

EDITORIAL

Molecular typing of micro-organisms: at the centre of diagnostics, genomics and pathogenesis of infectious diseases?

Introduction

Infection control relies increasingly on the recently introduced nucleic acid-mediated procedures for detection of microbial pathogens and new molecular technologies developed for microbial identification. Molecular typing or microbial DNA fingerprinting was introduced successfully during the 1990s. Both technologies contributed to the efficiency of the medical microbiology laboratory, but the same molecular technology is also enhancing our insight into the processes of microbial population genetics and pathogenesis of infectious diseases. The status of these concerted developments will be outlined from the perspective of the microbial epidemiologist. This presupposes an integrated framework where molecular typing of micro-organisms supports molecular diagnostic tests and comparative genomics generates data relevant to the study of microbial populations and host–pathogen interactions.

Current status of molecular typing of microbes

Molecular typing has irreversibly gained access to the clinical microbiology laboratory. Over the past decade, a methodology that was initially laborious, expensive and complex, has been adapted to accommodate the workflow in a high-throughput medical working environment. Molecular methods have evolved and various categories of nucleic acid-mediated approaches have been screened for clinical utility over the years [1]. This has now resulted in the general acceptance of a limited number of technologies, to date considered as ‘molecular typing standard procedures’. These methods involve PCR- and macrorestriction-mediated analysis of microbial DNA. For *Mycobacterium tuberculosis* for instance, both methods (*IS6110* probing and spoligo-typing) are well standardised, and now some are commercially available [2]. Pulsed-field gel electrophoresis (PFGE) is a broad-spectrum separation procedure for DNA macrorestriction fragments, effectively covering the requirements for typing a wide range of different microbial pathogens. The possibilities for genotyping have expanded to an extent that phenotyp-

ing procedures are considered to be largely confirmatory in nature [3].

It is interesting to note that the current typing technology that emerged from fundamental molecular biology research is finding its way back into these same laboratories. Typing procedures are essential for the definition of microbial population structures and the current emphasis on genomics in general and comparative genomics in particular requires the application of typing methodology [4]. The typing methods are used for strain comparisons, frequently as a tool in the search for determinants of genetic or functional differences. These differences are used to search for differences at the strain level, some of which can be exploited for diagnostic purposes and others that can be employed for the definition of virulence factors or other biologically interesting features [4] (Fig. 1). This latter development is extremely important, because the accessibility of large amounts of microbial sequence data allows for the immediate identification of typing differences down to the gene (expression) level. Some examples of the relevance and data-output of these approaches are highlighted below for a single bacterial species, *Campylobacter jejuni*.

Identification of new diagnostic features

The fact that typing or comparative genomics can be used for the development of diagnostic tools was demonstrated in the early 1990s for *C. jejuni* and other *Campylobacter* spp. [5]. From PCR fingerprints, species consensus fragments were identified on the basis of their identical electrophoretic behaviour (Fig. 1). Hybridisation-mediated analyses of these DNA fragments confirmed that these sub-genomic DNA molecules frequently harboured strictly species-specific sequences. Sequence elucidation revealed that one of the most reliable species-specific probes derived from a gene encoding a *C. jejuni* GTPase, based on which phylogenetic comparisons could be made to identify distinct species in even more detail [6]. Furthermore, the sequence heterogeneity thus documented among species could be used to improve diagnostic width and to develop and validate a reliable reversed hybridisa-

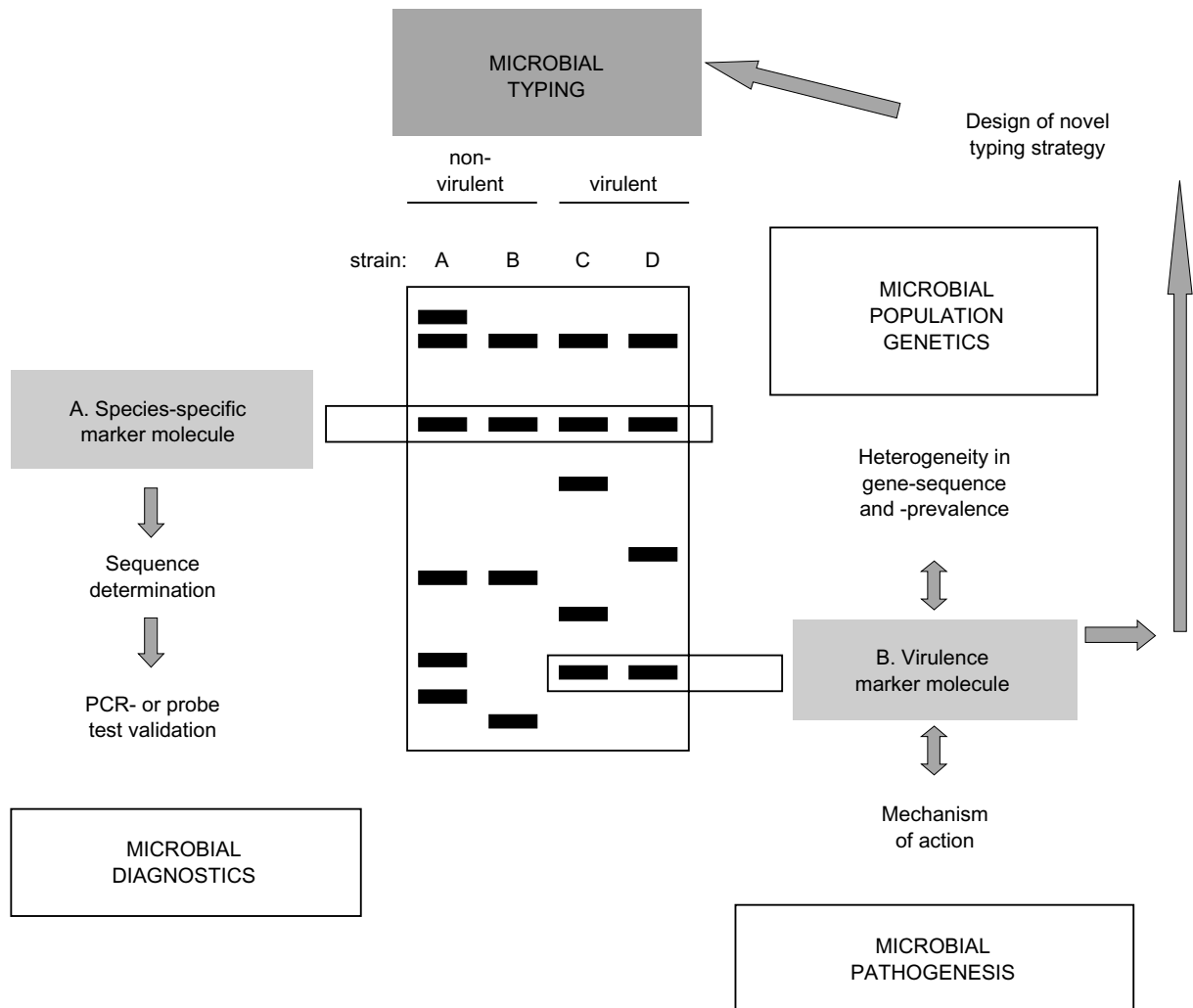


Fig. 1. Microbial typing in the centre of various microbiological research disciplines. When the fingerprint is studied, two important molecular markers can be discerned: (A) the species-specific DNA fragments and (B) differentially occurring, isolate-specific fragments. The latter type of fragments may be linked to phenotypes including virulence characteristics. Elucidation of the primary structure of the markers allows the development of novel diagnostic assays, whereas differentially presenting fragments may reveal population genetics- or pathogenesis-related features. The latter data may be translated into novel typing systems.

tion-mediated line probe assay (Innogenetics, Ghent, Belgium). This example shows that exploitation of DNA fingerprinting data can be used to develop species-specific DNA tests. Similarly, the same approaches can be used to identify DNA molecules capable of distinguishing among isolates of a single species, which can be translated in easy-to-perform microbial typing tests [7].

Identification of microbial virulence factors by comparative genomics

Comparative genomics in its simplest format compares restriction profiles of different isolates of a bacterial species. Already it was noted that these patterns varied among *C. jejuni* strains that showed evidence of antigenic shifts in serotype, thereby linking genotype and 'pathotype' [8]. Also, distinct virulence factors such as the major outer-membrane protein (MOMP) were shown to be variable among strains. This surface

protein, putatively involved in adaptation to various host environments, showed numerous mutations when genes from different strains were compared [9]. Identification by comparative genomics of distinct virulence factors is helpful for the characterisation of possible vaccine candidates, where typing identifies the antigenically stable domains of a vaccine candidate molecule.

The entire genome sequence has been elucidated for *C. jejuni* [10]. This facilitates the straightforward identification of differential presence or genetic polymorphism of those genes thought to be involved in pathogenesis. Investigations *in silico* precede the experiments to be performed in the laboratory. Several of such putatively important virulence genes of *C. jejuni* encode the bacterial lipo-oligosaccharide (LOS), in which surface-exposed polysaccharide ganglioside mimics may be present. These moieties can cause the infected host to produce antibodies that are autoreactive towards the gangliosides on peripheral nerves. This

can lead to post-infectious paralytic syndromes such as Guillain-Barré syndrome (GBS) or Miller-Fisher syndrome (MFS) [11]. Whether or not GBS and MFS are caused by a single type of *C. jejuni* is an epidemiologically important question. Recently, it was demonstrated that clonality of neuropathy-associated *C. jejuni* is not a prerequisite, not even in a relatively small geographic region [12]. This population genetics observation suggests that a unique set of genes (e.g., those involved in LOS biosynthesis) may be required for a bacterial isolate to have auto-immunogenic capacities.

PCR RFLP-mediated screening of the *C. jejuni* gene cluster involved in the biosynthesis of LOS in a large collection of GBS/MFS-associated strains led to the identification of a gene that seemed to be over-represented in the disease-related group compared with a set of control strains [13]. Based on whole genome sequence-assisted development of comparative genomics in a virulence-associated locus led to the identification of the first *C. jejuni* gene involved in polyneuropathy development. Moreover, the RFLP approach (which is straightforward gene typing) revealed a large number of putative typing targets; many of the LOS-associated genes showed significant sequence polymorphism. With the availability of multiple genome sequences for various other microbial species, more examples of this approach will emerge.

Generalisation and future developments

Many microbial whole genome sequences have been published recently and the database will continue to expand over the coming years. This provides a rich source for new typing targets. Recently, for instance, the whole genome sequence of a pair of type strains of methicillin-resistant *Staphylococcus aureus* was published [14]. Based on the straightforward 'in-silico' comparison of the genomes of these two strains, various novel virulence factor-encoding genes were discovered. These included a significant number of completely novel super-antigen genes. The current possibilities in the field of bacterial genomics will increasingly facilitate this type of study. This implies that molecular typing tools will often be based on preliminary computerised analyses of massive amounts of primary sequence data. Comparative genomics, as pursued in microbial typing, eventually will be based on genomics and will cover far greater stretches of sequence than the current procedures are capable of analysing. For *S. aureus*, the most 'dense' molecular epidemiological typing procedure, multi-locus sequence typing or MLST [15], has already provided interesting data on the population structure of this medically highly relevant species, showing that virulent clones are often the most widely dispersed ones [16]. With the help of modern nano-technology, including DNA chips, an additional impetus will be given: comparative data

will be generated in greater amounts and faster than ever [17]. As the same approaches can be used to monitor human response to infection, the same technology will also increase our knowledge of host-pathogen interplay [18].

These types of studies will link typing with important features of microbial biology, thus increasing our insight not only of the enormous diversity but at the same time identifying and highlighting bacterial, fungal and parasite characteristics that might provide distinct new policies for the continuous human battle against infectious diseases. It is proposed here that microbial typing, comparative genomics in modern speech, can and should play a central role in the future developments in diagnostic, genomic and pathogenesis-related microbiology research.

ALEX VAN BELKUM

Erasmus University Medical Center Rotterdam EMCR,
Department of Medical Microbiology & Infectious Diseases,
Dr Molewaterplein 40,
3015 GD Rotterdam,
The Netherlands
(e-mail: vanbelkum@bacl.azr.nl)

References

1. Goering RV. The molecular epidemiology of nosocomial infection: past, present and future. *Rev Med Microbiol* 2000; **11**: 145–152.
2. van Embden JDA, van Gorkom T, Kremer K, Jansen R, van der Zeijst BAM, Schouls LM. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J Bacteriol* 2000; **182**: 2393–2401.
3. van Belkum A. Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* strains: state of affairs and tomorrow's possibilities. *Microbial Drug Resist* 2000; **6**: 173–188.
4. van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. *Clin Microbiol Rev* 2001; **14**: 547–560.
5. Giesendorf BAJ, van Belkum A, Koeken A *et al.* Development of species-specific DNA probes for *Campylobacter jejuni*, *Campylobacter coli*, and *Campylobacter lari* by polymerase chain reaction fingerprinting. *J Clin Microbiol* 1993; **31**: 1541–1546.
6. van Doorn LJ, Verschuuren-van Haperen A, Burnens A *et al.* Rapid identification of thermotolerant *Campylobacter jejuni*, *Campylobacter coli*, *Campylobacter lari*, and *Campylobacter upsaliensis* from various geographic locations by a GTPase-based PCR-reverse hybridization assay. *J Clin Microbiol* 1999; **37**: 1790–1796.
7. van Leeuwen W, Libregts C, Schalk M, Veuskens J, Verbrugh H, van Belkum A. Binary typing of *Staphylococcus aureus* strains through reversed hybridization using digoxigenin-universal linkage system-labeled bacterial genomic DNA. *J Clin Microbiol* 2001; **39**: 328–331.
8. Mills SD, Kurjanczyk LA, Shames B, Hennessy JN, Penner JL. Antigenic shifts in serotype determinants of *Campylobacter coli* are accompanied by changes in the chromosomal DNA restriction endonuclease digestion pattern. *J Med Microbiol* 1991; **35**: 168–173.
9. Zhang Q, Meitzler JC, Huang S, Morishita T. Sequence polymorphism, predicted secondary structures, and surface-exposed conformational epitopes of *Campylobacter* major outer membrane protein. *Infect Immun* 2000; **68**: 5679–5689.
10. Parkhill J, Wren BW, Mungall K *et al.* The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000; **403**: 665–668.

11. Nachamkin I, Allos BM, Ho T. *Campylobacter* species and Guillain-Barré Syndrome. *Clin Microbiol Rev* 1998; **11**: 555–567.
12. Endtz HP, Ang CW, van den Braak N *et al.* Molecular characterization of *Campylobacter jejuni* from patients with Guillain-Barré and Miller Fisher syndromes. *J Clin Microbiol* 2000; **38**: 2297–2301.
13. van Belkum A, van den Braak N, Godschalk P *et al.* A *Campylobacter jejuni* gene associated with immune-mediated neuropathy. *Nat Med* 2001; **7**: 752–753.
14. Kuroda M, Ohta T, Uchiyama I *et al.* Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* 2001; **357**: 1225–1240.
15. Enright MC, Day NPJ, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol* 2000; **38**: 1008–1015.
16. Day NPJ, Moore CE, Enright MC *et al.* A link between virulence and ecological abundance in natural populations of *Staphylococcus aureus*. *Science* 2001; **292**: 114–116.
17. Young RA. Biomedical discovery with DNA arrays. *Cell* 2000; **102**: 9–15.
18. Diehn M, Relman DA. Comparing functional genomic datasets: lessons from DNA microarray analyses of host–pathogen interactions. *Curr Opin Microbiol* 2001; **4**: 95–101.