the variant we identified was the dominant one of the 7th infection-wave in Japan.

Intelli-OVI for expedited SARS-CoV-2 diagnosis and variant monitoring

To establish an efficient and rapid monitoring system for current and future emerging pathogens, we would like to propose the following strategy using SARS-CoV-2 as an example. IntelliPlex assay requires PCR amplification with biotinylated primer. To increase the speed of the assay, we would be able to use primers of SARS-CoV-2 kit v4 for TaqMan RT-qPCR for the first screening of SARS-CoV-2 infection (Fig. 6a). This necessitates an additional primer cost; however, we can directly analyse PCR product of positive clinical sample by Intelli-OVI system without performing PCR again. We analysed the clinical sample ($n = 42$) and confirmed RT-qPCR detection with primers of SARS-CoV-2 kit v4 was equivalent to that with primers of conventional RT-qPCR protocol (Supplementary Fig. 7a). Moreover, we optimised the minimum disc number required per probe to get MFI from each hybridised probe-target to reduce the overall detection cost per sample (Fig. 6b and Supplementary Fig. 7b, c). Furthermore, the Intelli-OVI method has proven efficient in identifying new variants that emerged with a new combination of mutations or with a novel mutation in the target region, thereby being useful as an initial screening method for SARS-CoV-2 variants and determining which specimens should be secondarily analysed using WGS (Fig. 6c). This approach would decrease the number of samples for WGS by selecting atypical samples for WGS, and thereby increase the cost-effectiveness of SARS-CoV-2 variant monitoring.

Discussion

In this study, altogether we have added 13 new DNA probes along with the IntelliPlex default probe panel (increasing the total probe number to 35) to detect wild/mutation sequences of 14 target positions. We examined over 250 clinical samples, from which we were able to extract multiple target information and successfully identify different virus variants. Theoretically, it is possible to detect 35 DNA targets simultaneously in clinical samples; however, due to the available targets among different variants analysed in this study, we were able to detect 14 targets for the first time using a probe-based detection method. This finding will provide the basis for higher specificity of rapid molecular diagnostics by identifying several biomarkers of the respective diseases or pathogens.

Since late 2020, SARS-CoV-2 evolution has been characterised by the emergence of a set of mutations that alter various aspects of virus biology, including pathogenicity, infectivity, and/or antigenicity. The spike protein plays a critical role in engaging the viral receptor ACE2 protein[32], and it is a major antibody target[33,34]. In addition, the spike protein is the SARS-CoV-2 component of mRNA- and adenovirus-based vaccines that have been used globally. Consequently, mutations in the spike protein are of particular interest because they can impact the effectiveness of the vaccines. More than 30 mutations have been found in the spike protein of the Omicron variant[35], which are associated with considerable viral escape from antibody neutralisation[36]. The fast-evolving virus poses several challenges. The conventional qPCR method, which employs TaqMan probes, is incapable of covering such a large number of mutations in Omicron variants. Although several new technologies for the detection of SARS-CoV-2 variants have been reported[37,38], including mass spectrometry (exploited for its multiplex capacity to call emergent variants in robust assays[39–41]), WGS remains the gold standard for determining viral lineages. However, we discovered that there are several advantages of the IntelliPlex assay when compared with the viral WGS. (1) Speed of the assay: the IntelliPlex system workflow consists of three steps, PCR, hybridisation/washing, and detection; the total running time is only 3–4 h, as opposed to WGS, which takes several days (Fig. 6b). (2) High sample throughput: the assay is performed in a 96-well plate; therefore, a larger number of clinical samples can be processed simultaneously than with other methods such as WGS. This property is quite advantageous, especially for a situation like the current SARS-CoV-2 Omicron variant pandemic. (3) Simplicity of the procedure: WGS requires the technical expertise of examiners and the skill of NGS data analysis. NGS data analysis for SARS-CoV-2 variant detection involves several analysis steps, which are hurdles for a basic clinical diagnostic laboratory. While the Intelli-OVI packages for objective variant detection also require the basic knowledge of data analysis, and can be executed with a simple command line to run each part of the algorithm. However, to be used for clinical diagnostic purposes, a simple variant detection software can be created combining the three algorithms used in Intelli-OVI that can be provided along with the IntelliPlex variant detection kit. (4) Low running cost: the cost of consumables for the IntelliPlex assay is three to four times lower than that of WGS. (5) Flexibility of the system: As demonstrated in this study, the mutation panel can be upgraded by adding new probes for the novel mutations associated with new variants. Once a new mutation is identified, either by WGS or from public data analysis, it is simple to design several probe candidates targeting the new mutation. As soon as the designed probes are available, it takes about 5–6 h to select the best candidate probe for the new target (Fig. 3d). Therefore, it is plausible to upgrade the system by obtaining sequence information of emerging viruses anywhere in the world, and to strengthen the surveillance system to monitor new variants even if they have not yet reached the scale of a pandemic in a particular region.

An evident limitation of the IntelliPlex assay is that sequence mismatch between designed DNA probes and new variants could inevitably cause impaired signals, which we observed for probe regions containing new mutations (Figs. 1e and 2c). However, mutations of different VOCs have

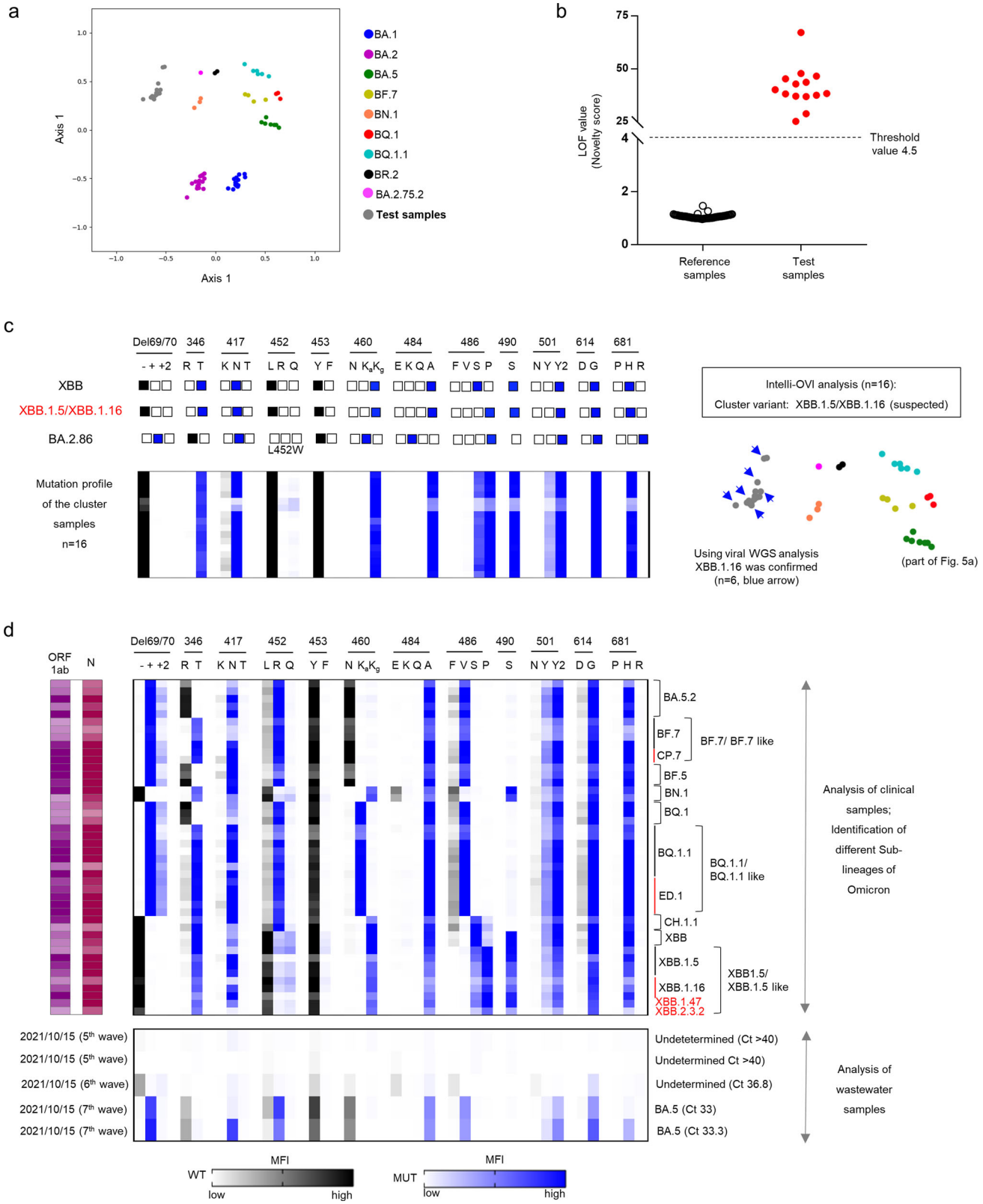

Fig. 5 | Application of Intelli-OVI for efficient monitoring of a SARS-CoV-2 cluster infection; and analysis of wastewater samples and more clinical samples of Omicron sub-lineages. a, **b** Intelli-OVI output for the SARS-CoV-2 cluster ($n = 16$): **a** cluster distribution in the MDS plot for the test samples compared to previously analysed samples; **b** outlier detection using LOF algorithm and comparison with the values from previously analysed samples as reference. **c** Analysis of the sequence of publicly available global dominant variants at the time of the cluster and identification of their mutation profile based on the v4 probe panel; comparison of the mutation profile heatmap for the test samples with the global dominant variants. **d** Mutation pattern heatmap for clinical samples ($n = 44$) from more diverse sub-lineages of Omicron and wastewater samples ($n = 5$) collected during different infection waves in Japan. The heatmap was generated based on the absolute MFI values of each target. MUT mutation, WT wild-type, WGS whole viral genome sequencing.

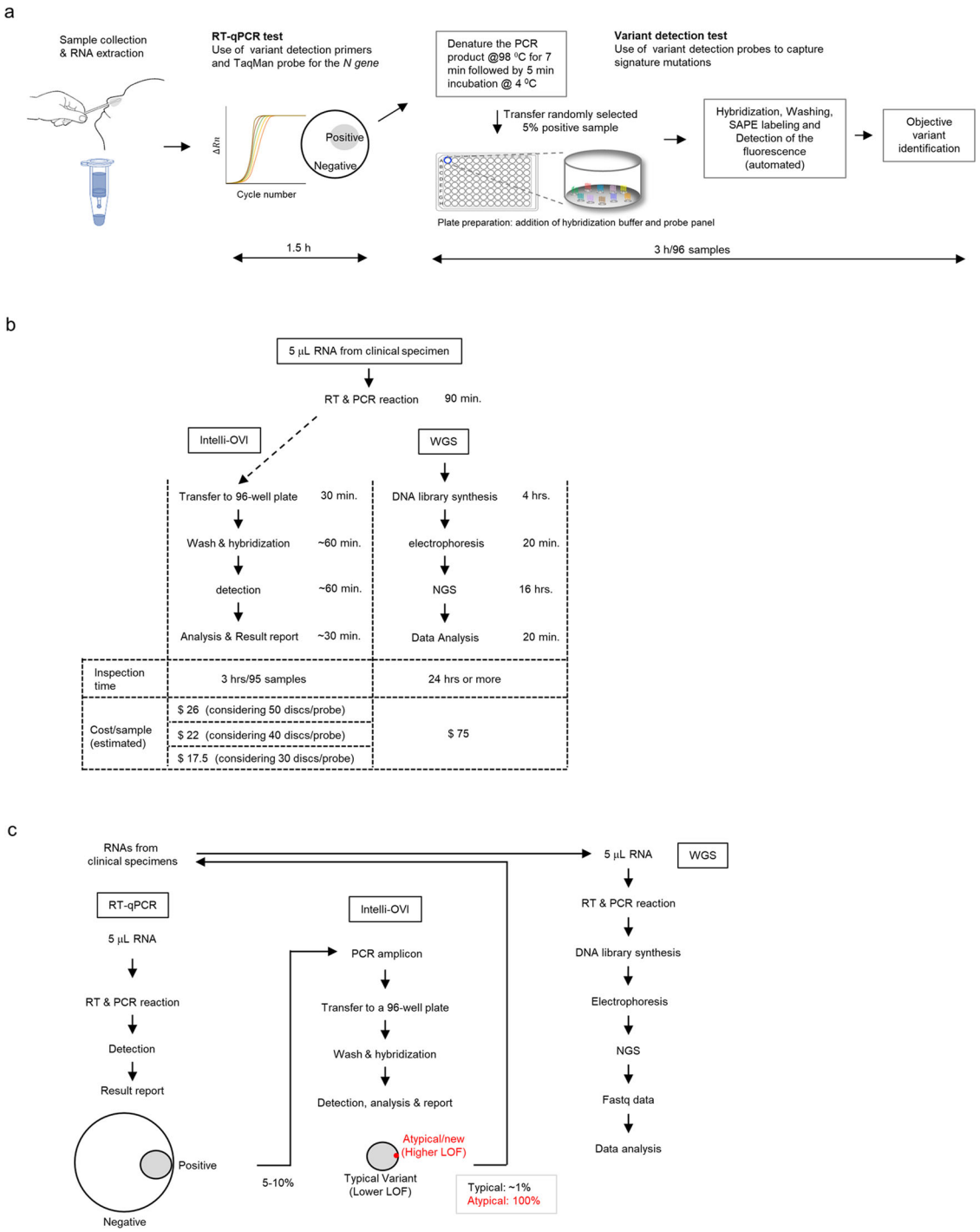

Fig. 6 | Strategy to expedite SARS-CoV-2 diagnosis and proposal for an efficient and cost-effective approach for variant monitoring. a Schematic illustration of an expedited approach for SARS-CoV-2 diagnosis combining TaqMan RT-qPCR and Intelli-OVI variant detection. **b** Comparison of the Intelli-OVI system with the WGS for the detection of the SARS-CoV-2 variant in terms of inspection time and cost. **c** The outline of a proposed efficient and cost-effective SARS-CoV-2 variant monitoring strategy introducing the Intelli-OVI system. The Intelli-OVI system could be an efficient system for routine variant monitoring. Once unknown cases are reported, WGS can be performed to characterise those samples along with a minimum percentage of known samples to look over the system performance. WGS whole viral genome sequencing, LOF local outlier factor, RT-qPCR reverse transcription-quantitative polymerase chain reaction.

been accumulating in the coding region of the spike protein. To date, several variants shared some common spike mutations, with emerging viruses acquiring new mutations in addition to the mutations already present in VOCs. There are 33 DNA probes targeting most of the representative mutations, which is likely to cover the majority of spike mutations in upcoming variants. Further, we showed that the addition of the bioinformatics approaches to the IntelliPlex data enabled us to identify new variants as different mutation phenotypes from known variants (Fig. 4 and Supplementary Fig. 6). This cannot be achieved by other methods, such as RT-qPCR, with a limited number of fluorescent probes. Thus, this is a strong advantage of multiple mutation panel data obtained by the πCode microdisc system.

Another possible application of the system is its use at quarantine stations in international airports or stations. With the emergence of dangerous pathogens of concern in global regions, it is essential to monitor these pathogens at international borders to prevent their spread from one country

to another. A noteworthy aspect of the system is its flexibility, conferred by sets of DNA probes targeting viral mutations or deletions. In our experience, the system can be upgraded within 1–2 weeks through the addition of new probes for novel virus mutations (Fig. 3d). Since the upgrade is not easy for end users to perform, the upgrade should be done in some core facility with the expertise of nucleic acid research like our team or producing company. After that, upgraded reagents can be distributed to end users. Considering the pathogenic characteristics and global dominance of new variants, regular upgrades of the primer and probe panel can be conducted at the manufacturer level (once or twice a year, depending on the outbreak situation). There are several different systems to be approved as clinical tests. In Japan, if the clinical test is approved as an in vitro diagnostic (IVD), frequent upgrades are challenging. Because it requires multiple stages of examination and takes a long time to be approved as an IVD. However, Intelli-OVI can be used as reagent for research use only same as WGS, which has also been upgraded several times according to accumulating mutation in the primer binding site for tiling PCR. Overall, prior to the upgrade, the research team needs to study both the publicly available sequences and their own repository to select new target regions for the design and validation of new primers and probes. In some circumstances, this may take over 2 weeks, depending on personnel, and the number of new targets.

IntelliPlex assay provides a result within 3–4 h, which is remarkably different from NGS-based pathogen identification. In theory, the IntelliPlex assay can rapidly monitor various targets simultaneously (e.g., SARS-CoV-2 variants). This will facilitate the ability to clearly distinguish viruses causing dangerous infections from those causing the common cold in people presenting with a high fever at a country border, owing to the multiple targets and timely nature of the assay. Again, this notion needs to be validated by further investigation.

One major limitation of Intelli-OVI is the requirement of upgrading the primers and probes once a mismatch appears in the primer/probe binding region in the new variant. Besides, it is necessary to perform WGS at a regular interval for a small number of samples to avoid overlooking new variants that do not have atypical mutations in the spike region, but instead harbour mutations in a different region of the viral genome. Another limitation is the difficulty in designing suitable candidate probes for certain target sequences, such as GC-rich regions where the probes can form secondary structures, although we experienced only one probe like that among 35 probes used in this study.

In conclusion, we demonstrated an efficient strategy to tackle a pandemic outbreak such as the current SARS-CoV-2 infection by employing interdisciplinary approaches, including nanomaterials, molecular biology, and bioinformatics technology. The flexibility and wide variation of πCode microdiscs enable the system to timely co-evolve with virus evolution. The system will be useful in establishing a sustainable world with people living with COVID-19 while also preparing for other emerging pathogens.

Data availability

The accession codes for all sequencing data (submitted to GISAID) utilized in this study, any other raw datasets, as well as, all source data for the figures in the main manuscript, including the numerical results underlying the graphs and charts, are available in the supplementary Data accompanying this paper (Supplementary Data 4, 5, and 6). Researchers and readers interested in accessing the data are encouraged to consult the supplementary material for comprehensive instructions on data retrieval and utilization. The corresponding author can be contacted via email for any further inquiries or assistance.

Code availability

The scripts for all analyses were written in Python programming language. The source code for each data analysis algorithm used in SARS-CoV-2 Intelli-OVI has been deposited to GitHub and is available from (https://github.com/satoulab).

Received: 22 June 2023; Accepted: 26 July 2024;
Published online: 09 August 2024

References

1. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
2. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
3. Dutta, A. COVID-19 waves: variant dynamics and control. *Sci. Rep.* **12**, 9332 (2022).
4. Thakur, V. et al. Waves and variants of SARS-CoV-2: understanding the causes and effect of the COVID-19 catastrophe. *Infection* **50**, 309–325 (2022).
5. Korber, B. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827 e819 (2020).
6. Faria, N. R. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021).
7. Motozono, C. et al. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host Microbe* **29**, 1124–1136 e1111 (2021).
8. Meng, B. et al. Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. *Nature* **603**, 706–714 (2022).
9. Grubaugh, N. D. et al. Tracking virus outbreaks in the twenty-first century. *Nat. Microbiol.* **4**, 10–19 (2019).
10. Neopane, P., Nypaver, J., Shrestha, R. & Beqaj, S. S. SARS-CoV-2 variants detection using TaqMan SARS-CoV-2 mutation panel molecular genotyping assays. *Infect. Drug Resist.* **14**, 4471–4479 (2021).
11. Chung, H. Y. et al. Emergency SARS-CoV-2 variants of concern: novel multiplex real-time RT-PCR assay for rapid detection and surveillance. *Microbiol. Spectr.* **10**, e0251321 (2022).
12. Oude Munnink, B. B. et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat. Med.* **26**, 1405–1410 (2020).
13. Frampton, D. et al. Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect. Dis.* **21**, 1246–1256 (2021).
14. Kao, J. H. et al. Clinical evaluation of IntelliPlex(TM) HCV genotyping kit for hepatitis C virus genotyping. *Diagn. Microbiol. Infect. Dis.* **94**, 344–348 (2019).
15. Suzuki, K. et al. HIV-1 viral blips are associated with repeated and increasingly high levels of cell-associated HIV-1 RNA transcriptional activity. *AIDS* **35**, 2095–2103 (2021).
16. Shirato, K. et al. Development of genetic diagnostic methods for detection for novel coronavirus 2019(nCoV-2019) in Japan. *Jpn. J. Infect. Dis.* **73**, 304–307 (2020).
17. Rajib, S. A. et al. A SARS-CoV-2 delta variant containing mutation in the probe binding region used for RT-qPCR test in Japan exhibited atypical PCR amplification and might induce false negative result. *J. Infect. Chemother.* **28**, 669–677 (2022).
18. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
19. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).
20. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
21. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
22. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
23. Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data – from vision to reality. *Eur. Surveill.* **22**, 30494 (2017).

24. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
25. Torii, S. et al. Establishment of a reverse genetics system for SARS-CoV-2 using circular polymerase extension reaction. *Cell Rep.* **35**, 109014 (2021).
26. Robert, T. M. & Friedman, T. J. *The elements of statistical learning, data mining, inference and prediction* 2nd edn (Springer, 2009).
27. Theodoridis, S. K. K. *Pattern Recognition.* (Academic Press, London, 1999).
28. Breunig, M. M., Kriegel, H. P., Ng, R. T. & Sander, J. LOF: identifying density-based local outliers. *Sigmod Rec.* **29**, 93–104 (2000).
29. Hossain, M. B. et al. Clone dynamics and its application for the diagnosis of enzootic bovine leukosis. *J. Virol.* **97**, e0154222 (2023).
30. Vanjur, L. et al. Non-langmuir kinetics of DNA surface hybridization. *Biophys. J.* **119**, 989–1001 (2020).
31. Cruz-Davalos, D. I. et al. Experimental conditions improving in-solution target enrichment for ancient DNA. *Mol. Ecol. Resour.* **17**, 508–522 (2017).
32. Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).
33. Piccoli, L. et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042.e1021 (2020).
34. Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).
35. Viana, R. et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679–686 (2022).
36. Planas, D. et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).
37. Lai, E. et al. A method for variant agnostic detection of SARS-CoV-2, rapid monitoring of circulating variants, and early detection of emergent variants such as Omicron. *J. Clin. Microbiol.* **60**, e0034222 (2022).
38. Kumar, M. et al. FnCas9-based CRISPR diagnostic for rapid and accurate detection of major SARS-CoV-2 variants on a paper strip. *eLife* **10**, e67130 (2021).
39. Hernandez, M. M. et al. A robust, highly multiplexed mass spectrometry assay to identify SARS-CoV-2 variants. *Microbiol Spectr.* **10**, e0173622 (2022).
40. Rybicka, M., Milosz, E. & Bielawski, K. P. Superiority of MALDI-TOF mass spectrometry over real-time PCR for SARS-CoV-2 RNA detection. *Viruses* https://doi.org/10.3390/v13050730 (2021).
41. Hernandez, M. M. et al. RT-PCR/MALDI-TOF mass spectrometry-based detection of SARS-CoV-2 in saliva specimens. *J. Med Virol.* **93**, 5481–5486 (2021).

Acknowledgements

We are grateful to Y. Matsuoka, N. Ninomiya, R. Fujito, R Irie, F Ishikawa, and Y Ichiyasu for their technical support and to E. Tanaka, A. Iemura, and T. Kashiwagi for their dedicated efforts to establish and run the PCR Center, Clinical Laboratory Center of Kumamoto City Medical Association. We are also grateful to Plex Bio Co., Ltd and Denka Co., Ltd for providing information about how to design probes when they developed the IntelliPlex SARS-CoV-2 variant detection kit version 2 (v2). However, we have upgraded the kit by ourselves by considering the information we obtained. We also thank Editage (www.editage.jp) for English language editing. This work was supported by research grants from the Japan Agency for Medical Research and Development (AMED) (JP22fk0410052, JP22wm0325015, JP22jm0210074, and JP22fk0410040) to Y.S. and JP23fk0410057 to K.S. Y.S. received research funds from Denka Co. Ltd. This work was also supported by the JSPS core-to-core program. The funders had no role in the study design, data collection, data interpretation, or discussion regarding submission for publication.

Author contributions

Study conception: Y.S. Methodology and formal analysis: M.B.H., Y.U., S.A.R., and Y.S. Investigation: M.B.H., S.A.R., A.R., M.T., B.J.Y.T., K. Sugata, M.N., M.K., H.T., R.K., K. Sadamasu, and Y.O. Data curation: M.B.H., Y.U., and Y.S. Resources: Y.O., M.N., R.K., K. Sadamasu, T.K., T.T., T.F., and K.Y. Writing/original draft: M.B.H. and Y.S. Writing/review and editing: Y.U., S.A.R., A.R., M.T., B.J.Y.T., K. Sugata, M.N., K. Sadamasu, Y.O., T.K., T.T., T.F., M.O., and K.Y. Supervision: T.K., M.O., K.Y., and Y.S. Project administration and funding acquisition: T.K., M.O., K.Y., and Y.S.

Competing interests

Y.S. received research funds from Denka Co. Ltd. The funders had no role in the study design, data collection, data interpretation, or discussion regarding submission for publication. All other authors have no competing interests to declare.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-024-00582-z.

Correspondence and requests for materials should be addressed to Yorifumi Satou.

Peer review information *Communications Medicine* thanks Matthew Hernandez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Valerie Salvatico. A peer review file is available.

[1]Division of Genomics and Transcriptomics, Joint Research Center for Human Retrovirus Infection, Kumamoto University, Kumamoto, Japan. [2]Department of Food Microbiology, Faculty of Nutrition and Food Science, Patuakhali Science and Technology University, Patuakhali, Bangladesh. [3]Department of Information and Communication Technology, Faculty of Engineering, University of Miyazaki, Miyazaki, Japan. [4]Department of Microbiology, Tokyo Metropolitan Institute of Public Health, Tokyo, Japan. [5]Clinical Laboratory Center of Kumamoto City Medical Association, Kumamoto, Japan. [6]Department of Medical Technology, Kumamoto Health Science University, Kumamoto, Japan. [7]Department of Microbiology and Immunology, Faculty of Medicine, Hokkaido University, Sapporo, Japan. [8]Institute for Vaccine Research and Development, HU-IVReD, Hokkaido University, Sapporo, Japan. [9]One Health Research Center, Hokkaido University, Sapporo, Japan. [10]AMED-CREST, Japan Agency for Medical Research and Development (AMED), Tokyo, Japan. [11]Laboratory of Virus Control, Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [12]Department of Life Sciences, Imperial College London, London, UK. [13]Collaboration Unit for Infection, Joint Research Center for Human Retrovirus Infection, Kumamoto University, Kumamoto, Japan. ✉e-mail: y-satou@kumamoto-u.ac.jp