High-throughput molecular typing of microbes using the Sequenom Massarray platform

David Whiley[A,B,E], *Ella Trembizki*[A,B], *Melanie Syrmis*[A,B], *Jenny Nakos*[C], *Cheryl Bletchly*[C], *Michael Nissen*[A,B,C], *Graeme Nimmo*[C,D] *and Theo P Sloots*[A,B,C]

[A]Queensland Paediatric Infectious Diseases Laboratory, Royal Children's Hospital, Brisbane, Qld 4029, Australia

[B]Queensland Children's Medical Research Institute, Royal Children's Hospital, The University of Queensland, Brisbane, Qld 4029, Australia

[C]Pathology Queensland Central Laboratory, Herston, Qld 4029, Australia

[D]Griffith University School of Medicine, Southport, Qld 4215, Australia

[E]Corresponding author. Tel: +61 7 3636 1623, Fax: +61 7 3636 1401, Email: d.whiley@uq.edu.au

The advent of newer technologies, including next-generation sequencing (NGS) and matrix assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF MS), have continued to drive a considerable renaissance in microbial molecular typing methods. The Sequenom Massarray iPLEX single nucleotide polymorphism (SNP) typing platform (Figure 1) combines MALDI-TOF MS with single-base extension PCR for high-throughput multiplex SNP characterisation[1]. While the method has primarily been used in the human genetics field, it is now beginning to find its way into the microbiology laboratory. In our hands, the platform has proved highly advantageous for large-scale microbial investigations, enabling multiple SNPs to be characterised for many hundreds of isolates in just over one working day, and at a fraction of the cost of conventional technologies.

In recent years, our laboratory has become increasingly involved in studies involving genotyping (typically multi-locus sequence typing; MLST) and characterisation of genetic mechanisms of antimicrobial resistance (AMR) for organisms including *Staphylococcus aureus*, *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae*. Originally we relied on classical DNA sequencing approaches that, while providing a rich data source, can be prohibitively expensive, exceeding $150 per isolate to provide both MLST and AMR data. With our expanding research interests and focus on larger scale studies (involving hundreds to thousands of isolates), assay cost has been at the forefront of our minds when selecting a suitable genotyping tool. For these reasons, we are now using the Sequenom Massarray iPLEX as our primary detection tool for large-scale studies, with each iPLEX reaction being used to characterise up to 20 SNPs for each test isolate[2,3].
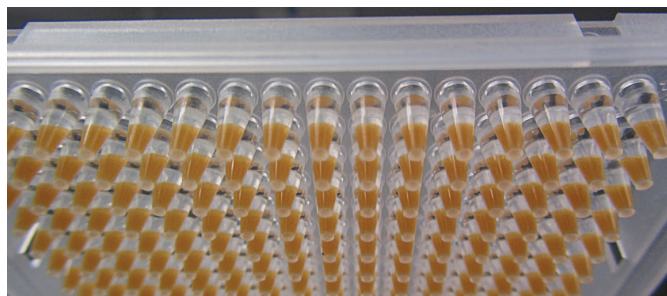
Figure 1. The Sequenom platform.

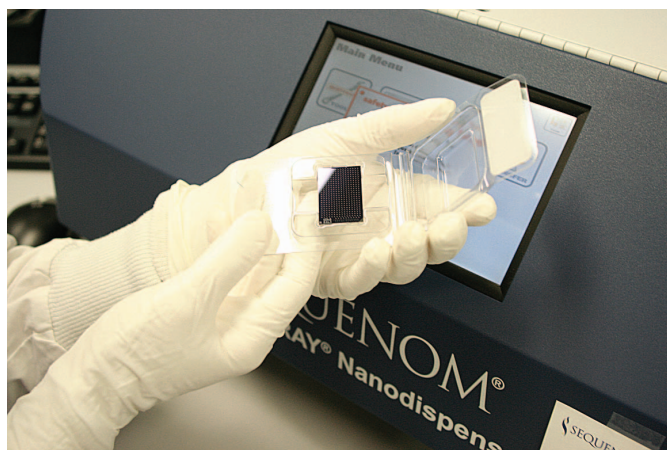Figure 2. Addition of resin to the 384-well reaction plate.

Figure 3. The SpectroCHIP.

Briefly, each iPLEX test involves four steps: (1) an initial conventional multiplex PCR reaction to amplify the genes of interest; (2) a shrimp alkaline phosphatase treatment to remove excess dNTPs; (3) an 'extension' multiplex PCR targeting the SNPs of interest and using terminator nucleotides; and (4) analysis, whereby samples are conditioned using a resin (Figure 2), dispensed onto a SpectroCHIP (Figure 2) and characterised using MALDI-TOF MS (Figure 3). The MALDI-TOF step is essentially very similar to the protein-based MALDI-TOF systems that are now increasingly being found in clinical microbiology laboratories for bacterial identification[4]. However, rather than targeting protein, the Sequenom measures the

mass of DNA, specifically that of the extension primers that are introduced at the extension PCR (step 3 above). The key design feature is ensuring mass separation of the extension primers, such that each extension primer has a unique mass irrespective of which terminator nucleotide is added to each primer during the extension PCR.

The main benefit of Sequenom iPLEX approach is cost, which is typically less than $10 per iPLEX reaction. For example, in our ongoing national study investigating AMR in *N. gonorrhoeae* we are able to genotype each isolate and identify all common AMR markers for under $20 per isolate (2 × iPLEX reactions). Additionally, the platform is ideally suited to high throughput applications; for example, 768 isolates (2 × 384 well plates) can be characterised in approximately 10 hours, with the potential for further increases in throughput with higher test numbers. An unexpected benefit was that we found the system can also be used on crude 'heat-denatured' isolates, removing the need for a costly commercial DNA extraction kit; for the above organisms we simply make a suspension of each isolate strain (1 to 2 McFarland) in sterile water, heat at 100°C for 10-minutes, vortex, briefly centrifuge and then use the supernatant for the initial PCR. Despite these benefits, there are also some limitations and other aspects of the technology that need to be considered:

(1) Overall for our studies we have found that approximately 2 to 3% of SNPs failed to be called by the platform. These problems are almost always caused by sequence variation in primer targets, poor isolate preparation or otherwise low DNA loads (further discussed below). In any event, we do not see this as a limitation affecting the broader utility of the iPLEX methods. Rather, this limitation is outweighed by the sheer throughput of the system, and that other methods, such as DNA sequencing or individual real-time PCR methods, could otherwise be applied to the small numbers of SNPs failing to type if warranted.

(2) Although our data shows that the technology is ideal for use on cultured isolates, initial data suggests it might be less suitable for use on nucleic acid extracts from non-cultured clinical samples where organism loads might be low. We recently trialled an iPLEX for typing of influenza A virus from clinical samples and found that while complete SNP profiles could be obtained at high viral load, SNP failure rates approached 50% at very low loads, e.g. samples that provided cycle threshold values of greater than 30 cycles in an influenza A real-time PCR (data not published). There are a number of potential problems that could have impacted on these results, which include: competitive inhibition between templates at low load; sequence variation in primer targets (a likely problem given the primers were targeting the highly variable hemagglutinin and neuraminidase genes); or otherwise was linked to the fact that we used a different RT-PCR mix for initial template amplification. That is, for RNA amplifications the Sequenom PCR reagents are replaced with a RT compatible PCR mix from another commercial supplier.

(3) For our studies we have used SNP-typing in place of MLST, whereby typing is based on a selection of 'informative' SNPs from the same house-keeping genes used for MLST. The main problem with this type of informative-SNP approach is that it is essentially a 'mini MLST', whereby the discriminatory power will always fall short of that of MLST based on DNA sequencing. For example, we recently developed a 20-SNP iPLEX method for

MLST-style characterisation of *P. aeruginosa*; analysis of the 1401 *P. aeruginosa* MLST types present on the MLST website (December 2012) showed that these were divided into 927 types using the 20-SNP profiles (data not published).

(4) For our AMR studies, we can only target recognised mutations and therefore might miss novel mutations that might otherwise be detected using DNA sequencing.

(5) Each iPLEX 10-hour test run uses approximately 2 hours hands-on time. While being time-efficient for large scale investigations, it would not be recommended for one-off small-scale studies involving relatively few isolates.

(6) Likewise, consideration needs to be given to the number of SNPs that are of interest. That is, small numbers of SNPs might just as easily be characterised using SYBR-green-based allele-specific PCR or melting curve analysis methods on a standard real-time PCR instrument.

(7) The instrument is not cheap when compared with a typical real-time PCR thermocycler and therefore might be beyond the reach of smaller laboratories. However, laboratories can still access the technology via numerous commercial biotech companies that now offer such services. In fact, for many of our studies, we simply conduct the initial primary conventional PCR in our own laboratory and then courier the PCR products to a company for the processing on the Sequenom.

Finally, it should be noted that genotyping technologies are constantly evolving and leading towards tools with greater discriminatory power that can be readily implemented into a routine clinical laboratory at minimal cost. Hence, there is now a considerable range of other technologies that may be available, depending on the end-use purpose. In particular, NGS is looking very promising, both in its ability to provide whole genome analyses, as well as high-through-put MLST-style data[5,6].

In summary, the Sequenom Massarray iPLEX platform provides a cost-effective, high throughput means of SNP characterisation and has provided a significant boost to our microbial characterisation work.

References

1. Leushner, J. and Chiu, N.H. (2000) Automated mass spectrometry: a revolutionary technology for clinical diagnostics. *Mol. Diagn.* **5**, 341–348.

2. Syrmis, M.W. *et al.* (2011) Comparison of a multiplexed MassARRAY system with real-time allele-specific PCR technology for genotyping of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* **17**, 1804–1810. doi:10.1111/j.1469-0691.2011.03521.x

3. Syrmis, M.W. *et al.* (2013) High-throughput single-nucleotide polymorphism-based typing of shared *Pseudomonas aeruginosa* strains in cystic fibrosis patients using the Sequenom iPLEX platform. *J. Med. Microbiol.* **62**, 734–740. doi:10.1099/jmm.0.055905-0

4. Wieser, A. *et al.* (2012) MALDI-TOF MS in microbiological diagnostics – identification of microorganisms and beyond (mini review). *Appl. Microbiol. Biotechnol.* **93**, 965–974.

5. Dunne, W.M., Jr *et al.* (2012) Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* **31**, 1719–1726. doi:10.1007/s10096-012-1641-7

6. Boers, S.A. *et al.* (2012) High-throughput multilocus sequence typing: bringing molecular typing to the next level. *PLoS ONE* **7**, e39630. doi:10.1371/journal.pone.0039630

Biographies

Associate Professor David Whiley is based at the Queensland Paediatric Infectious Diseases (QPID) Laboratory, Queensland Children's Medical Research Institute, Brisbane. His research is principally focussed on the development of novel molecular diagnostic and typing tools for infectious diseases.

Ella Trembizki is a PhD student at the QPID laboratory, Queensland Children's Medical Research Institute. Her project is aimed at investigating the emergence and spread of *Neisseria gonorrhoeae* antimicrobial resistance.

Dr Melanie Syrmis is a Senior Research Scientist at the QPID laboratory, Queensland Children's Medical Research Institute. Her research interests are molecular detection and genotyping of clinically relevant infectious agents in adult and paediatric populations.

Jenny Nakos is a scientist at the Molecular Diagnostic Unit within Pathology Queensland.

Cheryl Bletchly completed her BSc(Hons) at the University of Queensland. She spent 10 years at the Sir Albert Sakzewski Virus Research Centre as a research assistant while also completing a PhD on dengue virus. After 5 years in a public health virology laboratory she took up her current position as the Supervising Scientist of the Molecular Diagnostic Unit within Pathology Queensland.

Professor Nimmo is State Director of Microbiology for Pathology Queensland and Professor in the School of Medicine at Griffith University. His research interests include healthcare-associated and community-associated MRSA, the epidemiology of antimicrobial-resistant pathogens and molecular typing methods for healthcare-associated and community-associated pathogens.

Associate Professor Michael Nissen is the Director of Infectious Diseases and Clinical Unit Head of the Queensland Paediatric Infectious Diseases laboratory (Qpid) at the Royal Children's Hospital, Brisbane, Australia and serves as a practising Clinical Microbiologist for Pathology Queensland Central in Brisbane. He has research interests in the detection and clinical impact of antibiotic-resistant Gram-negative bacteria, as well as the epidemiology, molecular diagnosis and prevention of infectious diseases more generally, particularly respiratory pathogens.

Associate Professor Theo Sloots is the Unit Director of Research at the Queensland Paediatric Infectious Diseases (QPID) laboratory at the Queensland Children's Medical Research Institute, Brisbane. His laboratory conducts basic and applied research in infectious diseases, particularly respiratory and gastrointestinal infections, applies molecular technology to the diagnosis and characterisation of microorganisms, and investigates undiagnosed infectious disease in children.