Taxon-specific markers for the qualitative and quantitative detection of bacteria in human samples

Nicole Strittmatter[1], James McKenzie[1], Adam Burke[1], Tony Rickards[2], Monica Rebec[2], Zoltan Takats[1]

[1]Section of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, UK
[2]Department of Microbiology, Imperial College Healthcare NHS Trust, Charing Cross Hospital, London, UK

INTRODUCTION - DETECTION OF BACTERIA IN CLINICAL SAMPLES

Microorganisms play an important role as human commensals and pathogens. To ensure the most effective treatment for bacterial infections, the causative agent has to be identified as precisely and as rapidly as possible. In routine clinical microbiology laboratories identification of a human pathogen is routinely based on isolating the pathogen of a clinical sample on solid culturing medium and subsequently obtaining a pathogen ID. Identification results can be traditionally obtained by observing phenotypic characteristics including colony morphology, carbon source utilization and enzymatic activity to define genera and species. A more recent and paradigm changing development in microbiological laboratories was the introduction of MALDI-TOF MS which enabled the identification of bacteria based on their cellular protein patterns. This methodology shortened the time-demand needed for identification by at least 24hrs. However, a purified culture is still needed for reliable identification results. Although direct-on sample applications were reported for urine and positive blood cultures, the protocols are of increasing complexity compared to the identification of pure cultures.

Molecular genetic methods have the advantage of being independent of culture conditions and can potentially reduce the time-demand needed for identification, especially if the culturing step is completely eliminated. The sequencing of 16S rRNA relies on the highly conserved nature of the 16S rRNA which serves as a basis for current microbiological taxonomy. However, the method is rarely used in routine clinical microbiology settings due to the need for extensive sample preparation and associated costs.

Rapid evaporative ionization mass spectrometry (REIMS) was developed recently, specifically for real-time, in-vivo analysis of biological tissues in the surgical environment. The rationale for the development of the technique was to improve surgical margin assessment and identification of unknown tissues in the course of solid tumor resection interventions. Mass spectra obtained using REIMS feature mainly complex lipids of the cellular membranes and elicit high histological specificity in the case of mammalian tissues. It has been shown recently that this technology can be further applied to the analysis of microorganisms such as bacteria and fungi. 28 of the most important human bacterial pathogens and five *Candida* species have been shown to be distinguishable with >95% identification accuracy as compared to MALDI TOF MS. Furthermore, the technique was shown to distinguish seven different *Escherichia coli* strains independently of culture medium and culture age.

Although microbial REIMS phospholipid profiles were shown to be highly specific and reproducible, detection and identification of microorganisms based on their full lipidomic profile in a complex human matrix such as sputum or tissues is most likely not feasible. Therefore, an alternative strategy is proposed here to allow for the detection of bacteria in complex matrices. This strategy involves the determination of complex lipid markers that are specific for certain taxa of bacteria (phylum- down to species-level). Subsequently, these so-called taxonomical markers are used to detect bacteria in human matrices. As a main example, the visualization of bacteria in colorectal tissue specimens is demonstrated. This approach is particularly interesting as using other techniques in this case would either lead to loss of spatial information (as with molecular methods) or loss of the untargeted nature (as in case of immunohistochemistry).

MATERIALS AND METHODS

Rapid Evaporative Ionization Mass Spectrometry. A REIMS spectral database was compiled of pure cultures of a variety of human pathogens and commensals. As taxonomical markers have to be present in a variety of different culture conditions, different culture ages, culture media, and different culturing atmospheres were included into the microbial database. REIMS analysis was performed using a forceps-shaped sampling probe connected to a power-controlled Valleylab Force EZ electrosurgical generator that was used as a RF power supply at a setting of 60W. Rapid heating of the microbial biomass results in the production of an aerosol which is directly sampled by a Thermo Scientific Exactive Mass Spectrometer

operated in negative ion mode over the mass range of *m/z* 150-2000 using a 1.5m long polymer tubing connected to the mass spectrometer inlet capillary.

Determination of taxon-specific markers. Bacterial REIMS profiles were imported into Matlab environment where an ANOVA test followed by Tukey's honest significant difference test was performed to derive taxon-specific markers. For this, the spectral profiles of different groups of bacteria were compared to each other on all relevant taxonomical levels (species- up to kingdom-level) to determine whether certain types of signals could be assigned only to a certain group of bacteria. Fungi were included into the analysis as was the NCI60 cell line panel to cover the human lipidome.

Desorption Electrospray Ionization Mass Spectrometry Imaging. In DESI, a pneumatically-assisted electrospray is directed on the surface of a sample to allow for ambient surface sampling. DESI-MSI experiments were performed using an automated home-built DESI source mounted on a Thermo Scientific Exactive Mass Spectrometer. Mass spectra were acquired in negative ion mode over the mass range of *m/z* 200-1050. For imaging experiments, human colorectal tissue specimens (both healthy and cancerous) were cryo-sectioned to 10µm thickness and thaw-mounted on Superfrost glass slides. For DESI mass spectrometry imaging experiments, horizontal line scans were performed over the tissue surface and combined to form a two-dimensional spatially-resolved chemical map of the tissue section. For DESI experiments, methanol/water (95:5 *v/v*) was used as electrospray solvent at a flow rate of 1.5uL/min. Nitrogen was used as a nebulizing gas at a pressure of 7bar. The distance between DESI sprayer and sample was 1.5mm and the distance of inlet capillary and sprayer was set to 14mm.

RESULTS

This study was designed to determine the applicability of taxonomical markers and the proposed workflow to determine different types of bacteria present in a given sample. These samples can be composed of a mixture of different bacteria as in microbial communities or bacteria in a complex human matrix such as tissues. Taxon-specific markers can be found for a range of bacteria and on different taxonomical levels, however, there is not necessarily a specific marker available at every taxonomical level for a certain group of bacteria. The following selection of taxonomic markers could be established that showed specificity for certain groups of bacteria and that were absent in both human samples and fungal samples.

There are two different types of taxonomical markers, those that are continuously expressed under arbitrary culturing conditions and others that are produced from certain strains only or under certain conditions only and in variable amounts.

Very specific conserved taxonomical markers include

- Mycolic acids for bacteria belonging to the Corynebacterineae suborder such as *Mycobacterium* spp., *Corynebacterium* spp. and *Rhodococcus* spp.. From REIMS experiments and from literature it is known that the mycolic acid chain length is additionally very specific to the genus of bacteria within the Corynebacterineae suborder. It was further shown that mycolic acid patterns can be specific for certain species within a single genus. However, the REIMS dataset has to be further enlarged to validate species-specific mycolic acid patterns. The following mycolic acids have been detected from the corresponding genera:
    - *Mycobacterium* spp.: C77-C81 (even and odd numbered, 0-2 unsaturations)
    - *Corynebacterium* spp.: C28-C36 (even numbered, 0-2 unsaturations)
    - *Nocardia* spp.: C48-C56 (even numbered, 0-3 unsaturations)
    - *Rhodococcus* spp.: C28-C38 (even and odd numbered, 0-4 unsaturations)
- A variety of sphingolipid species were found to be specific for members of the Bacteroidetes. These sphingolipids include oxidized ceramides species, phosphoethanolamine dihydroceramides and C15:0-substituted phosphoglycerol dihydroceramides and dihydroceramide.

  Among those sphingolipid species, a series of galactosylated sphingolipids was found to be specific for *Bacteroides fragilis* (*Bacteroides fragilis* alpha-Galactosylceramides)
- Among bacteria, plasmalogens are highly specific for anaerobic bacteria such as *Clostridium* spp. and *Fusobacterium* spp.. This is due to the fact that aerobic bacteria lost the biochemical pathway required for plasmalogen synthesis. Humans however are able to synthesize plasmalogens (although via a different biochemical pathway from anaerobes) although these were generally found to have longer chain lengths than bacterial plasmalogens.

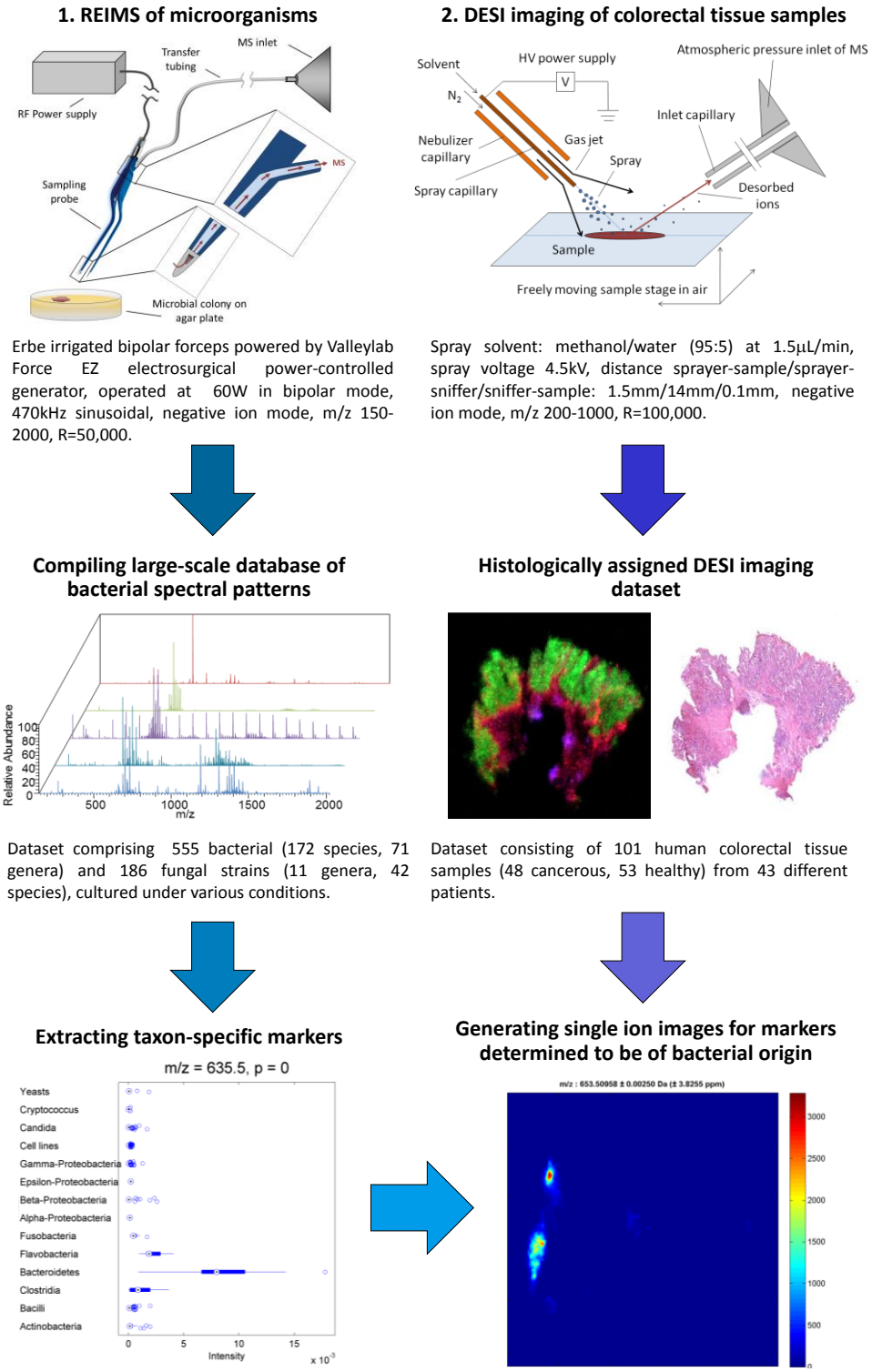

Figure 1. Workflow used in presented study. A REIMS spectral database is generated and used to derive taxon-specific markers for bacteria. DESI imaging of tissue sections is performed and bacterial markers are visualised in DESI imaging datasets.

Other less conserved taxonomical markers that are indicative of a certain group of bacteria but might show very different production amounts include for instance lipopeptides that are produced specifically by certain *Bacillus* species such as surfactin for *B. subtilis* and lichenysin for *B. licheniformis*. Production of these two molecules also enables straightforward differentiation of these otherwise very closely related bacteria. A further example includes PQS-derived quorum-sensing molecules and mono- and di-rhamnolipid species found for *Pseudomonas aeruginosa*.

Subsequently, to demonstrate the applicability of these taxonomical markers for the detection of bacteria in biological samples, we attempted to visualize the presence and distribution of bacteria in colorectal tissue specimens where bacteria are known to cover the mucosal membranes in the gut by generating single ion images for the taxonomical markers that were obtained in earlier steps. A variety of largely co-localized compounds were detected that could be attributed to the Bacteroidetes phylum as is displayed for an exemplary tissue specimen in Figure 2. The areas of strong bacterial presence in this cancerous tissue specimen could be correlated with necrotic tissue areas (see Figure 2) as determined by a histopathologist. The most intense bacterial presence was always found in necrotic tissue areas. A multitude of sphingolipid species could be detected in necrotic tissues whereas only the most intense bacterial peaks could be observed in the majority of non-necrotic samples. The molecules that are specific to *Bacteroidetes fragilis* could not be detected suggesting that the Bacteroidetes bacteria present do not contain a high amount of the opportunistic pathogen *B. fragilis*. Members of the Bacteroidetes phylum were reported in metagenomic studies to be accountable for up to 50% of the gut microbial community.

A mass spectrum of the necrotic tissue and the surrounding cancer tissue is shown in Figure 3 and displays markedly different phospholipid compositions with a variety of sphingolipid-derived taxonomic marker species clearly visible in the mass range of m/z 500-700.

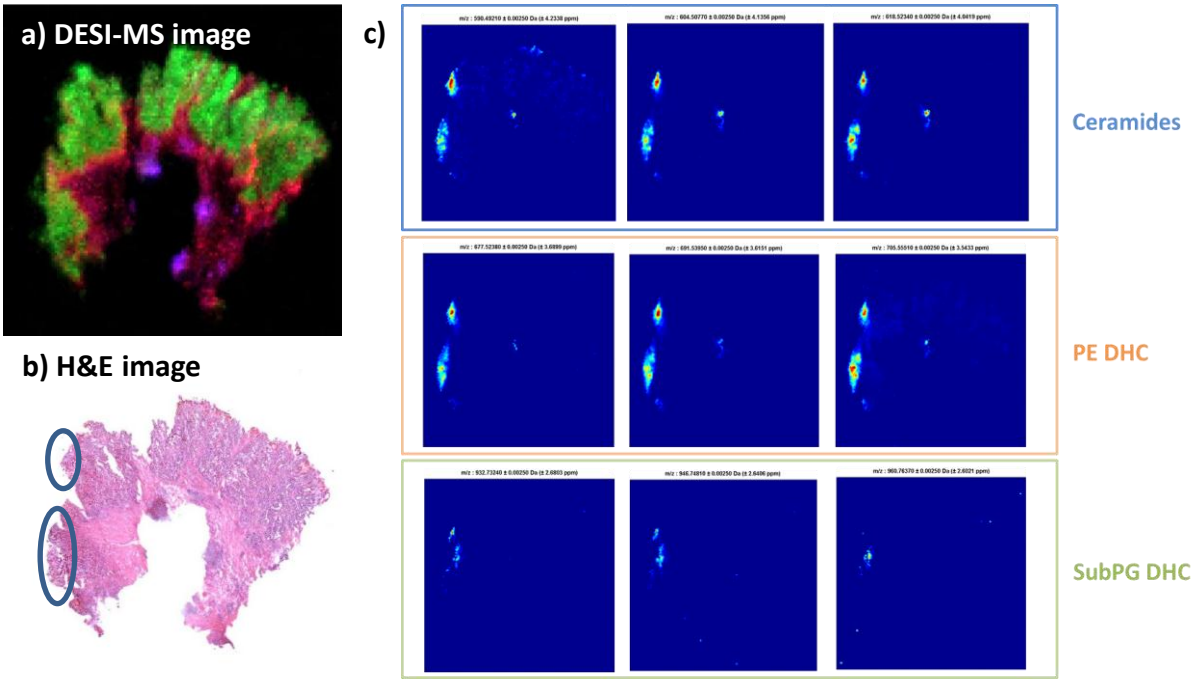

Figure 2. a) RGB DESI image displaying tissue morphology, b) corresponding H&E image, necrotic areas marked. c) Single ion images for homologous ceramides (top), phosphoethanolamine dihydroceramides (middle) and C15:0-substituted phosphoglycerol dihydroceramides (bottom).

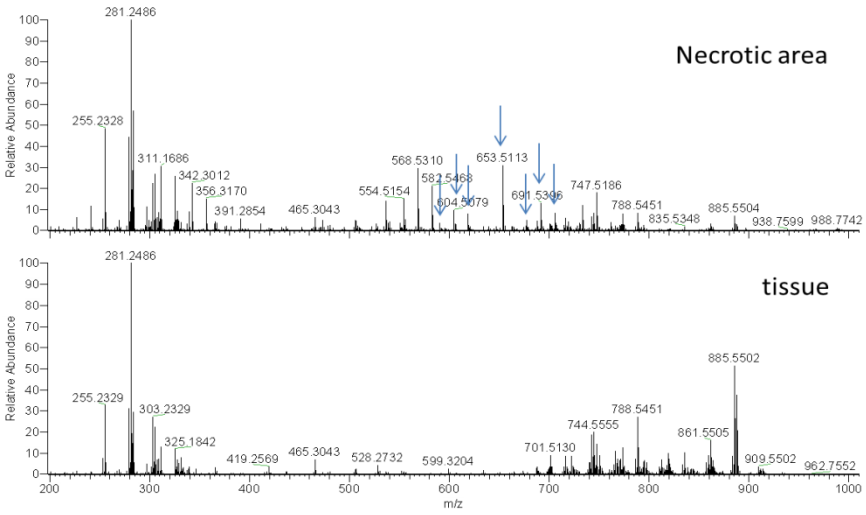

Figure 3. Full scan mass spectra for colorectal adenocarcinoma (bottom) and as obtained for necrotic area of same tissue section. Arrows indicate taxonomic markers.

CONCLUSIONS

This study demonstrates that molecular species differ significantly between microbial lipidomes and the human tissue lipidome. Based on a database of REIMS lipid profiles of

both microorganisms and human cell lines, taxon-specific markers were found for a variety of bacterial species on different taxonomic levels. These markers were shown to be absent in human lipidomes/metabolome and can thus be used to visualize the presence of bacteria in human samples such as shown for human colorectal tissues. It was further demonstrated that taxonomic markers derived by the REIMS technique can be used in conjunction with other mass spectrometric ionization techniques detecting lipid profiles such as DESI, which is more suitable to the direct analysis of mucosal sampling devices or mass spectrometry imaging. The presented approach can be extended to a wide variety of applications as the characterization of microbial communities during endoscopic interventions using REIMS, and for the detection of microbes in human sterile and non-sterile biofluid samples to facilitate appropriate antibiotic treatment.