Measuring phenotypic heterogeneity in isogenic bacterial populations using flow cytometry and Raman spectroscopy

Authors:

Cristina García-Timermans[a*], Peter Rubbens[b*], Jasmine Heyse[a], Frederiek-Maarten Kerckhof[a], Ruben Props[a], Andre G. Skirtach[c], Willem Waegeman[b], Nico Boon[a]

[a]*CMET, Center for Microbial Technology and Ecology, Ghent University, Gent, Belgium*
[b]*KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Gent, Belgium*
[c]*Nano-BioTechnology group, Department of Biotechnology, Ghent University, Gent, Belgium*
*These authors contributed equally to the manuscript

Corresponding authors: E-mail: Nico.Boon@UGent.be

Word count abstract: 194 Word count text: 5276

Abstract:

Investigating phenotypic heterogeneity can help to better understand and manage microbial communities. However, characterizing phenotypic heterogeneity remains a challenge, as there is no standardized analysis framework. Several optical tools are available, which often describe properties of the individual cell. In this work, we compare Raman spectroscopy and flow cytometry to study phenotypic heterogeneity in bacterial populations. The growth phase of *E. coli* populations was characterized using both technologies. Our findings show that flow cytometry detects and quantifies shifts in phenotypic heterogeneity at the population level due to its high-throughput nature. Raman spectroscopy, on the other hand, offers a much higher resolution at the single-cell level (i.e. more biochemical information is recorded). Therefore, it is capable of identifying distinct phenotypic populations when coupled with standardized data analysis. In addition, it provides information about biomolecules that are present, which can be linked to cell functionality. We propose an automated workflow to distinguish between bacterial phenotypic populations using Raman spectroscopy and validated this approach with an external dataset. We recommend to apply flow cytometry to characterize phenotypic heterogeneity at the population level, and Raman spectroscopy to perform a more in-depth analysis of heterogeneity at the single-cell level.

Importance:

Single-cell techniques are frequently applied tools to study phenotypic characteristics of bacterial populations. As flow cytometry and Raman spectroscopy gain popularity in the field, there is a need to understand their advantages and limitations, as well as to create a more standardized data analysis framework. Our work shows that flow cytometry allows to study and quantify shifts at the bacterial population level, but since its resolution is limited for microbial purposes, distinct phenotypic populations cannot be distinguished at the single-cell level. Raman spectroscopy, combined with appropriate data analysis, has sufficient resolving power at the single-cell level, enabling the identification of distinct phenotypic populations. As regions in a Raman spectrum are associated with specific (bio)molecules, it is possible to link these to the cell state and/or its function.

Keywords:

E. *coli*; Flow cytometry; phenotypic heterogeneity; Raman spectroscopy; single-cell technology; microbial ecology; growth phase

Introduction

Single-cell phenotypic differences arise even in genetically identical cultures (Govers et al. 2017). A definition of a phenotypic population is an observed cellular state within a given genetic and environmental background. It arises due to epigenetic variations, stochastic gene expression, cellular age or oscillations such as the cell cycle. This is one of the strategies that bacteria use to adapt to a changing environment, as well as to divide the labour within the community (Avery 2006; Ackermann 2015).

Phenotypic heterogeneity in laboratory cultures is well-documented. For example, it has been studied in bacterial subpopulations that could tolerate antibiotics (known as ‘persisters’) (Dhar and McKinney 2007), in the production of cytotoxin K in *Bacillus cereus* (Ceuppens, Boon, and Uyttendaele 2013) or in the differential expression of flagellin in *Salmonella* Typhimurium (Stewart et al. 2011). The challenge remains to find tools to measure and quantify this heterogeneity (i.e. phenotypic populations), in order to be able to link heterogeneity with bacterial functionality.

Several tools are available for the single-cell phenotyping (Davis and Isberg 2016). Imaging techniques are popular, but they require tagged bacterial cells or a probe to visualize the bacteria or the molecule of interest (Taniguchi et al. 2010; Anetzberger, Schell, and Jung 2012), making them less suitable for studying environmental communities. There are other techniques that do not require a probe, such as intrinsic fluorescence (Georgakoudi et al. 2007) or the detection of autofluorescent NAD(P)H (Bhattacharjee et al. 2017). However, the amount of information that can be gathered is limited compared to other techniques. Single-cell transcriptomics are also an option for phenotyping, but a few hundred cells are needed and only about 15 - 25% of the expressed mRNAs can be detected (Tang, Lao, and Surani 2011). This analysis is destructive, and it was found in *E. coli* that a single cells' protein and mRNA copy numbers are uncorrelated for any given gene (Taniguchi et al. 2010).

A more high-throughput option for single-cell analysis is flow cytometry, with the capacity to measure between hundreds to thousands of bacteria per second. Individual cells pass through a laser, after which detectors collect both information on the scattered laser light (forward scatter, FSC, and side scatter, SSC) and on fluorescent emissions of specific probes. As bacteria are highly diverse, there is no general antibody that can be used for all of them. Therefore, general nucleic acid stains are used (such as SYBR Green I or DAPI) (Koch et al. 2014). Upon collection of the data, a cytometric fingerprint of the population can be derived to infer inter- and intra-species variations (De Roy et al. 2012; Koch et al. 2014; Props et al. 2016). Different dyes can be combined to collect more information on the bacteria (Nebe-von-Caron et al. 2000), although multicolour experiments are still limited at present (Buysschaert et al. 2016).

Raman spectroscopy is another single-cell technology that does not require labelling and is non-destructive. The laser excites individual cells, which leads to inelastic scattering, which in turn is collected in the form of Raman spectra. It is less throughput compared to flow cytometry: even when enhancing the signal with metals (known as Surfaced-Enhanced Raman Spectroscopy or SERS), each cell takes 1-3 seconds to measure (Liu et al. 2016). The resulting spectrum contains biochemical information of the molecules that are present in bacteria – e.g. lipids, carbohydrates, nucleic acids and proteins- and can be used to classify bacteria (Huang et al. 2010; Goodacre et al. 1998). This information can be quantitative if an internal standard for the molecule(s) of interest is made (for example, (Cowcher, Xu, and Goodacre 2013) quantified the dipicolinate (DPA) biomarker for *Bacillus* spores; or (Samek et al. 2016) quantified polyhydroxyalkanoates produced by *Cupriavidus necator* H16). The Raman signal can be enhanced, for example, using metal nanostructures (called Surfaced-Enhanced Raman Spectroscopy) (Jarvis and Goodacre 2004; Yashchenok et al. 2013), by combining SERS with microchip devices (McIlvenna et al. 2016) or optical tweezers (Xie, Chen, and Li 2005), or by adding heavy water (D2O) to monitor its incorporation (Berry et al. 2015).

Both flow cytometry and Raman spectroscopy give rise to data, which need specific pre-processing and analysis (O'Neill et al. 2013; Saeys, Gassen, and Lambrecht 2016; García-Timermans et al. 2018; Ryabchykov, Guo, and Bocklitz 2018). While microbial flow cytometry is rather limited in its phenotypic resolution (i.e., only a few properties are measured per cell), Raman spectroscopy characterizes many more biochemical properties of bacterial cells. It therefore requires analysis of high-dimensional data, which can be challenging, but it allows to characterize phenotypic heterogeneity at a much higher resolution.

In this work, we analysed bacteria from nine phenotypic populations –with a different growth stage and/or from a different replicate- using flow cytometry and Raman spectroscopy. We demonstrate how these populations can be automatically retrieved using data-specific algorithms. We also illustrate how metabolic inference can be performed based on Raman spectroscopy in a data-driven way. Finally, the advantages and disadvantages of these tools for microbial phenotyping are discussed. We will motivate that, in its current form, microbial flow cytometry can be used to describe community-level dynamics, while Raman spectroscopy can be applied to describe single-cell heterogeneity. We include a recommendation for microbiologists on how to employ Raman spectroscopy and flow cytometry in future phenotyping studies.

Results

Throughout this paper, we define a 'phenotypic population' as a group of bacteria grown under the same environmental conditions (i.e. cells from the same biological replicate at a certain growth stage). This population will share morphological and/or metabolic traits that can be detected by flow cytometry and Raman spectroscopy. *E. coli DSM 2092* were measured in the lag, log and stationary phase. For every condition, triplicates of the cell culture were made. Thus, we expected to retrieve 9 phenotypic populations. As it will be argued in the discussion, this is not to say that there might not be other subpopulations in each of these 'phenotypic populations'.

Flow cytometry

Three biological replicates of *Escherichia coli* DSM 2092 were measured in the lag, log and stationary phase through flow cytometry (Figure S1). Data was analysed at two levels: (a) the single-cell level (i.e., cells were analysed as individual instances) (b) the cell population level (i.e., cytometric fingerprints were constructed to describe population dynamics) (Fig. 1). t-distributed stochastic neighbourhood embedding (t-SNE) was used to visualize the data at the single-cell level, whereas principal coordinate analysis (PCoA) was applied to visualize the differences of the phenotypic populations based on Bray-Curtis dissimilarities (Fig 1f). As a validation, t-SNE was performed on the population level (Fig 1e) and principal component analysis was performed on the single-cell level (Fig. S3).

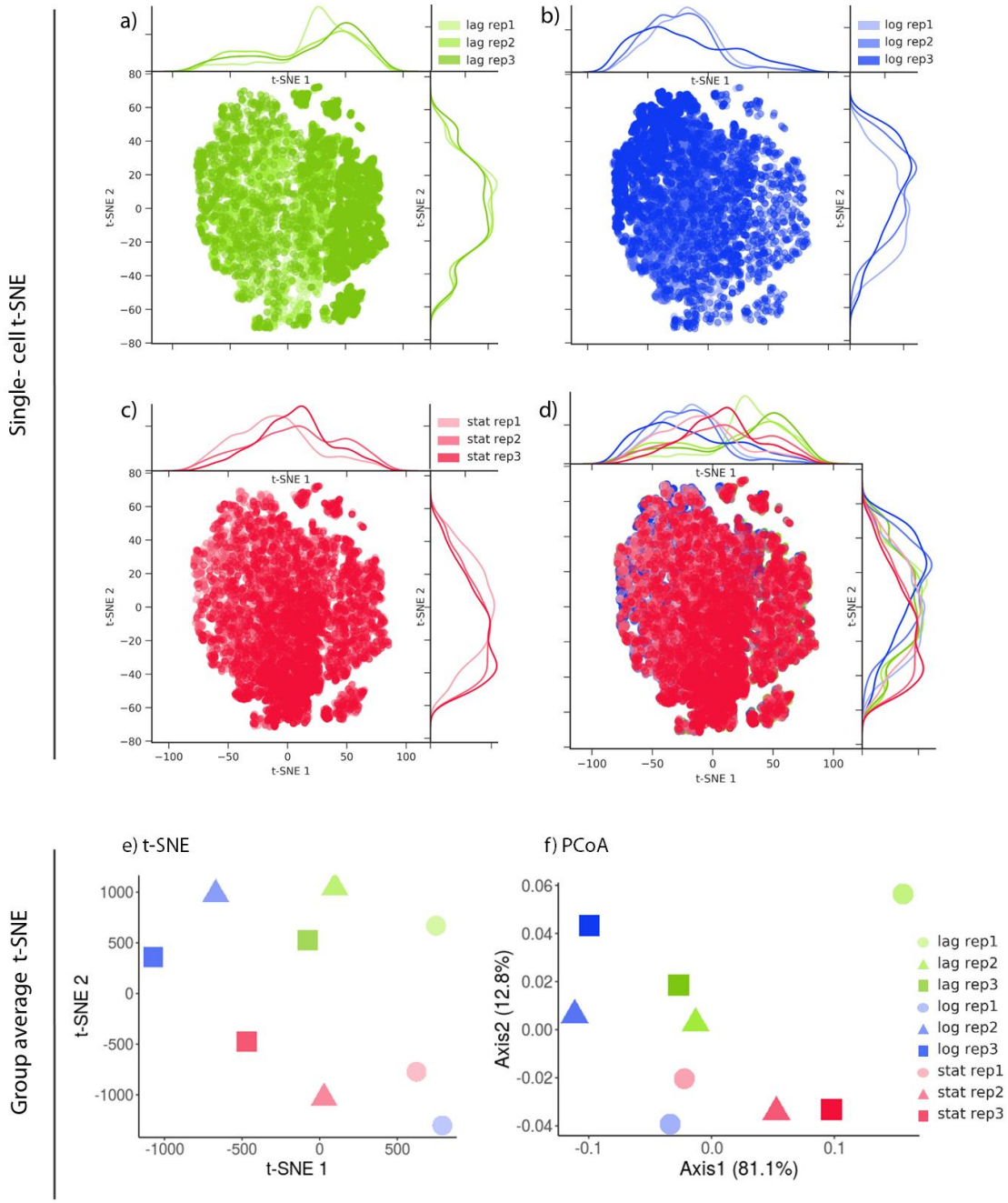

Fig. 1: *E. coli* measured with flow cytometry and analyzed at the single cell level (a-d) and population level (e and f). t-SNE was performed on the aggregation of all samples (d), and visualized separatedly for each growth phase, to allow for easier interpretation (a-c). Distributions on the side represent the t-SNE distributions separately visualised for each growth phase/replicate to allow for easier interpretation. (e-f) Visualization of cytometric fingerprints at the sample level, using t-SNE (e) or PCoA (f).

No separated subpopulations could be distinghuished based on cytometric single-cell data (Fig. 1a-d). Yet, shifts in the distribution of cells were clear, both between different growth phases and replicates, as can be seen from the marginal distributions. Therefore, by creating cytometric fingerprints, these differences could be quantified and visualized at the community

level (Fig. 1e-f). Differences between fingerprints were calculated using the Bray-Curtis dissimilarity. Average dissimilarities per growth phase and replicate were summarized in Table 1. The average Bray-Curtis dissimilarity between samples within the same growth phase is smaller compared to samples that originated from the same replicate, except for samples from the lag phase, for which comparable dissimilarities were retrieved. The stationary phase for replicate 1 was quite different from the other samples (Fig 1f), this makes the average distance of the samples of the stationary phase the highest (Table 1). t-SNE and PCoA were applied on the sample level as well, resulting in a grouping of the data according to mostly the growth phase (Figs. 1e-f). Bray-Curtis dissimilarities were used to perform PCoA.

Table 1: Average Bray-Curtis distance between the samples based on their growth phase or their replicate. A.U.= arbitrary units.

	Average distance between samples (A.U., n=3)	**Standard deviation (n=3)**
Lag phase	0.123	0.071
Log phase	0.080	0.018
Stationary phase	0.078	0.029
Replicate 1	0.136	0.080
Replicate 2	0.114	0.037
Replicate 3	0.136	0.051

Raman spectroscopy – clustering results

The samples used for flow cytometric analysis were fixed and analysed using label-free Raman spectroscopy following the protocol from (García-Timermans et al. 2018). To identify phenotypic populations, two clustering methods were used. First, using an agglomerative hierarchical clustering approach and secondly, using the PhenoGraph algorithm - a tool developed for the analysis of high-dimensional cytometry data. To determine the hierarchical clustering, the spectral contrast angle between samples was calculated (a measure of the spectra similarity). Then, phenotypic populations can be delineated by setting a threshold upon inspection of the resulting dendrogram after hierarchical clustering (Fig. 2). On the other hand,

PhenoGraph makes use of *k*-nearest-neighbours clustering, in order to determine groups of similar cells, and as such, phenotypic populations. In other words, *k* expresses the amount of local information that is included when cells are grouped according to similar spectra. *k* will therefore, in a similar way as the threshold used in hierarchical clustering, impact the number of phenotypic populations that are defined (Fig. 2&3).

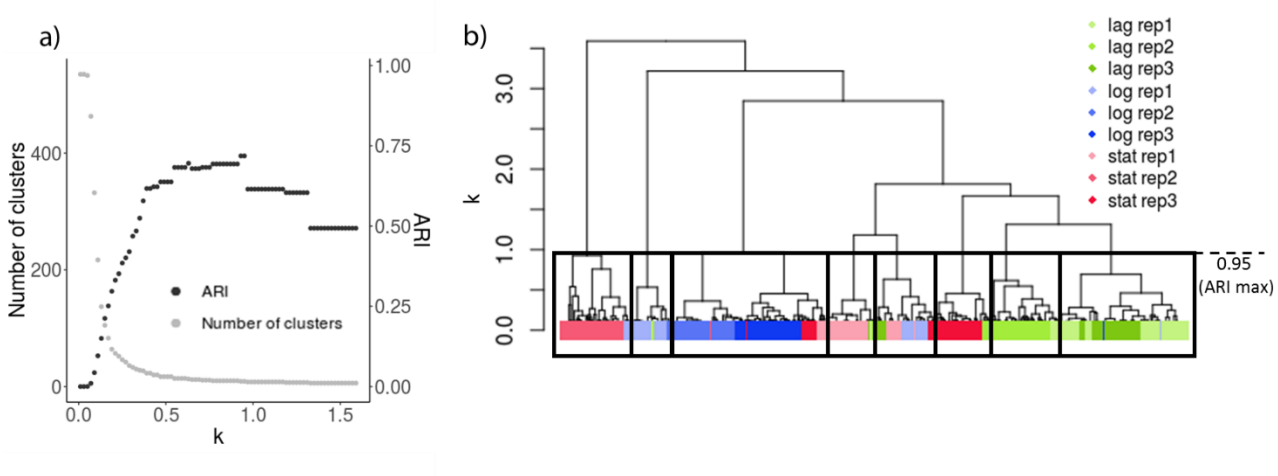

Fig. 2: Hierarchical clustering of Raman spectra derived from the *E. coli* culture. Cells were measured in the log, lag and stationary phase using Raman spectroscopy. (a) Left axis, grey: Visualization of the number of clusters we can obtain by cutting the cluster at different heights (*k*). Right axis, black: ARI, which quantifies how many cells are identified as the expected phenotypic population when clusters are made at different levels. (b) Dendrogram representing the distance (calculated as the spectral contrast angle) amongst the Raman spectra. The black boxes dived the phenotypic populations when the dendrogram is cut at *k*=0.9. This is the maximum adjusted Rand index (ARI), as calculated in Fig. 2b.

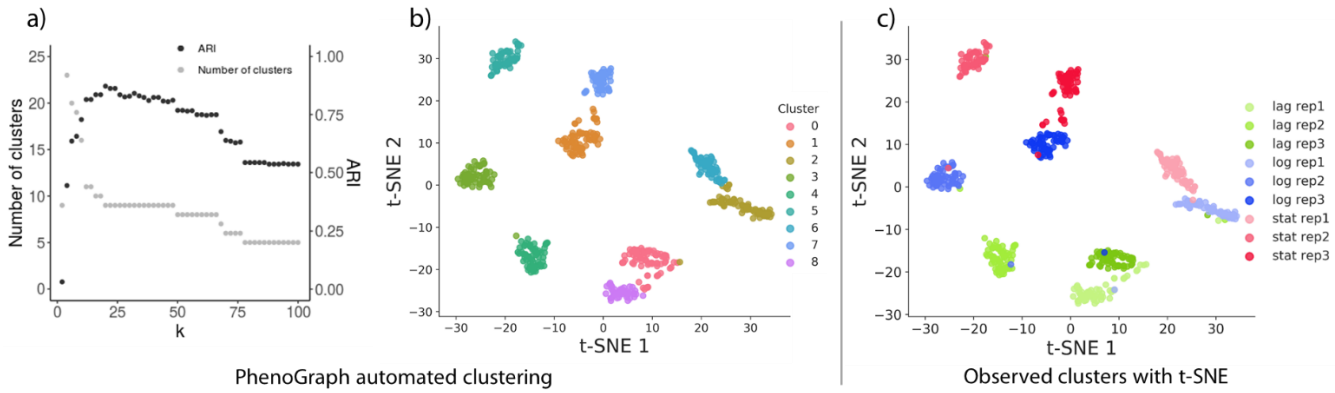

Fig. 3: PhenoGraph clustering of Raman spectra derived from the *E. coli* culture. a) Influence of hyperparameter *k* on the automated identification of phenotypic populations. Left axis, grey: Visualization of the number of clusters in

function of *k*. Right axis, black: Adjusted Rand index (ARI), which quantifies how many cells are identified as the expected phenotypic population when clusters are made at different levels (k). b) t-SNE visualization, colored according to PhenoGraph clustering with optimal ARI (*k*=22). c) t-SNE results colored according to growth phases and replicates.

The adjusted Rand index (ARI) was used to quantify similarity between the clusters that were predicted by Phenograph and the known phenotypic populations (i.e. growth phase and replicate). An ARI of 1 indicates perfect grouping of the data. The PhenoGraph algorithm resulted in a higher ARI as opposed to hierarchical clustering based on the spectral contrast angle (Fig. 2b vs. Fig. 3a). Inspecting the PhenoGraph results, there is a stable region for *k* that retrieves clustering according to both growth phase and replicate *(*i.e., nine clusters were found for *k* = 20, …, 48*)*. A value of *k* = 22 or 24 gave rise to an optimal clustering (Fig. 3a). Smaller *k* allowed to inspect phenotypic populations at smaller scales and investigate the heterogeneity accordingly. See for example the clustering results for *k* = 15, which resulted in eleven different groups of cells (Fig. S5). Additional clusters that emerged were the result of splitting two clusters into two smaller ones. Likewise, larger *k* will result in larger clusters. For example, for *k* = 100, data is grouped in five clusters (Fig. S5). Structure in the data is retained, as clusters are merged either according to growth phase (clusters 0, 2 and 3) or replicate (cluster 1).

Raman spectroscopy – tentative region assignment

The Boruta algorithm was used to associate the most distinctive regions in the Raman spectrum with cluster assignments according to the PhenoGraph algorithm. Regions were linked with different molecules based on a recent summary from Wang et al., 2016. In this way, metabolic associations could be inferred that contained predictive power as a function of different phenotypic populations. Cluster assignments, which resulted in an optimal ARI according to both hierarchical clustering and PhenoGraph, were used as target variables to associate relevant Raman regions with differences in phenotypic populations (Table 2).

Table 2: Tentative assignment of Raman spectra using the Boruta algorithm based on phenotypic identification using hierarchical clustering and PhenoGraph. The 10 highest ranked areas are shown. When there is no known compound in the spectral region, either the closest compound or a blank is shown. A.U.=arbitrary units.

Wavelength (cm -1)	Features importance		Tentative assignment
	PhenoGraph (k=22, 9 clusters) (A.U.)	hclust (h=1, 8 clusters) (A.U.)	
1042	9.8	9.8	Carbohydrates, Proline (1043)
971	9.5	9.7	v(C-C) wagging (971) // Phosphate monoester groups of phosphorylated proteins and cellular nucleic acids (970)
945	9	8.9	vs (CH3) of proteins (α-helix) (951)
1057	8.9	8.8	Lipids (1057) // Carbohydrates (1030-1130)
1294	8.5	8.5	CH2 deformation (1295)
1050	8.5	8.6	Nucleic acids, CO stretching; protein, C-N stretching, PHB (1054) // Carbohydrates (1030-1130)
1053	8.4	8.4	Nucleic acid (1054) // Carbohydrates (1030-1130)
1127	8.3	8.4	v(C-N), protein (1127) // Carbohydrates (1030-1130)
1046	8.1	8.2	v3PO4 3− (symmetric stretching vibration of v3PO43− of HA) (1044)
641	8	7.9	C-C twisting of tyrosine (642)

To understand how the molecules in Table 1 vary from one group to another, the distribution of intensities of these Raman regions was plotted for every growth phase (Fig. 4). A Wilcoxon rank sum test with a Benjamini–Hochberg correction was performed for the ten highest ranked variabels according to the Boruta algorithm (Fig. 4).

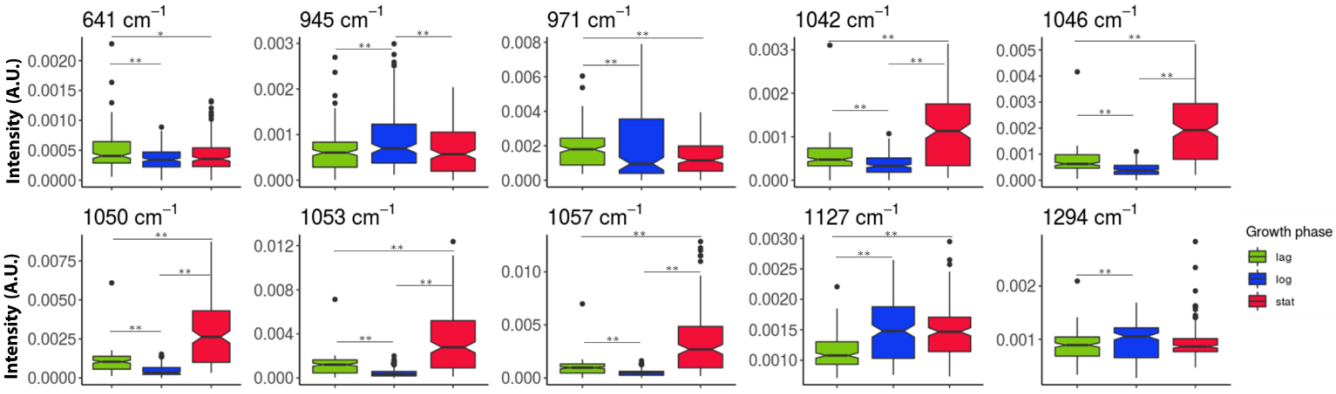

Fig. 4: Distribution of intensities of the most relevant regions associated with phenotypic populations according to the Boruta algorithm. Boxplots represent the growth phases (replicates were pooled together). Groups were made according to the growth phase: lag (green), logarithmic (blue) or stationary (red). For every spectral region, a Wilcoxon rank sum test was made, with a Benjamini-Hochberg correction (upon rejection of the null hypothesis). Groups with significantly different peaks are signalled with (*) ($p<0.05$) or (**) ($p<0.01$).

To better understand what regions of the Raman spectra (and therefore, what biomolecules) were making these phenotypic populations different, we first defined phenotypic populations at different levels (changing the *k* parameter in the PhenoGraph algorithm) and then used the Boruta algorithm to identify the most relevant regions.

When more phenotypic populations were distinguished (i.e. setting the value of k lower) more regions in the Raman spectrum were required to allow for a correct classification. As shown in Fig. 5, to find the phenotypic populations with different growth phases, 59% of the data is included (green); to find the biological replicates, it is 67%; and to find both categories, 77%. Although this result was expected, it was also noticed that a large number of Raman regions (48%) were relevant for all levels of classification.

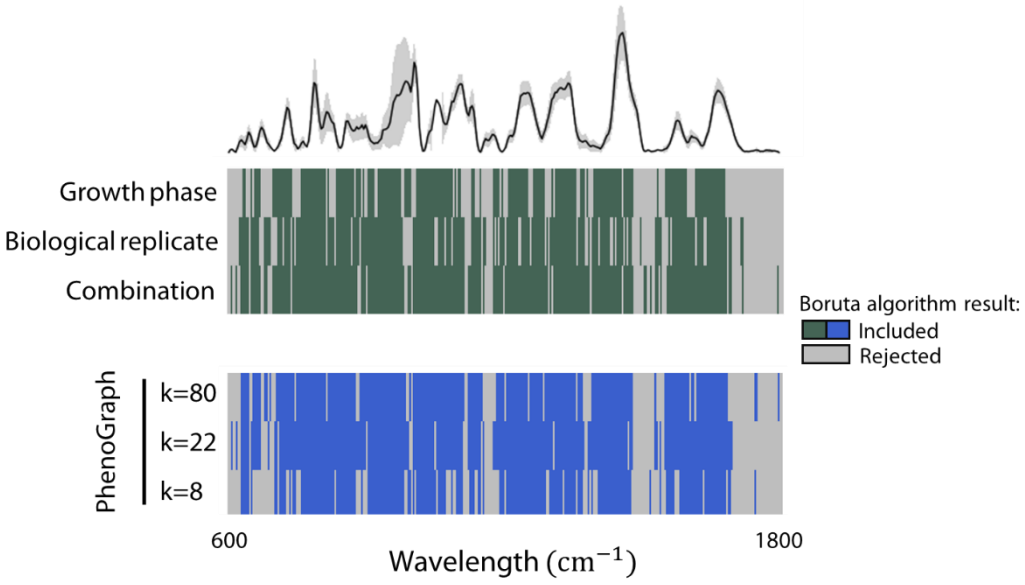

Fig. 5: Spectral regions relevant for phenotypic classification at different levels, according to the Boruta algorithm. The top heatmap shows the results for the growth phase and replicates. At the bottom, the results for PhenoGraph with the clustering hyperparameter k as 80, 22 or 8 are shown. In green and blue, spectral regions confirmed by the Boruta algorithm as relevant, in grey the rejected. The average of all spectra is also plotted; the grey areas in the average spectrum correspond to the standard deviation.

Validation of single-cell analysis of Raman spectra

To validate our pipeline to analyse microbial single-cell Raman data, the dataset from Teng et al., 2016 was used. In this work, *E. coli* was exposed to different chemicals (ethanol, antibiotics, n-butanol or heavy metals) and the spectra of the bacteria were measured at several time points after the treatment (5, 10, 20, 30 and 60 min). Three replicates of the cell culture were made for each treatment. Here we show the results for cells treated with ethanol (Fig. 6), representative for what is observed in the other groups (Fig S6). t-SNE was able to visualize groups of bacteria that received different treatments at different points in time. Furthermore, two subpopulations are seen in every group. They correspond to the replicates, where two replicate samples are separated, and the third replicate is either assigned to one of the two or divided amongst the two subpopulations. The optimal ARI is lower than the one reported for our own work, but still considerably higher than zero. This means that although the clusters

assigned according to PhenoGraph have a better match with the treatments induced in our own dataset compared to this one, the clustering is still meaningful.

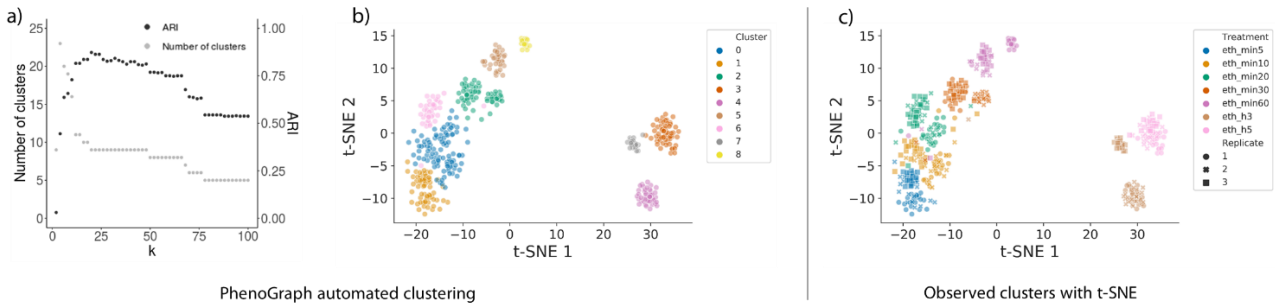

Fig. 6: External dataset clustered using t-SNE. *E.coli* treated with ethanol and measured at time points 5, 10, 20, 30 and 60 min. (a) Influence of hyperparameter *k* on the automated identification of clusters. Left axis, grey: Visualization of the number of clusters in function of *k*. Right axis, black: Adjusted Rand index (ARI), which quantifies how many cells are identified as the expected phenotype when clusters are made at different levels (k). (b) Using the maximum ARI, samples were automatically clustered using PhenoGraph. (c) t-SNE was performed on the data set and samples were labelled according to their treatment. The shapes represent the sample replicate.

Discussion

Flow cytometry quantifies population shifts

When using single-cell data from flow cytometry, the different phenotypic populations overlapped and did not form separate clusters, as shown by both the t-distributed stochastic neighbourhood embedding (t-SNE) and the principal component analysis (PCA) (Fig. 1a, SI Fig. 2). However, in the t-SNE plot, a consistent shift in the cells distribution could be observed in response to the different growth phases (Fig. 1a-d). In other words, gradual shifts in the structure of the phenotypic population, i.e. the phenotypic heterogeneity, could still be detected, although individual cells could not be separated according to growth phase or replicate. As flow cytometry is capable of rapidly measuring a large amount of cells, the differences in phenotypic heterogeneity at the population level could be described by constructing cytometric fingerprints (Koch et al. 2014; Props et al. 2016). The Bray-Curtis dissimilarity was used to quantify these differences at a sample level. The average Bray-Curtis dissimilarity showed that the effect of the replicates exceeded the effect of the growth phase (except for the lag phase). This implies that the differences of *E. coli* cells in different growth stages are small –comparable or smaller– to those the differences amongst replicates. The results from (Teng et al. 2016) analysed in Fig. 6 leads us to hypothesize that when the community is steered with a stressor, the effect in the fingerprint is bigger, making the replicate effect smaller.

When t-SNE and principal coordinate analysis (PCoA) were applied to these fingerprints, samples showed a clear differentiation based on both the growth phase and replicate (Fig. 1b, 4). In other words, gradual shifts could be detected in the flow cytometric data at the level of individual cells, while at the sample-level (i.e. the population distribution level), differences between communities could be quantified (e.g., using the Bray-Curtis dissimilarity) and separated accordingly.

It is worth noting that in this work the effect of using additional or more specific labels for cytometric analysis has not been explored, which might improve the resolution. It is possible

to add stains to target specific substrates (see the review of (Léonard et al. 2016) on the use of individual and double stains, and an example of a three-color analysis by (Barbesti et al. 2000)), but the number of markers describing microbial cells using flow cytometry will never be of the same order as that of Raman spectroscopy. In eukaryotic flow cytometry, where the tagging of specific antibodies is much more feasible, 19-parameter flow cytometry is routinely used (17 fluorescence and two scatters) (Perfetto, Chattopadhyay, and Roederer 2004) and 30-parameter flow cytometry has just recently been published (Mair and Prlic 2018). However, the dimensionality of cytometry data in these settings is still much lower than the number of variables derived from Raman spectroscopy. Even in the best-case scenario, the dimensionality of flow cytometry data cannot get close to the number of parameters that Raman spectroscopy exhibits. On the other hand, depending on the research question, a high-dimensional tool might not be needed. For example, biomolecules that are associated with a phenotypic population might be known, and there could be a dye available to highlight these molecules. In this case, flow cytometry could be more relevant for phenotyping than Raman spectroscopy, provided the proper parameters are chosen to differentiate among treatments. In this work, we compared phenotypic populations that are not differentiated by (a) specific, known molecule(s) but used a general marker to characterize the DNA content. This explains why flow cytometry did not have enough resolution to differentiate the phenotypic populations.

Flow cytometry is a high-throughput technique, able to rapidly measure hundreds to thousands of individual cells. By applying fingerprinting approaches to cytometry data, differences between microbial populations at the population level can be assessed and quantified. In other words, while current resolution at the single-cell level appears to be limited for microbial flow cytometry, due to its high-throughput nature, its power lies in the possibility to characterize and phenotypic heterogeneity at the population level and quantify shifts in phenotypic heterogeneity.

Raman spectroscopy detects phenotypic populations at the single-cell level

Raman spectroscopy is lower throughput for single-cell analysis, but is able to retrieve much more information per cell. This resolution allows to conduct research at the single-cell level. For instance, Raman spectroscopy has been used to detect stress-induced phenotypic populations (Teng et al. 2016), study antibiotic response (Athamneh et al. 2014) or determine antibiotic susceptibility (Novelli-Rousseau et al. 2018) and to discriminate between different *Acinetobacter* (Maquelin et al. 2006) or between different *E. coli* strains (Jarvis and Goodacre 2004). As mentioned in the introduction, it is possible to improve the speed of the measurements enhancing the Raman signal the of bacteria using metallic substrates (Surface-Enhanced Raman Spectroscopy, SERS), microfluidic chips (McIlvenna et al. 2016) or optical tweezers (Xie, Chen, and Li 2005). Another advantage of Raman spectroscopy is that it can be applied without the use of labels. This allows to analyse the biochemistry of samples even without knowing their nature. Raman spectroscopy also presents disadvantages. For instance, the Raman signals of certain compounds can be quite weak, making them difficult to detect or undetectable. Also, the Raman signal of certain compounds can be composed of several peaks, or be unknown. The background of samples can interfere with the Raman signal of bacteria. Also, the equipment can be quite costly, depending on the type of Raman spectroscope. There are several tools (such as the Raman Tool set (Candeloro et al. 2013) or the 'MicroRaman' package (Kerckhof et al. 2018) amongst others) that allow non-experts for a rapid analysis of Raman spectra. To combine the advantages of these two techniques, flow cytometers for measuring Raman scattering have been built. The first prototype has been developed to sort polymer microspheres (polystyrene/divinylbenzene) labelled with SERS tags (Surface-Enhanced Raman Scattering) (Watson et al. 2008). A newer version incorporates a microfluidic chip, and has been tested using unlabelled red blood cells (Perozziello et al. 2016).

Raman spectroscopy offers more parameters per cell compared to flow cytometry (hundreds versus typically three or four for microbial experiments). Thus, individual bacteria are described in a much larger multivariate space and can therefore be clustered into separate phenotypic

populations. This explains why bacterial subpopulations can be visualized at the single-cell level using t-SNE (Fig. 3). t-SNE results were confirmed with a PCA (Fig. S3).

As previously mentioned, the main downside of the use of label-free Raman spectroscopy is that the time of measurement is long: in this experiment, for single-cell label-free measurements, we used an acquisition time of 40 seconds per cell. Even when the acquisition time is lower –for instance, (Liu et al. 2016) reported a 1-3 seconds acquisition time to detect antibiotic susceptibility using surface-enhanced Raman spectroscopic biomarkers- the speed of Raman spectroscopy cannot match the high-throughput nature of flow cytometry for single-cell analysis.

Raman spectroscopy allows to detect differences in biomolecules from one sample to another

Raman spectroscopy allows to detect biomolecules present in different phenotypic populations. Therefore, after automated identification of phenotypic populations, one can use the phenotypic groups to perform a variable selection strategy to select important regions in a data-driven way. We illustrated this approach using the Boruta algorithm, which was recently evaluated as one of the state-of-the-art variable selection methods using Random Forests for omics datasets (Degenhardt, Seifert, and Szymczak 2017). We found that a majority of selected spectral regions were the same according to treatment and automated phenotypic population identification using PhenoGraph (Fig. 3). This information can be used to infer how phenotypic populations are different at the level of their metabolism. To do so, we have based ourselves on a recent literature survey summarizing associations between Raman regions and certain biological compounds (Teng et al. 2016). The ten most important regions in function of phenotypic identification are listed in Table 2, along with the distribution of their intensities (Fig. 4). These regions correspond to carbohydrates and nucleic acids, as well as some unknown regions. An increase in the carbohydrate band (peaks 1042, 1046, 1050, 1057 cm-1) was observed for the stationary phase. The band at 1053 cm-1 could also be a nucleic acid peak, expected at 1054 cm-1. Nevertheless, these assignments for the Raman bands are tentative

and based solely on a literature research, and thus proper validation of these results would have to be made in future experiments.

How to define a phenotypic population?

In this work, we have steered microbial communities towards a certain growth stage, expecting that they would express a certain phenotype that could be retrieved using flow cytometry and Raman spectroscopy. However, in each one of these isogenic populations there might be subpopulations, as can be seen in Figure 2b. We acknowledge the difficulty in defining what a phenotypic population is, and setting a threshold to determine when does one phenotypic population ends and another begins.

A similar problem exists in the area of bacterial taxonomy, where the similarity of 16S sequences is compared. In this case, an arbitrary threshold is set (e.g., 95% similarity at the genus level, 98.56% at the species level) (Stackebrandt and Goebel 1994; Kim et al. 2014). As (Beye et al. 2018) explains, this cut-off was meant to standardize the use of 16S rRNA gene amplicon sequencing, but it had to evolve; i.e., the first threshold for the species level has changed from 97%, to 98.7%, to the current 98.65%. Even now, it is argued that these thresholds are not applicable to multiple genera (Mysara et al. 2017). In the case of phenotypic populations, we propose a definition based on their similarity (after setting a similarity threshold) and their ecology (their relationship with one another and with their environment). Quantifying their similarity can be done in a data-driven way, by means of for example clustering, at the resolution that is required for the specific research. This operational definition allows to define phenotypic populations depending on the research question, as long as researchers motivate and validate their choice. However, using this operational definition means that results cannot be compared across experiments or labs. This is why we highlight the need to find a more standard way to define 'basic phenotypic units', that would allow to measure phenotypic traits and determine if bacteria belong to the same phenotypic population.

We propose to use algorithms -such as hierarchical clustering, t-SNE or PhenoGraph, applied throughout this paper- to define, visualize and characterize phenotypic populations. t-SNE is a well-known technique to visualize high-dimensional single-cell data, being commonly applied to visualize for example cytometry and single-cell RNA sequencing data (Amir et al. 2013; Andrews and Hemberg 2018). Our results confirm that it can be used as an 'off-the-shelf' visualization method to detect phenotypic populations in Raman data when applied to microorganisms.

Bacteria were grown in 9 different conditions (three replicate cultures of three growth stage conditions) to steer the same *E. coli* population to a different morphological and/or metabolic state –to steer them into 9 phenotypic populations. While hierarchichal clustering was able to find eight of these phenotypic populations, PhenoGraph was able to retrieve all nine of them, resulting in a higher ARI as well.

t-SNE and PhenoGraph were also applied to an external dataset from Teng et al., 2016, consisting of *E. coli* that had been treated with different agents, and measured at several time points. We showed that PhenoGraph was capable of differentiating the time points per treatment. Interestingly, two subpopulations were identified per treatment, although samples were measured in triplicate. These corresponded to two replicates, where the third was either assigned to one subpopulation or divided between both (Fig. 6, Fig. S6). Our group has previously shown how small technical variations can create subpopulations that have no biological meaning (García-Timermans et al. 2018), which might explain these findings.

Each algorithm has its own advantages and disadvantages. t-SNE is a highly effective technique to visualize high-dimensional single-cell data, which we confirmed for the Raman analysis of microorganisms. However, automated clustering of the data is not possible using t-SNE without additional algorithms. The PhenoGraph algorithm was originally developed for mass cytometry data (Levine et al. 2015), a variation of flow cytometry which makes use of heavy metal ion tags instead of fluorochromes, resulting in more observed variables but at a

lower acquistion speed (Spitzer and Nolan 2016). PhenoGraph was demonstrated to be highly effective for clustering purposes of single-cell Raman data, and returned a higher clustering performance compared to a more traditional hierarchical clustering approach. However, hierarchical clustering allows to inspect which cells are most similar to each other, a characteristic which is lost when using PhenoGraph. Therefore we want to reiterate that, as proposed by (Andrews and Hemberg 2018) for the analysis of single-cell RNA sequencing data, "[l]ikewise, no computational methods for dimensionality reduction, feature selection and unsupervised clustering will be optimal in all situations". The algorithm of choice depends on the needs of the user. If a researcher wants to visualize subpopulations, we recommend the use of t-SNE. If identification of phenotypic populations is needed in an automated way, PhenoGraph is more appropriate. To assess which individual cells are phenotypically closest, hierarchical clustering can be used. Further investigation of the analysis of Raman data is needed, but investigating additional algorithms specifically developed for high-dimensional single-cell data might further support the impact of the use of Raman spectroscopy. Therefore, we recommend to use flow cytometry to detect shifts in a population at the community level, whereas Raman spectroscopy is preferred to investigate changes at the single-cell level.

Conclusions

The results of this research suggest that:

- Flow cytometry is a more high-throughput technology than label-free Raman spectroscopy, but Raman describes bacterial cells in many more variables, without the need for staining.

- Flow cytometry can be applied to quantify differences in phenotypic heterogeneity at the population level, whereas Raman spectroscopy has sufficient discriminating power to identify separated phenotypic populations at the single-cell level.

- Raman spectroscopy provides the possibility to infer which metabolic properties define different phenotypic populations.

- Currently we propose to use t-SNE to visualize Raman data. State-of-the-art clustering algorithms, specifically developed for high-dimensional single-cell data, can assist in the automated identification of phenotypic populations.

- From a broader perspective, one can motivate that phenotypic populations depend on a similarity threshold, which can be set in clustering algorithms, and on their ecological niche. We therefore suggest that researchers try to include validation controls in their experimental setup, in order to motivate their algorithm with threshold of choice.

Materials & Methods

Cell culture

To determine the growth stages of the cell culture (lag, log and stationary phase), *Escherichia coli* DSM 2092 was grown in nutrient broth (NB, Oxoid, United Kingdom) at 28°C, 120 rpm shaking in three replicates. Cultures had an initial concentration of 10^6 cells/ml (measured with a BD Accuri C6 flow cytometer (BD Biosciences), following the protocol from Van Nevel 2013. The samples were incubated in the dark for 30 h at 28°C, during which optical density (OD, λ = 620 nm) measurements were automatically collected each hour using a microtiter plate reader (Tecan Infinite M200 Pro; Tecan UK, Reading, United Kingdom). The growth phases were assigned after fitting the results with the function *SummarizeGrowth()* from the 'growthcurver v0.30' R package (Sprouffske and Wagner 2016). Cells were harvested 1h, 7h30 and 24h after inoculation, corresponding accordingly to the lag, log and stationary phases of *E.coli* (see Fig. S1). Nutrient broth was included as a negative control.

Following the protocol from (García-Timermans et al. 2018), samples were harvested and fixed in formaldehyde 4% (Sigma- Aldrich) dissolved in PBS (protocol from Bio-Techno Ltd., Belgium). First, 1mL of the cell suspension was centrifuged for 5 min at room temperature and 1,957 x g. For the samples in the lag phase, up to 10 mL were suspended until a pellet could be seen. The supernatant was discarded and cells were suspended in filtered and cold PBS (4°C). The samples were again centrifuged at 1,957 x g for 5 min at room temperature. The supernatant was discarded and the pellet was re-suspended in filtered formaldehyde 4%. The cells were allowed to fix for 1h at room temperature (21°C)~~(21 °C)~~. Then, the samples were centrifuged at 1,957 x g for 5 min at room temperature and washed twice with cold PBS (4°C). Cells were stored at 4°C and analysed in the Raman spectroscope within the week.

Flow cytometry

Fresh samples taken at the lag, log and stationary phase were diluted in filtered PBS and stained with SYBR Green I 1% (Thermo Fisher) during 13 min at 37°C. They were measured with the flow cytometer BD Accuri C6 (BD Biosciences) following the protocol from (Van Nevel

et al. 2013). This resulted in a multivariate description of each cell by four fluorescence detectors (FL1: 533/30 nm, FL2: 585/40 nm, FL3: > 670 nm long pass, FL4: 675/25 nm), of which the FL1 detector was targeted by SYBR Green I, and two scatter detectors (forward scatter, FSC and side scatter, SSC).

Raman spectroscopy

Fixed samples were centrifuged at 1,957 x g for 5 min at room temperature and re-suspended in cold Milli-Q water (Merck-Millipore) (4°C). Then, a 5µL drop was allowed to dry until evaporation on a CaF2 slide (grade 13 mm diameter by 0.5 mm polished disc, Crystran Ltd). The metadata were collected following the guidelines from (García-Timermans et al. 2018) and can be found in Supplementary Information.

Data analysis

Data analysis was conducted using the program R (R Core Team 2018) and RStudio (RStudio team and RStudio 2016).

Flow cytometric data

PhenoGraph

PhenoGraph is a clustering algorithm specifically designed for the analysis of high-dimensional flow- or mass-cytometry data (Levine et al. 2015). It employs a two-step approach, in which for every cell its *k*-nearest cells of similar phenotypic populations are grouped together. This means that, if *N* denotes the number of cells, *N* neighbourhoods are created. Next, a weighted graph is created on these sets of cells. The weight between nodes scales with the number of neighbours that are shared. The Louvain community detection method is implemented to cluster the graph by maximizing the modularity of different groupings of the nodes (Blondel et al. 2008). The PhenoGraph algorithm was run with default settings, in which k was evaluated for different values between five and 100 (github.com/jacoblevine/PhenoGraph). Four channels (FL1-H, FL3-H, FSC-H and SSC-H) were selected to derive a phenotypic fingerprint for each sample. Bacteria were gated to differentiate from background noise as shown in Fig S2. As quality control, the stability of the FL1 signal over time was checked. A 128 x 128 binning grid was constructed for each pairwise combination of these channels (resulting in 6

in total). For each bin a kernel density estimation was performed (with a Gaussian kernel density bandwidth of 0.01). Then, all bins are concatenated to a one-dimensional vector, representing the cytometric fingerprint. Data was normalized to have zero mean and a standard deviation of one. At least 10.000 cells were measured per sample.

PCA

For the principal component analysis (PCA), the data of either the single bacteria, or the pulled information for every group were analysed with the function *fviz_pca_ind()* from the R package 'factoextra' (Kassambara and Mundt 2017).

Principal coordinate analysis (PCoA, also known as multidimensional scaling) was calculated based on the Bray-Curtis dissimilarities between all fingerprints. The function *beta_div_fcm()* from the R package 'PhenoFlow' was used (Props et al. 2016).

Raman spectroscopy

Raman spectra data

The Raman spectra were analysed in the 600-1800 cm^{-1} region, and baseline correction using the SNIP algorithm (ten iterations) and normalization were performed. The area under the curve (AUC) normalization was calculated with the MALDIquant package (v1.16.2) (Gibb and Strimmer 2012). These steps are implemented in the MicroRaman R package (Kerckhof et al. 2018).

Hierarchical clustering

To measure how dissimilar the samples were, we calculated the spectral contrast angle (Wan, Vidavsky, and Gross 2002) between individual cells based on Raman spectra. Then, clusters were determined in an agglomerative way, through Ward's method (ward.D2) from the *fastcluster* R package (Müllner 2018). Hierarchical clustering was implemented using the *hclust*() function from the stats package (R Core Team 2018).

t-distributed stochastic neighborhood embedding (t-SNE)

t-SNE is a dimensionality reduction technique developed for the visualization of high-dimensional data (Van Der Maaten and Hinton 2008). Flow cytometry and Raman spectral data were both visualized with the *TSNE()* function from the scikit-learn machine learning library (Pedregosa et al. 2011), v0.19.1). Principal component analysis was set as initialization method. TSNE was run with default settings unless reported otherwise.

Adjusted Rand Index

Clustering results from both hierarchical clustering and PhenoGraph were quantified by the Adjusted Rand Index (ARI) (Hubert and Arabie 1985). The ARI was calculated with the *adjusted_rand_score()* function from the scikit-learn machine learning library (v0.19.1) (Pedregosa et al. 2011). The Rand index is defined as the number of pairs of instances that are in the same group or in different groups based on two partitions, which is divided by the total number of pairs of instances. This index is then corrected for the expected index, which is based on random clustering in which the elements per cluster are shuffled between clusters. A value of 1 resembles the perfect match between cluster assignments and ground truth labels, a value of 0 resembles random clustering and a negative value (up to -1) resembles arbitrarily worse clustering.

Boruta variable selection

The Boruta variable selection extends on traditional variable selection using Random Forest based variable importance measures. The method includes shadow variables, which are copies of original variables that have been permutated. By performing a Random Forest analysis multiple times, one can decide by means of multiple hypothesis testing which variables are relevant with a certain significance level compared to the most relevant shadow variable (Kursa and Rudnicki 2010). It has been recently proposed as one of the most accurate and stable variable selection methods based on Random Forest based (Degenhardt, Seifert,

and Szymczak 2017). The Boruta algorithm from the *Boruta* R package was run, using the default settings (v6.0.0) (Kursa and Rudnicki 2010).

Statistical test on Boruta outcome

The ten most relevant regions for classification according to the Boruta algorithm were selected. The intensity of these peaks amongst the growth phases were compared with the Wilcoxon rank sum test with a Benjamini-Hochberg correction (upon rejection of the null hypothesis). The functions pairwise.wilcox.test() and p.adjust() from the R package stats v3.5.1 (R Core Team 2018) were used.

PCA

For the principal component analysis (PCA), the data of either the single bacteria, or the pooled information for every group were analysed with the function *fviz_pca_ind()* from the R package 'factoextra' (Kassambara and Mundt 2017). After the first run, two outliers were found using the qqPlot() from the R package 'distr' (Ruckdeschel, Peter ; Kohl, Matthias; Stabla, Thomas ; Camphausen 2006)(Fig. S3). PCA was applied again without these two points.

Data availability

Data and code to reproduce analysis is available on the following repository: https://github.com/CMET-Ugent/FCMvsRaman.

External dataset

We included the dataset from (Teng et al. 2016) in order to validate the generalizability of the PhenoGraph and t-SNE algorithms for the analysis of label-free bacterial Raman data. As described in their article, they tested the stress response of *E. coli* to six chemical stressors at different time intervals with label-free Raman spectroscopy: ethanol, antibiotics ampicillin and kanamycin, n-butanol or heavy metals Cu^{2+} ($CuSO_4$) and Cr^{6+} (K_2CrO_4). Tang et al. showed that each of these treatments resulted in a different phenotype. In other words, each treatment resulted in a unique Raman characterization of cells, which should group together upon

analysis. These treatments were therefore used as label according to which PhenoGraph or t-SNE should group the cells. Three biological replicates of the cell culture were made, and 20 cells were tested per replicate. Bacteria were sampled at different stages of the cell growth. The Raman spectra of the stressed cells were collected after the treatment (5, 10, 20, 30 and 60 min).

Data availability

Our code and data to reproduce the analysis is available at https://github.com/CMET-Ugent/FCMvsRaman.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

CGT and PR co-wrote the paper with contributions from JH, FMK, RP, AS, WW, and NB. CGT collected the data. PR, CGT and RP performed the data analysis. CGT, PR, JH, FMK, RP, WW and NB designed the study. All authors read and approved the final version of the manuscript.

Acknowledgements

The authors would like to thank the funding that made possible this research. CG is funded by Qindao Beibao Marine Science & Technology Co. Ltd., Qingdao West-coast economic new area, China. PR is funded by Special Research Fund (BOFSTA2015000501) from Ghent University. RP is funded by by Ghent University (BOFDOC2015000601). JH is funded by the Flemish Fund for Scientific research (FWO-Vlaanderen, 1S80618N). AGS acknowledges support of BOF UGent (BOF14/IOP/003, BAS094-18, 01IO3618) and FWO-Vlaanderen (G043219). This work was supported through the Geconcerteerde Onderzoeksactie (GOA) from Ghent University (BOF15/GOA/006) and the MICCAS project (project grant no.

3G020119 of the FWO Flanders). The authors would also like to thank Dmitry Khalenkow for his help setting up the Raman microscope.

Bibliography

Ackermann, Martin. 2015. “A Functional Perspective on Phenotypic Heterogeneity in Microorganisms.” *Nature Reviews Microbiology* 13 (8): 497–508. https://doi.org/10.1038/nrmicro3491.

Amir, El-ad David, Kara L Davis, Michelle D Tadmor, Erin F Simonds, Jacob H Levine, Sean C Bendall, Daniel K Shenfeld, Smita Krishnaswamy, Garry P Nolan, and Dana Pe’er. 2013. “ViSNE Enables Visualization of High Dimensional Single-Cell Data and Reveals Phenotypic Heterogeneity of Leukemia.” *Nature Biotechnology* 31 (6): 545–52. https://doi.org/10.1038/nbt.2594.

Andrews, Tallulah S., and Martin Hemberg. 2018. “Identifying Cell Populations with ScRNASeq.” *Molecular Aspects of Medicine* 59: 114–22. https://doi.org/10.1016/j.mam.2017.07.002.

Anetzberger, Claudia, Ursula Schell, and Kirsten Jung. 2012. “Single Cell Analysis of Vibrio Harveyi Uncovers Functional Heterogeneity in Response to Quorum Sensing Signals.” https://doi.org/10.1186/1471-2180-12-209.

Athamneh, A I M, R A Alajlouni, R S Wallace, M N Seleem, and R S Senger. 2014. “Phenotypic Profiling of Antibiotic Response Signatures in Escherichia Coli Using Raman Spectroscopy.” *Antimicrobial Agents and Chemotherapy* 58 (3): 1302–14. https://doi.org/10.1128/AAC.02098-13.

Avery, Simon V. 2006. “Microbial Cell Individuality and the Underlying Sources of Heterogeneity.” *Nature Reviews Microbiology* 4 (8): 577–87. https://doi.org/10.1038/nrmicro1460.

Barbesti, Silvia, Sandra Citterio, Massimo Labra, Maurizio Davide Baroni, Maria Grazia Neri, and Sergio Sgorbati. 2000. “Two and Three-Color Fluorescence Flow Cytometric Analysis of Immunoidentified Viable Bacteria.” *Cytometry* 40 (3): 214–18. https://doi.org/10.1002/1097-0320(20000701)40:3<214::AID-CYTO6>3.0.CO;2-M.

Berry, David, Esther Mader, Tae Kwon Lee, Dagmar Woebken, Yun Wang, Di Zhu, Marton Palatinszky, et al. 2015. “Tracking Heavy Water (D2O) Incorporation for Identifying and Sorting Active Microbial Cells.” *Proceedings of the National Academy of Sciences* 112 (2): E194–203. https://doi.org/10.1073/PNAS.1420406112.

Beye, M., N. Fahsi, D. Raoult, and P.-E. Fournier. 2018. “Careful Use of 16S RRNA Gene Sequence Similarity Values for the Identification of Mycobacterium Species.” *New Microbes and New Infections* 22 (March): 24–29. https://doi.org/10.1016/J.NMNI.2017.12.009.

Bhattacharjee, Arunima, Rupsa Datta, Enrico Gratton, and Allon I Hochbaum. 2017. “Metabolic Fingerprinting of Bacteria by Fluorescence Lifetime Imaging Microscopy.” *Scientific Reports* 7 (1): 3743. https://doi.org/10.1038/s41598-017-04032-w.

Blondel, Vincent D., Jean Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. “Fast Unfolding of Communities in Large Networks.” *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10). https://doi.org/10.1088/1742-5468/2008/10/P10008.

Buysschaert, Benjamin, Bo Byloos, Natalie Leys, Rob Van Houdt, and Nico Boon. 2016. “Reevaluating Multicolor Flow Cytometry to Assess Microbial Viability.” *Applied Microbiology and Biotechnology* 100 (21): 9037–51. https://doi.org/10.1007/s00253-016-7837-5.

Candeloro, Patrizio, Elisabetta Grande, Raffaella Raimondo, Daniele Di Mascolo, Francesco Gentile, Maria Laura Coluccio, Gerardo Perozziello, Natalia Malara, Marco Francardi, and Enzo Di Fabrizio. 2013. “Raman Database of Amino Acids Solutions: A Critical Study of Extended Multiplicative Signal Correction.” *The Analyst* 138 (24): 7331. https://doi.org/10.1039/c3an01665j.

Ceuppens, Siele, Nico Boon, and Mieke Uyttendaele. 2013. “Diversity of Bacillus Cereus Group Strains Is Reflected in Their Broad Range of Pathogenicity and Diverse Ecological Lifestyles.” *FEMS Microbiology Ecology*. https://doi.org/10.1111/1574-6941.12110.

Cowcher, David P., Yun Xu, and Royston Goodacre. 2013. “Portable, Quantitative Detection

of *Bacillus* Bacterial Spores Using Surface-Enhanced Raman Scattering." *Analytical Chemistry* 85 (6): 3297–3302. https://doi.org/10.1021/ac303657k.

Davis, Kimberly M, and Ralph R Isberg. 2016. "Defining Heterogeneity within Bacterial Populations via Single Cell Approaches." *BioEssays*. https://doi.org/10.1002/bies.201500121.

Degenhardt, Frauke, Stephan Seifert, and Silke Szymczak. 2017. "Evaluation of Variable Selection Methods for Random Forests and Omics Data Sets." *Briefings in Bioinformatics*, no. November: 1–12. https://doi.org/10.1093/bib/bbx124.

Dhar, Neeraj, and John D McKinney. 2007. "Microbial Phenotypic Heterogeneity and Antibiotic Tolerance." *Current Opinion in Microbiology* 10 (1): 30–38. https://doi.org/10.1016/J.MIB.2006.12.007.

García-Timermans, Cristina, Peter Rubbens, Frederiek Maarten Kerckhof, Benjamin Buysschaert, Dmitry Khalenkow, Willem Waegeman, Andre G. Skirtach, and Nico Boon. 2018. "Label-Free Raman Characterization of Bacteria Calls for Standardized Procedures." *Journal of Microbiological Methods* 151 (August): 69–75. https://doi.org/10.1016/j.mimet.2018.05.027.

Georgakoudi, Irene, Irene Tsai, Cherry Greiner, Cheryl Wong, Jordy DeFelice, and David Kaplan. 2007. "Intrinsic Fluorescence Changes Associated with the Conformational State of Silk Fibroin in Biomaterial Matrices." *Optics Express* 15 (3): 1043. https://doi.org/10.1364/OE.15.001043.

Gibb, Sebastian, and Korbinian Strimmer. 2012. "MALDIquant: A Versatile R Package for the Analysis of Mass Spectrometry Data," March. https://doi.org/10.1093/bioinformatics/bts447.

Goodacre, R., e. M Timmins, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell, and P. J. Rooney. 1998. "Rapid Identification of Urinary Tract Infection Bacteria Using Hyperspectral Whole-Organism Fingerprinting and Artificial Neural Networks." *Microbiology* 144 (5): 1157–70. https://doi.org/10.1099/00221287-144-5-1157.

Govers, Sander K., Antoine Adam, Hendrik Blockeel, and Abram Aertsen. 2017. "Rapid Phenotypic Individualization of Bacterial Sister Cells." *Scientific Reports*. https://doi.org/10.1038/s41598-017-08660-0.

Huang, Wei E., Mengqiu Li, Roger M. Jarvis, Royston Goodacre, and Steven A. Banwart. 2010. "Shining Light on the Microbial World: The Application of Raman Microspectroscopy." *Advances in Applied Microbiology* 70 (January): 153–86. https://doi.org/10.1016/S0065-2164(10)70005-8.

Hubert, Lawrence, and Phipps Arabie. 1985. "The Trasition from Bargaining to a Competitive Market." *Journal of Classification* 2: 193:218. https://doi.org/10.1007/BF01908075.

Jarvis, Roger M., and Royston Goodacre. 2004. "Discrimination of Bacteria Using Surface-Enhanced Raman Spectroscopy." *Analytical Chemistry* 76 (1): 40–47. https://doi.org/10.1021/ac034689c.

Kassambara, Alboukadel, and Fabian Mundt. 2017. "Factoextra: Extract and Visualize the Results of Multivariate Data Analyses." *R Package Version*. https://doi.org/https://CRAN.R-project.org/package=factoextra.

Kerckhof, Frederiek-Maarten, Benjamin Buysschaert, Dmitry Khalenkow, and Cristina García-Timermans. 2018. "MicroRaman." microraman: Phenotyping of microbial%0A Raman data.

Kim, M., H.-S. Oh, S.-C. Park, and J. Chun. 2014. "Towards a Taxonomic Coherence between Average Nucleotide Identity and 16S RRNA Gene Sequence Similarity for Species Demarcation of Prokaryotes." *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY* 64 (Pt 2): 346–51. https://doi.org/10.1099/ijs.0.059774-0.

Koch, Christin, Falk Harnisch, Uwe Schröder, and Susann Müller. 2014. "Cytometric Fingerprints: Evaluation of New Tools for Analyzing Microbial Community Dynamics." *Frontiers in Microbiology* 5: 273. https://doi.org/10.3389/fmicb.2014.00273.

Kursa, Miron B, and Witold R Rudnicki. 2010. "Feature Selection with the Boruta Package." *Journal Of Statistical Software* 36 (11): 1–13. https://doi.org/Vol. 36, Issue 11, Sep

2010.
Léonard, Lucie, Lynda Bouarab Chibane, Balkis Ouled Bouhedda, Pascal Degraeve, and Nadia Oulahal. 2016. “Recent Advances on Multi-Parameter Flow Cytometry to Characterize Antimicrobial Treatments.” *Frontiers in Microbiology* 7: 1225. https://doi.org/10.3389/fmicb.2016.01225.
Levine, Jacob H., Erin F. Simonds, Sean C. Bendall, Kara L. Davis, El Ad D Amir, Michelle D. Tadmor, Oren Litvin, et al. 2015. “Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells That Correlate with Prognosis.” *Cell* 162 (1): 184–97. https://doi.org/10.1016/j.cell.2015.05.047.
Liu, Chia-Ying, Yin-Yi Han, Po-Han Shih, Wei-Nan Lian, Huai-Hsien Wang, Chi-Hung Lin, Po-Ren Hsueh, Juen-Kai Wang, and Yuh-Lin Wang. 2016. “Rapid Bacterial Antibiotic Susceptibility Test Based on Simple Surface-Enhanced Raman Spectroscopic Biomarkers.” *Scientific Reports* 6 (March): 23375. https://doi.org/10.1038/srep23375.
Maaten, L J P Van Der, and G E Hinton. 2008. “Visualizing High-Dimensional Data Using t-Sne.” *Journal of Machine Learning Research* 9: 2579–2605. https://doi.org/10.1007/s10479-011-0841-3.
Mair, Florian, and Martin Prlic. 2018. “OMIP-044: 28-Color Immunophenotyping of the Human Dendritic Cell Compartment.” *Cytometry Part A* 93A: 402–5. https://doi.org/10.1002/cyto.a.23331.
Maquelin, Kees, Lenie Dijkshoorn, Tanny J K van der Reijden, and Gerwin J Puppels. 2006. “Rapid Epidemiological Analysis of Acinetobacter Strains by Raman Spectroscopy.” *Journal of Microbiological Methods* 64 (1): 126–31. https://doi.org/10.1016/j.mimet.2005.04.028.
McIlvenna, David, Wei E. Huang, Paul Davison, Andrew Glidle, Jon Cooper, and Huabing Yin. 2016. “Continuous Cell Sorting in a Flow Based on Single Cell Resonance Raman Spectra.” *Lab on a Chip* 16 (8): 1420–29. https://doi.org/10.1039/C6LC00251J.
Müllner, Daniel. 2018. “The Fastcluster Package: User’s Manual Version 1.1.25.” http://cran.r-project.org/web/packages/flashClust/ChangeLog.
Mysara, Mohamed, Peter Vandamme, Ruben Props, Frederiek-Maarten Kerckhof, Natalie Leys, Nico Boon, Jeroen Raes, and Pieter Monsieurs. 2017. “Reconciliation between Operational Taxonomic Units and Species Boundaries.” *FEMS Microbiology Ecology* 93 (4). https://doi.org/10.1093/femsec/fix029.
Nebe-von-Caron, G, P.J Stephens, C.J Hewitt, J.R Powell, and R.A Badley. 2000. “Analysis of Bacterial Function by Multi-Colour Fluorescence Flow Cytometry and Single Cell Sorting.” *Journal of Microbiological Methods* 42 (1): 97–114. https://doi.org/10.1016/S0167-7012(00)00181-0.
Nevel, Sam Van, Stefan Koetzsch, Hans Ulrich Weilenmann, Nico Boon, and Frederik Hammes. 2013. “Routine Bacterial Analysis with Automated Flow Cytometry.” *Journal of Microbiological Methods*. https://doi.org/10.1016/j.mimet.2013.05.007.
Novelli-Rousseau, A., I. Espagnon, D. Filiputti, O. Gal, A. Douet, F. Mallard, and Q. Josso. 2018. “Culture-Free Antibiotic-Susceptibility Determination From Single-Bacterium Raman Spectra.” *Scientific Reports* 8 (1): 3957. https://doi.org/10.1038/s41598-018-22392-9.
O’Neill, Kieran, Nima Aghaeepour, Josef Špidlen, and Ryan Brinkman. 2013. “Flow Cytometry Bioinformatics.” *PLoS Computational Biology* 9 (12). https://doi.org/10.1371/journal.pcbi.1003365.
Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. “Scikit-Learn: Machine Learning in Python.” *Journal of Machine Learning Research* 12 (Oct): 2825–30. http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html.
Perfetto, Stephen P., Pratip K. Chattopadhyay, and Mario Roederer. 2004. “Seventeen-Colour Flow Cytometry: Unravelling the Immune System.” *Nature Reviews Immunology* 4 (8): 648–55. https://doi.org/10.1038/nri1416.
Perozziello, Gerardo, Patrizio Candeloro, Antonio De Grazia, Francesco Esposito, Marco Allione, Maria Laura Coluccio, Rossana Tallerico, et al. 2016. “Microfluidic Device for

Continuous Single Cells Analysis via Raman Spectroscopy Enhanced by Integrated Plasmonic Nanodimers." *Optics Express* 24 (2): A180. https://doi.org/10.1364/OE.24.00A180.

Props, Ruben, Pieter Monsieurs, Mohamed Mysara, Lieven Clement, and Nico Boon. 2016. "Measuring the Biodiversity of Microbial Communities by Flow Cytometry." *Methods in Ecology and Evolution* 7 (11): 1376–85. https://doi.org/10.1111/2041-210X.12607.

R Core Team. 2018. "R: A Language and Environment for Statistical Computing." *R Foundation for Statistical Computing*. https://doi.org/10.1007/978-3-540-74686-7.

Roy, Karen De, Lieven Clement, Olivier Thas, Yingying Wang, and Nico Boon. 2012. "Flow Cytometry for Fast Microbial Community Fingerprinting." *Water Research* 46 (3): 907–19. https://doi.org/10.1016/J.WATRES.2011.11.076.

RStudio team, and RStudio. 2016. "RStudio: Integrated Development Environment for R."

Ruckdeschel, Peter ; Kohl, Matthias; Stabla, Thomas ; Camphausen, Florian. 2006. "Package 'Distr.'" http://distr.r-forge.r-project.org/.

Ryabchykov, Oleg, Shuxia Guo, and Thomas Bocklitz. 2018. "Analyzing Raman Spectroscopic Data." *Physical Sciences Reviews* 0 (0): 1–16. https://doi.org/10.1515/psr-2017-0043.

Saeys, Yvan, Sofie Van Gassen, and Bart N. Lambrecht. 2016. "Computational Flow Cytometry: Helping to Make Sense of High-Dimensional Immunology Data." *Nature Reviews Immunology*. https://doi.org/10.1038/nri.2016.56.

Samek, Ota, Stanislav Obruča, Martin Šiler, Petr Sedláček, Pavla Benešová, Dan Kučera, Ivana Márova, Jan Ježek, Silva Bernatová, and Pavel Zemánek. 2016. "Quantitative Raman Spectroscopy Analysis of Polyhydroxyalkanoates Produced by Cupriavidus Necator H16." *Sensors (Basel, Switzerland)* 16 (11). https://doi.org/10.3390/s16111808.

Spitzer, Matthew H, and Garry P Nolan. 2016. "Leading Edge Primer Mass Cytometry: Single Cells, Many Features." https://doi.org/10.1016/j.cell.2016.04.019.

Sprouffske, Kathleen, and Andreas Wagner. 2016. "Growthcurver: An R Package for Obtaining Interpretable Metrics from Microbial Growth Curves." *BMC Bioinformatics*. https://doi.org/10.1186/s12859-016-1016-7.

Stackebrandt, E., and B. M. Goebel. 1994. "Taxonomic Note: A Place for DNA-DNA Reassociation and 16S RRNA Sequence Analysis in the Present Species Definition in Bacteriology." *International Journal of Systematic and Evolutionary Microbiology* 44 (4): 846–49. https://doi.org/10.1016/S0140-6736(01)43317-4.

Stewart, Mary K, Lisa A Cummings, Matthew L Johnson, Alex B Berezow, and Brad T Cookson. 2011. "Regulation of Phenotypic Heterogeneity Permits Salmonella Evasion of the Host Caspase-1 Inflammatory Response." *Proceedings of the National Academy of Sciences of the United States of America* 108 (51): 20742–47. https://doi.org/10.1073/pnas.1108963108.

Tang, Fuchou, Kaiqin Lao, and M Azim Surani. 2011. "Development and Applications of Single-Cell Transcriptome Analysis." *Nature Methods* 8 (4 Suppl): S6-11. https://doi.org/10.1038/nmeth.1557.

Taniguchi, Yuichi, Paul J Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X Sunney Xie. 2010. "Quantifying E. Coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells." *Science (New York, N.Y.)* 329 (5991): 533–38. https://doi.org/10.1126/science.1188308.

Teng, Lin, Xian Wang, Xiaojun Wang, Honglei Gou, Lihui Ren, Tingting Wang, Yun Wang, Yuetong Ji, Wei E. Huang, and Jian Xu. 2016. "Label-Free, Rapid and Quantitative Phenotyping of Stress Response in E. Coli via Ramanome." *Scientific Reports* 6 (1): 34359. https://doi.org/10.1038/srep34359.

Wan, Katty X., Ilan Vidavsky, and Michael L. Gross. 2002. "Comparing Similar Spectra: From Similarity Index to Spectral Contrast Angle." *Journal of the American Society for Mass Spectrometry* 13 (1): 85–88. https://doi.org/10.1016/S1044-0305(01)00327-0.

Watson, Dakota A., Leif O. Brown, Daniel F. Gaskill, Mark Naivar, Steven W. Graves, Stephen K. Doorn, and John P. Nolan. 2008. "A Flow Cytometer for the Measurement of Raman Spectra." *Cytometry Part A* 73A (2): 119–28.

https://doi.org/10.1002/cyto.a.20520.
Xie, Changan, De Chen, and Yong-qing Li. 2005. “Raman Sorting and Identification of Single Living Micro-Organisms with Optical Tweezers.” *Optics Letters* 30 (14): 1800. https://doi.org/10.1364/OL.30.001800.
Yashchenok, Alexey, Admir Masic, Dmitry Gorin, Bong Sup Shim, Nicholas A. Kotov, Peter Fratzl, Helmuth Möhwald, and Andre Skirtach. 2013. “Nanoengineered Colloidal Probes for Raman-Based Detection of Biomolecules inside Living Cells.” *Small* 9 (3): 351–56. https://doi.org/10.1002/smll.201201494.

Supplementary Information

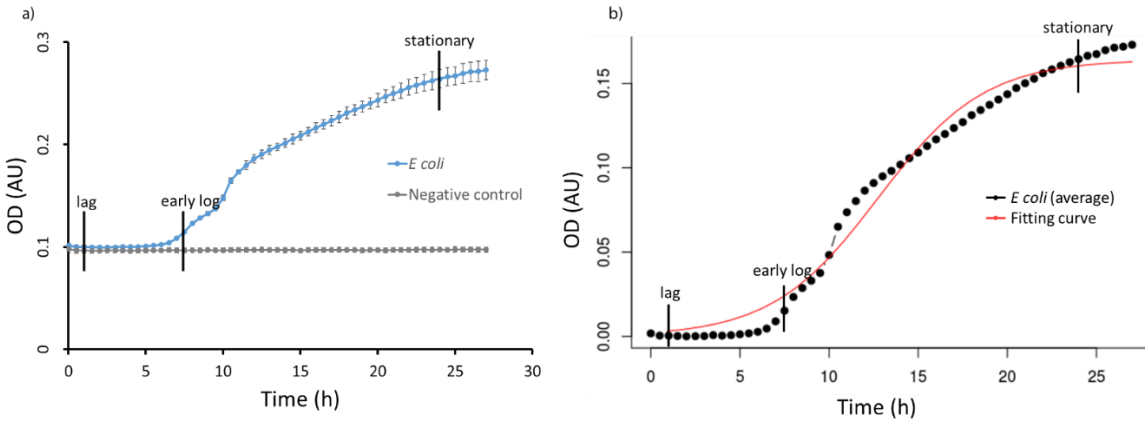

Fig S1: Growth curve for *E. coli* in nutrient broth at 28°C, 120 rpm shaking. Three replicates of the cell culture were made. (a) OD results. In blue, the results for the E. coli; in grey, the negative control (medium). (b) Fitting model for the *E. coli* OD results (after background substraction) to assign the lag, log and stationary phases.

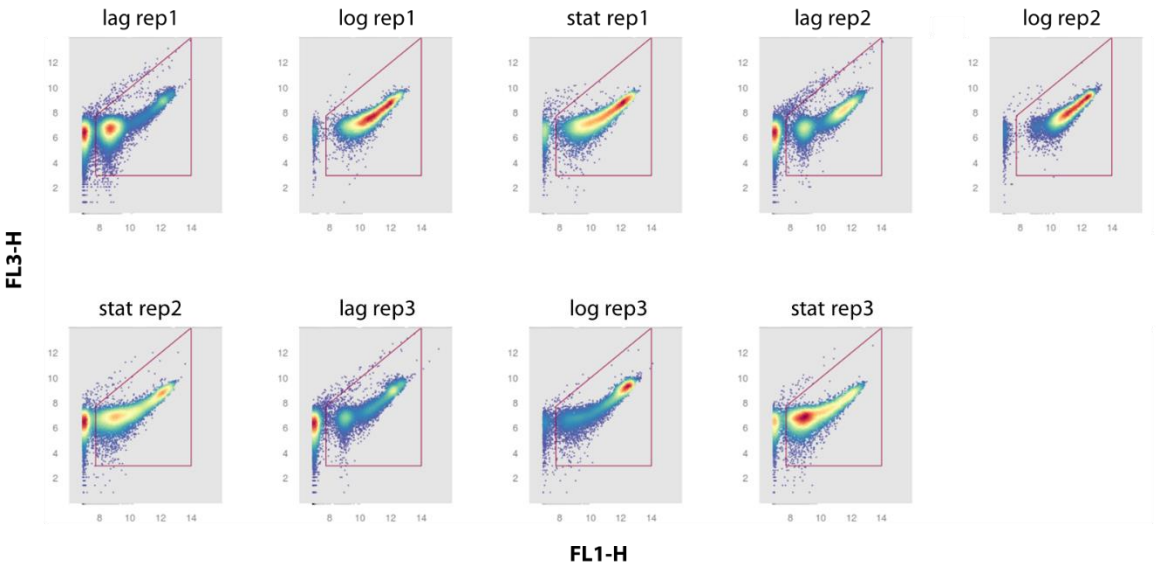

Fig S2: Gating strategy for the flow cytometric data. Arcsinh transformed data of FL1/FL3 in a density plot (see Materials & Methods).

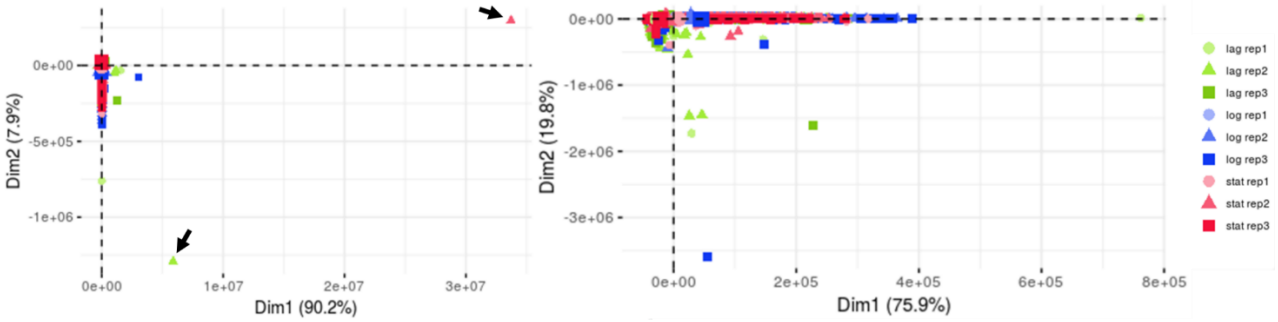

Fig S3: First two components of a PCA for single-cell flow cytometric data. On the left, the two outliers found using the qqPlot() function of the 'distr package are marked with an arrow. On the right, the same PCA was run without the two outliers. Colors correspond to growth phases, shapes to replicates.

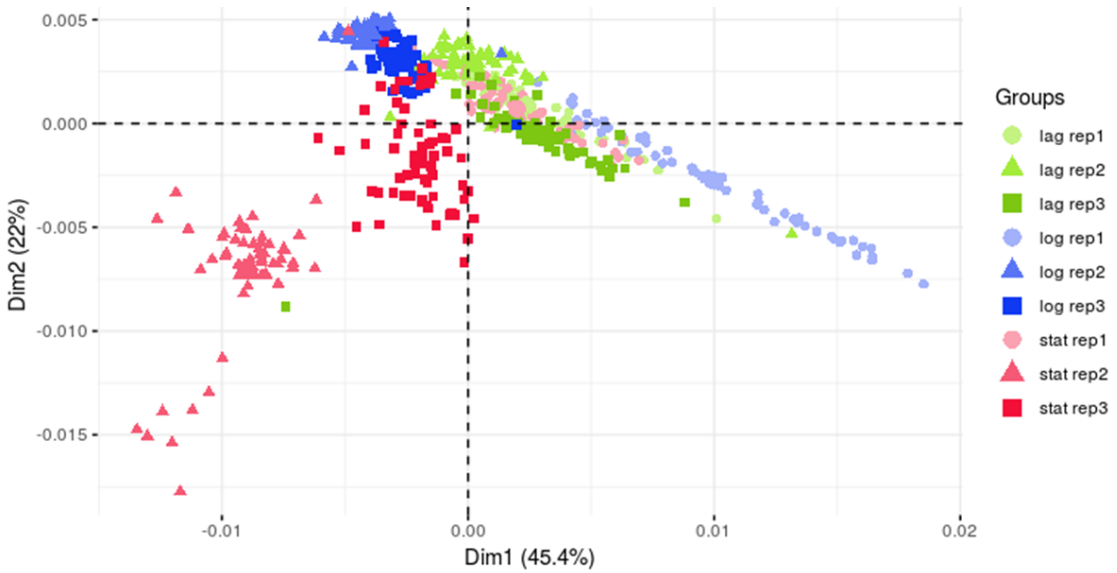

Fig S4: First two components of the PCA for Raman spectra of individual cells. Colors correspond to growth phases, shapes to replicates.

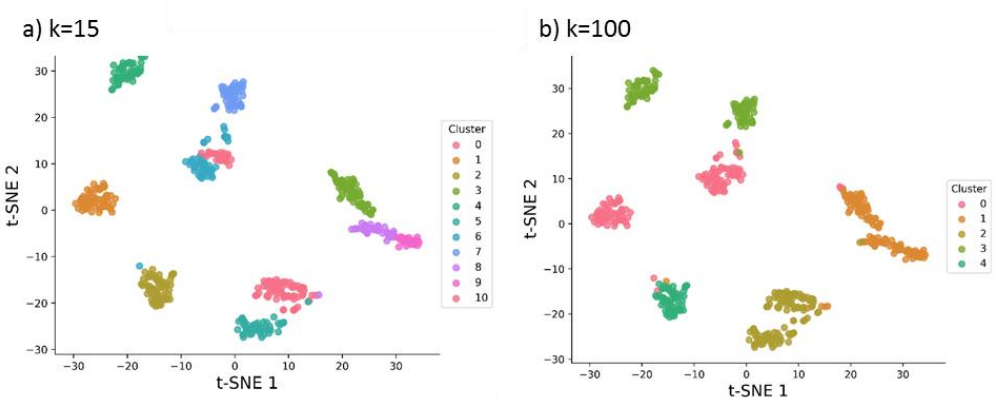

Fig S5: Visualization of PhenoGraph clustering results of the Raman data derived from the **E. coli** culture presented in Figure 3 for different k: (a) k=15, (b): k=100.

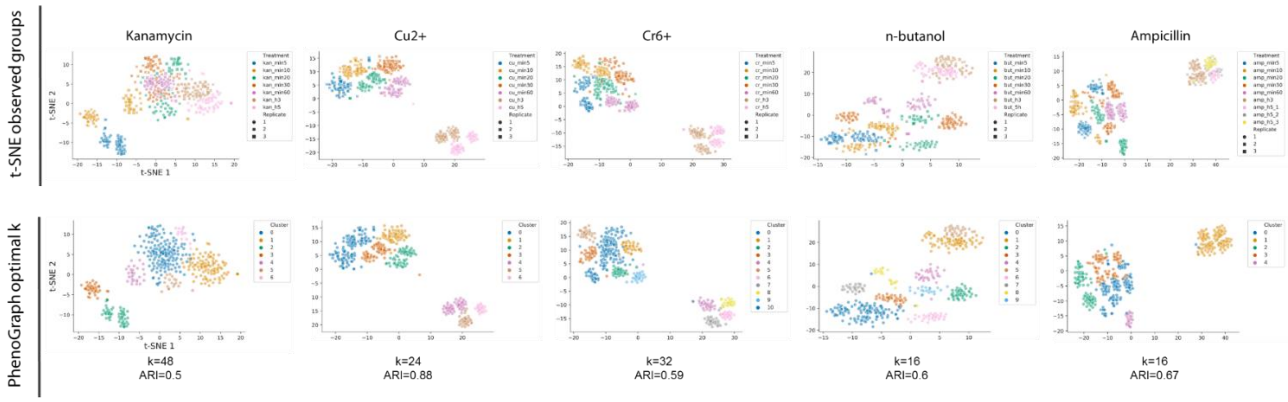

Fig. S6: Ramanone results for all the groups. On top, the t-SNE for the observed phenotypic groups. At the bottom, the PhenoGraph results after calculating the optimal number of clusters (highest ARI).

Table S1: Metadata aid for Raman spectra

Experiment overview:	
Hypothesis	*E. coli* in different growth stages and/or cells from different replicates will differ in their Raman spectra
Variable(s) tested	*E.coli* were fixed in log, lag and stationary phase. Three replicates of the cell culture.
Conclusions	Differences in the Raman spectra were found when cells differed in the growth stage and/or belonged to another culture
Quality control (internal/external)	Silica gel check
Samples and sample acquisition	
Material and source	*Escherichia coli DSM 2092*
Growing conditions/sampling	Cells were grown in LB, 25°C, 120 rpm shaking
Filename format: <Replicate number>_<Treatment>_<Cell number>	
Label in the samples	No label used
Fixation method	Filtered PFA 4%
Integration time	40 sec
Accumulations	1
Grid	300 –mm/g
Instrument	
Laser	785 nm excitation diode laser (Toptica). 175 mW of power before the objective.
Silica check (quality control)	Before objective/after objective or just the total laser power: stability regarding other days was checked
Objective used (magnification) / Numeric aperture (NA)	100x/0.9 NA (Nikon)
Camera	-70 °C cooled CCD camera (iDus 401 BR-DD, ANDOR)
Dry/water/oil objective	Dried samples
Model of spectroscope	WITec Alpha300R+
Other specifications (chromatic/flat field correction/other)	
Data analysis	
Background subtraction method (if used)	No. Measurements with cosmic rays were deleted
Normalization method (peak /min-max /area under-curve /other)	Area under the curve (‘Total Ion Count’)

Smoothing and interpolation (if done)	Baseline correction
Statistics/Machine learning algorithm	‘MicroRaman’ package (GitHub). Spectral contrast angle, ward.D2 dissimilarity and hierarchical clustering. Boruta
Accessibility	https://github.com/CMET-Ugent/FCMvsRaman.
Other relevant information	