Infrared spectroscopy coupled with a dispersion model for quantifying the real-time dynamics of kanamycin resistance in artificial microbiota

Naifu Jin[1], Maria Paraskevaidi[2], Kirk T. Semple[1], Francis L. Martin[2,*], Dayi Zhang[1,*]

[1]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK;
[2]School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK

Corresponding authors:

Dayi Zhang, Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK; Tel.: +44(0)1524 510288; Fax: +44(0)1524 510082, Email: d.zhang@lancaster.ac.uk

Francis L. Martin, School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK; Tel.: +44(0)1772 896482; Email: flmartin@uclan.ac.uk

Abstract

Over-usage of antibiotics leads to the widespread induction of antibiotic resistance genes (ARGs). Developing an approach to allow real-time monitoring and fast prediction of ARGs dynamics in clinical or environmental samples has become an urgent matter. Vibrational spectroscopy is potentially an ideal technique towards the characterization of the microbial composition of microbiota as it is non-destructive, high-throughput and label-free. Herein, we employed attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy and developed a spectrochemical tool to quantify the static and dynamic composition of kanamycin resistance in artificial microbiota to evaluate microbial antibiotic resistance. Second order differentiation was introduced in identifying the spectral biomarkers, and principal component analysis followed by linear discriminant analysis (PCA-LDA) was used for the multivariate analysis of the entire spectral features employed. The calculated results of the mathematical dispersion model coupled with PCA-LDA showed high similarity to the designed microbiota structure, with no significant difference (P >0.05) in the static treatments. Moreover, our model successfully predicted the dynamics of kanamycin resistance within artificial microbiota under kanamycin pressures. This work lends new insights into the potential role of spectrochemical analyses in investigating the existence and trends of antibiotic resistance in microbiota.

Keywords Antibiotic resistance, Artificial microbiota, ATR-FTIR spectroscopy, Kanamycin, Multivariate analysis, Spectrochemical

Introduction

Antibiotics have played a vital role in modern medicine contributing to a considerable reduction in childhood mortality and increasing life expectancy[1]. However, the increasing number of fatal infections caused by antibiotic-resistant bacteria is gradually developing into a global threat. The environment has become the primary "sink" for most applied antibiotics and their residues arising from human or animal excretion[1-3]. Since bacteria with antibiotic resistance genes (ARGs) can tolerate antibiotics, selection pressures from contaminated water or soil will boost the abundance of ARGs in the environment and increase the possibility of their spread through microbial species[4,5]. Therefore, real-time monitoring and quantification of ARGs or antibiotic-resistant bacteria is urgently required.

Besides measuring the concentration of antibiotics *via* chemical analysis, various biological analytical methods have been used to determine the presence, abundance and diversity of ARGs in the microbiota to capture a "static map" of their existence, *e.g.*, meta-sequence and quantitative polymerase chain reaction (qPCR)[6,7]. However, genetically identical cells from the same population have stochasticity in gene expression, meaning that there is significant variation in their molecular content and phenotype, even under similar environmental influences. Moreover, bacterial resistance to the antibiotics can also be affected and regulated epigenetically[8]. In combination, these factors provide an opportunity for phenotypic and cell-type diversity regardless of genotype[9]. This questions the reliability of determining ARGs abundance by molecular biological approaches in real-world situations, leading to the necessity of developing a phenotypic assay that depicts *in situ* dynamics of ARGs or microbial antibiotic resistance in environmental samples.

It is well accepted that genetic and epigenetic factors cannot be studied independently as a complete phenotype emerges from both together[10]. The spectrochemical analysis is an alternative approach to characterize the phenotypic features of organisms and has already demonstrated its ability to investigate clinical

samples, as well as to describe and identify bacterial species[11,12]. Previous studies indicates that spectroscopic techniques are capable of studying phenotypic features, at either population[13] or single-cell[14] level, such as diagnosing the distinct spectral signatures and metabolomes from isogenic cell lines[15]. However, the current techniques have limited application in characterizing ARGs under antibiotic pressures, mainly due to the lack of appropriate analytical models and well-trained databases. Recently, some studies using spectroscopic techniques have set out to investigate biological response to environmental stress, like nanomaterials[16,17] and antimicrobial reagents[18,19]. The introduction of spectroscopic techniques coupled with a suitable prediction model to characterize microbial composition may bring new insights in detecting the presence or even the dynamics of microbial antibiotic resistance in environmental microbiota in real-time, owing to its non-destructive, high-throughput and label-free character[20,21]. It also allows for *in situ* spectral measurements, helping in understanding the interactions between microbes and their physical environment.

Kanamycin is a subclass of aminoglycoside antibiotics, one of the most widely applied antibiotics in health and molecular biology[22]. Because of the well-established mechanisms of kanamycin resistance and characterized sequence[23], it was selected as the model antibiotic in the present study. Herein, we used attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy, coupled with the multivariate analysis and the dispersion indicator model, to quantify the kanamycin resistance within artificial microbiota and evaluate their phenotypic change associated with kanamycin resistance, from both static and dynamic perspectives. This work raises the potential feasibility of applying spectroscopic techniques to diagnose ARGs phenotypic dynamics in the microbial community *in situ*.

Experimental section

Sample preparation

The present study included two strains without kanamycin-resistant-gene, *Mycobacterium vanbaalenii* PYR-1 and *Escherichia coli* DH5α, and one

kanamycin-resistant strain *Acinetobacter baylyi* ADPWH_recA, which has a continuously expressed kanamycin resistance gene kan^R (from Mini-Tn5/Km[24], Genbank accession number: U32991.1) inserted into the *recA* gene in the chromosome of *A. baylyi* ADP1[25]. Before the experiment, they were all cultured in Luria-Bertani (LB) broth medium for 24 h at 30±2°C.

The three control groups contained pure *M. vanbaalenii* PYR-1, *E. coli* DH5α and *A. baylyi* ADPWH_recA, respectively. The artificial microbiotas were prepared for both static (M1 to M5) and dynamic (AM1 and AM2) experiments by gently mixing the cells in the compositions listed in Table 1. The optical density at 600 nm (OD_{600}) in each treatment was monitored continuously for 24 h by a multimode plate reader (FLUOstar Omega, Germany) to evaluate bacterial growth. For static tests, the cells were directly collected by centrifugation (4000 rpm for 5 min), washed three times with sterile deionized water to remove the residues of growth media and then suspended in 70% ethanol to fix the bacterial cells. For dynamic tests, all the artificial microbiotas were treated with kanamycin (final concentration 10 mg/L). After exposure for 4, 8, 12 or 24 h, the cells from microbiotas were harvested following the same procedure as above.

Table 1. The compositions of artificial microbiotas (volume ratio, v:v:v).

Treatments	Control			Static test					Dynamic test	
	M. vanbaalenii	***E. coli***	***A. baylyi***	**M_1**	**M_2**	**M_3**	**M_4**	**M_5**	**AM_1**	**AM_2**
M. vanbaalenii	100%	-	-	40%	30%	30%	15%	5%	40%	25%
E. coli	-	100%	-	50%	45%	20%	10%	5%	40%	25%
A. baylyi	-	-	100%	10%	25%	50%	75%	90%	20%	50%

ATR-FTIR spectroscopy

The washed cell pellets (minimal amount >5 μL) were applied onto Low-E slides for the interrogation by ATR-FTIR spectroscopy. A TENSOR 27 FTIR spectrometer (Bruker Optics Ltd., UK) equipped with a Helios ATR attachment (containing a

diamond internal IRE; incidence angle of the IR beam: 45°) was used. Instrument parameters were set at 32 scans and spatial resolution of 8 cm^{-1}. Before the measurement of a new sample, the crystal was cleaned with deionized water and background readings were retaken. A total of 30 spectra were randomly acquired for each treatment (3 replicates).

Computational analysis

The primary analysis methods employed in this study involved multivariate analysis and the dispersion indicator model. The initial data generated from ATR-FTIR spectroscopy were analyzed within MATLAB R2011a (*TheMathsWorks, Natick, MA, USA*) software, coupled with IRootLab toolbox (http://irootlab.googlecode.com)[26]. Unless otherwise stated, the acquired spectra were cut to the biochemical-cell fingerprint region (1800-900 cm^{-1}), rubberband baseline corrected and normalized to Amide I (1650 cm^{-1}). Second order differentiation baseline correction and vector normalization was also performed as an alternative mean to process the data (the number of the filter coefficients of the Savitzky-Golay smoothing/differentiation filter was 9). Principal component analysis followed by linear discriminant analysis (PCA-LDA) was subsequently applied to the pre-processed data to reduce the number of spectra to 10 uncorrelated principal components (PCs), which account for >99% of the total variance; LDA is a supervised technique coupled with PCA in order to maximize inter-class and minimize intra-class variance[21]. In addition, cluster vector approach was conducted to visualise the discriminating difference[21,27]. This method takes input from PCA-LDA to create a loadings vector for each category contributing to respective data points. The pseudo-spectra allow identifying which variables (or wavenumber) are responsible for variance in the data set related to the original spectra[21,27]. The detailed information of the dispersion indicator model was described in the Electronic Supplementary Information (ESI).

Biological analysis

The copy numbers of total bacterial 16S rRNA and targeted kanamycin resistance gene (kan^R) were determined by quantitative polymerase chain reaction (qPCR). For 16S rRNA, the primer pair set was 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3'), and the primer pair for kan^R was KanF (5'-TGTCATACCACTTGTCCGCC-3') and KanR

(5'-ATCGAGCTGTATGCGGAGTG-3'). The 20 μL qPCR system consisted of 2 μL of each primer, 1 μL DNA template, 5 μL molecular water and 10 μL iTaq™ Universal SYBR® Green Supermix (BioRad, USA). The relative abundance of *kanR* in each pure strain was calculated as the ratio of *kanR* copy numbers to 16S rRNA copy numbers (kanR/16S). The microbial kanamycin resistance within the artificial microbiota was calculated as the ratio of *A. baylyi* population to the total bacterial population.

Statistical analysis

One-way analysis of variance (ANOVA) with Tukey's post hoc test/or T-test was employed to examine the discriminating differences. All statistical analysis was carried out in GraphPad Prism 6.

Results and Discussion

Growth and kanamycin resistance gene of individual strains

All the three bacterial strains (*A. baylyi* ADPWH_recA, *M. vanbaalenii* PYR-1 and *E. coli* DH5α) had similar growth curves without kanamycin pressure (see ESI Figure S1A). Cultivated in 10 mg/L kanamycin, only *A. baylyi* ADPWH_recA maintained positive growth because of the expression of kan^R gene and resistance to kanamycin (see ESI, Figure S1B). Neither *M. vanbaalenii* PYR-1 nor *E. coli* DH5α grew post-exposure to 10 mg/L kanamycin. The results of qPCR further confirmed that the high relative abundance of kan^R gene (kanR/16S) were only found in *A. baylyi* ADPWH_recA (0.306 in medium without kanamycin and 0.275 in medium with 10 mg/L kanamycin respectively, no significant difference), whereas it was less than 0.001 or below the limit of detection for *M. vanbaalenii* PYR-1 or *E. coli* DH5α (see ESI, Figure S2). It was further proved that kanamycin resistance gene is only detectable in *A. baylyi* ADPWH_recA, but neither *M. vanbaalenii* PYR-1 nor *E. coli* DH5α, and the latter two cannot tolerate kanamycin pressure. The active group of kanamycin, 2-deoxystreptamine, impairs bacterial protein synthesis through binding to prokaryotic ribosomes 30S subunit [22]. The kan^R encoding neomycin phosphotransferase is an aminoglycoside-modifying enzyme, using ATP as donor to modify the hydroxyl functions of 2-deoxystreptamine and inhibit its binding to

ribosomes[28]. The *kan*^{R} gene is therefore a reliable molecular indicator in detecting the kanamycin resistance.

IR spectral fingerprints of individual strains and microbiotas

The IR spectral fingerprint region (1800 - 900 cm^{-1}) of the three strains and artificial microbiotas are shown in Figure 1. The representative peaks of the biochemical fingerprint include lipids (~1750 cm^{-1}), Amide I (~1650 cm^{-1}), Amide II (~1550 cm^{-1}), Amide III (~1260 cm^{-1}), carbohydrate (~1155 cm^{-1}), asymmetric phosphate stretching vibrations ($\nu_{as}PO_2^-$; ~1225 cm^{-1}), symmetric phosphate stretching vibrations ($\nu_sPO_2^-$; ~1080 cm^{-1}), glycogen (~1030 cm^{-1}) and protein phosphorylation (~970 cm^{-1})[20 21]. Past literatures[12,20,29,30] suggest the characteristic peaks given by the region can be used as biomarkers to characterize microbial cell types (even at subspecies level) and diagnose microbe-induced diseases.

However, the visual spectral differences with the mean spectra are almost identical regardless of the bacterial species or community composition. For this reason, we applied the cluster vectors after multivariate analysis (PCA-LDA) and the second order differentiation baseline correction to further reveal the underlying biochemical differences between each strain or microbiota. Based on the derived spectral biomarkers from PCA-LDA (Figure 1B), all the microbiota samples showed marked segregation (see ESI, Table S1). Characteristics associated with microbial composition were observed in particular wavenumber-absorbance intensities. For instance, the intensities at 980 cm^{-1} and 1740 cm^{-1} were increased with increasing ratio of ARGs but fluctuated in some artificial microbiotas, particularly for microbiota M3 (*M. vanbaalenii* PYR-1: *E. coli* DH5α: *A. baylyi* ADPWH_recA = 30%:20%:50%). Additionally, IR spectral analysis (Figure 2A) based on the second order differentiation baseline correction and vector normalization highlighted several key biomarkers. Two apparent shifts from ~1630 cm^{-1} to ~1640 cm^{-1} (Amide I) and from ~1222 cm^{-1} to ~1235 cm^{-1} ($\nu_{as}PO_2^-$) associated with *A. baylyi* were regarded as biomarkers for the presence of kanamycin resistance. These spectral alterations might be attributed to the upregulated activities of the *kan*^{R} encoding aminoglycoside *O*-phosphotransferase, which contributes to microbial resistance by inactivating kanamycin molecular via catalyzing ATP-dependent phosphorylation of specific aminoglycoside hydroxyl groups[31]. Some other weaker discriminations included the

polysaccharide fingerprint region (1000-1150 cm^{-1}) and the protein absorbance region (1500-1700 cm^{-1})[27]. These alterations were probably induced by the interference of extracellular polymeric substances (EPS) produced by different species[32-34] and resulted in the difficulties in distinguishing biomarkers from the PCA-LDA extracted peaks. Based on the previous studies[32,35,36], we speculate that these extracellular materials may interact with each other and generate new biochemical compositions within the communities, influencing the discriminating peaks obtained with spectrochemical interrogation.

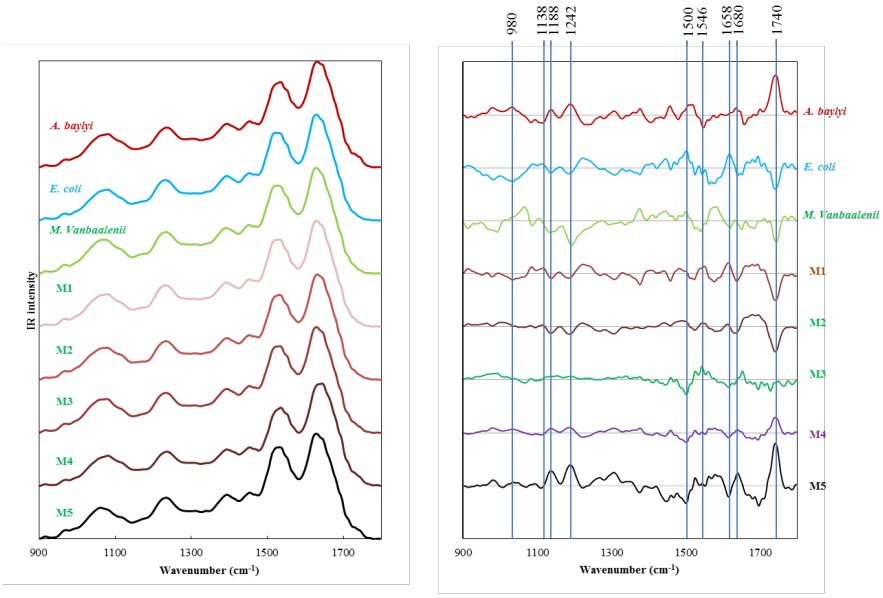


Figure 1. (A) Infrared spectra of *A. baylyi*, *M. vanbaalenii*, *E. coli* and five artificial microbiotas (M1-M5). (B) Cluster vector plots after PCA-LDA, indicating significant wavenumbers for the segregation between bacterial species and artificial microbiotas.

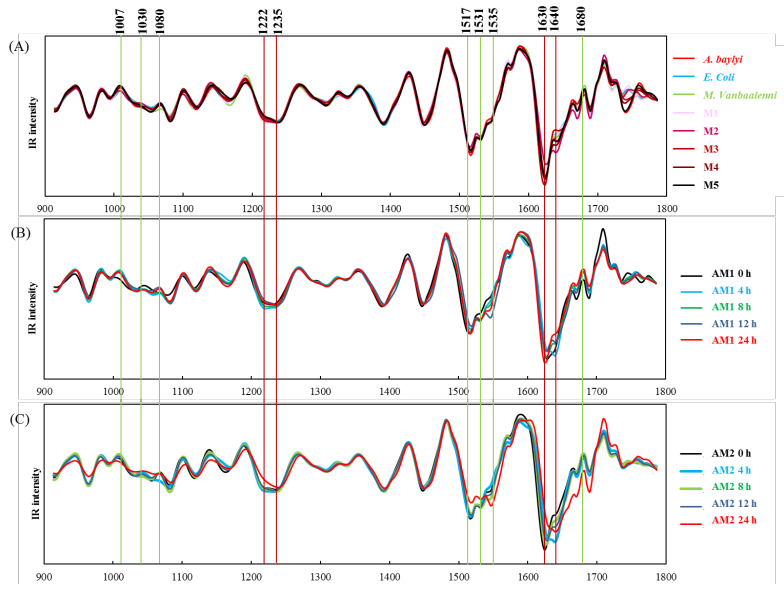


Figure 2. Class means spectra of pre-processed data based on second order differentiation baseline correction and vector normalization. (A) Processed spectra of *A. baylyi*, *M. vanbaalenii*, *E. coli* and five artificial microbiotas (M1-M5). (B) Processed spectra of AM1 at different time point in dynamic experiment. (C) Processed spectra of AM1 at different time point in dynamic experiment.

Predicting community composition in artificial microbiotas

Comparing to the IR spectra in the static tests, we observed identical spectral biomarkers in artificial community dynamics (Figure 2B and 2C) that the same shifts from ~1630 cm^{-1} to ~1640 cm^{-1} (Amide I) and from ~1222 cm^{-1} to ~1235 cm^{-1} ($\nu_{as}PO_2^-$) developed along with the time. The results indicated the consistent spectral biomarkers in both static and dynamic microbiotas in analyzing the phenotypic presence and abundance of kanamycin resistance gene in the targeted microbiota.

The PCA-LDA scores plot (Figure 3A) also illustrates a significant segregation of the different groups, associated with differing microbiota compositions. The control groups (*M. vanbaalenii*, *E. coli,* and *A. baylyi*) are clearly separated from each other. In contrast to *M. vanbaalenii* and *E. coli*, all the converted spectral values of *A. baylyi* are aligned as negative along linear discriminant one (LD1), likely attributed to its kanamycin resistance. Meanwhile, along with linear discriminant two (LD2), the group of *M. vanbaalenii* (Gram-positive bacteria) is located on the negative axis alone, separated from the other two groups (*E. coli* and *A. baylyi*), which are Gram-negative.

The five artificial microbiota samples (M1 to M5) are located inbetween, and their distances to the control groups are correlated with their community compositions.

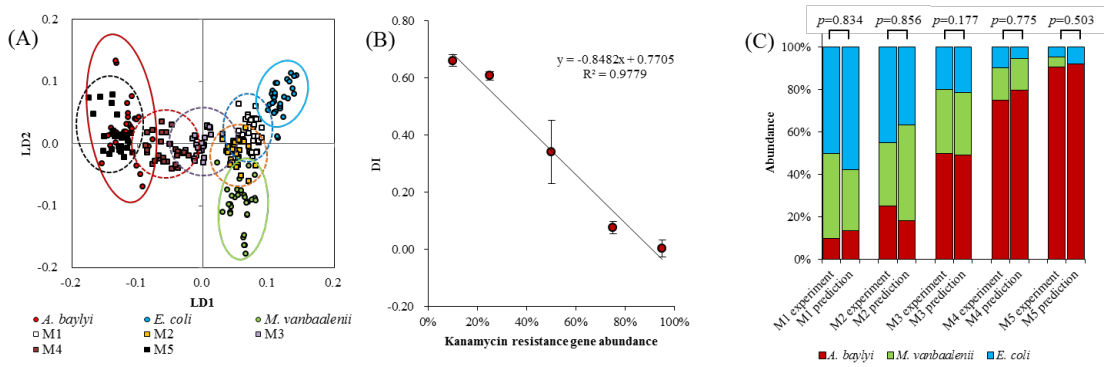


Figure 3. (A) Two-dimensional (LD1, LD2) scores plot after PCA-LDA of pure microbial strains and artificial microbiotas with different composition. (B) Correlation between kanamycin resistance gene abundance and group distance dispersion (D_I). (C) Comparison of artificial microbiota composition between experimental data and model prediction.

In order to predict the composition of artificial microbiota, the dispersion indicator model[37] was carried out by transferring the dispersion analysis from the IR spectral variables to the vectors (LD1 and LD2) and using D_I as the indicator, comparing to the ARGs gene copy numbers quantified by qPCR as reference. This method used the summarized spectral information from PCA-LDA which accounts for over 90% of spectral variations in the present study, and was more conclusive than the limited biomarkers from second order differentiation. Here, microbiotas with less abundance of *A. baylyi* were further separated from the *A. baylyi* group, but closer to those of *E. coli* and *M. vanbaalenii*, leading to an increasing D_I against the decreasing kanamycin resistance (kanamycin resistance genes in *A. baylyi*). Figure 3B illustrates the negative linear correlation between D_I and the abundance of *A. baylyi* (kanamycin resistance gene abundance) within the artificial microbiotas ($D_I = -0.8482 \times [kanamycin\ resistance\ gene] + 0.7705$). The high coefficient (R^2=0.9779) suggests a good linear regression of D_I against kanamycin resistance. The composition of each microbiota was, therefore, calculated from the D_I linear

regression based on PCA-LDA, as shown in Figure 3C. The results indicated that the predicted microbial compositions had high similarity to their theoretical structure with no significant differences found (P >0.05). The standard deviation of microbiota M3 (middle point in Figure 3B) was greater than the others, possibly attributing to their higher Shannon-Wiener index (1.02) than other microbiotas (0.35 to 0.94 for M1, M2, M4 and M5). Shannon-Wiener index represents the diversity of microbial community, and higher microbial diversity has been reported to increase complicated intracommunity interaction[32]. It might cause huge variation of microbial chemical composition, consequently leading to the difficulties in interrogating spectral biomarkers and significant standard deviation in data prediction.

Quantification of kanamycin resistance dynamics within microbiota

Figure 4A illustrates the PCA-LDA scores plot of microbiotas post-exposure to kanamycin, derived from the spectral dynamics of the artificial microbiotas (see ESI Figure S3). All the interrogated communities exhibit a dramatic shift from the original location as the exposure time increases. The *M. vanbaalenii* category moves towards a different direction when compared to *A. baylyi* and *E. coli*, which might be attributed to distinct cell structures between Gram-positive (*M. vanbaalenii*) and Gram-negative bacteria (*A. baylyi* and *E. coli*). Specifically, there is only one lipid bilayer in the membrane of Gram-positive bacteria, with a thick ring of peptidoglycan and teichoic acid[38,39]. On the other hand, the cell membrane of Gram-negative bacteria contains two lipid associated bilayers, which appear to increase the chance that the applied treatments influence their structure[38,39]. The artificial microbiotas, AM1 and AM2, follow similar trends as the *A. baylyi* and they come even closer to *A. baylyi* after extended exposure to the kanamycin antibiotic. After PCA-LDA, the most discriminating peaks were observed in Gram-negative bacteria and were attributed to lipids (~1750 cm^{-1}), $\nu_{as}PO_2^-$ (~1225 cm^{-1}) and $\nu_sPO_2^-$ (~1080 cm^{-1}). Kanamycin's antimicrobial mechanism is associated with aminoglycosides, interfering with aminoacyl-tRNA recognition at the ribosomal A site and disrupting protein expression[40]. Such a mechanism causes series of secondary effects, *e.g.*, membrane damage. Our results are consistent with previous findings showing that the damage is mainly linked to a broad range of alterations associated with the elements of membranes, *e.g.*, proteins, supported by derived peaks the protein absorbance region from 1500 to 1700 cm^{-1}, such as Amide II (~1517 cm^{-1}, ~1543 cm^{-1}) and Amide I

(~1650 cm^{-1}, ~1680 cm^{-1})[16,21,41,42].

Applying the linear D_1 regression model, we successfully predicted the dynamic abundance of *A. baylyi* and kanamycin resistance within the microbiotas under kanamycin antibiotic pressures. Both artificial microbiotas, AM1 (Figure 4B) and AM2 (Figure 4C), had defined community composition at 0 h, with *A. baylyi* (kanamycin resistance gene) accounting for 10% and 40% of the total population, respectively. Post-exposure to kanamycin, the ARGs abundance from qPCR results gradually increased to 85.0% in AM1 and 92.2% in AM2 after 12 h, which is explained by the competitive advantages of bacteria with kanamycin resistance gene in the community[43]. It therefore led to a faster growth of *A. baylyi* compared to other strains and subsequent dominancy of *A. baylyi* within the microbiota. From the dynamics of discriminant functions, the predicted ARGs abundance in both microbiotas fitted efficiently with experimental data (Figure 4B and 4C). The linear correlation at each time point did not show significant difference between predicted and experimental ARGs abundance (Figure 4D), with a Pearson correlation coefficient of 0.9487. The prediction *via* infrared spectroscopy coupled and multivariate analysis fitted the experimental data better at higher ARGs abundance, but was slightly lower than the qPCR results at low ARGs abundance, *e.g.*, 33% in Figure 4D, which might underestimate the ARGs abundance to some extent. These results not only prove that our model can be used for static community composition and abundance/dynamics of kanamycin resistance gene, but they also evaluate the impact of antibiotic pressure on kanamycin resistance gene transfer or dominancy.

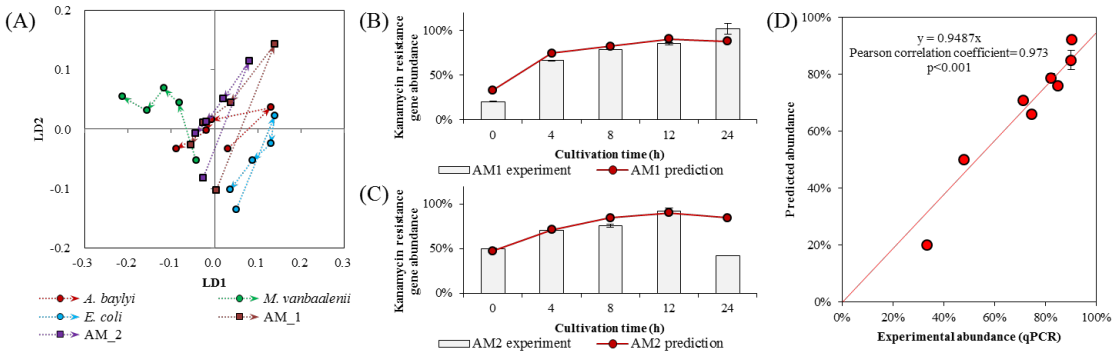


Figure 4. (A) Two-dimensional (LD1, LD2) scores plot after PCA-LDA of IR dynamics of artificial microbiotas. Dots along with the arrow point in each colour

refer to the measurement at 0, 4, 8, 12 and 24 h, respectively. The prediction of kanamycin resistance gene abundance is based on the dispersion among the classification groups in PCA-LDA for artificial microbiotas AM1 (B) and AM2 (C). (D) Regression correlation of kanamycin resistance gene abundance between experimental data *via* qPCR and model prediction.

It is worth mentioning that less dispersion is observed for *A. baylyi* after exposure because *A. baylyi* ADPWH_recA contains the kan^R kanamycin resistance gene, which is capable of tolerating kanamycin pressure. In the present study, the kan^R kanamycin resistance gene belongs to *npt* encoding neomycin phosphotransferase and shows high similarity to *addA* encoding aminoglycoside phosphotransferase (aminoglycoside kinase), which modifies the aminoglycosides by phosphoryl transfer, catalysing the phosphate addition from ATP to 3'-hydroxyl group[40]. By expressing kan^R, *A. baylyi* ADPWH_recA inactivates the interference of protein expression by kanamycin, achieves fast recovery from suppression, and minimizes spectral alterations as compared to others. It is confirmed by the presence of consistent shifts and discriminating biomarkers in *A. baylyi* postexposure to kanamycin, including Amide I (~1630 cm^{-1}, ~1640 cm^{-1}) and $\nu_{as}PO_2^-$ (~1222 cm^{-1}, ~1235 cm^{-1})[42].

An unexpected decline of kanamycin resistance gene was observed for AM2 artificial microbiota after 24 h exposure to kanamycin (42%, Figure 4C), but the predicted kanamycin resistance by D_I regression model remained close to 100%. It might be explained by the dramatically decreasing kanamycin concentration via the metabolism of aminoglycoside modifying enzyme and the change in microbial community structure. The functions of kan^R encoding aminoglycoside kinase are stabilizing a metaphosphate transition state and inactivating kanamycin[31], and the spectral alterations represent the alignment disruption of ß-phosphate and γ-phosphate by amide backbone. The declining kanamycin results in less inhibition on bacteria without kanamycin resistance gene (*M. vanbaalenii* and *E. coli*), and their growth and regeneration consequently reduce the abundance of A. *baylyi* and kan^R gene. Alternatively, the FTIR spectral alteration reflects such phenotypic changes of the whole microbiota under the low kanamycin exposure, illustrating the fact that the majority of microbial cells within the microbiota have the pseudo-resistance to

kanamycin. The spectrochemical interrogation therefore actually quantifies the microbial phenotypic antibiotic resistance rather than the ARGs abundance only.

Infrared spectroscopy has demonstrated the ability to diagnose the phenotypic alteration of the cellular components induced by kanamycin, hinting its potential possibility for the application to other members of the aminoglycoside family. Our findings indicate that this dispersion model coupled with PCA-LDA is a potential approach for monitoring the population dynamics within a microbiota in real-time. Additionally, the model applied in the present study summarizes the whole spectral information derived from the multivariate analysis, rather than only several biomarkers, showing its potential as a universal predicting tool for a broad spectrum of antibiotics based on well-trained databases. Though only successfully applied in the case of kanamycin through phosphotransferase resistance pathway, this technique is also feasible for detecting *N*-acetyltransferases and *O*-nucleotidyltransferases, which also belong to aminoglycoside-modifying enzymes assisted by acetyl-coenzyme A and ATP respectively[22], attributing to their similar anti-kanamycin mechanisms as kan^R encoding neomycin phosphotransferase. Future work should refer to more comprehensive range of antibiotics and their mechanisms including penicillin-class (*e.g.*, ampicillin and amoxicillin), which disrupts the synthesis of peptidoglycan layer and inhibits bacterial cell wall synthesis [44], and tetracycline, which inhibits the binding of aminoacyl-tRNA and suppresses protein expression[45]. For the urgent need to characterize antibiotic resistance in complex environmental microbiota with spectroscopy, the primary challenges are raised as the lack of routine protocols, reproducible computational analysis, and reliable database[10]. Validated in the artificial microbiota, our work provides the solutions for the first two barriers by distinguishing biomarkers representing antibiotic resistance from the numerous biological fingerprints. A well-built dataset along with robust analytical models coupled with spectroscopic methods are suggested to address the antibiotic resistance dynamics in real environmental samples.

The present study indicates that infrared spectroscopy, in conjunction with multivariate analysis, is a potential tool for diagnosing the phenotypic existence and dynamics of ARGs within microbial communities. Our work employed ATR-FTIR spectroscopy coupled with a dispersion model to quantify microbial kanamycin resistance, based on secondary derivative and PCA-LDA. This method not only

quantified the static community composition of the artificial microbiotas but also successfully predicted the population dynamics of microbial communities and kanamycin resistance under antibiotic pressure. We also suggest that spectroscopic techniques have great potential in real-time monitoring of microbiota of interest in medical or environmental fields; this would provide an excellent opportunity to visualize the vivid phenotypic transformation during a biological and biochemical process rather than only intermittent snap-shots.

Acknowledgements N.J. was funded by Chinese Academy of Sciences and China Scholarship Council. Research is supported by the Engineering and Physical Sciences Research Council in F.L.M.'s laboratory (EPSRC; grant no: EP/K023349/1) and National Natural Science Foundation of China in D.Z.'s laboratory (NFSC; grant no: 41301331).

References

(1) Blair, J. M.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. *Nat. Rev. Microbiol.* **2015**, *13*, 42-51.

(2) Chee-Sanford, J. C.; Aminov, R. I.; Krapac, I. J.; Garrigues-Jeanjean, N.; Mackie, R. I. *Appl. Environ. Microbiol.* **2001**, *67*, 1494-1502.

(3) Cantas, L.; Shah, S. Q. A.; Cavaco, L. M.; Manaia, C. M.; Walsh, F.; Popowska, M.; Garelick, H.; Burgmann, H.; Sorum, H. *Front. Microbiol.* **2013**, *4*.

(4) Potera, C. *Environ. Health Perspect.* **2013**, *121*, A255-A255.

(5) Smillie, C. S.; Smith, M. B.; Friedman, J.; Cordero, O. X.; David, L. A.; Alm, E. J. *Nature* **2011**, *480*, 241-244.

(6) Colomer-Lluch, M.; Imamovic, L.; Jofre, J.; Muniesa, M. *Antimicrob. Agents Chemother.* **2011**, *55*, 4908-4911.

(7) Riesenfeld, C. S.; Goodman, R. M.; Handelsman, J. *Environ. Microbiol.* **2004**, *6*, 981-989.

(8) Paraskevaidi, M.; Martin-Hirsch, P. L.; Kyrgiou, M.; Martin, F. L. *Mutagenesis* **2017**, *32*, 335-342.

(9) Kaern, M.; Elston, T. C.; Blake, W. J.; Collins, J. J. *Nat. Rev. Genet.* **2005**, *6*, 451-464.

(10) Jin, N. F.; Zhang, D. Y.; Martin, F. L. *Integr. Biol.* **2017**, *9*, 406-417.

(11) Naumann, D.; Helm, D.; Labischinski, H. *Nature* **1991**, *351*, 81-82.

(12) Dunn, W. B.; Ellis, D. I. *Trends Anal. Chem.* **2005**, *24*, 285-294.

(13) Freedman, B. G.; Zu, T. N. K.; Wallace, R. S.; Senger, R. S. *Biotechnol. J.* **2016**, *11*, 877-889.

(14) Sun, S. W.; Wang, X. T.; Gao, X.; Ren, L. H.; Su, X. Q.; Bu, D. B.; Ning, K. *BMC Bioinformatics* **2015**, *16*.

(15) Winnard, P. T.; Zhang, C.; Vesuna, F.; Kang, J. W.; Garry, J.; Dasari, R. R.; Barman, I.; Raman, V. *Oncotarget* **2017**, *8*, 20266-20287.

(16) Li, J. Y.; Strong, R.; Trevisan, J.; Fogarty, S. W.; Fullwood, N. J.; Jones, K. C.; Martin, F. L. *Environ. Sci. Technol.* **2013**, *47*, 10005-10011.

(17) Bankapur, A.; Krishnamurthy, R. S.; Zachariah, E.; Santhosh, C.; Chougule, B.; Praveen, B.; Valiathan, M.; Mathur, D. *PLoS One* **2012**, *7*.

(18) Tao, Y. F.; Wang, Y.; Huang, S.; Zhu, P. F.; Huang, W. E.; Ling, J. Q.; Xu, J. *Anal. Chem.* **2017**, *89*, 4108-4115.

(19) Siddhanta, S.; Paidi, S. K.; Bushley, K.; Prasad, R.; Barman, I. *Chemphyschem* **2017**, *18*, 72-78.

(20) Baker, M. J.; Trevisan, J.; Bassan, P.; Bhargava, R.; Butler, H. J.; Dorling, K. M.; Fielden, P. R.; Fogarty, S. W.; Fullwood, N. J.; Heys, K. A.; Hughes, C.; Lasch, P.; Martin-Hirsch, P. L.; Obinaju, B.; Sockalingum, G. D.; Sule-Suso, J.; Strong, R. J.; Walsh, M. J.; Wood, B. R.; Gardner, P.; Martin, F. L. *Nat. Protoc.* **2014**, *9*, 1771-1791.

(21) Martin, F. L.; Kelly, J. G.; Llabjani, V.; Martin-Hirsch, P. L.; Patel, II; Trevisan, J.; Fullwood, N. J.; Walsh, M. J. *Nat. Protoc.* **2010**, *5*, 1748-1760.

(22) Mingeot-Leclercq, M. P.; Glupczynski, Y.; Tulkens, P. M. *Antimicrob. Agents Chemother.* **1999**, *43*, 727-737.

(23) Sadovskaya, I.; Vinogradov, E.; Li, J. J.; Hachani, A.; Kowalska, K.; Filloux, A. *Glycobiology* **2010**, *20*, 895-904.

(24) Delorenzo, V.; Herrero, M.; Jakubzik, U.; Timmis, K. N. *J. Bacteriol.* **1990**, *172*, 6568-6572.

(25) Song, Y.; Li, G.; Thornton, S. F.; Thompson, I. P.; Banwart, S. A.; Lerner, D. N.; Huang, W. E. *Environ. Sci. Technol.* **2009**, *43*, 7931-7938.

(26) Trevisan, J.; Angelov, P. P.; Scott, A. D.; Carmichael, P. L.; Martin, F. L. *Bioinformatics* **2013**, *29*, 1095-1097.

(27) Butler, H. J.; McAinsh, M. R.; Adams, S.; Martin, F. L. *Anal. Methods* **2015**, *7*, 4059-4070.

(28) Yenofsky, R. L.; Fine, M.; Pellow, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 3435-3439.

(29) Bosch, A.; Serra, D.; Prieto, C.; Schmitt, J.; Naumann, D.; Yantorno, O. *Appl. Microbiol. Biotechnol.* **2006**, *71*, 736-747.

(30) Mariey, L.; Signolle, J. P.; Amiel, C.; Travert, J. *Vib. Spectrosc.* **2001**, *26*, 151-159.

(31) Wright, G. D. *Curr. Opin. Microbiol.* **1999**, *2*, 499-503.

(32) Flemming, H. C.; Wingender, J. *Nat. Rev. Microbiol.* **2010**, *8*, 623-633.

(33) Holman, H. Y. N.; Miles, R.; Hao, Z.; Wozei, E.; Anderson, L. M.; Yang, H. *Anal. Chem.* **2009**, *81*, 8564-8570.

(34) Stewart, P. S. *Int. J. Med. Microbiol.* **2002**, *292*, 107-113.

(35) Hoiby, N.; Bjarnsholt, T.; Givskov, M.; Molin, S.; Ciofu, O. *Int. J. Antimicrob. Agents* **2010**, *35*, 322-332.

(36) Karunakaran, E.; Mukherjee, J.; Ramalingam, B.; Biggs, C. A. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 1869-1881.

(37) Li, H. B.; Martin, F. L.; Zhang, D. Y. *Anal. Chem.* **2017**, *89*, 3909-3918.

(38) Ede, S. M.; Hafner, L. M.; Fredericks, P. M. *Appl. Spectrosc.* **2004**, *58*, 317-322.

(39) Morones, J. R.; Elechiguerra, J. L.; Camacho, A.; Holt, K.; Kouri, J. B.; Ramirez, J. T.; Yacaman, M. J. *Nanotechnology* **2005**, *16*, 2346-2353.

(40) Boehr, D. D.; Thompson, P. R.; Wright, G. D. *J. Biol. Chem.* **2001**, *276*, 23929-23936.

(41) Heys, K. A.; Riding, M. J.; Strong, R. J.; Shore, R. F.; Pereira, M. G.; Jones, K. C.; Semple, K. T.; Martin, F. L. *Analyst* **2014**, *139*, 896-905.

(42) Movasaghi, Z.; Rehman, S.; Rehman, I. U. *Appl. Spectrosc. Rev.* **2008**, *43*, 134-179.

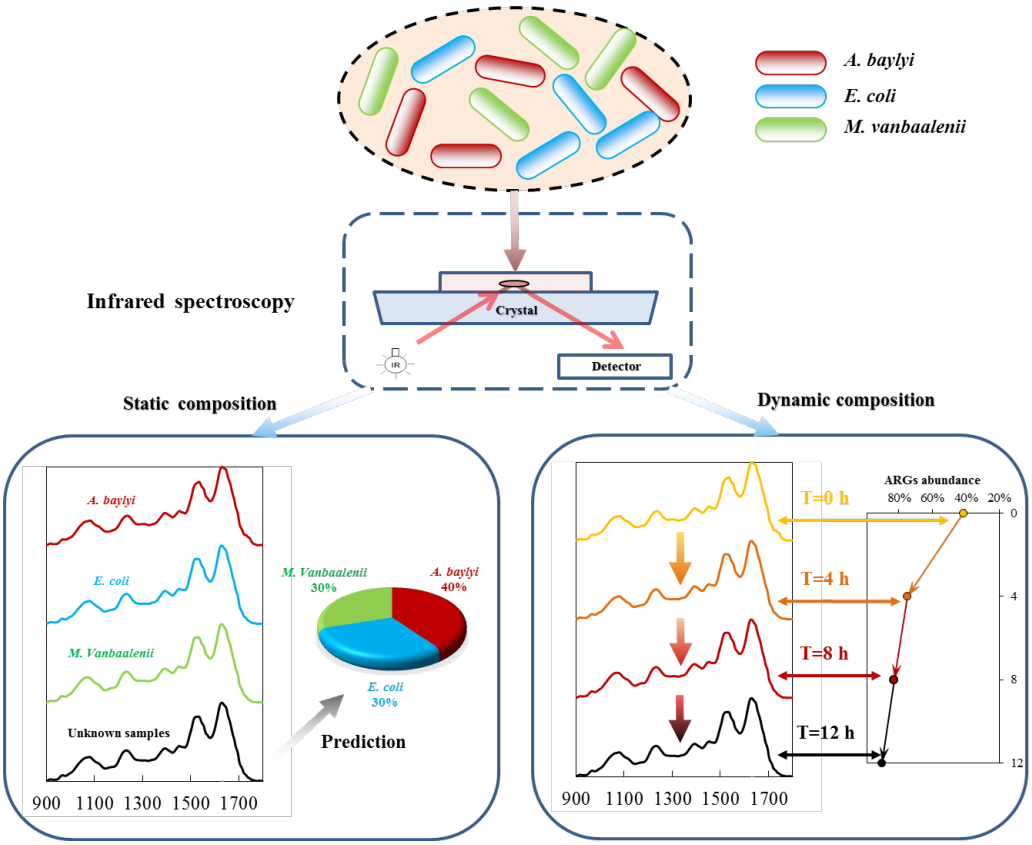
(43) Hibbing, M. E.; Fuqua, C.; Parsek, M. R.; Peterson, S. B. *Nat. Rev. Microbiol.* **2010**, *8*, 15-25.

(44) Strominger, J. L.; Park, J. T.; Thompson, R. E. *J. Biol. Chem.* **1959**, *234*, 3263-3268.

(45) Connell, S. R.; Trieber, C. A.; Dinos, G. P.; Einfeldt, E.; Taylor, D. E.; Nierhaus, K. H. *EMBO J.* **2003**, *22*, 945-953.

507

Table of Content Graphic



508

509

510

511 For TOC only

512