Talanta 182 (2018) 164–170

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Rapid and robust analytical protocol for *E. coli* STEC bacteria subspecies differentiation using whole cell MALDI mass spectrometry

Kevin Mclean[a,*], Javier Palarea-Albaladejo[b], Carol G. Currie[a], Lisa H.J. Imrie[c], Erin D.T. Manson[d], Douglas Fraser-Pitt[e], Frank Wright[b], Colin J. Alexander[b], Kevin G.J. Pollock[f], Lesley Allison[g], Mary Hanson[g], David G.E. Smith[h]

[a] *Moredun Research Institute, Pentlands Science Park, Penicuik, Mid Lothian, UK*
[b] *Biomathematics and Statistics Scotland (BioSS), James Clerk Maxwell Building, Peter Guthrie Tait Road, The King's Buildings, Edinburgh, Scotland EH9 3FD, UK*
[c] *University of Edinburgh, CH Waddington Building, Max Born Crescent, Kings Buildings, University of Edinburgh, Scotland EH9 3BF, UK*
[d] *University of Glasgow, Glasgow Polyomics, Wolfson Wohl Cancer Research Centre, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1QH, UK*
[e] *Novabiotics Limited, Cruickshank Building, Aberdeen AB21 9TR, UK*
[f] *Health Protection Scotland, NHS National Services Scotland, 5 Cadogan St, Glasgow G2 6QE, UK*
[g] *Scottish E. coli O157/VTEC Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, UK*
[h] *Heriot-Watt University, William Perkin Building, Currie, Edinburgh EH14 4AL, UK*

ARTICLE INFO

Keywords:
MALDI-MS
STEC
Mass spectrometry
Spectral processing
Escherichia coli
Shigatoxigenic *E. coli*

ABSTRACT

Whole cell MALDI is regularly used for the identification of bacteria to species level in clinical Microbiology laboratories. However, there remains a need to rapidly characterize and differentiate isolates below the species level to support outbreak management. We describe the implementation of a modified preparative approach for MALDI-MS combined with a custom analytical computational pipeline as a rapid procedure for subtyping Shigatoxigenic *E. coli* (STEC) and accurately identifying strain-specifying biomarkers. The technique was able to differentiate *E. coli* O157:H7 from other STEC. Within O157 serotype O157:H7 isolates were readily distinguishable from Sorbitol Fermenting O157 isolates. Overall, nine homogeneous groups of isolates were distinguished, each exhibiting distinct profiles of defining mass spectra features. This offers a robust analytical tool useable in reference/diagnostic public health scenarios.

1. Introduction

Recent technological advances in the field have seen matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) being used increasingly as a rapid and reliable method for bacterial identification in clinical microbiology laboratories [1,2]. Accurate identification of pathogens requires interrogation and access of a suitable and proprietary database, provided through analysis packages such as BioTyper (Bruker Daltonics), vitek-ms (Biomerieux) or s10-germs [3,4]. These MALDI-MS-based methods have been validated for identification of several gastro-intestinal pathogens including STEC, *Salmonella*, *Campylobacter* and *Listeria*. However, there is still a need for tools which can rapidly define pathogens below species level, e.g. serogroup or serotype. Such tools would have particular utility in managing outbreak situations, supporting decision making and targeting public health efforts, for instance in differentiating *E. coli* O157:H7 from other STEC serotypes. Sub-species notification and serotype differentiation can be achieved if relevant strain data are present in the reference database [5,6]. Currently preparation of bacterial samples for MALDI analysis is typically a multistep process and proprietary software and access to the accompanying database is required. While direct application of bacterial colonies to the MALDI plate can yield rapid identifications, this is not suitable when working with organisms that require biological containment (e.g. Containment level 3 in UK). We have developed a protocol which offers considerable savings in time and reagents which kills the bacteria and allows their removal from containment laboratory for analysis in a routine mass spectrometry lab. This protocol can be used in conjunction with proprietary software such as Bruker Biotyper to gain robust species identifications. Additionally, we have developed a custom pipeline for robust processing mass spectra from raw data

* Corresponding author.
E-mail addresses: kevin.mclean@moredun.ac.uk (K. Mclean), javier.palarea@bioss.ac.uk (J. Palarea-Albaladejo), carol.currie@moredun.ac.uk (C.G. Currie), limrie@staffmail.ed.ac.uk (L.H.J. Imrie), erin.manson@glasgow.ac.uk (E.D.T. Manson), doug@novabiotics.co.uk (D. Fraser-Pitt), frank.wright@bioss.ac.uk (F. Wright), colin.alexander@hutton.ac.uk (C.J. Alexander), kevin.pollock@nhs.net (K.G.J. Pollock), Lesley.Allison@luht.scot.nhs.uk (L. Allison), Mary.Hanson@luht.scot.nhs.uk (M. Hanson), david.smith@hw.ac.uk (D.G.E. Smith).

https://doi.org/10.1016/j.talanta.2018.01.055
Received 13 October 2017; Received in revised form 17 January 2018; Accepted 20 January 2018

which includes identification of low quality mass spectra and produces processed data which is compatible with a suite of statistical and phylogenetic analytical tools to produce epidemiologically informative outputs.

The protocol has been applied to a panel of enterohemorrhagic *E. coli* (EHEC) isolates to demonstrate its utility in differentiating within a group of closely-related serotypes. *E. coli* O157:H7 is the prototypical enterohemorrhagic *E. coli* (EHEC), having been identified since the 1980s as the cause of serious outbreaks of hemorrhagic colitis associated with the consumption of a variety of animal and plant products [7–9]. The association of *E. coli* O157:H7 with cattle and other ruminants has been verified in numerous studies [10,11] and these continue to be recognised as the main reservoir for these bacteria.

Since first recognition as a human pathogen, *E. coli* O157:H7 and other EHEC have remained major threats to global public health and outbreaks are regularly reported. EHEC are characterized by the carriage of one or more bacteriophages encoding potent Shiga-like toxins that are responsible for the most serious consequences of infection. In man, toxin release in the gut can lead to localised or systemic tissue damage resulting in the cardinal signs of EHEC infection most usually presenting as acute abdominal pain and bloody diarrhoea which may progress to hemorrhagic colitis (HC) and life-threatening hemolytic uremic syndrome (HUS) [12,13].

The most commonly reported STEC infections in many countries, including the UK, are caused by *E. coli* O157:H7. However, an increasing number of non-O157 STEC strains such as *E. coli* O26, O111 and O103 are increasingly being reported worldwide, likely due to heightened awareness of their disease-causing potential and improved laboratory methodology for their detection. These organisms also represent substantial public health threats, causing sporadic cases of infection and outbreaks resulting in HUS. Sorbitol-fermenting STEC O157 have also been of concern because of an association with more severe infection [14–16]. Additionally, novel STEC strains can arise unpredictably. The largest European STEC outbreak occurred in Germany in 2011 and resulted in a disproportionately high number of cases progressing to HUS resulting from infection with a Shiga toxin-producing *E. coli* O104:H4 with enteroaggregative properties [17,18]. Given the potential severity of infection, the persistent appearance of sporadic cases and the unpredictability and likelihood of outbreaks, the need for rapid, reliable detection, isolation and confirmatory tests for serotype and subtype is imperative. Recent reports have begun to demonstrate the potential capabilities of MALDI-MS in differentiating selected phylogroups, pathotypes and/or serotypes of *E. coli* [19–21] and, even, the novel O104 STEC.

As a step towards development and deployment of novel methods for identification and differentiation of STEC including emergent strains at serotype level, we assessed the use of MALDI-MS to differentiate and identify strain types within the STEC/EHEC family. We assessed the ability of a modified protocol and analytical pipeline for its utility and rapidity in (i) differentiating *E. coli* O157:H7 from other EHEC/STEC and (ii) differentiating within serotype O157. We demonstrate that these goals can be achieved and this approach can provide rapid identification and differentiation of strains. As sporadic cases and outbreaks persist in human populations this offers an analytical tool useable in reference/diagnostic public health scenarios to support management of outbreaks.

2. Materials and methods

2.1. Bacterial samples, culture and preparation

E. coli isolates were chosen for this analysis, representing a range of STEC and including some well characterized isolates from international origins. The majority of isolates (summarised in Table 1; Supplementary Table 1 provides further information for each isolate) were obtained via Scottish *E. coli* O157/VTEC Reference Laboratory (SERL) and comprised isolates from single case-patients or from epidemiologically-linked samples through the routine screening of suspect fecal samples submitted to SERL. This set of isolates covered the various O-serogroups that occur in Scotland and are predominated by *E. coli* O157:H7; to extend the range of serotypes, some isolates have been sourced from elsewhere. Three *E. coli* O157:H7 isolates of US origin (Walla Walla 1 [22], TT12A [23] and EDL933 [24]) and the *E. coli* O157:H7 Sakai strain from Japan [25] were also included in the analysis. Isolation and identification of *E. coli* was carried out at SERL following routine, in-house protocols [26] and were stored on Protect™ vials.

Table 1
STEC isolate panel summary information.

STEC isolate panel	
Serotype	No of isolates
O157:H7	32
O157: NM (SF)	7
O26	16
O111	11
O103	4
O5	3
O113	4
O-other	9
O-undefined	6
total	92

For preparation of bacterial suspensions for MALDI-MS analyses, isolates were cultured in a containment laboratory (as required in the UK) on nutrient agar plates; typically three plates per isolate. From each isolate, multiple samples were prepared by pooling 10 colonies into 150 µl 50% v/v acetonitrile (ACN): 0.1% v/v trifluoroacetic acid (TFA) in water and mixing vigorously on a vortex mixer. To ensure the procedure completely inactivated cells the suspensions were cultured overnight and no growth was observed. Inactivated bacterial samples were then removed from containment.

2.2. Maldi-Tof-MS

A 1 µl aliquot of bacterial suspension was spotted onto the plate and overlaid with 1 µl of matrix (α-Cyano-4-hydroxycinnamic acid - CHCA), 10 mg ml^{-1} in 50% v/v ACN: 0.1% v/v TFA in water on a 384 place aluminium target plate and allowed to dry. Several percentages of TFA in the matrix solvent were assessed and 0.1% TFA gave the most optimal combination peak number and signal intensity. For each isolate, 4 technical replicates from each of 5 biological replicates (each biological replicate represented by 10 colonies) were spotted on to the target plate giving 20 spots per isolate. Samples were analyzed with a Bruker Daltonics Ultraflex II mass spectrometer. Mass spectra were acquired in linear, positive mode over a mass range of 2000–20000 Da. The accelerating voltage was 25 kV, the laser frequency was 50 Hz and the delayed extraction time was 100 ns. One hundred shots were collected from 10 different sites on each spot and accumulated to give 1000 shots for each spot. To assess reproducibility, three isolates were cultured, processed and examined in replicate on six separate occasions – this preliminary analysis showed high reproducibility of *m/z* spectra within individual isolates. The instrument was externally calibrated using protein calibrants, Insulin (MW 5734.51), Ubiquitin I (MW 8565.76), Cytochrome C (MW 12360.97) and Myoglobin (MW 16952.30).

2.3. Data analysis

The raw mass spectrometry data were processed through an in-house bioinformatics pipeline implemented on the open-source R system for statistical computing [27]. The pipeline is based on the MALDIquant package [28] along with our novel processing routines