Raman spectra identify heterogeneous vancomycin resistance and transcriptomes of *Staphylococcus aureus*

Ken-ichiro F. Kamei,[1*] Reiko Okura,[1] Koseki J. Kobayashi-Kirschvink,[2,3] Yuki Katayama,[4] Yuichi Wakamoto[1,5,6*]

[1]Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo 153-8902, Japan

[2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[3]Laser Biomedical Research Center, G. R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4]KatayamaKikai Co., Ltd, Fukui 918-8014, Japan

[5]Research Center for Complex System Biology, The University of Tokyo, Tokyo 153-8902, Japan

[6]Universal Biology Institute, The University of Tokyo, Tokyo 153-8902, Japan

*To whom correspondence should be addressed; E-mail: kenichiro_kamei@cell.c.u-tokyo.ac.jp (K.F.K.), cwaka@mail.ecc.u-tokyo.ac.jp (Y.W.).

Staphylococcus aureus **is a pathogenic bacterium that has caused multiple epidemics linked with the emergence of new antibiotic resistance. Vancomycin is the first-line antibiotic to treat methicillin-resistant *S. aureus* (MRSA) infection. However, several types of vancomycin nonsusceptible MRSA strains have been recognized to date. Although rapid assessment of their resistance levels and underlying molecular profiles are crucial for preventing their spread and counteracting resistance, the broad resistance spectrum and the diversity of genetic changes have impeded the practice. Here, we show that Raman spectra obtained from single cells can elucidate vancomycin resistance levels and global transcriptomic profiles of different MRSA strains. Dimension reduction and classification of Raman spectra reveal common cellular state changes linked with vancomycin susceptibility and differences among the strains with equivalent resistance. Furthermore, Raman spectra predicted more significant phenotypic heterogeneity in a population of a vancomycin-intermediate strain, which was confirmed by single-cell growth analysis. Therefore, these results demonstrate that antibiotic resistance, gene expression profiles, and phenotypic heterogeneity can be interrogated spectroscopically without laborious assays and sample preparation. Overall, our findings underscore the potential of Raman spectroscopy as a powerful and versatile tool in the ongoing battle against antibiotic resistance.**

Introduction

The emergence and spread of antibiotic-resistant bacteria is a huge threat to public health worldwide. *Staphylococcus aureus* is known for its ability to acquire antibiotic resistance and has caused several epidemic waves since the introduction of antibiotics (*1, 2*). Vancomycin was introduced in the 1980s to treat severe methicillin-resistant *S. aureus* (MRSA) infections (*3*). However, various vancomycin-nonsusceptible MRSA strains have been identified clinically (*3–9*), which are often associated with treatment failures (*10–13*).

Vancomycin-nonsusceptible *S. aureus* strains fall into three categories: vancomycin-resistant *S. aureus* (VRSA), vancomycin intermediately-resistant *S. aureus* (VISA), and heterogeneous vancomycin intermediately-resistant *S. aureus* (hVISA) (*3, 8, 9*). VRSA is the category with the highest vancomycin resistance, which requires a vancomycin concentration of 16 μg/mL or more to inhibit growth *in vitro*. The minimum inhibitory concentration (MIC) for VISA is lower than that of VRSA, but higher than that of vancomycin-susceptible MRSA (VSSA) strains ($\leq$2 μg/mL). hVISA strains are often judged as VSSA by the standard MIC test, but they contain small subpopulations with the VISA-level vancomycin resistance and could lead to treatment failures (*5, 14, 15*). The hVISA strains are considered to be in intermediate evolutionary steps from VSSA to VISA (*15, 16*). As VISA prevalence has been increasing recently (*8, 9*), it is critical to identify the resistant phenotypes of infected bacteria swiftly and treat infections with appropriate drug regimens.

The methicillin resistance of *S. aureus* is predominantly caused by the acquisition of the *mecA* or *mecC* gene on a mobile genetic element (*2, 3, 15, 17, 18*). However, the mechanisms behind vancomycin-nonsusceptibility are not yet fully understood and are believed to be diverse (*3, 9, 15, 16, 19–24*). Genetic analyses have suggested that multi-step mutations at various genetic loci can increase vancomycin resistance levels (*15, 16, 22, 23*). Consequently, it remains

challenging to identify the vancomycin resistance levels through targeted gene analysis.

Despite the diverse mutational profiles, it is suggested that physiological characteristic changes such as cell wall thickening and slow growth are omnipresent amongst vancomycin-nonsusceptible MRSA strains (*19, 20, 24–26*). Therefore, analyzing their global molecular composition and gene expression profiles might allow us to identify their vancomycin-resistance levels. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) combined with a machine-learning classification can distinguish different pathogenic bacteria and has also been applied to identify MRSA and different phenotypes of vancomycin-nonsusceptible MRSA strains (*27–29*). However, MALDI-TOF MS requires culturing cells to obtain microbial biomass (for bacteria, 10^4 to 10^6 colony forming unit) for analysis (*27*). Furthermore, identifying the comprehensive molecular signatures of targetted cells based on the MALDI-TOF MS spectra remains intractable. In addition, due to the population-based approach of the analysis, it is intractable to unravel single-cell-level phenotypic heterogeneity within an infected bacterial population.

Raman microscopy also provides spectral information reflecting the comprehensive molecular composition of single cells (*30–34*) and has been applied to distinguish different pathogenic bacterial species (*35*). Moreover, it has been demonstrated that Raman microscopy can distinguish MRSA from non-MRSA strains (*36*). A Raman spectrum from a cell is the linear superposition of the spectra of its constituent biomolecules. Due to the complexity of molecular compositions of cells and severe spectral overlaps, it is impossible to directly decompose a cellular Raman spectrum and determine the abundances of different cellular components except for limited types of molecules with prominent and strong Raman spectral peaks (*30–32*). However, recent studies have shown that global changes in the entire fingerprint region of cellular Raman spectra correlate with global gene expression profiles. Consequently, transcriptomic and proteomic profiles of cells under different culture conditions can be inferred from

Raman spectra thanks to the intrinsic low-dimensional constraint of global gene expression patterns (*33, 37, 38*). Therefore, even though it is impossible to directly assign Raman spectral peaks to diverse molecular species, cellular Raman spectra can provide valuable insights into the molecular underpinnings behind different vancomycin-resistance levels.

This report demonstrates that Raman spectra of different vancomycin-nonsusceptible MRSA strains can predict their vancomycin-resistance phenotypes and transcriptomic profiles. We also show that the Raman spectra and growth rates of individual cells of an hVISA strain are more diverse than those of VSSA and VISA strains, which confirms greater phenotypic heterogeneity among clonal hVISA cells.

Results

Raman spectra distinguish cellular states of MRSA strains with distinct vancomycin susceptibility. In this study, we analyzed nine MRSA strains. Of these, three strains (N315, Mu3, and Mu50) are clinical isolates with distinct susceptibility to vancomycin (*5*), while the other six (ΔIP, ΔIP1–ΔIP5) are strains created by introducing specific mutations or by selection under rifampicin exposure (*16*). Vancomycin susceptibility tests classify N315 and ΔIP as VSSA, Mu3, ΔIP1, and ΔIP2 as hVISA, and Mu50 and ΔIP3–ΔIP5 as VISA (Fig. 1A) (*16*).

We cultured these strains in brain heart infusion (BHI) medium and sampled cells in exponential phase for Raman measurements. Raman spectra of 48 single cells per strain were acquired using a custom-built Raman microscope (Fig. 1B, 1C and S1). To eliminate systematic errors in weak spontaneous Raman signals, the spectra were acquired in three biological replicates on different dates. We focused on the fingerprint region (700 to 1,800 cm^{-1}) of the Raman spectra in the analysis, where the signals from biomolecules such as proteins and metabolites are mainly observed.

We applied principal-component linear discriminant analysis (PC-LDA) to the Raman spec-

tra to determine whether they could distinguish differences in cellular states among the strains (*33, 38, 39*). The Raman spectra were classified in a low-dimensional (eight-dimensional) space (Fig. 1D). In this LDA space, each point represents a Raman spectrum from an individual cell. The results showed that the points for the cells in each strain were closely positioned and clustered in this LDA space (Fig. 1D). Visualizing the eight-dimensional Raman spectra by embedding them in a two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) confirmed the cluster separation (Fig. 1E). Therefore, the Raman spectra can distinguish the distinct cellular states of the nine MRSA strains.

Notably, the positions along the first LDA axis (LDA1) correlated with their vancomycin resistance levels (Pearson correlation coefficient $r = 0.81 \pm 0.15$, Fig. 1F and 1G). Therefore, PC-LDA detects the Raman spectral changes correlated with the vancomycin resistance levels as the dominant differences. Additionally, the second LDA axis (LDA2) recognizes the differences between the clinically isolated vancomycin-nonsusceptible strains (Mu3 and Mu50) and the constructed vancomycin-nonsusceptible strains (ΔIP1–ΔIP5) (Fig. 1F). Therefore, the result shows that, even among the MRSA strains with equivalent vancomycin resistance levels, there exist differences in their molecular profiles detectable by Raman spectra.

Raman spectral differences among the MRSA strains are explainable by their transcriptomes We next investigated whether the differences in Raman spectra of nine MRSA strains can be explained by their gene expression profiles. To test this, we gathered transcriptome data of these strains cultured in the BHI medium from public databases (*16, 21, 26*) and examined their correspondences with the Raman data. We hypothesized a linear correspondence

$$\mathbf{r}'_s = A\mathbf{t}_s, \tag{1}$$

where $\mathbf{r}'_s$ is the population average of the eight-dimensional Raman points of strain *s* (Fig. 1F), and $\mathbf{t}_s$ is the 2,628-dimensional transcriptome vector of the same strain containing the expres-

sion levels of 2,628 genes. A is the strain-independent matrix that maps the transcriptomes onto the Raman LDA space.

We conducted leave-one-out cross-validation (LOOCV) to verify the linear correspondence between Raman and transcriptome data (*33*). First, we estimated matrix A by using the Raman and transcriptome data of eight strains through partial least squares regression (PLS-R) (*40*). We then applied this estimated matrix to the transcriptome data of the excluded strain (test strain) to predict an eight-dimensional Raman point, represented as $\hat{\mathbf{r}}'_s$, which was then compared to the actual population average point, represented as $\mathbf{r}'_s$ (Fig. 2A). We repeated this process by excluding one strain at a time and evaluated the overall estimation error using predicted residual error sum of squares (PRESS), calculated as $\mathrm{PRESS}_{\mathbf{r}'} = \sum_s \|\hat{\mathbf{r}}'_s - \mathbf{r}_s\|^2$. A small PRESS value validates the assumption of linear correspondence between the Raman and transcriptome data.

We quantitatively evaluated the resulting $\mathrm{PRESS}_{\mathbf{r}'}$ value through a permutation test (*33*). We created 1,000 datasets in which we randomly permuted the correspondences between the Raman and transcriptome data, calculated the $\mathrm{PRESS}_{\mathbf{r}'}$ values for these false datasets, and compared them with the original $\mathrm{PRESS}_{\mathbf{r}'}$ value. The result confirmed a significantly small overall estimation error ($\mathrm{PRESS}_{\mathbf{r}'} = 4.46$, $p = 9.99 \times 10^{-4}$, permutation test, Fig. 2B). This result therefore suggests that the observed Raman spectral differences in the LDA space can be explained by the differences in the gene expression profiles of these MRSA strains.

Global transcriptome profiles can be inferred from low-dimensional Raman data. In the previous section, the matrix A was estimated using the Raman and transcriptome data of eight strains through PLS-R and used to predict the LDA Raman point of a test strain from its transcriptome data. Conversely, multiplying the LDA Raman data of a test strain $\mathbf{r}'_s$ by the Moore-Penrose pseudo-inverse of the estimated matrix A gives an estimate of its transcriptome profile $\hat{\mathbf{t}}_s$. Following this procedure, we calculated the transcriptome profiles of the nine strains from

their Raman spectra and compared them with the actual transcriptome profiles. The result shows that the transcriptome profile estimated from the population average of the LDA Raman data of each strain agrees well with the measured profile (Fig. 3A and 3B).

Noticing that transcriptome profiles of different strains are also strongly correlated (Fig. S2), we calculated the PRESS for the transcriptome data as $\mathrm{PRESS}_{\mathbf{t}} = \sum_s \|\hat{\mathbf{t}}_s - \mathbf{t}_s\|^2$ to quantify the overall estimation error. A permutation test with randomly permuted 1,000 datasets found a significantly small overall estimation error ($\mathrm{PRESS}_{\mathbf{t}} = 1.07 \times 10^4$, $p = 9.99 \times 10^{-4}$, Fig. 3C). We also quantified the overall estimation error using other distance metrics and consistently found significantly small estimation errors (Table S1).

Therefore, differences in the Raman spectra of the MRSA strains contain the information that allows the prediction of global gene expression differences of these strains. Note that the transcriptome inference was performed based on the low-dimensional LDA Raman data; the high precision of transcriptome estimation suggests a constraint in the global transcriptomic changes caused by the genotypic differences among the MRSA strains.

The clinically isolated hVISA Mu3 strain is physiologically heterogeneous. Previous research has demonstrated that the hVISA Mu3 strain contains subpopulations that have the same level of resistance to vancomycin as VISA (*5*). This suggests that the *S. aureus* strains with the hVISA phenotype are physiologically more diverse among clonal cells than those of VSSA (N315) and VISA (Mu50) strains.

In line with this picture, we noticed that the Raman spectra of the Mu3 strain were more heterogeneous among individual cells than those of the clinically isolated VSSA and VISA strains. We evaluated the standard deviations of cellular Raman spectra at different wavenumbers among the clinically isolated MRSA strains (N315, Mu3, and Mu50), finding that the variations were consistently larger in Mu3 than those of N315 and Mu50 (Fig. 4A). This result suggests that

the Mu3 cells are more heterogeneous in their molecular composition at the single-cell level.

To ask whether the heterogeneity recognized by their Raman spectra is physiologically relevant, we observed the growth and microcolony formation of the three clinically isolated strains on the BHI agar by time-lapse microscopy (Movies S1–S3). As expected, the growth curves of Mu3 were more heterogeneous than those of N315 and Mu50 among the microcolonies grown from different individual cells (Fig. 4B–4D). The coefficient of variation (CV) of the growth rate of Mu3 was 36%, significantly higher than those of N315 and Mu50, which were both 18% (Fig. 4E–4G). The mean growth rate of Mu50 was smaller than that of N315 (0.47 h^{-1} for N315 and 0.26 h^{-1} for Mu50), which corroborates the slow-growing trait of the VISA strain. Although the median of the growth rate of Mu3 is close to that of N315, Mu3's growth rate distribution is skewed with a heavy left tail, indicating the presence of subpopulations in the Mu3 strain growing as slowly as the Mu50 strain (Fig. 4E–4G).

The fact that cell-to-cell heterogeneity of Mu3 is consistently larger in both Raman spectra and growth supports that this hVISA strain is physiologically more diverse among genetically identical cells. Furthermore, the results demonstrate that phenotypic heterogeneity in cellular population can be interrogated spectroscopically by evaluating their Raman spectral heterogeneity.

"Single-cell MIC" suggests heterogeneous vancomycin resistance in Mu3 strain. The previous section showed that Mu3 clonal cells were phenotypically diverse. This implies that vancomycin resistance level is also diverse at the single-cell level in a clonal hVISA Mu3 population.

To examine variations in vancomycin resistance level within single strains, we attempted to quantitatively evaluate vancomycin resistance of single cells from cellular Raman spectra. Although MIC is a measure of antibiotic resistance defined for cell populations determinable

by bulk measurements, we here extended the concept of this measure and calculated virtual MIC values for single cells as a measure of antibiotic resistance of each cell. Specifically, we calculated a coefficient vector that transforms a Raman spectrum to an MIC value against vancomycin by simply applying a linear regression method to a dataset consisting of (normal) population-level MIC values and strain-wise average of randomly-selected single-cell Raman spectra (Fig. S3A). We subsequently multiplied each of the remaining single-cell Raman spectra by the coefficient vector and obtained "single-cell MIC" values (Fig. S3A). Repeating the step of random sampling of spectra produced distributions of single-cell MIC (Fig. S3B and S3C), the median of which correlated with the normal population-level MIC (Pearson correlation coefficient $r = 0.79$, Fig. 4H). The standard deviation of the logarithm of single-cell MIC was significantly larger in the Mu3 strain than those of the other strains (Fig. 4I).

These results suggest that the Mu3 strain is phenotypically more heterogeneous even in its resistance to vancomycin compared to the other VSSA and VISA strains. The observed heterogeneity of single-cell MIC might account for the existence of subpopulations with VISA-level vancomycin resistance in the hVISA Mu3 populations.

DISCUSSION

Vancomycin resistance undermines the treatment efficacy of MRSA infection. Inappropriate administration of vancomycin could result in generating and enhancing the resistance of infected bacteria. Therefore, it is important to quickly detect vancomycin resistance to make the right choices in drug regimens and prevent the emergence and spread of vancomycin-nonsusceptible MRSA strains.

This study shows that the global transcriptome profiles and vancomycin resistance levels of the MRSA strains can be identified by their Raman spectral patterns (Fig. 1G, 3A, and 3B). Cellular Raman spectra can be obtained by exposing cells to light without the need for staining or tagging. Furthermore, unlike the MALDI-TOF MS-based method, it is not necessary to culture cells to target bacterial cells for Raman measurements, as long as they are recognizable in specimens. As a result, Raman microscopy has significant potential for quickly uncovering antibiotic resistance of infected bacteria in patients.

Previous studies have shown that cellular Raman spectra obtained under different culture conditions can be used to infer transcriptomic and proteomic differences (*33, 37, 38*). The present study shows that transcriptomic differences among the MRSA strains with different genotypes under identical culture conditions can also be inferred from their Raman spectral patterns (Fig. 3A and 3B). Therefore, Raman-based inference of gene expression profiles applies not only to external environmental differences but also to internal genotypic differences.

The Mu3 strain was found to be phenotypically heterogeneous in both growth and Raman spectra. This result suggests that cellular Raman spectra can interrogate cell-to-cell heterogeneity within a cellular population. In the case of hVISA, such phenotypic heterogeneity could be linked to resistance of a subpopulation under vancomycin exposures. In line with this view, single-cell vancomycin resistance metric estimated from single-cell Raman spectra

showed high heterogeneity for the Mu3 strain. Therefore, single-cell Raman analysis may provide crucial information about the potential risk of resistance that cannot be assessed by a single population-level resistance metric such as MIC.

While our study focused on vancomycin resistance in MRSA, the broader applicability of this method to detect drug resistance in other bacterial and fungal pathogens represents an exciting avenue for future research. Exploring the adaptability of Raman spectroscopy to various pathogenic strains could significantly enhance our capacity to combat antibiotic resistance.

MATERIALS AND METHODS

Bacterial strains. Nine *S. aureus* strains were used in this study. Three strains, N315, Mu3, and Mu50, are clinical isolates (*5*). The other six strains, ΔIP, ΔIP1, ΔIP2, ΔIP3, ΔIP4, and ΔIP5, were constructed from N315 by introducing mutations or by selection under rifampin exposures (*16*). The vancomycin resistant phenotypes of the strains are VSSA for N315 and ΔIP; hVISA for Mu3, ΔIP1, and ΔIP2; and VISA for Mu50, ΔIP3, ΔIP4, and ΔIP5 (*16*).

Raman measurement and preprocessing of Raman signals. To prepare samples for Raman measurements, *S. aureus* cells were inoculated from a glycerol stock into $3\,\mathrm{mL}$ BHI medium (Eiken Chemical Co., Ltd.) and cultured with shaking (Taitec BR-32FP, $200\,\mathrm{r\,min^{-1}}$) at $37\,^{\circ}\mathrm{C}$ overnight. A $30\,\mu\mathrm{L}$ overnight culture was inoculated into $3\,\mathrm{mL}$ BHI medium and grown to an exponential phase (OD_{600} = 2.0–2.5) with shaking at $37\,^{\circ}\mathrm{C}$. $1\,\mathrm{mL}$ of cell suspension was taken in an eppentube, and the cells were washed with phosphate-buffered saline (PBS) twice. The cells were fixed with $4\,\%$ (w/w) paraformaldehyde for $15\,\mathrm{min}$ at room temperature. The fixed cells were washed twice with $0.1\,\mathrm{M}$ glycine. After substituting the liquid by $1\,\mathrm{mL}$ PBS, the cell suspensions were stored at $4\,^{\circ}\mathrm{C}$ until Raman measurement (max. one day).

Immediately before the Raman measurement, the liquid of cell suspension was substituted by $1\,\mathrm{mL}$ distilled water. $5\,\mu\mathrm{L}$ of the cell suspension was placed on a synthetic quartz slide glass (TOSHIN RIKO Co., LTD.) and dried.

Raman spectra of the cells were measured using a custom-built Raman microscope with $532\,\mathrm{nm}$ excitation wavelength (Gem 532, Laser Quantum). Laser power at the sample plane was approximately $17.5\,\mathrm{mW}$. Raman spectra were obtained from single cells with exposure time of $5\,\mathrm{s}$. For each strain, 16 cells were measured per replicate. Three biological replicates were measured on different days. Consequently, we obtained Raman spectra from 48 cells for each strain. Background Raman spectra were also measured in regions adjacent to target

cells. Raman signals were detected by an sCMOS camera (ORCA-Flash4.0 V2, Hamamatsu Photonics) water-cooled at $15\,^{\circ}\mathrm{C}$. The optical setup was the same as that used for the Raman measurement of *Escherichia coli* in (*33*).

An sCMOS-specific noise reduction filter developed in (*33*) was applied to the obtained images. Then, the region where the spectra were recorded was cropped, and pixel counts were summed along the direction perpendicular to the wavenumber. Next, background Raman spectra were subtracted from single-cell Raman spectra, and the region of 632 to 1862 cm^{-1} was cropped. Subsequently, each spectrum was smoothed by the Savitzky-Golay filter, and normalized by subtracting the average and dividing it by the standard deviation.

Dimension reduction of cellular Raman spectra. Principal component analysis (PCA) was first applied to the normalized spectra to reduce noise. Principal components that explained $98\,\%$ of the variance were used. After PCA, linear discriminant analysis (LDA) was applied. LDA is a supervised method, which calculates the most discriminatory bases by maximizing the ratio of the between-group (in the case of this research, between-strain) variances to the sum of within-group (within-strain) variances. The dimension was reduced to $n - 1$ by LDA, where $n\ (= 9)$ was the number of groups (strains).

Linearly connecting LDA Raman spectra with transcriptomes. A linear regression method, partial least squares regression (PLS-R), was conducted using the dimension-reduced LDA Raman spectra and the transcriptome data. The transcriptomes of the nine *S. aureus* strains were obtained from the public database (NCBI GEO, accession number GSE43643, GSE42218, and GSE50826).

The dimensions of the transcriptomes were reduced to $n - 3$ in PLS-R. To evaluate the linearity between the LDA Raman spectra and transcriptomes, leave-one-out cross-validation (LOOCV) was performed. We calculated the predicted residual error sum of squares (PRESS)

as estimation errors. Statistical significance of the correspondence was tested with permutation test. See (*33*) for the detail of PLS-R, LOOCV, and permutation test.

Heterogeneity analysis of Raman spectra of the clinically isolated strains. To make the intensities of the average spectra approximately equal among the three strains, three outlier single-cell Raman spectra were removed for each strain. Then, the standard deviation of the remaining 45 normalized spectra was calculated at each wavenumber for each strain.

"Single-cell MIC" estimation from Raman spectra. As described in the previous section, the intensities of the average spectra were made approximately equal among the three clinically isolated strains. Similarly, the intensities of the average spectra of the six constructed strains were also made approximately the same as those of the three clinically isolated strains by removing outlier single-cell Raman spectra. Half of the spectra were randomly selected and the average of the normalized spectra was calculated for each strain. Then, a linear regression method, PLS-R, was applied on the mean spectra and the logarithm of MIC reported in (*16*) (Fig. S3A). Multiplying the obtained regression coefficient vector and the remaining single-cell normalized Raman spectra, which had not been selected for the training, produced logarithm of "single-cell MIC" values for each strain (Fig. S3A). The random sampling step and the subsequent calculations were repeated 1,000 times, and a distribution of "single-cell MIC" values were obtained for each strain (Fig. S3B and S3C). For example, if a certain strain has 40 individual cells' Raman spectra in total, 20 "single-cell MIC" values are obtained for this strain per one random sampling. Thus, 1,000 repetition of this calculation process produces a distribution composed of 20,000 "single-cell MIC" values for this strain. The distribution of each strain shown in Fig. S3B is obtained in this manner.

Time-lapse microscopy. To prepare samples for time-lapse microscopy, *S. aureus* cells were inoculated from a glycerol stock into $2\,\mathrm{mL}$ BHI medium and cultured with shaking ($200\,\mathrm{r\,min^{-1}}$) at $37\,^{\circ}\mathrm{C}$ overnight. A $15\,\mu\mathrm{L}$ overnight culture was inoculated into $3\,\mathrm{mL}$ BHI medium and grown to exponential phase (OD_{600} = 2.4–2.5) with shaking at $37\,^{\circ}\mathrm{C}$. $1\,\mu\mathrm{L}$ of cell suspension was taken on a coverslip (Neo No.1, $24\,\mathrm{mm}\times60\,\mathrm{mm}$, Matsunami) and covered by a BHI agarose pad ($20\,\mathrm{mm}\times25\,\mathrm{mm}\times7\,\mathrm{mm}$ (thickness)) solidified with $1.5\,\%$ low-melt agarose (GeneMate LowMelt Agarose, BioExpress). The assembled sample was placed on a microscope XY stage and covered by a transparent plastic dish to prevent the agarose from drying.

We used Nikon Eclipse Ti-E inverted microscope equipped with digital sCMOS camera (ORCA-Flash4.0 V3, Hamamatsu Photonics), phase-contrast objective (Plan Apo λ, 100×, NA1.45). We obtained phase-contrast images from 50 XY positions with $1\,\mathrm{min}$ interval in each experiment.

The acquired time-lapse images were analyzed using ImageJ (*41*). We applied Mexican Hat Filter to find outlines of microclonies of *S. aureus* and measured the areas to quantify the growth.

Heterogeneity analysis of microcolony growth of the clinically isolated strains. Growth rates were calculated by taking logarithm to 2 of relative colony sizes at $t = 3\,\mathrm{h}$ to those at $t = 0\,\mathrm{h}$ and dividing them by $3\,\mathrm{h}$.

Data analysis. All the data analyses in this study were conducted using MATLAB (R2019a and R2023b).

Declaration of interests

YK is the president and representative director of KatayamaKikai Co., Ltd. YW and KJKK are listed as inventors of patents (JP6993682 and US10,379,052 B2) filed by The University of Tokyo. The other authors declare no competing interests.

Acknowledgments

We thank Tetsuo Yamaguchi for reading the manuscript and providing comments. This work was supported by JST CREST Grant Number JPMJCR1927 (Y.W.); JST ERATO Grant Number JPMJER1902 (Y.W.); JSPS KAKENHI Grant Numbers 19J22448 (K.F.K.); and Takeda Science Foundation Award (Y.K. and Y.W.).

References

1. H. F. Chambers, F. R. Deleo, *Nature Reviews Microbiology* **7**, 629 (2009).

2. H. Grundmann, M. A. de Sousa, J. Boyce, E. Tiemersma, *The Lancet* **368**, 874 (2006).

3. Y. Cong, S. Yang, X. Rao, *Journal of Advanced Research* **21**, 169 (2020).

4. K. Hiramatsu, *et al.*, *Journal of Antimicrobial Chemotherapy* **40**, 135 (1997).

5. K. Hiramatsu, *et al.*, *The Lancet* **350**, 1670 (1997).

6. S. Chang, *et al.*, *New England Journal of Medicine* **348**, 1342 (2003).

7. L. M. Weigel, *et al.*, *Science* **302**, 1569 (2003).

8. S. Zhang, X. Sun, W. Chang, Y. Dai, X. Ma, *PLOS ONE* **10**, e0136082 (2015).

9. A. Shariati, *et al.*, *Scientific Reports* **10**, 1 (2020).

10. P. G. Charles, P. B. Ward, P. D. Johnson, B. P. Howden, M. L. Grayson, *Clinical Infectious Diseases* **38**, 448 (2004).

11. H. M. Neoh, *et al.*, *Annals of Clinical Microbiology and Antimicrobials* **6**, 1 (2007).

12. Y. Maor, *et al.*, *Journal of Clinical Microbiology* **45**, 1511 (2007).

13. R. K. Fong, J. Low, T. H. Koh, A. Kurup, *European Journal of Clinical Microbiology and Infectious Diseases* **28**, 983 (2009).

14. Y. Katayama, H. Murakami-Kuroda, L. Cui, K. Hiramatsu, *Antimicrobial Agents and Chemotherapy* **53**, 3190 (2009).

15. K. Hiramatsu, *et al.*, *Journal of Infection and Chemotherapy* **20**, 593 (2014).

16. Y. Katayama, M. Sekine, T. Hishinuma, Y. Aiba, K. Hiramatsu, *Antimicrobial Agents and Chemotherapy* **60**, 3730 (2016).

17. T. Ito, *et al.*, *Antimicrobial Agents and Chemotherapy* **45**, 1323 (2001).

18. K. Hiramatsu, L. Cui, M. Kuroda, T. Ito, *Trends in Microbiology* **9**, 486 (2001).

19. H. Hanaki, *et al.*, *Journal of Antimicrobial Chemotherapy* **42**, 199 (1998).

20. K. Sieradzki, A. Tomasz, *Journal of Bacteriology* **185**, 7103 (2003).

21. Y. Aiba, *et al.*, *Antimicrobial Agents and Chemotherapy* **57**, 4861 (2013).

22. S. Gardete, A. Tomasz, *The Journal of Clinical Investigation* **124**, 2836 (2014).

23. Q. Hu, H. Peng, X. Rao, *Frontiers in Microbiology* **7**, 221405 (2016).

24. L. Cui, H. Murakami, K. Kuwahara-Arai, H. Hanaki, K. Hiramatsu, *Antimicrobial Agents and Chemotherapy* **44**, 2276 (2000).

25. L. Cui, *et al.*, *Journal of Clinical Microbiology* **41**, 5 (2003).

26. M. Saito, *et al.*, *Antimicrobial agents and chemotherapy* **58**, 5024 (2014).

27. A. Wieser, L. Schneider, J. Jung, S. Schubert, *Applied Microbiology and Biotechnology* **93**, 965 (2012).

28. C. A. Mather, B. J. Werth, S. Sivagnanam, D. J. SenGupta, S. M. Butler-Wu, *Journal of Clinical Microbiology* **54**, 883 (2016).

29. K. Asakura, *et al.*, *PLoS ONE* **13**, 1 (2018).

30. I. W. Schie, T. Huser, *Applied Spectroscopy* **67**, 813 (2013).

31. H. J. Butler, *et al.*, *Nature Protocols* **11**, 664 (2016).

32. R. Smith, K. L. Wright, L. Ashton, *Analyst* **141**, 3590 (2016).

33. K. J. Kobayashi-Kirschvink, *et al.*, *Cell systems* **7**, 104 (2018).

34. A. Germond, *et al.*, *Communications Biology* **1**, 85 (2018).

35. C. S. Ho, *et al.*, *Nature Communications* **10** (2019).

36. A. Pistiki, *et al.*, *Microbiology Spectrum* **10** (2022).

37. K. J. Kobayashi-Kirschvink, *et al.*, *bioRxiv* p. 2021.11.30.470655 (2022).

38. K. F. Kamei, *et al.*, *bioRxiv* p. 2023.05.09.539921 (2023).

39. R. Goodacre, *et al.*, *Microbiology* **144**, 1157 (1998).

40. P. S. Gromski, *et al.*, *Analytica Chimica Acta* **879**, 10 (2015).

41. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, *Nature Methods* **9**, 671 (2012).

Figure legends

Figure 1 Single-cell Raman spectra of MRSA strains and their correlation with vancomycin resistance. **(A)** Nine MRSA strains analyzed in this study. Three strains are clinical isolates, while the other six are strains constructed in a laboratory (*16*). Vancomycin susceptibility of these nine strains are classified into VSSA, hVISA, and VISA. Colors and markers in this figure are used in some of the other figures throughout this paper. **(B)** Schematic figure representing our Raman measurements. Raman spectra were obtained from single cells placed on a synthetic quartz slide glass by focusing laser light on each targeted cell. **(C)** Representative single-cell Raman spectra. One is from ΔIP (top), and the other is from N315 (bottom). **(D)** Cellular Raman spectra in LDA space. LDA reduces the dimensionality of the spectra to $8\ (= 9 - 1)$. Each point represents a spectrum from a single cell. Each ellipse shows the 95% concentration ellipse. Projection to the plane spanned by the first two axes (LDA1 and LDA2) is shown. **(E)** t-SNE visualization of the eight-dimensional LDA space. **(F)** The population average of the single-cell Raman spectra of each strain in the LDA space. The arrows indicate the changes of the averages along the strain construction and selection steps in (*16*). **(G)** Correlation between the positions along the LDA1 axis and vancomycin resistance levels. Pearson correlation coefficient, $r = 0.81 \pm 0.15$.

Figure 2 Transcriptome profiles of different MRSA strains can explain their Raman spectral differences. **(A)** Estimation of Raman spectra in the LDA space from transcriptomes. Thick colored points represent Raman spectra estimated from the transcriptomes assuming the linear correspondence (Eq. 1). **(B)** Histogram of $\mathrm{PRESS}_{\mathbf{r}'}$ values of 1,000 randomly permuted data. The $\mathrm{PRESS}_{\mathbf{r}'}$ of the original dataset is 4.46. The p-value was 9.99×10^{-4} (no occurrence of smaller $\mathrm{PRESS}_{\mathbf{r}'}$ was observed in the 1,000 permuted datasets).

Figure 3 Inference of global transcriptome profiles from cellular Raman spectra. (A and B) Estimation of transcriptomes from Raman spectra in the LDA space. In (A), grey points represent the measured RNA abundance (average of the replicate measurements; error bar, SE), which are sorted along the horizontal axis. Green points represent the RNA abundance estimated from Raman spectra. (B) is scatterplots between the measured and the estimated RNA abundances. Straight lines represent $x = y$. **(C)** Histogram of $\mathrm{PRESS_t}$ values of 1,000 randomly permuted data. The $\mathrm{PRESS_t}$ of the original dataset is 1.07×10^4. The p-value was 9.99×10^{-4} (there was no occurrence of smaller $\mathrm{PRESS_t}$ in the 1,000 permuted datasets).

Figure 4 Heterogeneity of clinically isolated hVISA strain Mu3. (A) Comparison of standard deviations of normalized single-cell Raman spectra among the clinically isolated VSSA, hVISA, and VISA strains. **(B–D)** Microcolony growth of the three clinically isolated strains on BHI agar. Single lines represent relative size growths of different single colonies. **(E–G)** Growth rate distributions of the microcolonies on BHI agar shown in Fig. 4B–4D. **(H)** Scatterplot between normal MIC and the median of "single-cell MIC" estimated from single-cell Raman spectra. The straight line represents $x = y$. Pearson correlation coefficient $r = 0.79$. **(I)** Comparison of standard deviations of the logarithm of "single-cell MIC."

Fig. 1

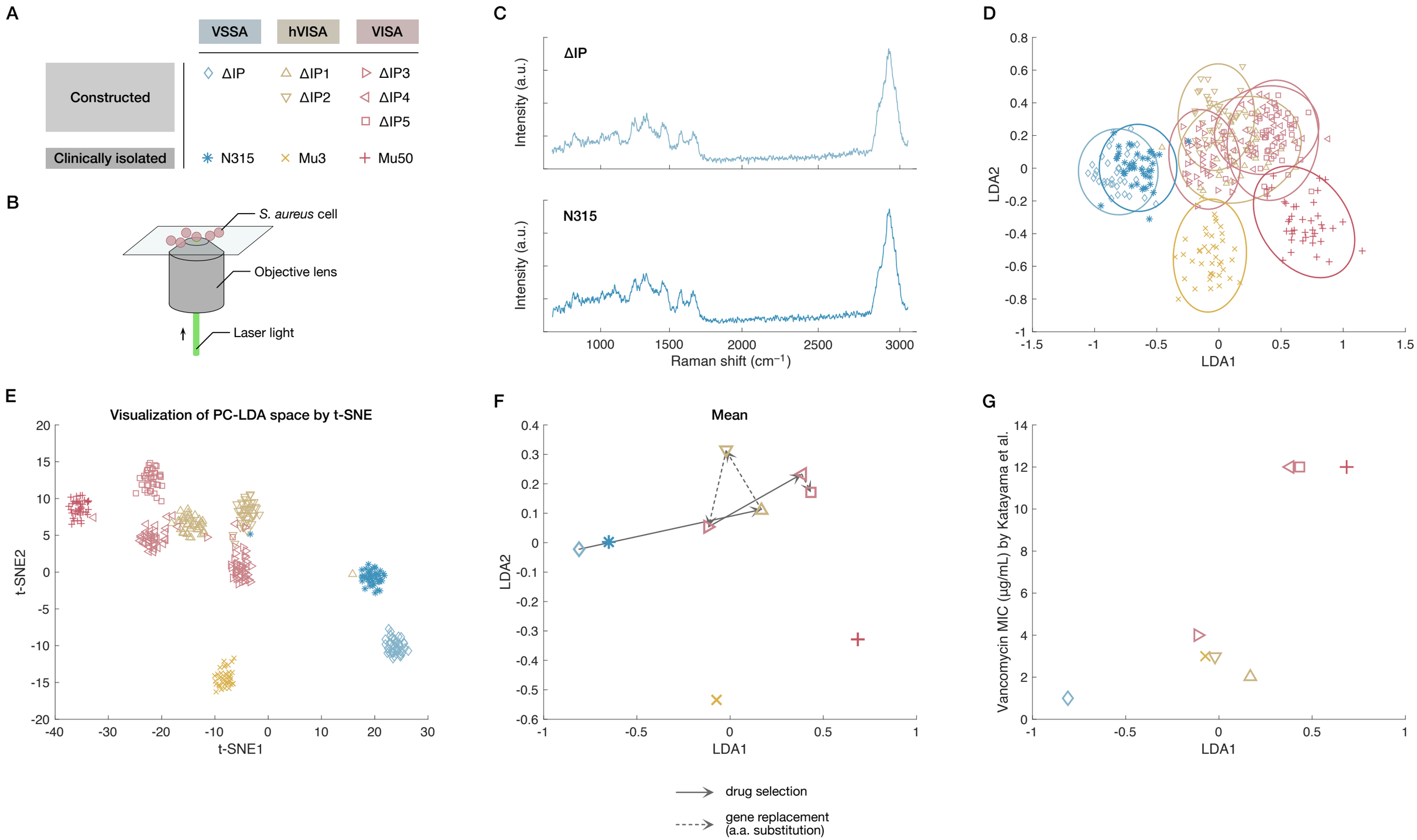
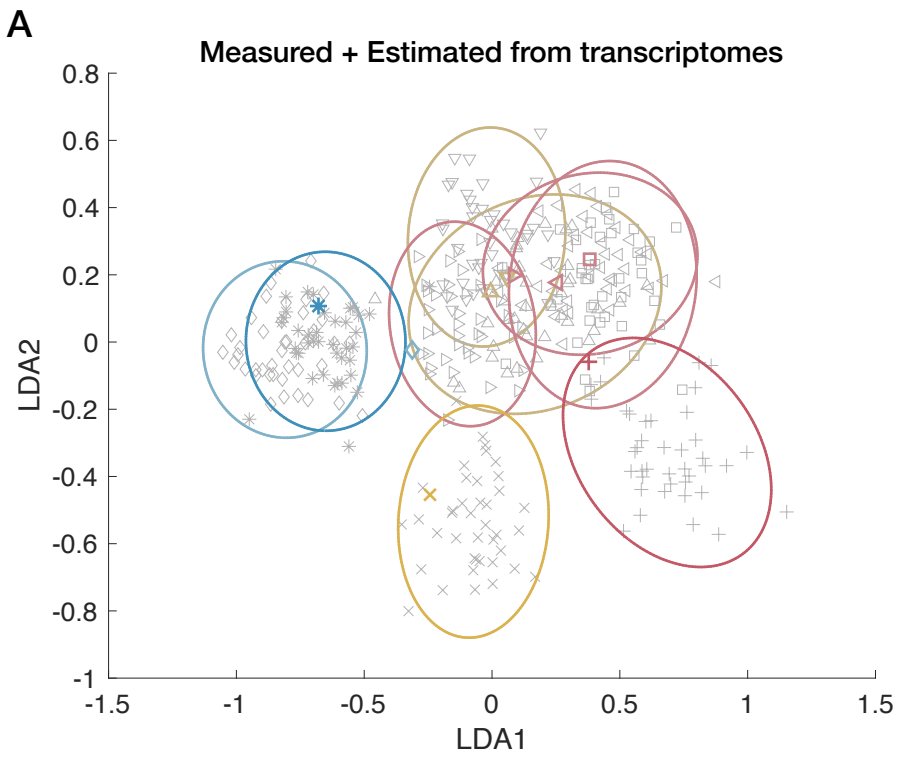
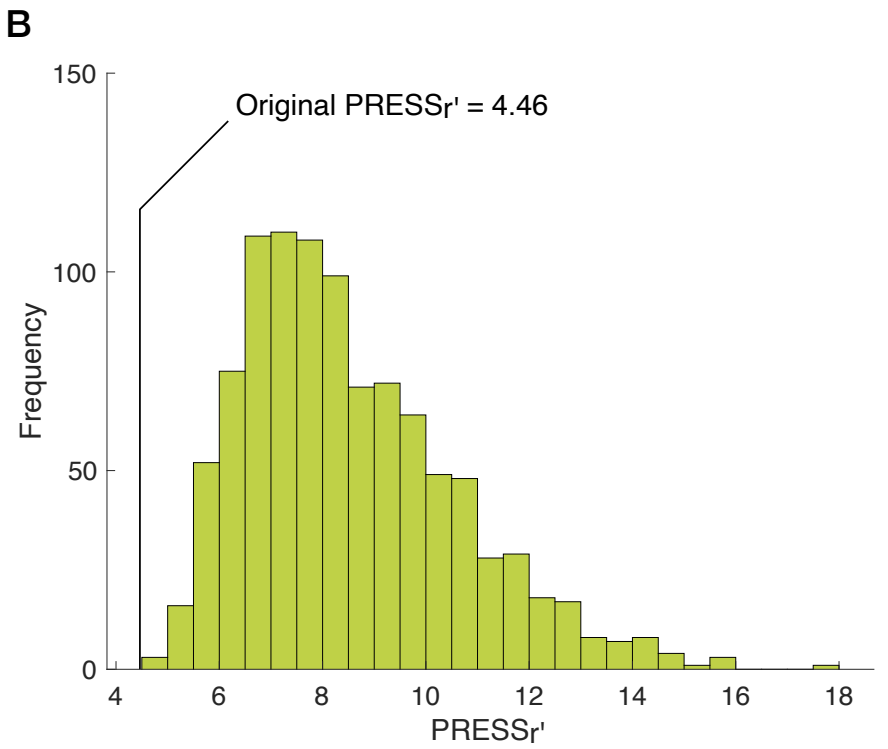

A
VSSA
hVISA
VISA
Constructed
ΔIP
ΔIP1
ΔIP2
ΔIP3
ΔIP4
ΔIP5
Clinically isolated
N315
Mu3
Mu50
B
S. aureus cell
Objective lens
Laser light
C
ΔIP
N315
Intensity (a.u.)
Raman shift (cm⁻¹)
D
LDA1
LDA2
E
Visualization of PC-LDA space by t-SNE
t-SNE1
t-SNE2
F
Mean
LDA1
LDA2
G
Vancomycin MIC (μg/mL) by Katayama et al.
LDA1
drug selection
gene replacement
(a.a. substitution)

Fig. 2

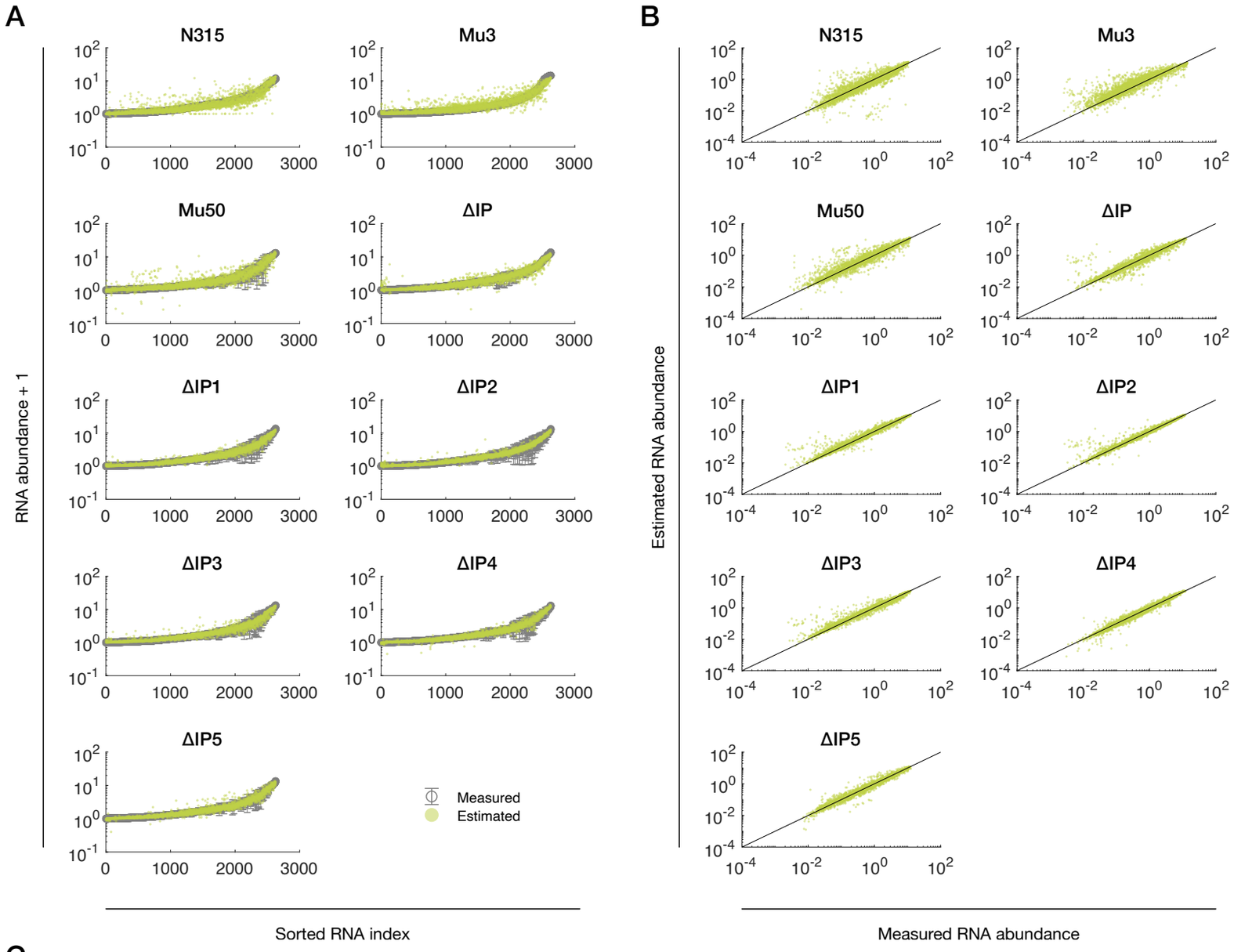




Gray points measured (same as Fig. 1C)

Colored points estimated from transcriptomes

Fig. 3


A
N315
Mu3
Mu50
ΔIP
ΔIP1
ΔIP2
ΔIP3
ΔIP4
ΔIP5
RNA abundance + 1
Sorted RNA index
Measured
Estimated
B
Estimated RNA abundance
Measured RNA abundance

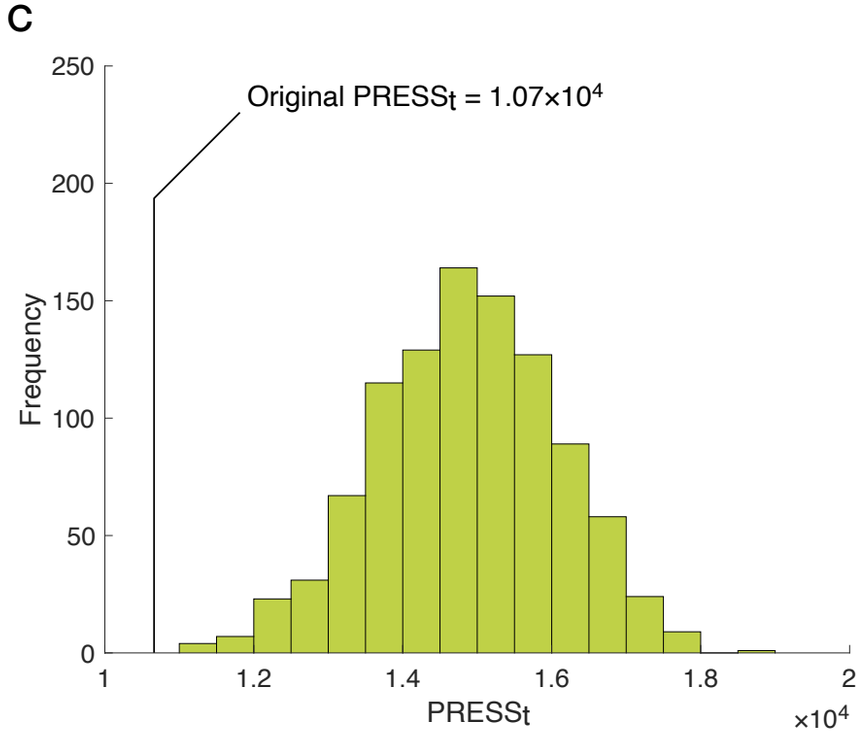
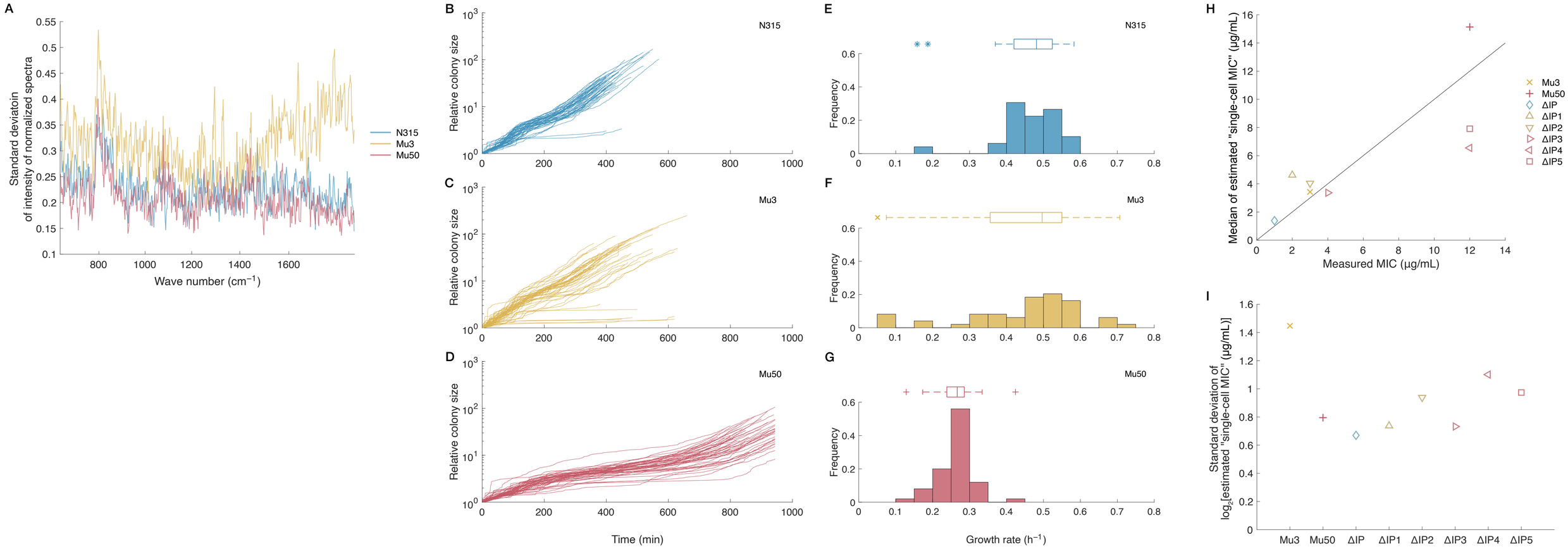

C
Original PRESSt = 1.07×10^4
Frequency
PRESSt
×10^4

Fig. 4


A
Standard deviatoin of intensity of normalized spectra
Wave number (cm^{-1})
N315
Mu3
Mu50
B
C
D
Relative colony size
Time (min)
E
F
G
Frequency
Growth rate (h^{-1})
H
Median of estimated "single-cell MIC" (μg/mL)
Measured MIC (μg/mL)
Mu3
Mu50
ΔIP
ΔIP1
ΔIP2
ΔIP3
ΔIP4
ΔIP5
I
Standard deviation of log$_2$[estimated "single-cell MIC" (μg/mL)]