Ivano de Filippis *Editor*

Molecular Typing in Bacterial Infections, Volume II

Second Edition

Springer

Molecular Typing in Bacterial Infections, Volume II

Ivano de Filippis
Editor

Molecular Typing in Bacterial Infections, Volume II

Second Edition

Springer

Editor
Ivano de Filippis
Instituto Nacional de Controle de
Qualidade em Saude (INCQS), Fundação Oswaldo Cruz (FIOCRUZ)
Rio de Janeiro, RJ, Brazil

ISBN 978-3-030-83216-2 ISBN 978-3-030-83217-9 (eBook)
https://doi.org/10.1007/978-3-030-83217-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Part III Vector-Borne and Biosafety Level 3 Pathogens

Contributors

Roland T. Ashford Department of Bacteriology, Animal and Plant Health Agency (APHA), Addlestone, Surrey, UK

G. Balakrish Nair Rajiv Gandhi Centre for Biotechnology, Trivandrum, India

Sébastien Boutin Department of Infectious Diseases, Medical Microbiology and Hygiene, Heidelberg University Hospital, Heidelberg, Germany

Marcelo L. L. Brandão Laboratory of Microbiology Control, Department of Quality Control, Institute of Technology of Imunobiologicals from Oswaldo Cruz Foundation (Bio-Manguinhos/Fiocruz), Rio de Janeiro, RJ, Brazil

Marina C. Claros Institute of Medical Microbiology and Infectious Epidemiology, University of Leipzig, Leipzig, Germany

Zaida C. Claros Department of Cardiothoracic and Vascular Surgery, German Heart Center Berlin, Berlin, Germany

Frances M. Colles The Peter Medawar Building for Pathogen Research, Department of Zoology, University of Oxford, Oxford, Brazil

Emilyn Costa Conceição Laboratório de Bacteriologia e Bioensaios em Micobactérias, Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Laboratório de Biologia Molecular Aplicada a Micobactérias, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Georg Conrads Division of Oral Microbiology and Immunology, Department of Operative and Preventive Dentistry & Periodontology, RWTH Aachen University Hospital, Aachen, Germany

Sally J. Cutler School of Health, Sport & Bioscience, University of East London, London, UK

Maria Cristina da Silva Lourenço Laboratório de Bacteriologia e Bioensaios em Micobactérias, Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Bhabatosh Das Infection and Immunity Division, Translational Health Science and Technology Institute, Faridabad, India

Gerly Anne de Castro Brito Department of Morphology, Faculty of Medicine, Federal University of Ceará, Fortaleza, Brazil

Eliane de Oliveira Ferreira Federal University of Rio de Janeiro, Institute of Microbiology Paulo de Góes, Laboratory of Anaerobic Biology, Rio de Janeiro, Brazil

Rafae Silva Duarte Laboratório de Micobactérias, Departamento de Microbiologia Médica, Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Adriana R. Faria Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Akinyemi M. Fasemore University of Würzburg, Würzburg, Germany

Stephan Forsythe Foodmicrobe.com, Nottingham, UK

Dimitrios Frangoulidis Bundeswehr Medical Service Headquarters VI-2, Medical Intelligence & Information (MI2), Munich, Germany

Cláudia Gomes Department of Microbiology, New York University Grossman School of Medicine, New York, NY, USA

Sophie J. Hedges Department of Comparative Biomedical Science, The Royal Veterinary College, London, UK

Junyan Liu School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

Zerong Lu School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

JoAnn M. McClure Centre for Antimicrobial Resistance, Alberta Health Services, Alberta Public Laboratories, University of Calgary, Calgary, AB, Canada

Vânia Lúcia C. Merquior Departamento de Microbiologia, Imunologia e Parasitologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

Asish K. Mukhopadhyay Division of Bacteriology, National Institute of Cholera and Enteric Diseases, Kolkata, India

Ankur Mutreja Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK

Ranjan K. Nandy Division of Bacteriology, National Institute of Cholera and Enteric Diseases, Kolkata, India

Dennis Nurjadi Department of Infectious Diseases, Medical Microbiology and Hygiene, Heidelberg University Hospital, Heidelberg, Germany

Maria J. Pons Laboratorio de Microbiología Molecular y Genómica Bacteriana, Universidad Científica del Sur, Lima, Peru

Thandavarayan Ramamurthy Division of Bacteriology, National Institute of Cholera and Enteric Diseases, Kolkata, India

Joaquim Ruiz Laboratorio de Microbiología Molecular y Genómica Bacteriana, Universidad Científica del Sur, Lima, Peru

Stephanie S. R. Souza Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Departamento de Microbiologia, Imunologia e Parasitologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Philip Noel Suffys Laboratório de Biologia Molecular Aplicada a Micobactérias, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Lúcia M. Teixeira Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Mathias C. Walter Bundeswehr Institute of Microbiology, Munich, Germany

Adrian M. Whatmore Department of Bacteriology, Animal and Plant Health Agency (APHA), Addlestone, Surrey, UK

Ruirui Xu School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

Zhenbo Xu School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

Home Economics Technology, Rajamangala University of Technology Phra Nakhon, Bangkok, Thailand

Research Institute for Food Nutrition and Human Health, Guangzhou, China

Department of Laboratory Medicine, the Second Affiliated Hospital of Shantou University Medical Collegey, Shantou, China

Research Center of Translational Medicine, Second Affiliated Hospital of Shantou University Medical College, Shantou, China

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD, USA

Sareh Said Yekta-Michael Department of Operative and Preventive Dentistry & Periodontology, RWTH Aachen University Hospital, Aachen, Germany

Guangchao Yu Center of Clinical Laboratory Medicin, First Affiliated Hospital of Jinan University, Guangzhou, China

Kunyan Zhang Centre for Antimicrobial Resistance, Alberta Health Services, Alberta Public Laboratories, University of Calgary, Calgary, AB, Canada

Part I

Gastrointestinal Pathogens

Campylobacter

1

Sophie J. Hedges and Frances M. Colles

1.1 Introduction

Campylobacter is a motile Gram-negative, curved rod-shaped bacterium and is one of the most common causes of human gastroenteritis (campylobacteriosis) globally [1]. *Campylobacter* has only been recognized as causing human disease relatively recently, aided by the development of specific and selective agar during the 1970s [2]. The bacterium also needs a microaerobic atmosphere containing 5% O_2, 10% CO_2 and 85% N_2 to grow [3]. Despite this, it is believed the pathogen was first described in infant stools as a cause of 'cholera infantum' as early as 1880 in Germany by Theodor Escherich [4]. *Campylobacter* normally causes a self-limiting infection in humans, with antimicrobial therapy not usually administered unless more severe symptoms persist. Occasionally, serious complications such as Guillain-Barré syndrome (GBS) can occur, which in addition to the high numbers of human infections, and the prominence of severe diarrhoea in developing countries [5], means that *Campylobacter* is viewed as an important public health hazard by the WHO and in many countries [6].

There are 40 species and subspecies of *Campylobacter* recognized to date [7, 8], of which *C. jejuni* accounts for 90% of campylobacteriosis in humans, with *C. coli* and occasionally *C. lari* and *C. upsaliensis* causing the majority of the remaining cases [9]. There is increasing evidence that *C. concisus* has a role in inflammatory bowel disease (IBD) in humans [8]. In animals, *Campylobacter* species are often identified as part of the microbiota of healthy individuals; however, *C. jejuni*, *C. coli*,

S. J. Hedges
Department of Comparative Biomedical Sciences, The Royal Veterinary College, London, UK

F. M. Colles (✉)
The Peter Medawar Building for Pathogen Research, Department of Zoology, University of Oxford, Oxford, UK
e-mail: frances.colles@zoo.ox.ac.uk

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_1

C. fetus subspecies *fetus* and *C. fetus* subspecies *venerealis* are recognized as veterinary pathogens, together with sporadic incidence of other *Campylobacter* species [10]. In the UK and New Zealand, *C. fetus* subspecies *fetus* is one of the leading causes of abortion outbreaks in sheep, cattle and goats, although in the USA, there has been a switch to a particular clone of *C. jejuni* as the primary cause over the past decade [10, 11]. *C. fetus* subspecies *venerealis*, also a cause of infertility and abortion in ruminants, is more closely associated with cattle. *Campylobacter hepaticus* was named as a new species in 2016 and is recognized as a cause of spotty liver disease in laying hens [12, 13].

A subset of *Campylobacter* species, including *C. jejuni*, *C. coli*, *C. upsaliensis* and *C. lari*, are thermophilic with an optimal growing temperature of 42 °C, the body temperature of birds. The thermophilic bacteria can be isolated from a range of sources including humans, chickens [14], ruminants [15], pigs [16] and wild birds [17–19] and also from environment sources such as water and sand from bathing beaches [20–24]. In clinical samples, *C. jejuni* and *C. coli* are most commonly isolated from patient stool samples, but in addition to other species such as *C. fetus* and *C. hyointestinalis* [25, 26], they are occasionally isolated from blood cultures, particularly in immunocompromised patients [10]. Most *Campylobacter* infections are sporadic [3, 27], with outbreaks occasionally occurring in association with sources such as contaminated or untreated milk or water [3].

Typing of *Campylobacter* has traditionally been used to differentiate the species, as well as to cluster strains within the same species. Typing can also be used to attribute the primary source of infection, monitor outbreaks and associate relatedness of different strains to understand the pathogen's epidemiology and history. Within clinical laboratories, identification is often limited to genus level as treatment, if deemed necessary, is not affected by species. In recent years, thanks to improved nucleotide sequencing technology, more can be done to investigate the largely sporadic nature of human disease and the increase in antibiotic resistance [28]. Fluoroquinolone resistance is a particular problem, primarily due to the prescription of fluoroquinolones in clinical cases, as well as in animal agriculture [29]. Typing methods can be used to trace antibiotic resistance in the food chain and prove valuable in predicting antibiotic resistance in clinical disease [30]. In this chapter, we describe different phenotypic and genotypic methods that have been developed over the years in the typing of *Campylobacter* and discuss how they have been advanced to give scientists a greater understanding of *Campylobacter* and the role it plays in human disease.

1.2 Early Typing Methods

Early techniques before the age of nucleotide sequencing relied heavily upon visual cues and simple tests to differentiate between bacterial species. It was not until later on that different strains of *Campylobacter* could be reliably identified through PCR and sequencing techniques. Typing methods of *Campylobacter* may require culturing of the isolates, which, in itself, can identify between different species. The need for quick, cheap and accurate methods for the identification of *Campylobacter* spp.

within clinical laboratories is a necessary importance; however, it is also critical in assessing biosecurity of farms and abattoirs, for example. As *C. jejuni* and *C. coli* have been identified as the primary species causing human disease, typing methods have largely focused on these species.

1.3 Biochemical Tests

Detection of *Campylobacter* within samples containing low *Campylobacter* numbers and high levels of background microflora often relies on pre-enrichment followed by selective culture. Early typing methods combined growth characteristics with simple biochemical tests to identify different bacterial genera and *Campylobacter* species [31]. For example, *Campylobacter* can be differentiated from phenotypically similar *Arcobacter* species by their inability to grow at 30 °C [32] and from other bacterial genera by their sensitivity to NaCl at greater than 2%w/v [3]. Indoxyl acetate hydrolysis (IAH) tests, reliant on a colour change to dark blue within 5–10 min, have been developed for the differentiation of *Campylobacter* and the related *Helicobacter* and *Wolinella* genera [33]. *Campylobacter* are Gram-negative bacteria that are oxidase and catalase positive [34].

C. jejuni and *C. coli* can traditionally be separated by the hippurate hydrolysis test, which tests for the presence of the glycine, one of the products of sodium hippurate hydrolysis catalysed by the hippurate hydrolysis enzyme [35]. *C. jejuni* is hippurate positive, causing the ninhydrin test indicator to form a deep purple colour, whilst *C. coli* is hippurate negative, and the test indicator remains clear in colour. This test, however, is subjective and has relatively high levels of false-positive and false-negative results [36], in part due to the emergence of atypical catalase-negative and hippurate-negative *C. jejuni* [37]. IAH tests are a quick way to differentiate IAH-negative *C. lari* and *C. fetus* subsp. *fetus* from IAH-positive *C. coli* and *C. jejuni* [33, 38]. Resistance to nalidixic acid and cephalothin has been used to differentiate between some *Campylobacter* species, e.g. *C. jejuni/C. coli*, *C. fetus*, *C. lari* and *C. upsaliensis* [39]. The commercially available API *Campylobacter* system (bioMérieux) consists of 11 enzyme and conventional biochemical tests and 9 assimilation and inhibition mini-strip tests that can also be used to differentiate between *Campylobacter* species [40]. More recently, the Vitek®2 microbial identification system (bioMérieux), which uses colorimetric technology and increased automation, was developed to identify *C. jejuni*, *C. coli* and *C. fetus* using the NH ID card. In a study of 1906 *Campylobacter* organisms, the API *Campylobacter* system was able to identify 94.4% of *C. jejuni* strains and 73.8% of *C. coli* strains, and the Vitek®2 system identified 89.6% of *C. jejuni* and 87.7% of *C. coli* strains. Both systems, however, showed poor sensitivity in the identification of other *Campylobacter* species and related *Epsilobacteria*, and confirmatory tests are recommended [41]. Further biochemical tests are rather limited for *Campylobacter* species since they are relatively biochemically inert [33]; those that are used can be subjective in their interpretation, and they do not allow for detailed epidemiological analysis.

1.4 Phenotypic Methods

Serotyping and phage typing methods are based on phenotypic properties of bacteria, for example, proteins expressed by the organism as surface structures [42]. Serotyping was developed in the 1980s [43] and divides *Campylobacter* into groups (serotypes) based on the response with antiserum [44, 45]. Two serotyping schemes were developed for *Campylobacter*: the Penner scheme based on passive haemagglutination of heat-stable antigens and the Lior scheme based on direct bacterial agglutination of heat-labile antigens [45]. A combination of both schemes was found to give best discrimination, but it is still not known exactly which antigens are detected [46]. Phage typing for *Campylobacter* was developed in the 1980s and 1990s; one study using 754 *Campylobacter* isolates from mixed sources identified 46 different phage types, using 19 typing phages [47]. Both the serotyping and phage typing methods were found to have low discriminatory power and poor reproducibility and were limited by the availability of antisera which are costly to produce. Since *Campylobacter* can undergo genetic recombination, it has been found that strains belonging to the same serotype are not necessarily of the same genotype [46].

Genotyping methods are more commonly used than phenotyping methods for *Campylobacter* these days, due to the advantages of increased accuracy, with fewer strains being untypable.

1.5 Enzyme-Based Methods

After the development of phenotypic methods for *Campylobacter* speciation and subtyping, methods using restriction enzymes and analysis of gel electrophoretic patterns shortly followed. Although widely used, one of the main drawbacks to these methods is the poor reproducibility and lack of portability of data between laboratories.

1.6 Pulsed-Field Gel Electrophoresis

Rapid pulsed-field gel electrophoresis (PFGE) can be used in the subtyping of *C. jejuni* and *C. coli* [48] and is often used as a reference typing method as the entire genome is analysed to create the restriction profiles [49]. The whole bacterial chromosome is digested with a rare-cutting restriction enzyme to yield a moderate number of DNA fragments which are separated using special electrophoresis conditions and stained, and the differences in the number and size of the bands are used in comparison analysis. PFGE is labour intensive and takes a long time to conduct. Not only this, but the resultant bands give way for more subjective results than sequence-based methods. Although widely used for tracing the source of infection, the genetic basis for the differing PFGE profiles is generally poorly understood, and changes in multiple bands may not necessarily be independent of each other [50]. Due to the

limited resolution of PFGE methods, the gel profiles could overestimate the relationship between the different *Campylobacter* isolates, or indeed, a clonal relationship could exist between strains with different profiles [51–53].

1.7 Restriction Fragment Length Polymorphism of the *flaA* Gene (*fla*A-RFLP)

The flagellin subunit of *Campylobacter* is encoded by two highly homologous genes *flaA* and *flaB*, arranged in tandem [46, 54]. Having both conserved and variable regions of nucleotide sequence, they are suitable targets for restriction fragment length polymorphism (RFLP) typing. Similar to PFGE, this method has most commonly been used for the *flaA* gene and amplifies and restricts a region of the gene found to vary between strains. A fingerprint profile is created, and the level of discrimination between strains is dependent upon which restriction enzyme was used. The method has been used for a number of *Campylobacter* species [46] and has advantages in its low complexity and low operating cost [55]. One major downfall of this typing method, however, is the recombination that occurs within and between the two flagellin loci which causes limitations for global or long-term epidemiological studies [54]. The method is not reliable for the speciation of *C. jejuni* and *C. coli* which have been found to share the same *flaA* types using genomic studies [56].

1.8 Commercial ELISA Kits

Several enzyme immunoassays (EIAs) are also available for the detection of *C. jejuni* and *C. coli* in clinical samples. A recent study by Granato and colleagues compared three commercially available kits with culture-based techniques and found that the three immunoassays had sensitivities that ranged from 98.5 to 99.3% and specificities that ranged from 98.0 to 98.2%, whilst standard culture had a sensitivity of 94.1% [57]. Whilst these methods can usefully avoid the need for pathogen-specific culture, Bojanić and colleagues [58] note that cross-reaction between closely related species has not been extensively tested and that sensitivity can vary with respect to species. In addition, results can be influenced by the consistency of the faecal sample used.

1.9 Molecular Methods

1.9.1 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry

This method uses mass spectrometry to analyse whole bacterial cells for intact proteins and differentiate between different species of *Campylobacter* using 'species-identifying' biomarker ions (SIBIs) [59]. The resulting spectra can be compared

with the MALDI Biotyper database (Bruker) in order to identify the bacterial species, and further discriminatory tests can be performed to confirm the results [41]. Advantages to the method include speed (results can be obtained within an hour), low running costs, minimal growth requirements and high levels of accuracy reported up to 100% of *Campylobacter* isolates tested [41]. In addition, the MALDI-TOF method allows mixed cultures containing multiple species of *Campylobacter* to be detected more easily than previously described methods [59]. The main disadvantages to the method are lack of antibiotic resistance typing [41] and inability to perform fine typing.

1.9.2 PCR-Based Methods

The use of PCR- and qPCR-based methods with specific primers allows for a rapid and more reliable differentiation of *Campylobacter* species compared to observations by eye [60]. The technique is considered to be more accurate than detection by culture which is the gold standard, although it is worth noting that PCR methods allow for detection of bacteria which may be dead or non-viable for culture purposes [61]. The main advantages of PCR and qPCR as a tool for bacterial typing are that they provide a quick and high-throughput detection of multiple genes [62] and different bacterial species simultaneously, with qPCR additionally giving information on quantification. Specific binding of primers to target DNA means that detection and identification are potentially highly accurate; however, nonspecific binding to DNA can occur, and methods need to be optimized [60]. Care must be taken to ensure that DNA extracted from sample material does not contain PCR inhibitors, for example, the high urea content in faecal samples from birds can prove problematic, and a number of different approaches may need to be tested. Accuracy of detection can be further increased by using nested primers or by using qPCR with a probe. A number of different gene targets have been used for the detection and speciation of *Campylobacter*, and new systems continue to be developed and evaluated depending on identification requirements, *Campylobacter* species and sample type. Generic 16S primers have been used for the detection of *Campylobacter* species in a range of samples; other gene targets used to differentiate between *C. jejuni* and *C. coli* include *hipO*, *glyA*, *mapA* and *ceuE* [61, 63].

1.9.3 Sequence-Based Methods

1.9.3.1 MLST

Multi-locus sequence typing (MLST) has been used for the characterization of many different bacteria including *C. jejuni* and *C. coli* [64]. The *C. jejuni* and *C. coli* MLST scheme targets fragments of seven housekeeping genes, each approximately 500 bp in length and originally designed to be compatible with Sanger sequencing. For each of the loci, every unique nucleotide sequence is assigned an allele number using the PubMLST *Campylobacter* database (https://pubmlst.org/campylobacter)

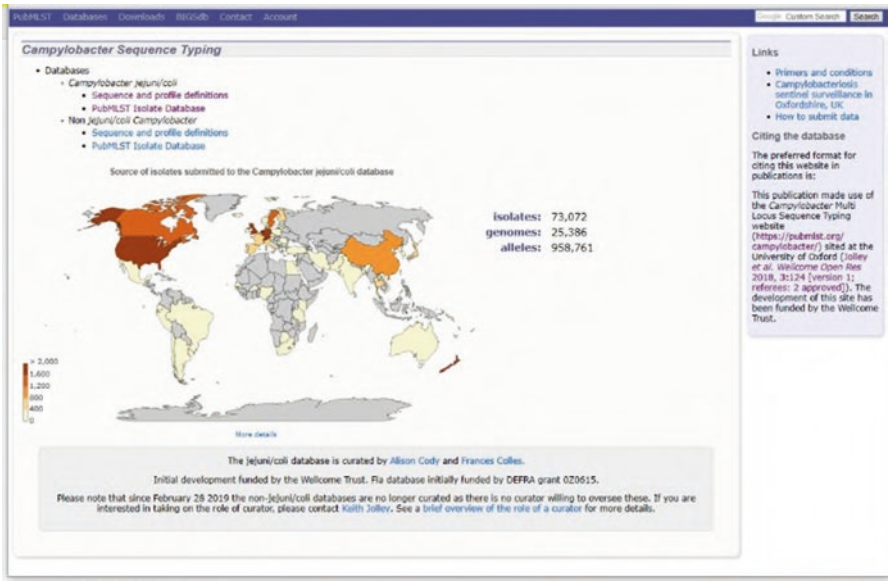

Fig. 1.1 The homepage of the PubMLST *Campylobacter* database (https://pubmlst.org/campylobacter/) giving links to sequence data and isolate metadata for both *C. jejuni/coli* and non-*C. jejuni/coli* typing schemes. The distribution of *C. jejuni/coli* isolates submitted to the database by country is shown, together with the current number of isolates, genomes and alleles that have been defined. If possible I would like to swap this image for an updated screen shot of the database.

(Fig. 1.1), therefore reducing a long sequence down into seven-digit 'allelic profile' which is then assigned to a sequence type (ST). These STs can then be clustered into closely related groups termed 'clonal complexes' (CCs) named after the central genotype [65]. MLST schemes exist for different *Campylobacter* species [64, 66], and additional antigen typing loci such as *flaA*, *flaB* and *porA* that are more variable than housekeeping genes have been used for fine typing, for example, in the investigation of diffuse outbreaks [67]. These original schemes still remain compatible with next-generation sequencing technology and whole-genome sequencing data today, and seven-locus MLST can readily be performed 'in silico'.

By sequencing isolates in this way, certain CCs have been seen to be associated with different host sources. For example, ST-21 and ST-45 CC are widespread, multi-host bacterial strains, whereas ST-61 CC is strongly associated with ruminant, ST-257 CC with chicken and ST-682 CC with wild bird sources [68] (Fig. 1.2). Using this information, the source of human infection can be attributed allowing for estimation of how *Campylobacter* survives through the food chain [69]. The results of MLST are highly discriminative, and the method can not only differentiate between different *Campylobacter* species but also identify individual strain types. Since the method is reliant on the DNA extraction from pure cultures, care must be taken to ensure that this is the case. Mixed cultures are most readily identified by mixed or indeterminate peaks on chromatograms produced by Sanger sequencing instruments or by unusually large, poor-quality genomes with multiple allele

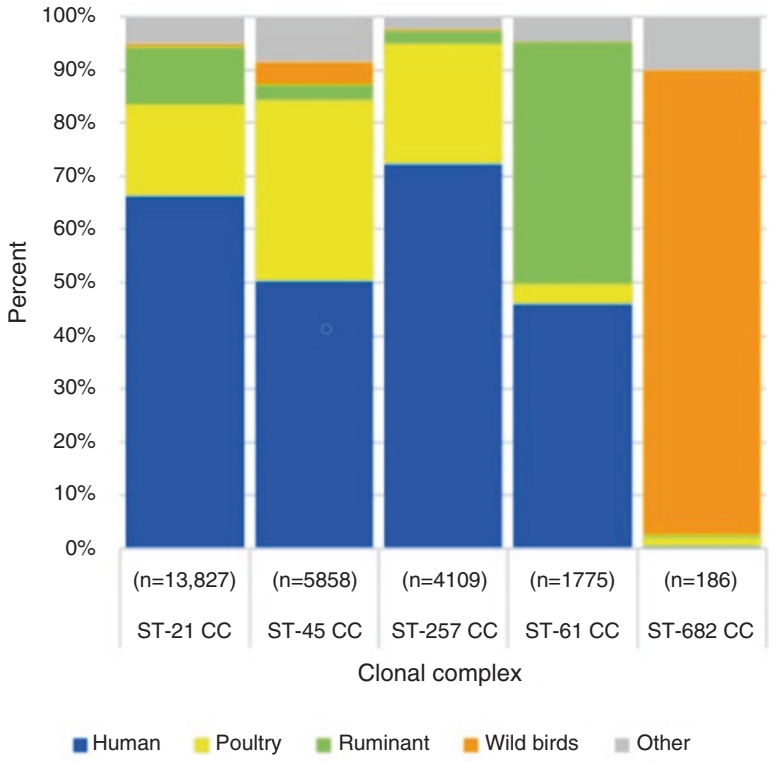

Fig. 1.2 The source distribution of *Campylobacter* isolates on the PubMLST *Campylobacter* database), shown for selected clonal complexes. Studies have found ST-21 and ST-45 complexes to have broad distribution, ST-257 complex to be poultry associated, ST-61 complex to be ruminant associated and ST-682 complex to be associated with wild birds

assignments using next-generation sequencing technology. However, alleles can be exchanged between *C. jejuni* and *C. coli* [70], and indeed, the very commonly isolated ST-61 consists of six alleles from *C. jejuni* and the uncA 17 allele from *C. coli.*

Both sequence data and metadata for *Campylobacter* isolates can be uploaded to a central, online and curated database (PubMLST https://pubmlst.org/campylobacter/ [71]) which has a global user base. There are more than 100,000 *Campylobacter* spp. uploaded to the *Campylobacter* PubMLST website (as of September 2021) which can be filtered by attributes such as species, source or country as well as by ST, CC or other combinations of sequence data. From this database, nucleotide sequence data can be downloaded in the correct format for further analysis, e.g. Interactive Tree of Life (iTOL) [72], MEGA [73], GrapeTree [74] and other tree drawing programmes which are easily performed by using tabs on the database.

1.9.3.2 rMLST

Extending the original gene-by-gene MLST approach, ribosomal MLST (rMLST) was designed as a potential means by which the whole bacterial domain could be typed, using 53 ribosomal protein subunit (rps) genes contributing to a function that is essential to all. Using the rMLST database (https://pubmlst.org/rmlst), new sequence data from an rps gene can be compared against an ever increasing numbers of genomes, more than 400,000 genomes, of which 43,500 are *Campylobacter* at the time of writing, in order to obtain a species identification. It should be noted that *Campylobacter* do not contain the rpmD gene, nor do they contain paralogous rps loci, and that identification is based on 52 of the 53 rps genes. Although species identification is comparable and often better than seven-locus MLST, rMLST does not give sufficient resolution for identifying genetic diversity within outbreaks [75]. An example of *Campylobacter* speciation based upon rMLST and visualized using a minimum spanning tree produced using the GrapeTree plug-in is shown in Fig. 1.3.

1.9.3.3 cgMLST

Core genome MLST (cgMLST) identifies loci which are present within most members of a population, balancing the high resolution needed for typing with comparability across a large number of strains [75]. Therefore, the greater the genetic

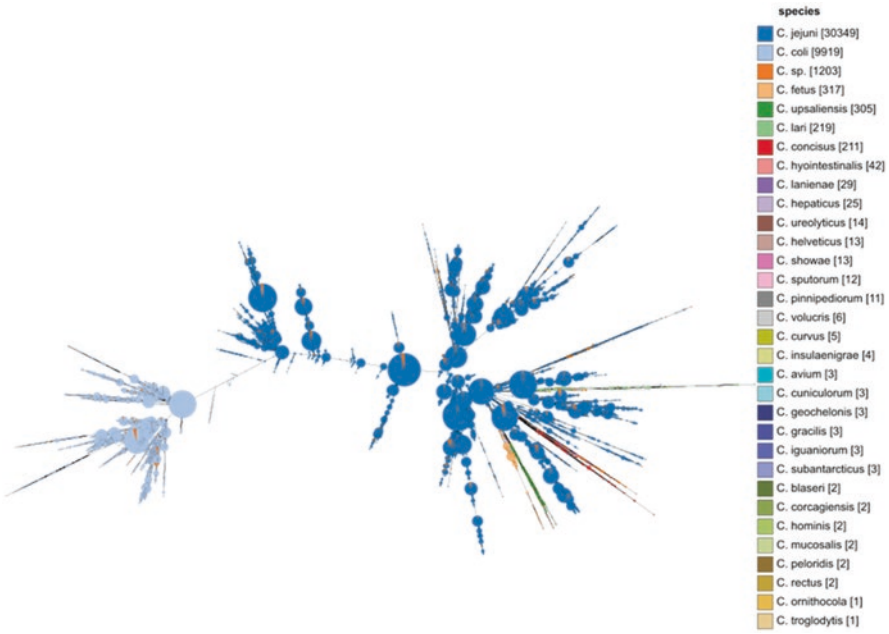

Fig. 1.3 Minimum spanning tree showing the relationships between *Campylobacter* species using ribosomal MLST (rMLST). Those isolates submitted to the PubMLST database as *Campylobacter* sp. (shown in orange) have been assigned to species clusters based upon rMLST data, and the method can be used as a tool for species identification. The tree was produced using the GrapeTree (Zhou et al. 2018) plug-in on the rMLST database (https://pubmlst.org/rmlst)

diversity within a selected population, the smaller the size of the core genome. The core genome version 1.0 scheme for *C. jejuni* and *C. coli* incorporated into the PubMLST database is based upon a representative collection of 2472 UK isolates from clinical disease and animal and environmental sources, of which 89.3% were *C. jejuni* and 10.7% *C. coli* [75]. Genes were identified according to the reannotation of *C. jejuni* NCTC 11168, and a subset of 1343 loci (excluding paralogous loci) from a total of 1643 were selected for the core genome scheme, being present in more than 95% of the isolates tested. The scheme was validated using a further 1478 isolates and proved efficient in the identification and resolution of closely related isolates from an outbreak situation. It should be noted that this core genome scheme has been developed with clinical *C. jejuni* and *C. coli* isolates in mind, and whilst it was able to identify *C. coli* from the three different clades based upon MLST, for example [70], further specific core genome schemes could be designed for more detailed investigation. Isolates can be assigned a cgMLST sequence type using the PubMLST *Campylobacter* database, and the threshold at which isolates differ from each other within their core genome can be set (e.g. 5, 10, and 25 different loci) in order to look for closely related clusters. However, it is down to the user to define how closely related individuals are, which therefore leads to someone analysing a dataset with a 25 loci difference having a much greater margin for inaccuracy than someone using the same dataset with a 5 loci difference. This raises similar issues to SNP analysis, in defining where the 'actual' accuracy for cg analysis lies. It should be remembered that by using a 'core' MLST scheme, there is still some information being lost from WGS data.

1.9.3.4 wgMLST

Despite the multilayered approach to using MLST as typing schemes, there can still be a requirement for greater resolution of multi-locus methods to identify bacterial species, for strain typing or for more detailed investigation of antibiotic resistance mechanisms, for example. Within recent years, low-cost and highly parallel next-generation sequencing have facilitated whole-genome sequencing (WGS) of a range of bacterial species including *Campylobacter*. By increasing the number of loci sequenced, whole-genome MLST (wgMLST) [76] using the full complement of annotated loci for a given isolate can be used for single-clone investigation. Increasing the number of loci used in genomic comparisons between isolates allows for a higher level of discrimination compared to cgMLST schemes which use only loci that are core to a high percentage of isolates. Using the PubMLST *Campylobacter* database, any number of schemes can be created to include core or accessory genome loci and allow gene-by-gene comparison of different isolates, drilling down to single nucleotide polymorphisms (SNPs) if required. The gene-by-gene approach is both backwards and forwards compatible: 16S rRNA and seven-locus MLST information can be extracted from WGS data and compared to data derived from Sanger methods [77].

1.9.4 Parallel Sequencing Methods

Parallel or 'deep' sequencing refers to increasing the number of times each region of DNA is sequenced to reduce the number of sequencing errors. By sequencing isolates multiple times, SNPs can be separated from errors, and rare sequence types can be detected. Based upon techniques used to study the microbiome, a deep sequencing method has recently been developed for *C. jejuni* and *C. coli*, in order to screen for the presence of multiple strain types within complex samples [78]. The method is culture independent and so removes the potential bias for preferential culture of particular strains in the laboratory. Briefly, DNA is extracted from samples and PCR used to amplify a short section of *Campylobacter*-specific nucleotide sequence – in this case a 405–473 bp fragment '*porAf2*' of the *porA* loci, which is then deep sequenced in order to look for diversity. As with other sequencing methods, individual nucleotide sequences can be assigned an allele number using the PubMLST *Campylobacter* database. The *porA* locus was selected based upon its historical use for fine typing in association with MLST, but the method could similarly be extended for use with other targets of interest. The parallel sequencing technique provides a way of testing for multiple strains of *Campylobacter* amongst a large number of samples in a cost-effective way and allows for the detection of rare sequence types which may be present at a thousandth frequency of other types. It would be practically impossible to identify such rare or poorly culturable *Campylobacter* strains using conventional methods. Although it can be assumed that different *porAf2* alleles represent different *Campylobacter* strains, additional sequence data from other loci are required to determine whether strains with the same *porAf2* are truly the same or different. Thus, the parallel sequencing method is envisaged to be a tool to determine the depth of strain diversity within a sample that is used alongside techniques such as cgMLST (requiring culture of isolates) that provide high levels of discrimination between strains.

1.10 Concluding Remarks

This chapter looks at the most commonly applied methods for typing *Campylobacter* spp. and how they have progressed since the organism's discovery. Each of the methods described has its own strengths and weaknesses, and all would serve useful in different scientific settings.

With constant advances in sequencing technologies, there is an exponential increase in the amount of genomic data surrounding the study of *Campylobacter* throughout the food chain, from farms to retail samples to human clinical samples. Because of this increase in data, there is a requirement of accurately curated online databases to store both sequencing and isolate metadata in a user-friendly and appropriate way, ensuring that all of this information can be freely exchanged between research groups.

Unfortunately, the typing of *Campylobacter* is often overlooked in clinical research as, traditionally, methods have been difficult and expensive to perform,

with no current influence over the treatment of human cases. However, with the increase in antibiotic resistance within certain strains, the use of *Campylobacter* typing would greatly assist in understanding the global acquisition and distribution of resistance. Similarly, more accurate typing and widely sampled isolate collections will greatly improve source attribution modelling of human infection and hopefully lead to more effective intervention strategies.

1.11 Summary

Campylobacter is a prolific human and veterinary pathogen, but its importance has only been recognized in relatively recent times, thanks to improved culture techniques that accommodate its complex requirements. This chapter describes early methods of distinguishing some of the species and subspecies of *Campylobacter* which currently number at least 40, as well as present-day advances, including molecular and nucleotide sequence-based typing methods. Accurate typing of *Campylobacter*, and more specifically clusters of strains, is of particular importance for tracing the epidemiology of sporadic infection and for investigating the reasons behind a global increase in antibiotic resistance.

References

1. Friedman CJ, Neiman J, Wegener HC, Tauxe RV (2000) Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialised nations. Campylobacter. ASM Press, pp 121–38
2. Skirrow M (1977) *Campylobacter* enteritis: a "new" disease. Br Med J 2(6078):9–11
3. Silva J, Leite D, Fernandes M, Mena C, Gibbs PA, Teixeira P (2011) *Campylobacter s*pp. as a foodborne pathogen: a review. Front Microbiol 2:200
4. Moore J, Matsuda M (2002) The history of *Campylobacter*: taxonomy and nomenclature. Ir Vet J 55(10):495–501
5. Platts-Mills JA, Kosek M (2014) Update on the burden of *Campylobacter* in developing countries. Curr Opin Infect Dis 27(5):444–450
6. Devleesschauwer B, Bouwknegt M, Mangen M-JJ, Havelaar AH (2017) Health and economic burden of *Campylobacter*. In: Campylobacter. Elsevier, pp 27–40
7. Man SM (2011) The clinical importance of emerging *Campylobacter* species. Nat Rev Gastroenterol Hepatol 8(12):669
8. Liu F, Ma R, Wang Y, Zhang L (2018) The clinical importance of *Campylobacter concisus* and other human hosted *Campylobacter* species. Front Cell Infect Microbiol 8:243
9. Cody AJ, McCarthy NM, Wimalarathna HL, Colles FM, Clark L, Bowler IC et al (2012) A longitudinal 6-year study of the molecular epidemiology of clinical *Campylobacter* isolates in Oxfordshire, United Kingdom. J Clin Microbiol 50(10):3193–3201
10. Sahin O, Yaeger M, Wu Z, Zhang Q (2017) *Campylobacter*-associated diseases in animals. Annu Rev Anim Biosci 5:21–42
11. Van Bergen MA, Dingle KE, Maiden MC, Newell DG, van der Graaf-Van Bloois L, van Putten JP et al (2005) Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J Clin Microbiol 43(12):5888–5898
12. Crawshaw T (2019) A review of the novel thermophilic *Campylobacter*, *Campylobacter hepaticus*, a pathogen of poultry. Transbound Emerg Dis

13. Van TTH, Gor M-C, Anwar A, Scott PC, Moore RJ (2017) *Campylobacter hepaticus*, the cause of spotty liver disease in chickens, is present throughout the small intestine and caeca of infected birds. Vet Microbiol 207:226–230
14. Skarp C, Hänninen M-L, Rautelin H (2016) Campylobacteriosis: the role of poultry meat. Clin Microbiol Infect 22(2):103–109
15. Stanley K, Jones K (2003) Cattle and sheep farms as reservoirs of C*ampylobacter*. J Appl Microbiol 94:104–113
16. Shange N, Gouws P, Hoffman LC (2019) *Campylobacter* and *Arcobacter* species in food-producing animals: prevalence at primary production and during slaughter. World J Microbiol Biotechnol 35(9):146
17. Colles FM, Ali JS, Sheppard SK, McCarthy ND, Maiden MC (2011) *Campylobacter* populations in wild and domesticated mallard ducks (*Anas platyrhynchos*). Environ Microbiol Rep 3(5):574–580
18. Griekspoor P, Colles FM, McCarthy ND, Hansbro PM, Ashhurst-Smith C, Olsen B et al (2013) Marked host specificity and lack of phylogeographic population structure of *Campylobacter jejuni* in wild birds. Mol Ecol 22(5):1463–1472
19. Cody AJ, McCarthy ND, Bray JE, Wimalarathna HM, Colles FM, Jansen van Rensburg MJ et al (2015) Wild bird-associated *Campylobacter jejuni* isolates are a consistent source of human disease, in Oxfordshire, United Kingdom. Environ Microbiol Rep 7(5):782–788
20. Khan IU, Hill S, Nowak E, Edge TA (2013) Effect of incubation temperature on the detection of thermophilic *Campylobacter* species from freshwater beaches, nearby wastewater effluents, and bird fecal droppings. Appl Environ Microbiol 79(24):7639–7645
21. Bolton F, Surman S, Martin K, Wareing D, Humphrey T (1999) Presence of *Campylobacter* and *Salmonella* in sand from bathing beaches. Epidemiol Infect 122(1):7–13
22. Colles F, Jones K, Harding R, Maiden M (2003) Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. Appl Environ Microbiol 69(12):7409–7413
23. Pitkänen T (2013) Review of *Campylobacter* spp. in drinking and environmental waters. J Microbiol Methods 95(1):39–47
24. Jones K, Obiri-Danso K (1998) Non-compliance of beaches with the EU directives of bathing water quality: evidence of non-point sources of pollution in Morecambe Bay. J Appl Microbiol 85(S1):101S–107S
25. Edmonds P, Patton C, Griffin P, Barrett T, Schmid G, Baker C et al (1987) *Campylobacter hyointestinalis* associated with human gastrointestinal disease in the United States. J Clin Microbiol 25(4):685–691
26. Patrick ME, Gilbert MJ, Blaser MJ, Tauxe RV, Wagenaar JA, Fitzgerald C (2013) Human infections with new subspecies of *Campylobacter fetus*. Emerg Infect Dis 19(10):1678
27. ECDC (2013) ECDC publishes the annual epidemiological report 2012. Eurosurveillance 18(10):20418
28. Sproston EL, Wimalarathna HM, Sheppard SK (2018) Trends in fluoroquinolone resistance in *Campylobacter*. Microb Genom 4(8)
29. McCrackin M, Helke KL, Galloway AM, Poole AZ, Salgado CD, Marriott BP (2016) Effect of antimicrobial use in agricultural animals on drug-resistant foodborne campylobacteriosis in humans: a systematic literature review. Crit Rev Food Sci Nutr 56(13):2115–2132
30. Duarte A, Seliwiorstow T, Miller WG, De Zutter L, Uyttendaele M, Dierick K et al (2016) Discriminative power of *Campylobacter* phenotypic and genotypic typing methods. J Microbiol Methods 125:33–39
31. Acuff G (1992) Media, reagents, and stains. In: Compendium of methods for the microbiological examination of foods, vol 3, pp 1093–1208
32. Bhunia AK (2008) *Campylobacter* and *Arcobacter*. In: Foodborne microbial pathogens: mechanisms and pathogenesis, pp 217–226
33. Popovic-Uroic T, Patton CM, Nicholson MA, Kiehlbauch JA (1990) Evaluation of the indoxyl acetate hydrolysis test for rapid differentiation of *Campylobacter*, *Helicobacter*, and *Wolinella* species. J Clin Microbiol 28(10):2335–2339

34. Bolton F, Wareing D, Skirrow M, Hutchinson D (1992) Identification and biotyping of Campylobacters. In: Identification methods in applied and environmental microbiology. Blackwell Scientific Publications, Oxford, pp 151–161
35. Adzitey F, Corry J (2011) A comparison between hippurate hydrolysis and multiplex PCR for differentiating *Campylobacter coli* and *Campylobacter jejuni*. Trop Life Sci Res 22(1):91
36. Nakari U-M, Puhakka A, Siitonen A (2008) Correct identification and discrimination between *Campylobacter jejuni* and *C. coli* by a standardized hippurate test and species-specific polymerase chain reaction. Eur J Clin Microbiol Infect Dis 27(7):513–518
37. Hodge D, Borczyk A, Wat L (1990) Evaluation of the indoxyl acetate hydrolysis test for the differentiation of Campylobacters. J Clin Microbiol 28(6):1482–1483
38. Steinbrueckner B, Haerter G, Pelz K, Kist M (1999) Routine identification of *Campylobacter jejuni* and *Campylobacter coli* from human stool samples. FEMS Microbiol Lett 179(2):227–232
39. Quinn P, Carter M, Markey B, Carter G (1994) *Campylobacter* species. Clin Vet Microbiol:268–272
40. Huysmans MB, Turnidge JD, Williams JH (1995) Evaluation of API Campy in comparison with conventional methods for identification of thermophilic Campylobacters. J Clin Microbiol 33(12):3345–3346
41. Martiny D, Dediste A, Debruyne L, Vlaes L, Haddou NB, Vandamme P et al (2011) Accuracy of the API Campy system, the Vitek 2 Neisseria–haemophilus card and matrix-assisted laser desorption ionization time-of-flight mass spectrometry for the identification of *Campylobacter* and related organisms. Clin Microbiol Infect 17(7):1001–1006
42. Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl Environ Microb 51(5):873
43. Penner JL, Hennessy J (1980) Passive hemagglutination technique for serotyping *Campylobacter fetus* subsp. jejuni on the basis of soluble heat-stable antigens. J Clin Microbiol 12(6):732–737
44. Oza A, Thwaites R, Wareing D, Bolton F, Frost J (2002) Detection of heat-stable antigens of *Campylobacter jejuni* and *C. coli* by direct agglutination and passive hemagglutination. J Clin Microbiol 40(3):996–1000
45. McKay D, Fletcher J, Cooper P, Thomson-Carter FM (2001) Comparison of two methods for Serotyping *Campylobacter* spp. J Clin Microbiol 39(5):1917–1921
46. Wassenaar TM, Newell DG (2000) Genotyping of *Campylobacter* spp. Appl Environ Microbiol 66(1):1–9
47. Khakhria R, Lior H (1992) Extended phage-typing scheme for *Campylobacter jejuni* and *Campylobacter coli*. Epidemiol Infect 108(3):403–414
48. Ribot EM, Fitzgerald C, Kubota K, Swaminathan B, Barrett TJ (2001) Rapid pulsed-field gel electrophoresis protocol for subtyping of *Campylobacter jejuni*. J Clin Microbiol 39(5):1889–1894
49. Maslow J (1993) Application of pulsed-field gel electrophoresis to molecular epidemiology. Diagnostic Mol Microbiol:563–572
50. Maiden MC (2006) Multilocus sequence typing of bacteria. Annu Rev Microbiol 60:561–588
51. On SL, McCarthy N, Miller WG, Gilpin BJ (2008) Molecular epidemiology of *Campylobacter* species. In: Campylobacter, 3rd edn. American Society of Microbiology, pp 191–211
52. Scott AE, Timms AR, Connerton PL, Carrillo CL, Radzum KA, Connerton IF (2007) Genome dynamics of *Campylobacter jejuni* in response to bacteriophage predation. PLoS Pathog 3(8):e119
53. Barton C, Ng L-K, Tyler SD, Clark CG (2007) Temperate bacteriophages affect pulsed-field gel electrophoresis patterns of *Campylobacter jejuni*. J Clin Microbiol 45(2):386–391
54. Harrington CS, Thomson-Carter FM, Carter PE (1997) Evidence for recombination in the flagellin locus of *Campylobacter jejuni:* implications for the flagellin gene typing scheme. J Clin Microbiol 35(9):2386–2392

55. Harrington CS, Moran L, Ridley A, Newell D, Madden R (2003) Inter-laboratory evaluation of three flagellin PCR/RFLP methods for typing *Campylobacter jejuni* and *C. coli:* the CAMPYNET experience. J Appl Microbiol 95(6):1321–1333
56. Cody AJ, Maiden MJC, Dingle KE (2009) Genetic diversity and stability of the *porA* allele as a genetic marker in human *Campylobacter* infection. Microbiology (Reading) 155:4145–4154
57. Granato PA, Chen L, Holiday I, Rawling RA, Novak-Weekley SM, Quinlan T et al (2010) Comparison of premier CAMPY enzyme immunoassay (EIA), ProSpecT Campylobacter EIA, and ImmunoCard STAT! CAMPY tests with culture for laboratory diagnosis of *Campylobacter* enteric infections. J Clin Microbiol 48(11):4022–4027
58. Bojanić K, Midwinter AC, Marshall JC, Rogers LE, Biggs PJ, Acke E (2016) Variation in the limit-of-detection of the ProSpecT Campylobacter microplate enzyme immunoassay in stools spiked with emerging *Campylobacter* species. J Microbiol Methods 127:236–241
59. Mandrell RE, Harden LA, Bates A, Miller WG, Haddon WF, Fagerquist CK (2005) Speciation of *Campylobacter coli*, *C. jejuni*, *C. helveticus*, *C. lari*, *C. sputorum*, and *C. upsaliensis* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol 71(10):6292–6307
60. De Boer P, Rahaoui H, Leer R, Montijn R, Van der Vossen J (2015) Real-time PCR detection of *Campylobacter* spp.: a comparison to classic culturing and enrichment. Food Microbiol 51:96–100
61. Ricke SC, Feye KM, Chaney E, Shi Z, Pavlidis HO, Yang Y (2018) Developments in rapid detection methods for the detection of foodborne *Campylobacter* in the United States. Front Microbiol 9:3280
62. Klein D (2002) Quantification using real-time PCR technology: applications and limitations. Trends Mol Med 8(6):257–260
63. Best EL, Powell EJ, Swift C, Grant KA, Frost JA (2003) Applicability of a rapid duplex real-time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture plates. FEMS Microbiol Lett 229(2):237–241
64. Dingle K, Colles F, Wareing D, Ure R, Fox A, Bolton F et al (2001) Multilocus sequence typing system for *Campylobacter jejuni*. J Clin Microbiol 39(1):14–23
65. Dingle KE, Colles FM, Ure R, Wagenaar JA, Duim B, Bolton FJ et al (2002) Molecular characterization of *Campylobacter jejuni* clones: a basis for epidemiologic investigation. Emerg Infect Dis 8(9):949
66. Miller WG, On SL, Wang G, Fontanoz S, Lastovica AJ, Mandrell RE (2005) Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari, C. upsaliensis*, and *C. helveticus*. J Clin Microbiol 43(5):2315–2329
67. Dingle KE, McCarthy ND, Cody AJ, Peto TE, Maiden MC (2008) Extended sequence typing of *Campylobacter* spp., United Kingdom. Emerg Infect Dis 14(10):1620
68. Colles F, McCarthy N, Howe J, Devereux C, Gosler A, Maiden M (2009) Dynamics of *Campylobacter* colonization of a natural host, *Sturnus vulgaris* (European starling). Environ Microbiol 11(1):258–267
69. Sheppard SK, Dallas JF, Strachan NJ, MacRae M, McCarthy ND, Wilson DJ et al (2009) *Campylobacter* genotyping to determine the source of human infection. Clin Infect Dis 48(8):1072–1078
70. Sheppard SK, McCarthy ND, Falush D, Maiden MC (2008) Convergence of *Campylobacter* species: implications for bacterial evolution. Science 320(5873):237–239
71. Jolley KA, Bray JE, Maiden MC (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3
72. Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44(W1):W242–W2W5
73. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549
74. Zhou Z, Alikhan N-F, Sergeant MJ, Luhmann N, Vaz C, Francisco AP et al (2018) GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res 28(9):1395–1404

75. Cody AJ, Bray JE, Jolley KA, McCarthy ND, Maiden MCJ (2017) Core genome multilocus sequence typing scheme for stable, comparative analyses of *Campylobacter jejuni* and *C. coli* human disease isolates. J Clin Microbiol 55(7):2086–2097
76. Maiden MCJ, van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA et al (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 11(10):728–736
77. Cody AJ, McCarthy ND, van Rensburg MJ, Isinkaye T, Bentley SD, Parkhill J et al (2013) Real-time genomic epidemiological evaluation of human *Campylobacter* isolates by use of whole-genome multilocus sequence typing. J Clin Microbiol 51(8):2526–2534
78. Colles FM, Preston SG, Barfod KK, Flammer PG, Maiden MCJ, Smith AL (2019) Parallel sequencing of *porA* reveals a complex pattern of *Campylobacter* genotypes that differs between broiler and broiler. Sci Rep 9

Clostridioides difficile

2

Eliane de Oliveira Ferreira and Gerly Anne de Castro Brito

2.1 Introduction

Initially called *Bacillus difficilis*, due to the difficulty in its cultivation, the species is currently known as *Clostridioides difficile* (*C. difficile*). It was renamed after a new categorization of microorganisms belonging to the genus *Clostridium*, based on 16S ribosomal RNA. Ezaki [1] proposed a regrouping of *C. difficile*, due to its 94.7% similarity in the gene sequence with *Clostridium mangenotii*, belonging to the *Peptostreptococcaceae* family. However, this new reclassification generated conflict due to the various acronyms adopted over the years to refer to diseases associated with *C. difficile*. Therefore, Lawson et al. [2] proposed a new modification in the nomenclature of both the *Clostridium difficile* and the *Clostridium mangenotii*, for *Clostridioides*, with *C. difficile* as the main representative of this genus.

C. difficile is a gram-positive, anaerobic spore-forming bacterium considered as the leading cause of antibiotic-associated diarrhea and pseudomembranous colitis. *C. difficile* is largely known as a nosocomial infectious agent in industrialized countries [3], but its finding in underdeveloped and developing countries is still limited. *C. difficile* was originally identified as part of the fecal flora of healthy newborn infants by Hall and O'Toole in 1935 [4] and described as a "thick, extremely mobile rod with terminal and subterminal spores." The organism grows slowly and is difficult to isolate in pure culture. Its presence in the stool of healthy neonates suggested that *C. difficile* was a not a pathogen, even though it produced toxins in broth

E. de Oliveira Ferreira
Federal University of Rio de Janeiro, Institute of Microbiology Paulo de Góes, Laboratory of Anaerobic Biology, Rio de Janeiro, Brazil

G. A. de Castro Brito (✉)
Department of Morphology, Faculty of Medicine, Federal University of Ceará, Fortaleza, Brazil

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_2

culture and has shown effect in animals (rabbits and guinea pigs). After its original description, in the 1970s, the antibiotic-associated pseudomembranous colitis (PMC) became more common with the introduction into clinical practice of broad-spectrum antibiotics. Hence, with the association of clindamycin and lincomycin treatment with PMC, the term "clindamycin colitis" started to be used [5]. In 1978, *C. difficile* was finally recognized as a source of a cytotoxin in the stool of patients with PMC [6], and in 1979, George et al. formulated a selective media, named cycloserine-cefoxitin-fructose agar (CCFA), allowing the isolation of *C. difficile*. During the past decades since its rediscovery, a great deal has been learned about the pathophysiology, epidemiology, and management of *C. difficile*, and the infection became known as *Clostridioides difficile*-associated illness/diarrhea (CDAD), but currently, the term *Clostridioides difficile* infection (CDI) is more acceptable [7].

The colonization in humans is facilitated by the establishment of dysbiosis (alteration of the intestinal microbiome) specially caused by the use of antimicrobials. This pathogen causes enteric infections, which can range from asymptomatic colonization to diarrheal episodes and to severe conditions of the disease, such as fulminant colitis, colectomy, toxic megacolon, and death [8]. The use of broad-spectrum antimicrobials, length of hospital stay, advanced age (> 65 years), and comorbidities are some of the factors that can increase the risk of individuals for the development of CDIs [9]. Patients transplanted and using immunosuppressive medications, individuals with inflammatory bowel disease and irritable bowel syndrome, and those using proton pump inhibitors may also be more likely to develop CDI [10]. It is of great concern the fact that the COVID-19 caused by the new coronavirus SarsCov-2 might bring up two important risk factors for the development of CDIs, the prolonged stay of patients in the intensive care unit and use of corticoids and antimicrobials, for instance, azithromycin, to control bacterial pneumonia and other infections. Thus, during the COVID-19 era, the world might experience a marked increase in the number of acute diarrheal diseases [11, 12].

2.2 Virulence Factors

Several virulence factors have been identified and related to *C. difficile* outbreaks and epidemics in the world, such as sporulation rates, antimicrobial resistance profile, adhesins, and biofilm production [13–18]. Although these multiple attributes have been associated with the severity of the CDI, including colonization, tissue damage, and formation of the pseudomembranous and fulminant cases of the disease, the pathogenesis of *C. difficile* is mainly due to the ability of some strains to produce two large toxins, which include toxins A (TcdA) and B (TcdB) [19]. The genes encoding toxins are in a pathogenic locus (PaLoc) of 19.6 kb, which encode toxin A (*tcd*A) and toxin B (*tcd*B) [20]. TcdA and TcdB are analogous to other large clostridial toxins, responsible for glycosylating small guanosine triphosphatases in the Rho and Ras families when endocytosed into epithelial cells, leading to cytoskeleton disassemble, disruption of tight junctions, and ultimately cell death [21]. In

this same locus are the *tcd*R, which encodes an alternative sigma factor that specifically directs transcription from the toxin promoters as its own promoter; the *tcd*C, which encodes an antagonist of TcdR; and the *tcd*E, a member of the class I family of holins, which includes the lambda phage S protein, necessary for toxin secretion [22, 23]. Outside the PaLoc is the binary toxin, known as *C. difficile* transferase (CDT), encoded by two genes, *cdt*A and *cdt*B. The CDT is adenosine diphosphate (ADP)-ribosyltransferase that causes actin cytoskeletal disruption, allowing the formation of microtubule protrusions on the surface of host cells that are thought to increase *C. difficile* adherence. CDT is typically produced by the major hypervirulent strains and has been associated with more severe disease, higher mortality, and elevated risk of recurrence in clinical studies [24] [25, 26]. CDT-expressing strains are not so rare anymore, becoming progressively common over the last 10 years, paralleling the overall increase in incidence and severity of CDI, and now account for up to 20% of isolates in the hospital units [27].

2.3 Epidemiology

CDI epidemiology changed dramatically after 2000 with a huge increase in the disease incidence and the number of severe cases of CDIs worldwide. In North America, for instance, the number of cases doubled from 1996 to 2003. The outnumbered cases were attributed to the appearance of certain types, or ribotypes, of *C. difficile*, such as the epidemic strain B1/NAP1/027 [28], which rapidly migrated to Europe. In Latin America, CDI cases associated with the epidemic strain were reported in Costa Rica, México, Panama, Chile, and Colombia [29]. In Brazil, however, the epidemic ribotype 027 strain has never been reported. However, a Clade 2 strain, phylogenetically related to the epidemic strain NAP1/027, with hypervirulent feature was isolated in northeast of Brazil [30].

The clone B1/NAP1/027 is epidemic (hypervirulent) due to the production of high concentrations of TcdA and TcdB and binary toxins and resistance to the fluoroquinolone and presents a high sporulation rate, which increases the potential for transmission [31]. At the time, an effort was made to identify the strain responsible for so many outbreaks and high levels of morbidity and mortality. The terminology created to the epidemic strain reflects the different techniques that were used at the time to its characterization, like class BI, according to the restriction endonuclease analysis (REA); NAP1 (North American pulsotype 1), based on the profile obtained from pulsed-field gel electrophoresis (PFGE); and ribotype 027, using the PCR ribotyping technique, which is based on the amplification of sequences in the intergenic space of the 16S–23S ribosomal RNA [28]. With the increase in cases related to the RT027 strain and the appearance of outbreaks caused by NAP1/RT027 and other ribotypes, for instance, 078/020, isolated from community infections (CA-CDI), the scientific community began to consider that other hosts could be harboring *C. difficile*, such as animals and the environment. When *C. difficile* was isolated in farm animals (cattle, pigs, and chickens); pets (dogs and cats), with a

prevalence of 39%; and wild animals, there was a solid proof of the zoonotic potential of *C. difficile*, these animals being considered a reservoir for virulent ribotypes of this pathogen [32–34]. In recent years, community-acquired CDI (CA-CDI) has also been increasing, and at least 600 *C. difficile* ribotypes are circulating worldwide [4, 35].

Rapid diagnosis of CDI is desirable to allow hasty isolation and treatment of patients, reducing potential patient-to-patient transmission and hospital staying for those affected. In addition, *C. difficile* strain typing can identify outbreaks within a hospital or the wider community. Over the past 35 years, the complexity and diversity of *C. difficile* bring a significant challenge for effective typing of clinical isolates, especially with concerns to associating genotypes with virulence factors and clinical samples [36] [37].

2.4 Molecular Assays for *C. difficile* Detection

The detection of *C. difficile* cytotoxin in stools from patients with diarrhea or PMC by testing in cell lines (VERO, HeLa, and McCoy cells) was considered a gold standard assay [38]; however, due to its difficulty and long processing time, rapid enzyme immunoassay (EIA) kits for detecting glutamate dehydrogenase (GDH) or toxins became more appreciated. EIA is rapid and easy to perform and does not require technical training or special equipment. EIA test to detect glutamate dehydrogenase (GDH), an enzyme produced specifically by *C. difficile*, remains a useful screening test, but sensitivities vary among available toxin EIA tests. Since this antigen is present in both toxigenic and nontoxigenic strains, GDH immunoassays lack specificity and must be combined with another (usually toxin) test [39]. The low specificity of GDH EIA test and low sensitivity of toxin EIA tests have driven the continuing search for sensitive and rapid methods for CDI diagnosis [40]. Hence, the molecular methods, such as the nucleic acid amplification tests (NAATs), started to become popular since the 1990s, because they were rapid and sensitive [41]. The first NAAT to become approved for *C. difficile* diagnosis, the BD Diagnostics BD Max Cdiff Assay, was only approved in 2009 by the American agency Food and Drug Administration (FDA). Although NAATs are vastly used for clinical analysis, some bias was observed: (i) NAATs can detect toxigenic *C. difficile* strains in colonized patients, but not necessarily the strain that is causing CDI; (ii) the heterogenicity of the strains can result in false-negative tests; (iii) there is TcdB variant strains; and (iv) some strains carry the *tcd*C gene base-pair deletion in the position 117 as the NAP1/RT027 [42], resulting in a false-positive result for the epidemic strain. While NAATs continue to be used by some laboratories and some researchers are convinced that they are necessary for CDI diagnosis, others think that two methodologies should be adopted in parallel or series for CDI diagnosis.

2.5 *C. difficile* Typing Methodologies

These molecular methodologies are used to characterize and compare the circulating strains and to identify emerging strains and those responsible for outbreaks worldwide. There are several typing methods available for *C. difficile*, including restriction endonuclease analysis (REA) [43], pulsed-field gel electrophoresis (PFGE) [44], PCR ribotyping [45], multilocus sequence typing (MLST) [46], repetitive-element PCR typing (rep-PCR) [47], toxinotyping [48], multilocus variable-number tandem-repeat analysis (MLVA) [49], surface-layer protein A-encoding gene (*slp*A) typing [50], and whole-genome sequencing (WGS) [51]. Apart from the most used methodologies for typing *C. difficile*, there are other but very promising techniques, including the CRISPR-Cas (clustered regularly interspaced short palindromic repeats (CRISPR)) [52] and the DNA microarray-based genotyping [53] for typing *C. difficile* that will also be mentioned here. A summary of *C. difficile* typing methodologies is shown in Table 2.1.

Table 2.1 *Clostridioides difficile* typing methods

Typing methods	Technique	Discriminatory capacity	High yield	References
Restriction endonuclease analysis – REA	Whole-genome DNA, which is frequently digested by restriction enzyme; detection by gel electrophoreses	High	No	Huber et al. (2013) [54]; Sambol et al. (2016) [43]
Pulsed field gel electrophoresis – PFGE	Resolves large fragments of DNA generated from whole-genome macro restriction with an infrequently cutting restriction enzyme; detection by gel electrophoresis	High	No	Corkill et al. (2000) [44]
PCR-ribotyping (agarose-base)	Based on the size variation of the 16S-23S rDNA intergenic spacer regions (ISR); detection by gel electrophoresis	Moderate	Yes	Stubbs et al. (1999) [45]
Multilocus sequence typing – MLST	Involves the partial amplification and sequencing of usually housekeeping genes	Low	Yes	Griffiths et al. (2010) [46]
Repetitive sequence-based PCR typing – Rep-PCR	Based on polymorphisms of repetitive elements that exist in multiple copies in the genome	High	Yes	Northey et al. (2005) [47]
Toxinotyping	Achieved by PCR amplification, followed by a restriction enzyme digestion of 10 regions of the PaLoc	Low	No	Rupnik and Jazenik (2016) [48]

(continued)

Table 2.1 (continued)

Typing methods	Technique	Discriminatory capacity	High yield	References
Multilocus variable-number tandem-repeat analysis – MLVA	Reached by using a multiplex PCR with primers designed to aim different VNTR regions in the genome	High	Yes	Manzoor et al. (2011) [49]
*slp*AST typing	Sequencing of the variable region of the *slpA* gene by using PCR amplification followed by restriction fragment length polymorphism (RFLP) analysis and DNA sequencing	High	No	Karjalainen et al. (2002) [50]
Whole-genome sequencing	Exposes the complete DNA of an organism at a single time (provides genetic variation)	High	Yes	Sebaihia et al. (2006) [70]
CRISPR-Cas/ CRISPR	Based on the occurrence and diversity of clustered regularly interspaced short palindromic repeats and associated genes	High	Yes	Andersen et al. (2016) [52]
DNA microarray	Relies on complementary DNA fragments (cDNA) from a sample to hybridize with synthetic DNA sequences	High	Yes	Gawlik et al. (2015) [53]

2.5.1 Restriction Endonuclease Analysis (REA)

Restriction endonuclease analysis (REA) of *C. difficile* uses whole-genome DNA, which is frequently digested by restriction enzyme HindIII. In contrast to PFGE, the digestion fragments are separated by standard electrophoresis, either on agarose or polyacrylamide gels. REA has high discriminatory power and stability, but the method is time-consuming and difficult to interpret, and the data are difficult to exchange between laboratories [43, 54].

2.5.2 Pulsed-Field Gel Electrophoresis (PFGE)

For tracking global circulating strains, pulsed-field gel electrophoresis (PFGE) is the method generally preferred by the North Americans and the Centers for Disease Control and Prevention (CDC). PFGE resolves large fragments of DNA generated from whole-genome macro-restriction with an infrequently cutting restriction enzyme for *C. difficile Sma*I or *Sac*II to cleave bacterial DNA at different restriction sites. The use of these infrequently cutting restriction enzymes limits the number of restriction fragments (to between 7 and 20) and ensures that they are relatively large

[44]. Generally, the frequency of cutting is inversely proportional to the number of nucleotides in the recognition site. This procedure separates the large fragments of DNA generated based on size using a pulsed-field electrophoresis gel with resulting electrophoresis patterns that are highly discriminatory. In the North America, strains are named according to the pulsotype followed by a number (e.g., North American pulsotype 1 (NAP1)). The CDC's PulseNet program can be accessed in the https://www.cdc.gov/pulsenet/pathogens/pfge.html address.

2.5.3 PCR Ribotyping

The *C. difficile* PCR ribotyping is nowadays the dominant typing method in Europe and Australia. Bacterial rRNA (rrn) operons are usually organized in the order 16S rrnA-ISR-23S rrnA-ISR-5S rrnA, and their copy numbers can range between 1 and 15. This method is based on the size variation of the 16S–23S rDNA intergenic spacer regions (ISRs). The *C. difficile* PCR ribotyping is possible because this pathogen has large intraspecific diversity in the ISRs [55]. In 1999, Bidet et al. [56] enhanced the reading of the banding patterns by selecting a partial sequence of the rRNA genes (16S–23S) and the intergenic spacer region with a new set of primers located closer to this intergenic spacer region. Now the method can be performed with agarose gel-based electrophoresis.

As in the PFGE methodology, the PCR ribotyping method follows a nomenclature, which was established by the Public Health Laboratory Service Anaerobe Reference Unit, Cardiff, in England for *C. difficile*. The nomenclature is designated by a three-digit number starting from 001 (e.g., PCR ribotype 027). Presently, the collection of existing PCR ribotypes and the delegation of new ones are performed by the Health Protection Agency-funded *C. difficile* Ribotyping Network (CDRN) in Leeds, England, which has more than 600 different PCR ribotypes in the CDRN database [57].

2.5.4 Multilocus Sequence Typing (MLST)

This method for typing *C. difficile* involves the partial amplification and sequencing of usually seven housekeeping genes (*ad*K, *atp*A, *dxr*, *gly*A, *rec*A, *sod*A, and *tpi*) and is the second most used method after PCR ribotyping. The internal fragments [58] of each gene (450–500 bp) are sequenced on both strands (forward and reverse) using an automated DNA sequencing equipment. For each housekeeping gene, the different sequences present within a bacterial species are assigned as unique alleles, and for each isolate, the alleles at each of the seven loci define the allelic profile or sequence type (ST) [59]. The data obtained are unequivocal, and the allelic profiles of the isolates can easily be compared to those in a large central database (University of Oxford: http://pubmlst.org/cdifficile) and therefore can be compared between laboratories. Depending on the housekeeping genes, there are two MLST databases [46, 60].

2.5.5 Repetitive Sequence-Based PCR Typing (Rep-PCR)

This method was proposed for the first time by Northey et al. [47], and typing is based on polymorphisms of repetitive elements that exist in multiple copies in the *C. difficile* genome. Specific repetitive PCR primers complement these repetitive sequences, and the amplified DNA fragments give a genomic fingerprint that can be used for subspecies discrimination. The DiversiLab system (bioMérieux, Marcy-l'Etoile, France) is an automated rep-PCR typing method that has a high discriminatory power when compared to traditional PCR ribotyping, and this high resolution may be useful for investigating outbreaks within a hospital [61].

2.5.6 Toxinotyping

Toxinotyping is a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)-based method for distinguishing *C. difficile* strains, positive for the TcdA and TcdB toxins (PaLoc chromosomal region), according to changes in their toxin genes when compared to the reference strain VPI 10463 (ATCC® 43255) [62, 63]. Basically, this methodology is achieved by PCR amplification, followed by a restriction enzyme digestion of ten regions of the PaLoc. With that, 31 different toxinotypes were identified and designated by Roman numerals from I to XXXI. For instance, when strains present toxin genes similar to VPI 10463, they are classified as toxinotype 0 (e.g., *C. difficile* toxinotype XIV is *tcd*A$^+$ and *tcd*B$^+$, CDT$^+$; B1 [7] and A3 [2]; *tcd*C [1]). Strains with changes in toxin genes are grouped into variant toxinotypes I to XXXI (e.g., *C. difficile* toxinotype XXXI *tcd*A$^-$ and *tcd*B$^+$, CDT$^+$; B1 [5] and A3 [negative]; *tcd*C [negative]). Although toxinotyping does not have a good discriminatory power when compared to PFGE and PCR ribotyping, it has very high reproducibility [64].

2.5.7 Multilocus Variable-Number Tandem-Repeat Analysis (MLVA)

MLVA typing method uses the variation in the number of tandem-repeat DNA, naturally found in sequences in many different loci in the genome. Therefore, the lengths of the variable number of tandem repeat (VNTR) regions are determined to distinguish among the strains. The technique is reached by using a multiplex PCR with primers designed to aim different VNTR regions in the genome. The band profile is visualized by electrophoresis or automated fragment analysis on a sequencer; and the amplicon size is used to calculate the number of repeat units of each locus. The calculated numbers of repeats of the VNTR loci (alleles) are combined, and this provides the MLVA profile. Each unique MLVA profile is given an MLVA type designation. The MLVA profile can be used for the comparison and clustering of the bacteria [49]. MLVA allows outbreaks to be tracked more efficiently than PCR ribotyping, and it has the prospective to define phylogenetic relationships. Several MLVA schemes have been described for typing *C. difficile*.

2.5.8 Surface-Layer Protein A-Encoding Gene Typing (*slp*AST) as a Substitute for Serotyping

Serotyping distinguishes ten major *C. difficile* variations (A, B, C, D, F, G, H, I, K, and X) based on the bacterial surface antigens, by slide agglutination or enzyme-linked immunosorbent assay, using rabbit antisera. Strains of the serogroup A have a flagellar antigen in common that is responsible for cross-agglutination on slides but can be divided into 20 subgroups (A1–A20) by polyacrylamide gel electrophoresis [50].

The *C. difficile* presents a surface-layer protein A (SlpA), an immunodominant protein, encoded by the *slp*A gene. The *slp*A typing of isolates is performed by the sequencing of the variable region of the *slpA* gene by using PCR amplification followed by restriction fragment length polymorphism (RFLP) analysis and DNA sequencing. Apart from the serogroup A, the *slp*A nucleotide sequences were identical within a given serogroup and different between serogroups. S-layer-based typing may therefore have greater relevance to vaccine development than conventional genotyping and could also be suitable for explaining the differences in the epidemiology of CDI [65–67].

2.5.9 Whole-Genome Sequencing (WGS)

Whole-genome sequencing (WGS) exposes the complete DNA of an organism at a single time and provides the most complete pool of an individual's genetic variation [68, 69]. The Sanger sequencing, the same place where the first *C. difficile* was sequenced [70], and Roche 454 and Illumina next-generation sequencing technologies have been applied to study the evolutionary dynamics of *C. difficile* at low cost. Even though there is an enthusiasm for adopting WGS-based methods, it is still high cost and inaccessible to most laboratories [54]. Another disadvantage is the deficiency of a standardized analysis scheme for WGS data, which are obstacles to the wide-scale adoption of this method in the world, and the complexity of analysis, because *C. difficile* has a low core genome SNP occurrence rate among clinical isolates [71–73].

2.5.10 CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR))

The bacterial phylogenetic analysis, CRISPR-Cas – clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated sequences (cas) – system, is an appreciated genetic target for high-resolution typing and micro-evolution bacterial [74]. The CRISPR-Cas loci have been found in *C. difficile* [75], but its occurrence and diversity are still not well explored [69, 76]. *C. difficile* species carry prophages, which impact their behavior, and also mobile genetic elements, known to possibly confer antibiotic resistance [77, 78]. Andersen et al. [52] have assessed

the potential of CRISPR-based phylogeny and high-resolution genotyping of *C. difficile* strains, and it seems to be a promising future methodology. The authors highlight that CRISPR-Cas might be a valuable methodology for genotyping of *C. difficile* isolates and also provide insights into the micro-evolutionary events that occur between closely related strains and the evolutionary course in their genomes.

2.5.11 DNA Microarray-Based Genotyping

The DNA microarray relies on complementary DNA fragments (cDNA) from a sample to hybridize with synthetic DNA sequences (ssDNA; specific A, G, T, C, combinations). The synthetic fragments (oligonucleotides) are the probes, and they will match, if present in the sample, with the complementary DNA target (cDNA). Once the cDNA and the ssDNA are bound, they form a hybrid double helix (dsDNA hybrid molecule). The efficacy of this methodology has been proved to be reliable for detecting other pathogens, such as *Staphylococcus aureus* MRSA, in hospitals [79].

Galik et al. [80] created microarray including the following genes: the surface layer protein *slp*A gene; toxin genes (*tcd*A and *tcd*B, *cdt*A and *cdt*B); two alleles of the A component ($cdtA_{R20291}$ and $cdtA_{630}$) of the binary toxin; ubiquitous resistance markers (*bcr*A, encoding the bacitracin ATP-binding cassette transporter; *lmr*B, associated with lincomycin/clindamycin resistance; and *vat*A (synonym sat) encoding a virginiamycin/streptogramin A acetyltransferase); genes related to antimicrobial resistance, e.g., *cat* (chloramphenicol acetyltransferase), *erm*(B) (RNA methyltransferase, conferring resistance to macrolides and clindamycin), and *tet*(M), encoding tetracycline resistance; other markers, e.g., genes *vnc*S and *vex*P1 encoding a histidine kinase and a permease; and other genes, e.g., septum formation initiation protein (*div*C), flagellin subunit C (*fli*C), and cell wall proteins 66 and 84. The authors suggest that the microarray-based assay permitted a rapid and high-throughput genotyping of clinical *C. difficile* isolates (n = 234), including toxin gene detection and strain assignment. The methodology employed by them could also correlate with MLST-derived clades. Another benefit is that the process can be made within half a day being more rapid than ribotyping.

2.6 Conclusions and Perspectives

Clostridioides difficile is an important nosocomial pathogen increasingly observed in the community and in different reservoirs, such as animals and the environment. Over the years, CDIs have become one of the most noteworthy threats to hospitalized and immunocompromised patients. Its significance on asymptomatic host is gradually recognized, not only as a cause of false-positive clinical testing but also as a source of new infections within hospitals and other healthcare environments. During decades, the transmissibility of *C. difficile* has been studied by using a variety of methods, although recently there is a tendency to use rapid and low-cost

techniques in diagnosis. However, precise molecular typing techniques have arisen to improve the diagnosis, treatment, and understanding of CDI epidemiology.

2.7 Summary

C. difficile is a gram-positive, anaerobic spore-forming bacterium considered as the leading cause of antibiotic-associated diarrhea and pseudomembranous colitis. *C. difficile* is a foremost problem in hospitals, where cases can be caused by community-acquired strains, as well as by nosocomial spread. The epidemiology of *C. difficile* has changed considerably along the years with the variety of strains (ribotypes) from different sources; hence, an accurate diagnosis of *Clostridioides difficile* infection (CDI) is important not only for patient care but also for epidemiology and disease research purposes. Here we make a review of the precise molecular typing techniques used to better understanding of CDI epidemiology.

Acknowledgments Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) grant # E-26/211.554/2019 (Programa Rede de Pesquisa em Saúde) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

1. Ezaki T (2009) Bergey's manual of systematic bacteriology, 2nd edn. Springer, New York
2. Lawson PA, Citron DM, Tyrrell KL, Finegold SM (2016) Reclassification of Clostridium difficile as Clostridioides difficile (Hall and O'Toole 1935) Prevot 1938. Anaerobe 40:95–99. https://doi.org/10.1016/j.anaerobe.2016.06.008
3. Warny M, Pepin J, Fang A et al (2005) Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe. Lancet 366:1079–1084. https://doi.org/10.1016/S0140-6736(05)67420-X
4. Rupnik M, Wilcox MH, Gerding DN (2009) Clostridium difficile infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 7:526–536. https://doi.org/10.1038/nrmicro2164
5. Gerding DN (2009) Clostridium difficile 30 years on: what has, or has not, changed and why? Int J Antimicrob Agents 33(Suppl 1):S2–S8. https://doi.org/10.1016/S0924-8579(09)70008-1
6. Bartlett JG, Chang TW, Moon N, Onderdonk AB (1978) Antibiotic-induced lethal enterocolitis in hamsters: studies with eleven agents and evidence to support the pathogenic role of toxin-producing Clostridia. Am J Vet Res 39:1525–1530
7. George WL, Sutter VL, Citron D, Finegold SM (1979) Selective and differential medium for isolation of Clostridium difficile. J Clin Microbiol 9:214–219
8. Deneve C, Delomenie C, Barc M-C et al (2008) Antibiotics involved in Clostridium difficile-associated disease increase colonization factor gene expression. J Med Microbiol 57:732–738. https://doi.org/10.1099/jmm.0.47676-0
9. Stiefelhagen P (2007) Severe diarrhea in antibiotic therapy. Test for Clostridium difficile! MMW Fortschr Med 149:22
10. Feuerstadt P, Das R, Brandt LJ (2014) The evolution of urban C. difficile infection (CDI): CDI in 2009-2011 is less severe and has better outcomes than CDI in 2006-2008. Am J Gastroenterol 109:1265–1276. https://doi.org/10.1038/ajg.2014.167

11. Zuo T, Zhang F, Lui GCY et al (2020) Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. Gastroenterology. https://doi.org/10.1053/j.gastro.2020.05.048
12. Hsu J (2020) How covid-19 is accelerating the threat of antimicrobial resistance. BMJ 369:m1983. https://doi.org/10.1136/bmj.m1983
13. Awad MM, Johanesen PA, Carter GP et al (2014) Clostridium difficile virulence factors: insights into an anaerobic spore-forming pathogen. Gut Microbes 5:579–593. https://doi.org/10.4161/19490976.2014.969632
14. Kovacs-Simon A, Leuzzi R, Kasendra M et al (2014) Lipoprotein CD0873 is a novel adhesin of Clostridium difficile. J Infect Dis 210:274–284. https://doi.org/10.1093/infdis/jiu070
15. Pechine S, Deneve-Larrazet C, Collignon A (2016) Clostridium difficile Adhesins. Methods Mol Biol 1476:91–101. https://doi.org/10.1007/978-1-4939-6361-4_7
16. Eveillard M, Fourel V, Barc MC et al (1993) Identification and characterization of adhesive factors of Clostridium difficile involved in adhesion to human colonic enterocyte-like Caco-2 and mucus-secreting HT29 cells in culture. Mol Microbiol 7:371–381
17. Stevenson E, Minton NP, Kuehne SA (2015) The role of flagella in Clostridium difficile pathogenicity. Trends Microbiol 23:275–282. https://doi.org/10.1016/j.tim.2015.01.004
18. Janoir C (2016) Virulence factors of Clostridium difficile and their role during infection. Anaerobe 37:13–24. https://doi.org/10.1016/j.anaerobe.2015.10.009
19. Rupnik M, Wilcox MH, Gerding DN (2009) Clostridium difficile infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 7:526–536
20. Bilverstone TW, Kinsmore NL, Minton NP, Kuehne SA (2017) Development of Clostridium difficile R20291DeltaPaLoc model strains and in vitro methodologies reveals CdtR is required for the production of CDT to cytotoxic levels. Anaerobe 44:51–54. https://doi.org/10.1016/j.anaerobe.2017.01.009
21. Di Bella S, Ascenzi P, Siarakas S et al (2016) Clostridium difficile toxins A and B: insights into pathogenic properties and extraintestinal effects. Toxins (Basel) 8. https://doi.org/10.3390/toxins8050134
22. Wang IN, Smith DL, Young R (2000) Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol 54:799–825. https://doi.org/10.1146/annurev.micro.54.1.799
23. Govind R, Dupuy B (2012) Secretion of Clostridium difficile toxins A and B requires the holin-like protein TcdE. PLoS Pathog 8:e1002727. https://doi.org/10.1371/journal.ppat.1002727
24. Gerding DN, Johnson S, Rupnik M, Aktories K (2013) Clostridium difficile binary toxin CDT: mechanism, epidemiology, and potential clinical importance. Gut Microbes 5:37–41. https://doi.org/10.4161/gmic.26854
25. Lim SK, Stuart RL, Mackin KE et al (2014) Emergence of a ribotype 244 strain of Clostridium difficile associated with severe disease and related to the epidemic ribotype 027 strain. Clin Infect Dis 58:1723–1730. https://doi.org/10.1093/cid/ciu203
26. Pechine S, Collignon A (2016) Immune responses induced by Clostridium difficile. Anaerobe 41:68–78. https://doi.org/10.1016/j.anaerobe.2016.04.014
27. Cowardin CA, Buonomo EL, Saleh MM et al (2016) The binary toxin CDT enhances Clostridium difficile virulence by suppressing protective colonic eosinophilia. Nat Microbiol 1:16108. https://doi.org/10.1038/nmicrobiol.2016.108
28. O'Connor JR, Johnson S, Gerding DN (2009) Clostridium difficile infection caused by the epidemic BI/NAP1/027 strain. Gastroenterology 136:1913–1924. https://doi.org/10.1053/j.gastro.2009.02.073
29. Guerrero-Araya E, Meneses C, Castro-Nallar E et al (2020) Origin, genomic diversity and microevolution of the clostridium difficile b1/nap1/rt027/st01 strain in costa Rica, Chile, Honduras and Mexico. Microb Genom 6:1–9. https://doi.org/10.1099/mgen.0.000355
30. Costa CL, Lopez-Urena D, de Oliveira AT et al (2016) A MLST clade 2 Clostridium difficile strain with a variant TcdB induces severe inflammatory and oxidative response associated with mucosal disruption. Anaerobe 40:76–84. https://doi.org/10.1016/j.anaerobe.2016.06.005
31. Dupont HL, Garey KW (2010) Clostridium difficile infection: an emerging epidemic with more questions than answers. Future Microbiol 5:1153–1156

32. Weese JS, Finley R, Reid-Smith RR et al (2010) Evaluation of Clostridium difficile in dogs and the household environment. Epidemiol Infect 138:1100–1104. https://doi.org/10.1017/S0950268809991312
33. Jardine CM, Reid-Smith RJ, Rousseau J, Weese JS (2013) Detection of Clostridium difficile in small and medium-sized wild Mammals in Southern Ontario, Canada. J Wildl Dis 49:418–421. https://doi.org/10.7589/2012-04-120
34. Rainha K, Fernandes Ferreira R, Trindade CNR et al (2019) Characterization of Clostridioides difficile ribotypes in domestic dogs in Rio de Janeiro, Brazil. Anaerobe 58:22–29. https://doi.org/10.1016/j.anaerobe.2019.06.007
35. Deshpande A, Pasupuleti V, Thota P et al (2013) Community-associated Clostridium difficile infection and antibiotics: a meta-analysis. J Antimicrob Chemother 68:1951–1961. https://doi.org/10.1093/jac/dkt129
36. Sirard S, Valiquette L, Fortier L-C (2011) Lack of association between clinical outcome of Clostridium difficile infections, strain type, and virulence-associated phenotypes. J Clin Microbiol 49:4040–4046. https://doi.org/10.1128/JCM.05053-11
37. Knetsch CW, Lawley TD, Hensgens MP et al (2013) Current application and future perspectives of molecular typing methods to study Clostridium difficile infections. Euro Surveill 18:20381
38. Asha NJ, Wilcox MH (2002) Laboratory diagnosis of Clostridium perfringens antibiotic-associated diarrhoea. J Med Microbiol 51:891–894. https://doi.org/10.1099/0022-1317-51-10-891
39. Martinez-Melendez A, Camacho-Ortiz A, Morfin-Otero R et al (2017) Current knowledge on the laboratory diagnosis of Clostridium difficile infection. World J Gastroenterol 23:1552–1567. https://doi.org/10.3748/wjg.v23.i9.1552
40. Seo JY, Jeong JH, Kim KH et al (2017) Laboratory diagnosis of Clostridium difficile infection: comparison of Techlab C. diff Quik Chek complete, Xpert C. difficile, and multistep algorithmic approach. J Clin Lab Anal. https://doi.org/10.1002/jcla.22135
41. Kumar S, Pollok R, Muscat I, Planche T (2017) Diagnosis and outcome of Clostridium difficile infection by toxin enzyme immunoassay and polymerase chain reaction in an island population. J Gastroenterol Hepatol 32:415–419. https://doi.org/10.1111/jgh.13504
42. Saito R, Usui Y, Ayibieke A et al (2019) Hypervirulent clade 2, ribotype 019/sequence type 67 Clostridioides difficile strain from Japan. Gut Pathog 11. https://doi.org/10.1186/s13099-019-0336-3
43. Sambol SP, Johnson S, Gerding DN (2016) Restriction endonuclease analysis typing of clostridium difficile isolates. Methods Mol Biol 1476:1–13. https://doi.org/10.1007/978-1-4939-6361-4_1
44. Corkill JE, Graham R, Hart CA, Stubbs S (2000) Pulsed-field gel electrophoresis of degradation-sensitive DNAs from Clostridium difficile PCR ribotype 1 strains. J Clin Microbiol 38:2791–2792
45. Stubbs SL, Brazier JS, O'Neill GL, Duerden BI (1999) PCR targeted to the 16S-23S rRNA gene intergenic spacer region of Clostridium difficile and construction of a library consisting of 116 different PCR ribotypes. J Clin Microbiol 37:461–463
46. Griffiths D, Fawley W, Kachrimanidou M et al (2010) Multilocus sequence typing of clostridium difficile. J Clin Microbiol 48:770–778. https://doi.org/10.1128/JCM.01796-09
47. Northey G, Gal M, Rahmati A, Brazier JS (2005) Subtyping of Clostridium difficile PCR ribotype 001 by REP-PCR and PFGE. J Med Microbiol 54:543–547. https://doi.org/10.1099/jmm.0.45989-0
48. Rupnik M, Janezic S (2016) An update on Clostridium difficile toxinotyping. J Clin Microbiol 54:13–18. https://doi.org/10.1128/JCM.02083-15
49. Manzoor SE, Tanner HE, Marriott CL et al (2011) Extended multilocus variable-number tandem-repeat analysis of Clostridium difficile correlates exactly with ribotyping and enables identification of hospital transmission. J Clin Microbiol 49:3523–3530. https://doi.org/10.1128/JCM.00546-11

50. Karjalainen T, Saumier N, Barc M-C et al (2002) Clostridium difficile genotyping based on slpA variable region in S-layer gene sequence: an alternative to serotyping. J Clin Microbiol 40:2452–2458
51. Eyre DW, Cule ML, Wilson DJ et al (2013) Diverse sources of C. difficile infection identified on whole-genome sequencing. N Engl J Med 369:1195–1205. https://doi.org/10.1056/NEJMoa1216064
52. Andersen JM, Shoup M, Robinson C et al (2016) CRISPR diversity and microevolution in Clostridium difficile. Genome Biol Evol 8:2841–2855. https://doi.org/10.1093/gbe/evw203
53. Gawlik D, Slickers P, Engelmann I et al (2015) DNA-microarray-based genotyping of Clostridium difficile. BMC Microbiol 15:158. https://doi.org/10.1186/s12866-015-0489-2
54. Huber CA, Foster NF, Riley TV, Paterson DL (2013) Challenges for standardization of Clostridium difficile typing methods. J Clin Microbiol 51:2810–2814. https://doi.org/10.1128/JCM.00143-13
55. Sadeghifard N, Gurtler V, Beer M, Seviour RJ (2006) The mosaic nature of intergenic 16S-23S rRNA spacer regions suggests rRNA operon copy number variation in Clostridium difficile strains. Appl Environ Microbiol 72:7311–7323. https://doi.org/10.1128/AEM.01179-06
56. Bidet P, Barbut F, Lalande V et al (1999) Development of a new PCR-ribotyping method for Clostridium difficile based on ribosomal RNA gene sequencing. FEMS Microbiol Lett 175:261–266
57. Knight DR, Elliott B, Chang BJ et al (2015) Diversity and evolution in the genome of Clostridium difficile. Clin Microbiol Rev 28:721–741. https://doi.org/10.1128/CMR.00127-14
58. Lemee L, Dhalluin A, Pestel-Caron M et al (2004) Multilocus sequence typing analysis of human and animal Clostridium difficile isolates of various toxigenic types. J Clin Microbiol 42:2609–2617. https://doi.org/10.1128/JCM.42.6.2609-2617.2004
59. Muñoz M, Ríos-Chaparro DI, Patarroyo MA, Ramírez JD (2017) Determining Clostridium difficile intra-taxa diversity by mining multilocus sequence typing databases. BMC Microbiol 17:1–12. https://doi.org/10.1186/s12866-017-0969-7
60. Lemee L, Dhalluin A, Testelin S et al (2004) Multiplex PCR targeting tpi (triose phosphate isomerase), tcdA (toxin a), and tcdB (toxin B) genes for toxigenic culture of Clostridium difficile. J Clin Microbiol 42:5710–5714. https://doi.org/10.1128/JCM.42.12.5710-5714.2004
61. Pasanen T, Kotila SM, Horsma J et al (2011) Comparison of repetitive extragenic palindromic sequence-based PCR with PCR ribotyping and pulsed-field gel electrophoresis in studying the clonality of Clostridium difficile. Clin Microbiol Infect 17:166–175. https://doi.org/10.1111/j.1469-0691.2010.03221.x
62. Rupnik M, Braun V, Soehn F et al (1997) Characterization of polymorphisms in the toxin a and B genes of Clostridium difficile. FEMS Microbiol Lett 148:197–202
63. Rupnik M, Brazier JS, Duerden BI et al (2001) Comparison of toxinotyping and PCR ribotyping of Clostridium difficile strains and description of novel toxinotypes. Microbiology 147:439–447. https://doi.org/10.1099/00221287-147-2-439
64. Kuijper EJ, van den Berg RJ, Brazier JS (2009) Comparison of molecular typing methods applied to Clostridium difficile. Methods Mol Biol 551:159–171. https://doi.org/10.1007/978-1-60327-999-4_13
65. Dingle KE, Didelot X, Ansari MA et al (2013) Recombinational switching of the Clostridium difficile S-layer and a novel glycosylation gene cluster revealed by large-scale whole-genome sequencing. J Infect Dis 207:675–686. https://doi.org/10.1093/infdis/jis734
66. Kato H, Kato H, Nakamura M et al (2009) Rapid analysis of Clostridium difficile strains recovered from hospitalized patients by using the slpA sequence typing system. J Infect Chemother 15:199–202. https://doi.org/10.1007/s10156-009-0681-0
67. Kato H, Kato H, Ito Y et al (2010) Typing of Clostridium difficile isolates endemic in Japan by sequencing of slpA and its application to direct typing. J Med Microbiol 59:556–562. https://doi.org/10.1099/jmm.0.016147-0
68. Ng PC, Kirkness EF (2010) Whole genome sequencing BT – genetic variation: methods and protocols. In: Barnes RM, Breen G (eds) Methods in molecular biology (Clifton, N.J.). Humana Press, Totowa, pp 215–226

69. He M, Miyajima F, Roberts P et al (2013) Emergence and global spread of epidemic healthcare-associated Clostridium difficile. Nat Genet 45:109–113. https://doi.org/10.1038/ng.2478
70. Sebaihia M, Thomson NR (2006) Colonic irritation. Nat Rev Microbiol 4:882–883
71. Eyre DW, Cule ML, Griffiths D et al (2013) Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in Clostridium difficile transmission. PLoS Comput Biol 9:e1003059. https://doi.org/10.1371/journal.pcbi.1003059
72. Knetsch CW, Connor TR, Mutreja A et al (2014) Whole genome sequencing reveals potential spread of Clostridium difficile between humans and farm animals in the Netherlands, 2002 to 2011. Euro Surveill 19:20954
73. Steglich M, Nitsche A, von Muller L et al (2015) Tracing the spread of Clostridium difficile Ribotype 027 in Germany based on bacterial genome sequences. PLoS One 10:e0139811. https://doi.org/10.1371/journal.pone.0139811
74. Shariat N, Dudley EG (2014) CRISPRs: molecular signatures used for pathogen subtyping. Appl Environ Microbiol 80:430–439
75. Sebaihia M, Wren BW, Mullany P et al (2006) The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. Nat Genet 38:779–786. https://doi.org/10.1038/ng1830
76. Mullany P, Allan E, Roberts AP (2015) Mobile genetic elements in Clostridium difficile and their role in genome function. Res Microbiol:1–7. https://doi.org/10.1016/j.resmic.2014.12.005
77. Sekulovic O, Garneau JR, Neron A, Fortier L-C (2014) Characterization of temperate phages infecting Clostridium difficile isolates of human and animal origins. Appl Environ Microbiol 80:2555–2563. https://doi.org/10.1128/AEM.00237-14
78. Sekulovic O, Fortier L-C (2016) Characterization of functional prophages in Clostridium difficile. Methods Mol Biol 1476:143–165. https://doi.org/10.1007/978-1-4939-6361-4_11
79. Dobson R (2008) Report names hospitals with highest proportions of deaths related to MRSA and C difficile. BMJ 336:1211
80. Gawlik D, Slickers P, Engelmann I et al (2015) DNA-microarray-based genotyping of Clostridium difficile. BMC Microbiol 15:158. https://doi.org/10.1186/s12866-015-0489-2

Cronobacter

3

Marcelo L. L. Brandão and Stephan Forsythe

3.1 Introduction to the *Cronobacter* Genus

It is important for clinical laboratories, food manufacturers and regulatory authorities that robust and reliable typing schemes are readily available for the emergent bacterial pathogen *Cronobacter*. However, this chapter will initially review the various revisions in our taxonomic understanding of the genus. This is essential in order to understand the accuracy and limitations of various typing methods which are currently available. Overall, the molecular typing of *Cronobacter* spp. has considerably advanced our understanding of the taxonomy, ecology, epidemiology and virulence of this organism. Consequently, we now have reliable and robust methods for accurate microbial source tracking for use both in the food production environment and in epidemiological analysis.

The *Cronobacter* genus belongs to the bacterial class *Gammaproteobacteria* and is within the family *Enterobacteriaceae.* The genus is composed of seven species: *Cronobacter sakazakii*, *C. malonaticus*, *C. turicensis*, *C. muytjensii*, *C. dublinensis*, *C. universalis* and *C. condimenti* [1–3]. A summary timeline of *Cronobacter* taxonomy reclassifications is shown in Fig. 3.1.

The first documented isolation of what would become known as *Cronobacter* spp. was from a can of dried milk in 1950 [4]. In 1980, Farmer et al. [5] proposed the name *Enterobacter sakazakii* for what had been known as 'yellow-pigmented *E. cloacae*', in honour of the Japanese bacteriologist Riichi Sakazaki.

M. L. L. Brandão (✉)
Laboratory of Microbiology Control, Department of Quality Control, Institute of Technology of Immunobiologicals from Oswaldo Cruz Foundation (Bio-Manguinhos/Fiocruz), Rio de Janeiro, RJ, Brazil
e-mail: marcelo.brandao@bio.fiocruz.br

S. Forsythe
Foodmicrobe.com, Nottingham, UK

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_3

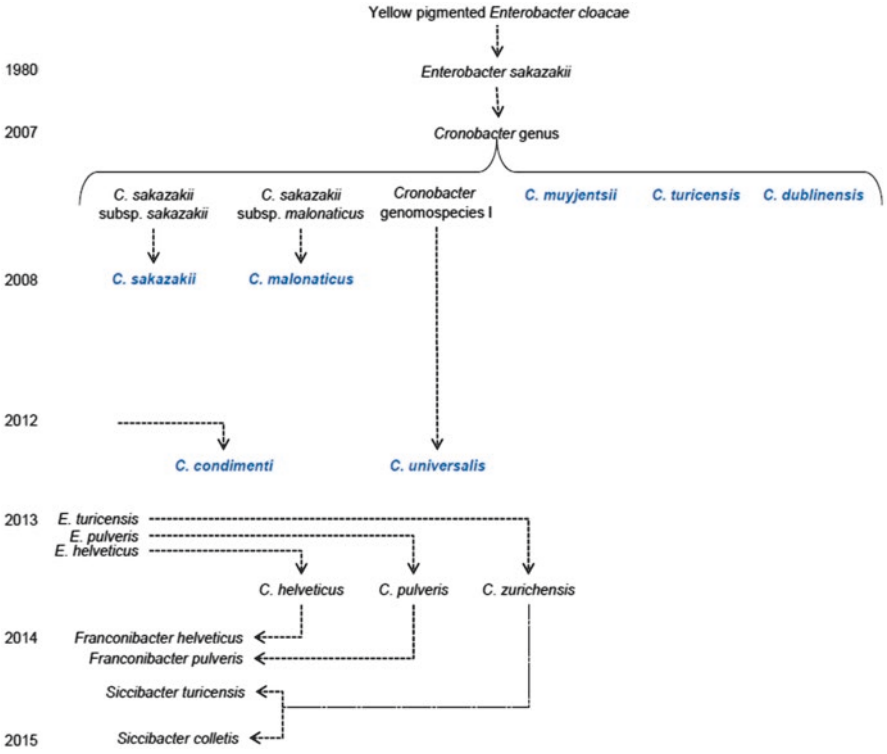

Fig. 3.1 Summary timeline of *Cronobacter* taxonomy reclassifications

This name continued in usage until 2007 when the first description of the *Cronobacter* genus was made [1]. Unfortunately, as the strains had been identified using phenotyping, rather than DNA-based methods, it is uncertain which particular species of *Cronobacter* was referred to in publications prior to 2007. Some strains which would have been misidentified as *E. sakazakii* would have been from the closely related genera such as *Enterobacter*, *Franconibacter* and *Siccibacter* [7–9].

The 16S rDNA sequence analysis, also known as '16S rRNA gene sequence analysis', is a major tool for supporting bacterial taxonomic groupings and identification of bacterial strains. 16S rDNA sequencing has greatly contributed to our initial understanding of the diversity within the *Cronobacter* genus. It has been used to differentiate *Cronobacter* from the closely related and biochemically similar genera already referred to above. However, distinguishing between members within the *Cronobacter* genus is problematic due to the high inter-species similarity of the 16S rRNA gene sequences, which ranges from 97.8% to 99.7%. This is why *C. malonaticus* was initially proposed as a subspecies of *C. sakazakii* (Fig. 3.1) [1]. The overlap of biochemical profiles and a poor correlation between genotypic and phenotypic identification also caused confusion when trying to speciate isolates using biotyping profiles [1, 9–11].

Subsequently, further revisions to the *Cronobacter* taxonomy have used more reliable techniques based on whole genome analysis along with average nucleotide identity (ANI), an in silico alternative to laboratory DNA-DNA hybridization determination for determining species boundary.

Unfortunately, in 2013, Brady et al. [12] used four loci (*atpD*, *gyrB*, *infB* and *rpoB*) to support their proposed reclassification of *Enterobacter helveticus*, *Enterobacter pulveris* and *Enterobacter turicensis* as three new *Cronobacter* species (*Cronobacter helveticus*, *Cronobacter pulveris* and *Cronobacter zurichensis*). However their use of four loci was less than normally used for defining new species and was in error. Stephan et al. [13] undertook more detailed analysis and clarified that these three former *Enterobacter* species instead should be reclassified into two new genera: *Franconibacter* and *Siccibacter*. These latter genera are closely related but separate from *Cronobacter.* This differentiation is important, as *Franconibacter* and *Siccibacter* can be co-isolated from the same samples as *Cronobacter*. In 2015, Jackson et al. [14] described a new species *Siccibacter colletis* from a strain previously identified as *Siccibacter turicensis* (formerly *Cronobacter zurichensis*). The relatedness of *Franconibacter* and *Siccibacter* to *Cronobacter* species using ribosomal multilocus sequence typing is shown in Fig. 3.2.

Another problem is that despite the revised taxonomies being published in the *International Journal of Systematic and Evolutionary Microbiology* which normally equates to accepted status, the reclassification has not yet been updated in the 'List of Prokaryotic Names with Standing in Nomenclature' (LPSN bacterio.net, http://www.bacterio.net/cronobacter.html; last access: 10/18/2019), creating confusion for some researchers.

The *Cronobacter* genus includes facultative anaerobic, Gram-negative, oxidase-negative, catalase-positive, non-spore-forming rods which are generally motile and able to reduce nitrate to nitrite, show positive reaction in the Voges-Proskauer test

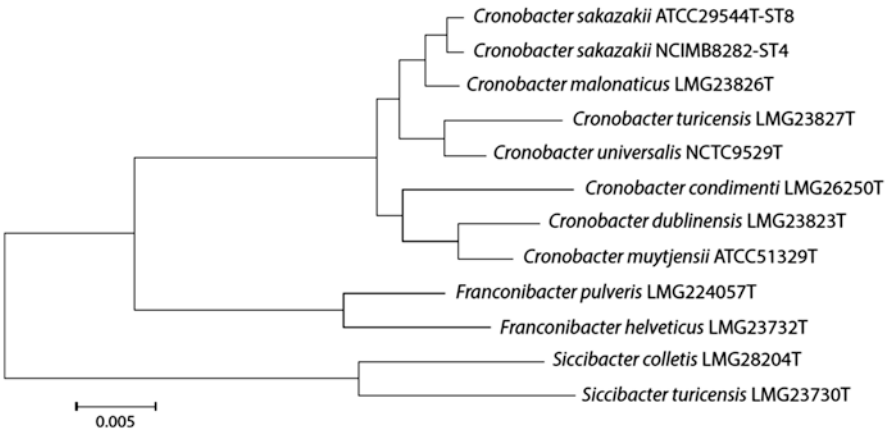

Fig. 3.2 Phylogenetic analysis of the *Cronobacter* genus and closely related organisms using ribosomal multilocus sequence typing (53-loci; 21,195-nt concatenated sequence). The tree was drawn using MEGA 6.05 (http://www.megasoftware.net/) with 1000 bootstrap replicates [15]

and are negative for the methyl red test [1–3]. The capability to endophytically colonize maize roots was demonstrated for several *Cronobacter* strains, providing evidence that plants may be the natural habitat of *Cronobacter* spp. [16]. *Cronobacter* produces capsular material that may facilitate its attachment to plant surfaces, biofilm formation and persistence under desiccated conditions, which may explain the organism's occurrence in food ingredients, and its environmental fitness which contributes to its survival during food processing [17–19]. Although *C. condimenti* has not been isolated from clinical cases of infection, all *Cronobacter* species are considered potential bacterial pathogens [20].

3.2 *Cronobacter* Pathogenicity and Virulence

3.2.1 *Cronobacter* Infections

Cronobacter infections occur in all age groups, albeit with a greater incidence in the very young and elderly, who are generally more immunocompromised [21–24]. The majority of *Cronobacter* spp. infections occur in the adult population but are less severe and therefore have not received so much attention as the outbreaks and cases among neonates [22, 25, 26]. *Cronobacter* infections can result in severe clinical presentations of necrotizing enterocolitis, bacteraemia and meningitis. These can be fatal, and even if the infant survives meningitis, they are likely to be neurologically damaged for life [15, 27–30]. Friedemann [24] reported the lethality of *Cronobacter* meningitis, bacteraemia and NEC to be 41.9% ($P < 0.0001$), <10% and 19.0% ($P < 0.05$), respectively, based on 120–150 microbiologically *Cronobacter*-confirmed neonatal infections between 2000 and 2008.

Contaminated powdered infant formula (PIF) has been epidemiologically linked with many neonatal infections by *Cronobacter* [21, 31–34]. Unlike commercially available 'ready-to-feed' liquid formula, PIF is not a sterile product and must conform to national and international microbiological criteria [35]. It should be noted, however, that such neonatal infections are rare and not all have been associated with the ingestion of reconstituted formula. Some neonatal infections have been associated with exclusively breast-fed infants [36, 37].

Most cases of infections described in the literature are from developed countries, while the situation seems to be considerably different in other countries [24, 27, 38]. This variation can be attributed to sociocultural differences in infant feeding practices, since the use of PIF in neonate feeding is less frequent in developing countries [38, 39]. In addition, many countries do not have a surveillance system for reporting cases of *Cronobacter* spp. infections, which leads to an underestimation of their actual incidence [21].

Infections in older age groups are principally bacteraemias, urosepsis and wound infections. *Cronobacter* has been isolated from different types of clinical syndromes, including pulmonary infections, urinary tract infections and acute cholecystitis [22, 23, 40]. However, in many of these cases, the identification of

Cronobacter as the aetiological agent of the infection is uncertain since other micro-organisms were also identified at the site of infection [38].

Yong et al. [41] reported the first outbreak of acute gastroenteritis strongly related to *C. sakazakii* in healthy adults. This occurred in a local senior high school in China in October 2016. A case-control study including 70 case-patients and 295 controls indicated a strong association between eating supper at school canteen of the outbreak onset and age. The authors identified two different species of *Cronobacter* spp. (*C. sakazakii* and *C. malonaticus*) in the patients' samples. The *C. sakazakii* strains S2 from one patient's rectal swab sample and S4 from a potential contaminated food formed a tightly clustered group using whole genome sequencing and were both identified as sequence typing (ST) 73. The dish of knotted thin sheets of bean curd with braised pork, from which the *C. sakazakii* strain S4 was identified, was only served on 24 October and was shown to be strongly related to the outbreak in the case-control study.

Neonatal infections are predominantly caused by *C. sakazakii*, with *C. malonaticus* attributed to virtually all the remaining cases [15, 23, 25, 38, 42]. There has been one reported case due to *C. turicensis*; however, the isolate was from blood and not from the site of infection and therefore may not have been the causative agent as neonates can be colonized by more than one strain. No other species have been associated with neonatal or infant infections. *C. malonaticus* is more associated with adult infections than other *Cronobacter* species [15, 20, 43].

According to Forsythe [15], *Cronobacter* species can be grouped according to their clinical relevance as group 1, comprising *C. sakazakii* and *C. malonaticus*, which form the majority of clinical isolates in all age groups, and group 2, comprising *C. turicensis* and *C. universalis*, which have been rarely reported, and the other three species (*C. dublinensis*, *C. muytjensii* and *C. condimenti*) are primarily environmental commensals and are probably of little or no clinical significance (Fig. 3.3).

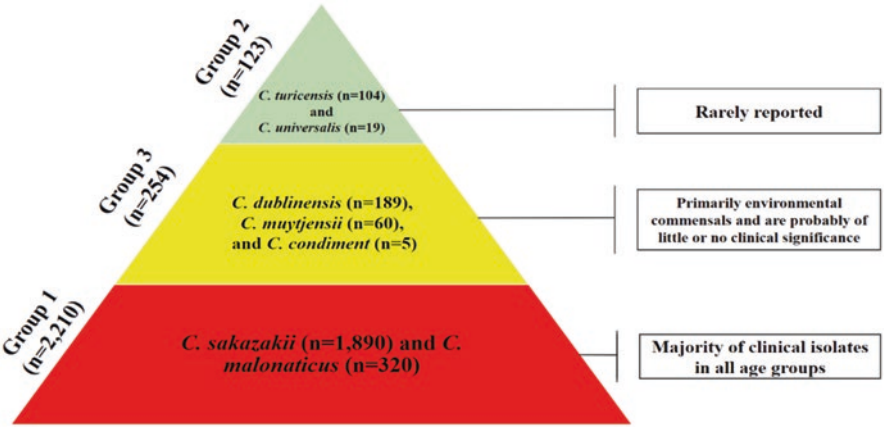

Fig. 3.3 Grouping of *Cronobacter* species according to their clinical relevance. The number of strains deposited in the PubMLST *Cronobacter* database is indicated (last accessed: 09/27/2019)

A number of outbreaks of *Cronobacter* spp. have been reported in neonatal intensive care units [6, 24, 31, 32, 44]. Many of these cases have been directly linked to reconstituted PIF, which may have been contaminated intrinsically or during preparation and administration [32, 44]. A common feature in some of these outbreaks is the opportunity for temperature abuse of the prepared feed, which would permit bacterial growth [6].

Asymptomatic humans (infants and adults) can carry *Cronobacter* [23, 45, 46]. The bacterium has been isolated from mouths of stroke patients [45] and faeces [46]. Also, it has been recovered from the feeding tubes of neonates fed breast milk or ready-to-feed formula, and not PIF [47]. Holy et al. [23] evaluated the data on the incidence of *Cronobacter* spp. from hospital records for the records for 2005-2011 and observed that the majority of *Cronobacter* spp. isolates (n = 91) were from throat swabs (61), followed by urine (5), tracheal aspirates (5), bronchoalveolar lavage (4), cannulae (4) and sputum (3) samples. It reveals a high recovery (63.7% of strains, n = 91) of the organism from children, 1–14 years of age.

Given the human carriage and environmental occurrence of *Cronobacter*, a wide range of plausible sources of the organism need to be investigated during an outbreak and not just the use of PIF. Bowen et al. [36] reported a case of *C. sakazakii* infection caused by the consumption of extrinsically contaminated expressed human milk that led to meningitis, brain necrosis and marked developmental delays in a female infant. Similarly, McMullan et al. [37] described a *C. sakazakii* infection case resulting from the consumption of contaminated expressed breast milk confirmed by WGS, highlighting the potential risks associated with storage and acquisition of expressed breast milk.

It should be noted that infants can be colonized by more than one strain of *Cronobacter*, and therefore, multiple isolates need to be characterized in epidemiological investigations [6, 48]. Recovery of different *Cronobacter* species and strains from the same food sample has also been reported [49–52]. These facts together show the importance of picking multiple colonies from primary isolation plates from food, clinical and environmental samples, especially in cases of outbreak investigations. These findings reinforce the importance of robust and reliable typing schemes readily available for *Cronobacter* spp.

3.2.2 Virulence Traits in *Cronobacter* spp.

The sequencing of *Cronobacter* genomes has revealed an array of plausible virulence traits including adhesins, outer membrane proteins (OMPs), sialic acid utilization (*nanAKT*), proteolytic enzymes (*zpx*), efflux systems (*ibeB*), iron uptake mechanisms, haemolysins (*hly*), plasminogen activator (*cpa*), siderophore-interacting protein (*sip*) and type VI secretion systems (T6SS) [53–57]. Other possible virulence determinants include superoxide dismutase (*sodA*) for macrophage survival [58], flagella [59] and enterotoxin production [60]. The bacterium can attach to intestinal cells and survive in macrophages [58]. OmpA and OmpX possibly have a role in the organism penetrating the blood-brain barrier, though the

mechanism leading to the destruction of the brain cells is unknown and could, in part, be a host response [61]. As will be considered later (Sect. 7.2), the composition of the capsule layer may relate to the clinical symptoms, with neonatal meningitis being associated with strains carrying the K2:CA2:Cell^{+} capsule profile.

Based on virulence factors already described in the literature, Singh et al. [62] proposed a possible model for *Cronobacter sakazakii* infection and pathogenesis (Fig. 3.4).

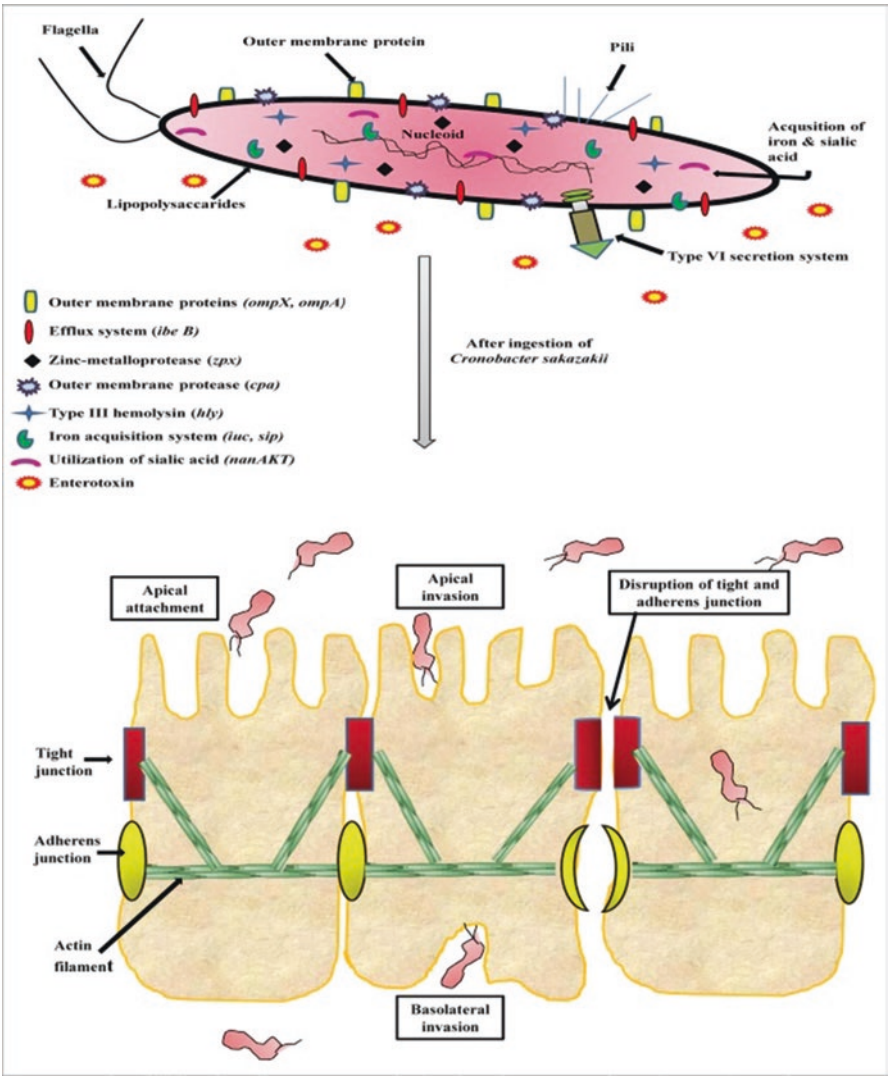

Fig. 3.4 Proposed model for *Cronobacter sakazakii* infection and pathogenesis. The pathogen encodes several pathogenicity-associated factors engaged in imperative processes including adhere to host surfaces, transmigration across, invasion into and disrupt the intestinal barrier within intestinal epithelial cells. Adapted from Singh et al. [62]

3.3 *Cronobacter* in Food Industry

A summary timeline of *Cronobacter* events important for the preventions and control in powdered infant formulae industries is presented in Fig. 3.5.

Before 2002, various outbreaks in neonatal intensive care units and sporadic cases had been reported, and PIF had been identified as a vehicle of contamination in some cases [5, 31, 32]. Consequently, in 2002, the International Commission on Microbiological Specifications for Foods [63] ranked *Cronobacter* spp. (then known as *E. sakazakii*) as a 'severe hazard for restricted populations, life threatening or substantial chronic sequelae or long duration'.

Following the outbreak at the University of Tennessee [32], the Food and Agriculture Organization of the United Nations and the World Health Organization (FAO/WHO) started its three risk assessment evaluations of *Cronobacter* spp. in PIF, powdered follow-up infant formula (FUF) and other infant foods [21, 33, 34]. Their summary recommendations were:

- The use of internationally validated detection and molecular typing methods for *Cronobacter* spp. and other relevant microorganisms should be promoted.
- Investigation and reporting of sources and vehicles, including PIF, and of infection by *Cronobacter* spp. and other relevant microorganisms should be promoted.
- Research should be promoted to gain a better understanding of the ecology, taxonomy, virulence and other characteristics of *Cronobacter* spp. and on ways to reduce its levels in reconstituted PIF.

In 2008, the microbiological criteria applying to PIF were revised by Codex Alimentarius Commission [35], and new criteria were defined for PIF, formula for special medical purposes and human milk fortifiers (Table 3.1).

These criteria are not applied for formula commonly known as 'follow-on formula' or 'follow-up formula' which are used at the weaning stage (>6 months of age). Although there is microbiological evidence of the frequent isolation of

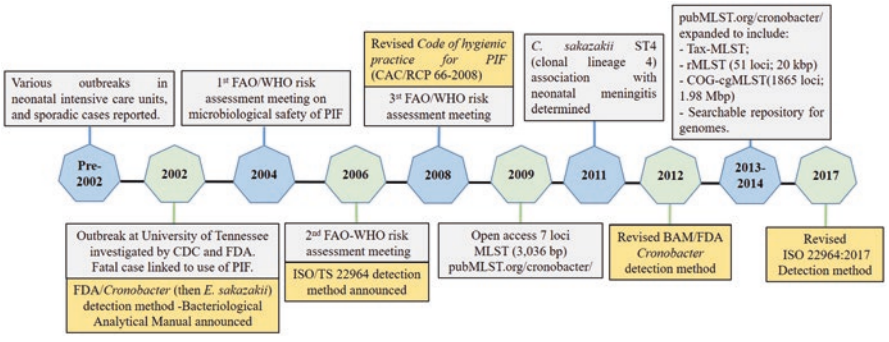

Fig. 3.5 Summary timeline of *Cronobacter* events important for the preventions and control in powdered infant formulae industries

Table 3.1 Codex Alimentarius Commission microbiological criteria for powdered infant formula, formula for special medical purposes and human milk fortifiers (Annex I – CAC, 2008)

Microorganisms	n	c	m	Class plan
Enterobacter sakazakii (*Cronobacter* species)	30	0	0/10 g	2
Salmonella	60	0	0/25 g	2

Cronobacter spp. from such formula and weaning foods, there has been insufficient epidemiological evidence to support the need for additional microbiological testing by manufacturers [34].

After the publication of the Codex Alimentarius Commission [35], the prevalence of *Cronobacter* in PIF and FUF has been determined in samples from many countries, and the reported values vary between 0–12.0% and 0–12.8%, respectively [38, 64–69].

3.4 *Cronobacter* Isolation Methods in PIF and Other Foods

As *Cronobacter* has only been reported at low numbers (<1 cfu/g) in PIF, large volumes need to be tested such that any cells present will be recovered from the desiccated stressed state [21, 33]. Subsequently a resuscitation stage is used as a first stage in isolation. The procedure is primarily for presence/absence testing of *Cronobacter* spp. in PIF, though it can be modified for enumeration by taking multiple samples of different sizes [70].

The two most common international standard methods used for isolation of *Cronobacter* spp. in PIF, FUF and other food products are:

1. The US Food and Drug Administration (USFDA) method for isolation and enumeration of *Cronobacter* from formula samples and confirmation of *Cronobacter* using PCR screening and selective media [70]. A flowchart of the complete procedure is illustrated in Fig. 3.6.
2. The International Organization for Standardization method ISO 22964:2017 – microbiology of the food chain, horizontal method for the detection of *Cronobacter* spp. This method is applicable to (a) food products and ingredients intended for human consumption and the feeding of animals and (b) and environmental samples in the area of food production and food handling. A flowchart of the complete procedure is illustrated in Fig. 3.7.

3.5 *Cronobacter* Identification Methods

Biochemical tests which distinguish *Cronobacter* from closely related organisms, especially those that may be isolated from similar materials, have been described [14]. However, phenotyping is prone to subjectivity and operator bias, and although centralized databases are available with standardized kits, common commercial kits

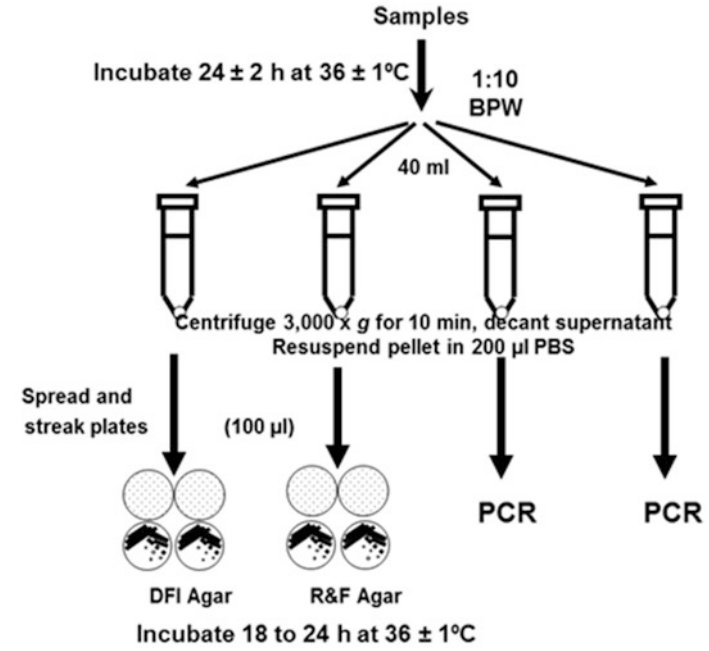

Fig. 3.6 Flowchart of the complete procedure for *Cronobacter* isolation method [70]

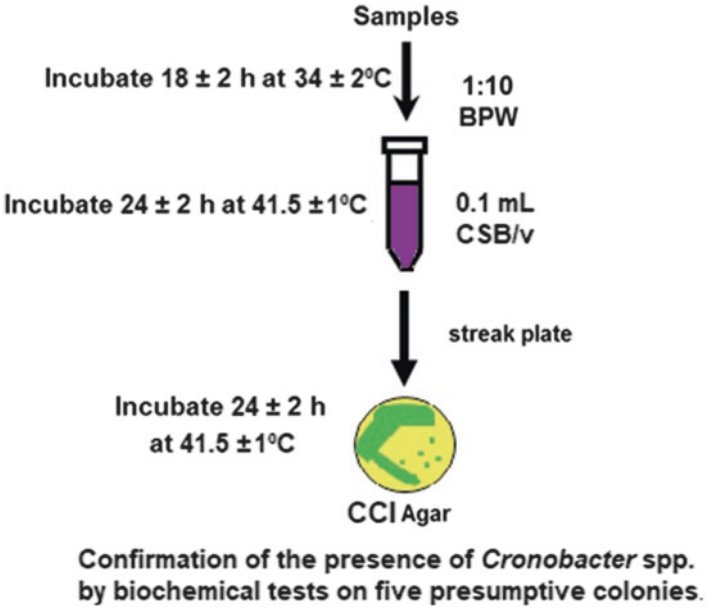

Fig. 3.7 Flowchart of the complete procedure for *Cronobacter* isolation method [71]

for *Enterobacteriaceae* have a poor reliability for *Cronobacter*. Jackson and Forsythe [72] reported that in their study of over 250 strains, the databases supporting the commercial API20E and ID32E kits were only 82.3% and 43.2% reliable, respectively. In addition, false-positive results using Vitek 2.0 have previously been reported due to the misidentification of strains of *Franconibacter helveticus* and *Franconibacter pulveris* as *Cronobacter* [72]. The revised ISO method [71] does not advocate the use of any particular kits but instead recommends specific individual biochemical tests.

Due to the limitations described above, PCR probe-based identification methods were developed and used for the identification and speciation of *Cronobacter* isolates. However, although PCR probes are useful for small-scale studies, their application is limited, as they often have not been validated against a robust *Cronobacter* strain collection of the seven species that represent the diversity of the organism. Since many PCR probe methods were developed prior to the *Cronobacter* taxonomic revisions, they were not necessarily validated using closely related strains from the *Franconibacter* and *Siccibacter* genera [72, 73]. Consequently false-positive and false-negative results occur. One reported *Cronobacter* outbreak originally based on phenotyping and *rpoB* PCR probe was reinvestigated using DNA sequence-based methods (MLST and whole genome sequencing) and was found to be *E. hormaechei* and *E. cloacae* [14]. A summary of these methods and their limitations are presented in Table 3.2.

3.6 Molecular Typing of *Cronobacter* Isolates

3.6.1 Pulsed-Field Gel Electrophoresis (PFGE) Analysis of *Cronobacter* Strains

Brengi et al. [86] proposed a PFGE protocol with *Xba*I as the first choice restriction enzyme and *Spe*I being used for confirmation if deemed appropriate. This protocol was adopted by the CDC PulseNet and has been used for outbreak investigations and source tracking in food production facilities. Mullane et al. [87] used PFGE analysis to show the long-term colonization of a milk protein manufacturing plant, recovering three different pulsotypes of *Cronobacter* over a period of 10 months. PFGE has been also used to profile *Cronobacter* strains isolated from the environment of milk powder and infant formula processing plnts, including air filters [88]. The method has also been applied to *Cronobacter* isolates from various food samples, such as raw dried pasta [89], spices [90], cereal kernels, flour and flour-based products [91].

PFGE has been regarded for many years as an essential tool for microbial source tracking. However, PulseNet lists a number of limitations:

Time-consuming.
Requires a high level of skill.
Does not work for everything (i.e. clonal patterns).

Table 3.2 Comparison of the most common PCR probe-based profiling methods used for *Cronobacter* identification in genus and species loci

		Gene targets	Comments	References
Genus loci	Ribosomal DNA (rDNA)	16S rRNA	–	Iversen et al. [9]
			Real-time PCR method	Malorny and Wagner [74]
		23S rRNA	–	Derzelle et al. [75]
		tRNAGlu	–	Hassan et al. [76]
		16S–23S rRNA		Liu et al. [77]
		FISH	–	Iversen et al. [1]
	1,6-a-glucosidase	*gluA*	Cawthorn et al. [78] reported false-positive results using EsAgf/EsAgf primers (1.680 bp fragment)	Iversen et al. [79]
	MMS operon	*dnaG*	Real-time PCR included in BAM/FDA method [70] applied to the presumptive *Cronobacter* isolates. Brandao et al. [51] reported a false-negative result for a *C. malonaticus* strain, and Vasconcellos et al. [52] reported a false-positive result for an *Enterobacter* spp. strain	Chen et al. [70]
	Zinc-containing metalloprotease	*Zpx*	–	Jaradat et al. [80]
	Outer membrane protein A	*ompA*	Jackson and Forsythe [72] evaluated in silico PCR targeting *ompA* which predicted that amplification would only occur with *Cronobacter* species, and this method may be a feasible alternative to biochemical phenotyping	Nair and Venkitanarayanan [81]
Species loci	β-Subunit of RNA	*rpoB*	Silva et al. [82] reported that the PCR was unable to identify *Cronobacter* isolates to the species level, due to non-amplification and discordant results to *fusA* allele sequences	Stoop et al. [83], Lehner et al. [84]
	Diguanylate cyclase-encoding	*cgcA*	Multiplex PCR for identifying *Cronobacter* species. However, Jackson et al. [73] reported that the *cgcA* target cannot be sufficient for identification of all *Cronobacter* spp., since whole genome sequence studies show that the gene is absent from *C. condimenti* and a wide range of *C. sakazakii* sequence types Vasconcellos et al. [52] and Silva et al. [82] reported that the multiplex PCR was unable to identify *Cronobacter* isolates to the species level, due to non-amplification, unspecific amplifications or discordant results to *fusA* allele sequences	Carter et al. [85]

Pattern results vary from person to person.
Cannot optimize separation in every part of the gel at the same time.
Bands are bands, not sequences.
Do not really know if bands of the same size are the same pieces of DNA.
Bands are not independent.
Change in one restriction site can mean more than one band change.
'Relatedness' should be used as a guide, not true phylogenetic measure.
Some strains are untypable by PFGE.

For these reasons, PulseNet is transitioning towards using whole genome sequence-based analysis and core-genome MLST [92].

The issue referred to by PulseNet of not distinguishing strains with clonal patterns is very pertinent to *Cronobacter* isolates, though often overlooked. Brandao et al. [93] used PFGE for typing three *C. malonaticus* strains isolated from blood culture from an outbreak with only restriction enzyme *Spe*I. The strains were clustered in the same clonal group and could not be differentiated. Further studies using MLST revealed they were ST394 (n = 1) and ST440 (n = 2) [18]. Similarly, Caubilla-Barron et al. [6] reported *C. sakazakii* strains from neonates with necrotizing enterocolitis were not distinguishable by PFGE, but were later shown by whole genome analysis to differ due to latent phage [6].

In summary, the PFGE method has been invaluable for many years. However, it has limitations and is being replaced by DNA sequence-based methods. With respect to *Cronobacter*, PFGE neither speciates isolates nor determines the relatedness of strains. In addition, due to intrinsic DNase activity, some strains do not give profiles and are therefore non-typable [25, 88]. Finally, PFGE may not differentiate between unrelated Cronobacter strains from different countries and different years. This is because of the high level of clonality, particularly within *C. sakazakii* and *C. malonaticus* [1, 25].

3.6.2 PCR-Based Serotyping of *Cronobacter* spp.

The *Cronobacter* capsule is composed of up to five compounds: O-antigen, K-antigen, colanic acid, cellulose and the *Enterobacteriaceae* common antigen [94, 95]. The compositions of these vary and are therefore plausible sites for typing schemes.

The O-antigen, also known as lipopolysaccharide (LPS), is anchored in the outer membrane and elicits an immunological response. The O-antigen (somatic antigen) is a very important component of the cell walls of Gram-negative bacteria, with its variations being responsible for the various serotypes in bacterial species. The chemical structure of the LPS for a number of *Cronobacter* strains has been determined [8, 96, 97].

PCR-based O-serotyping schemes have been designed for *Cronobacter* spp. [98–101]. The serotyping method uses long-range PCR, to generate a PCR

product (size range 9.8–14.8 kbp) which is then restricted using *Mbo*II followed by separation by gel electrophoresis. The restriction fragment length polymorphism (RFLP) banding patterns can be further analysed to generate gene clusters. Unfortunately, the method has been marred due to initial species misidentifications and discrepancies between the initial protocols [98, 99]. Subsequently, Yan et al. [101] proposed a harmonized scheme describing 14 serotypes which covered 6 *Cronobacter* species, but not *C. condimenti*. Afterwards, using modified primers, Blažková et al. [102] identified further serotypes according to their banding patterns.

To date, the *Cronobacter* serotyping scheme has *C. sakazakii* O1–O4, O6 and O7, *C. malonaticus* O1–O5, *C. muytjensii* O1–O3, *C. turicensis* O1–O4, *C. dublinensis* O1–O4, *C. universalis* O1 and *C. condimenti* O1 (Table 3.3). The majority of serotypes (60% of the genus) are *C. sakazakii* O:1 and O:2, and therefore, serotyping does not offer discrimination between a significant portion of isolates. The typing scheme is not complete with 5–20% of strains not giving any PCR products using the current PCR probes [99, 101]. In addition, the same serotype occurs in different *Cronobacter* species and other *Enterobacteriaceae* such as *Citrobacter koseri*, *Franconibacter pulveris* and *E. coli* [103]. This is shown in Table 3.3 where serotypes are ascribed to non-target organisms, for example, *C. universalis* having Csak O:7 and Cmal O1 serotypes. While the method has revealed the considerable diversity of the lipopolysaccharide in *Cronobacter* spp., there are practical issues that have yet to be reduced to a manageable stepwise analysis of *Cronobacter* strains. Currently the method requires a prior knowledge of the *Cronobacter* species in order to select the appropriate PCR primer pairs. In total 15 primer pairs are required to cover 6 of the 7 *Cronobacter* species. Another major limitation is that a large number of strains (up to 20%) cannot be serotyped yet using the current primers described in the literature [97, 99, 101].

3.6.3 Capsular Profiling

Characterization of the capsule composition of Gram-negative bacterial pathogens can be important in developing profiling schemes and has a role in virulence of the organism, for example, *E. coli* K1 and neonatal meningitis, where ‘K’ refers to the capsular antigen. Whole genome sequencing has revealed the variation in the K-antigen, colanic acid and cellulose synthesis genes for the capsule, abbreviated to K, CA and Cell, respectively. There are essentially two variants of the K-antigen (K1 and K2) and two variants of the colanic acid synthesis gene cluster differing in the presence/absence of *galE* (CA1 and CA2), and some strains do not encode for cellulose synthesis (Cell+/−). It is notable that capsular profiling of *Cronobacter* has revealed the only trait which associates a specific profile (K2:CA2:Cell^{+}) with clinical symptoms [94, 95]. In fact, this profile is the primary ‘virulence trait’ for neonatal meningitis which has been identified. *Cronobacter* cases of neonatal

Table 3.3 Summary of *Cronobacter* and non-*Cronobacter* serotype isolates in the PubMLST *Cronobacter* database (last accessed 05/02/19)

Species	Serotype	Number of strains with both serotype and ST	Number of STs
C. sakazakii	Csak O:1	233	41
	Csak O:2	363	50
	Csak O:3	47	18
	Csak O:4	67	12
	Csak O:6	3	3
	Csak O:7	21	11
	Not found	10	6
C. malonaticus	Cmal O:1	22	15
	Cmal O:2	80	21
	Cmal O:3	3	3
	Cmal O:4	0	
	Cmal O:5	5	4
	Ctur O:2	1	1
	Not found	6	4
C. muytjensii	Cmuyt O:1	4	4
	Cmuyt O:2	3	2
	Cmuyt O:3	1	1
C. turicensis	Ctur O:1	16	9
	Ctur O:2	1	1
	Ctur O:3	13	11
	Ctur O:4	2	1
	Csak O:5	1	1
	Not found	4	4
C. dublinensis	Cdub O:1	33	27
	Cdub O:2	26	20
	Cdub O:3	0	
	Cdub O:4	0	
	Not found	11	10
C. universalis	Cuni O:1	6	4
	Cmal O:1, Csak O:7	1	1
C. condimenti	Ccon O:1	[1][a]	1
Citrobacter koseri	Cmal O:1	1	1
	Cmal O:1, Csak O:4	3	1
Franconibacter pulveris	Csak O:4	1	1

[a]Serotype defined based on *gnd* and *galF* loci analysis, not PCR serotyping

meningitis are primarily associated with *C. sakazakii* CC4 which differs from most other *C. sakazakii* STs in encoding for K2:CA2:Cell^{+}. This profile is also found from rare cases of neonatal meningitis caused by *C. malonaticus*, and therefore the capsule profile is not clonally linked, but virulence linked.

3.6.4 Standardized Seven-Loci Multilocus Sequence Typing (MLST)

Typing schemes are required for both epidemiological and environmental investigations of *Cronobacter* and inevitably require the direct comparison of isolates between laboratories around the world. Consequently, standardized DNA sequence-based methods supported by a centralized database are more suitable, as they enable direct unambiguous comparison between isolates being typed in different locations [104].

Suitable DNA sequence typing-based methods for *Cronobacter* started with seven-loci multilocus sequence typing (MLST) [105]. This can be laboratory or whole genome based. More refined typing using whole genome sequences include ribosomal MLST (53 loci), core-genome MLST (1836 loci) and clustered regularly interspersed short palindromic repeat (CRISPR)-*cas* gene array profiling. These methods are supported by a curated, open-access database (http://pubmlst.org/cronobacter), enabling international collaboration and surveillance. The single-nucleotide polymorphism (SNP) analysis can also be applied when whole genome of the strains is available.

The standardized seven-loci PubMLST *Cronobacter* scheme is based on the housekeeping genes for ATP synthase b chain (*atpD*), elongation factor G (*fusA*), glutaminyl-tRNA synthetase (*glnS*), glutamate synthase large subunit (*gltB*), DNA gyrase subunit B (*gyrB*), translation initiation factor IF-2 (*infB*) and phosphoenolpyruvate synthase A (*ppsA*). The protocols and database for >2800 strains of *Cronobacter* and closely related genera are curated and have open access (www.pubMLST.org/cronobacter). The scheme has the potential to distinguish $>7 \times 10^{10}$ different genotypes, therefore making it highly unlikely to obtain identical STs by chance [105]. Within the MLST scheme, none of the *fusA* profiles are shared between two or more species. Consequently, sequencing this locus (438 nt in length) can be used to define the species of *Cronobacter* isolates [43]. When concatenated together, the seven sequences provided 3036 nucleotides for phylogenetic analysis (MLSA). This has shown concordance with our current understanding in the taxonomic groups within the *Cronobacter* genera and closely related organisms and agrees with whole genome sequence analysis [55].

The PubMLST *Cronobacter* site (www.pubMLST.org/cronobacter) is essentially in two sections: protocols and databases. The main protocols repeat the procedures as given in Baldwin et al. [105]; plus there are alternative primer pairs if difficulties are experienced which were designed more recently following whole genome sequence analysis (Jackson, pers. comm.). Essentially the protocols give the initial PCR primer sets and inner sequencing primers for each allele. These normally work across the seven *Cronobacter* species and related organisms: *Franconibacter*, *Siccibacter*, *Citrobacter* and *Enterobacter.* Occasionally a primer may fail, and it is recommended to use the alternative primer sets.

The sequence information are submitted to the database curator (Prof Steve Forsythe) and stored on a central database (http://www.pubMLST/cronobacter). These can either be laboratory-generated sequences for each locus or whole genome

sequence (FASTA file format) from which the seven loci are automatically extracted from by the curator, therefore making the technique electronically portable and reproducible in a laboratory in any part of the world. Currently the database contains the sequences for over 2272 *Cronobacter* strains, including >600 whole genomes. There are >600 defined STs (Table 3.4), considerably more than the 26 of the PCR-based serotyping scheme (Table 3.3).

The majority of strains was isolated from environmental, food and clinical sources are *C. sakazakii*, which has been the main species of study to date (Table 3.4). The PubMLST *Cronobacter* profile database enables searches for single and multiple loci queries, BLAST, similarity searches and download options, for the profiles of all the identified sequence types in the scheme, as well as tools for comparative genomic analysis. The database hosts a plethora of information on the isolates that have been typed, including sequence data and background information of the strains. All this data is stored with open access, enabling convenient downloading and analysis for researchers working on the organism across the world. The portal also has provided links for analysing the sequences with tools to compare alleles and sequence types (STs), detect the presence of linkage and recombination in the population dataset, establish relationships between related STs or isolates by constructing phylogenetic trees or minimum spanning trees, etc. [19].

The frequency of each ST varies with many STs being present at low frequency, and others are more prevalent (e.g. *C. sakazakii* ST4 strains). According to the origins of the analysed strain, the prevalent STs are usually isolated over multiple years and diverse geographical locations. A major advantage of MLST is the further analysis that can be undertaken, unlike most other typing methods. For example, SplitsTree [106], minimum spanning tree [107] and eBURST [108] analysis can be used to show the relatedness of STs and distribution according to country and isolation site. The genetic relationships among all strains deposited in PubMLST *Cronobacter* database (www.pubMLST.org/cronobacter) (last accessed 09.26.2019) are illustrated with a minimum spanning tree constructed using GrapeTree using a categorical coefficient and graphing [107] analysis of 53 clonal complexes (CCs) (Fig. 3.8). The CC4 of *C. sakazakii* is the central genotype, and it was identified as a stable clonal lineage associated with neonatal meningitis [20, 43, 109, 110]. Similarly, *C. sakazakii* ST12 is associated with necrotizing enterocolitis and *C. malonaticus* CC7 with adult infection.

Researchers can compare their MLST profiles with those already in the database. For example, the eBURST algorithm is used to characterize the clonality in the MLST dataset of a bacterial population, by using only the allelic profiles and ST and not the nucleotide sequences. It identifies the most frequent and persistent ST in the population (the one associated with the maximum number of isolates) which is then identified as the 'founder clone'. This founder is then linked to its nearest possible neighbour which is the ST with one allele difference, therefore known as a single-locus variant (SLV). These links result in the formation of clusters of closely linked STs within the population, known as CC. The SLV can also then be further linked to its closest neighbours identified as double-locus variants (DLVs) (differing at two alleles compared to the founder ST) or triple-locus variants or TLVs (differing at

Table 3.4 Summary of *Cronobacter* isolates in the PubMLST *Cronobacter* database (last accessed 14/10/19)

Species	Number of strains (%)	Number of STs[a]	Number of genomes	Earliest isolate	Countries	Clinical	Infant formula[b]	Food and ingredients	Environmental[c]	Others
C. sakazakii	1668 (73.4)	294	484	1950	37	255 (15.8)[d]	404 (25.1)	593 (36.8)	340 (21.1)	18 (1.1)
C. malonaticus	282 (12.4)	110	86	1973	23	67 (23.5	44 (15.4)	129 (45.3)	25 (8.8)	20 (7.0)
C. dublinensis	163 (7.2)	115	52	1956	13	5 (3.1)	11 (6.7)	98 (60.1)	46 (28.2)	3 (1.8)
C. turicensis	92 (4.1)	54	24	1970	14	10 (10.9)	6 (6.5)	38 (41.3)	29 (31.5)	9 (9.8)
C. muytjensii	44 (1.9)	29	14	1988	9	0 (0)	5 (11.4)	21 (47.7)	1 (2.3)	17 (38.6)
C. universalis	19 (0.8)	11	9	1956	8	3 (15.8)	0 (0)	8 (42.1)	5 (26.3)	3 (15.8)
C. condimenti	4 (0.2)	1	3	2010	3	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)
Total	2272	614	672		37[e]	340	470	891	446	70

[a]Sequence type; [b]include powdered infant formula (target age < 6 months) and follow-up formula (target age > 6 months); [c]include environment of food production facilities as well as water and soil sources; [d]percentage of species; [e]unique countries

Fig. 3.8 Genetic relationships among all strains deposited in PubMLST *Cronobacter* database (www.pubMLST.org/cronobacter) (last accessed 09.26.2019). A minimum spanning tree was constructed using GrapeTree analysis of 53 clonal complexes (CC) and 2797 strains. The size of the circle is proportional to the number of the strains. Strains are distributed according to the CC with different colours identified in the figure

three alleles compared to the founder ST). As a result, the overall population snapshot of the dataset in a graphical output represents a forest composed of a number of unrooted trees as well as isolated STs. Each ST in the eBURST algorithm output is denoted by a circle, where the diameter of the circle is proportional to the frequency of the ST [108]. The definition of a 'clonal complex' (CC) depends on the threshold value that is decided upon. The PubMLST *Cronobacter* database by default uses the founder SLV and DLV cluster to define a CC.

Taking into account the stability and portable nature of the MLST scheme, this database has played an important role in our understanding of *Cronobacter* and is expandable due to the WGS repository for future purposes.

3.6.5 Ribosomal MLST (53 Loci)

As an increasing number of bacterial genomes are sequenced, it is plausible that the optimal loci profile will vary in the coming years. Consequently, published *Cronobacter* genomes are now centrally accessible from the *Cronobacter* MLST database, and the BIGSdb facility has been integrated in the 'Genome Comparator' such that users can select loci from over 500 sequenced genomes for population analysis [104].

The whole genomes of *Cronobacter* strains are accessible from the PubMLST *Cronobacter* database which is linked to the ribosomal MLST (rMLST) database as described by Jolley et al. [111]. This can be used to automatically generate the

rMLST profile of the *Cronobacter* isolate according to the 53 *rps* genes [42] through the comparative genomic analysis facility within the database. The corresponding DNA sequences can be concatenated together to generate a phylogenetic tree for the *Cronobacter* genus (Fig. 3.4).

3.6.6 Core-Genome MLST (1836 Loci)

Similar to generating a genotype from a WGS, it is also possible to generate a genotype profile using a core-genome approach. Forsythe et al. [42] demonstrated its application across the *Cronobacter* genus for 1865 loci to profile *Cronobacter* genomes in the database and also generate a phylogenetic analysis across the genome. This was achieved using the comparative genomic analysis facility within the PubMLST *Cronobacter* database. The cg-MLST analysis used *C. sakazakii* ES15 as the reference genome since it is well annotated, but the user can choose their own genome for comparative analysis if they wish. The user can either profile all annotated genes or choose their own interest area, for example, biochemical pathways, adhesins and virulence traits. The output is an Excel™ spreadsheet, and no bioinformatic knowledge or specialist software is required, thus making the approach highly accessible with no usage costs.

3.6.7 CRISPR (Clustered Regularly Interspersed Short Palindromic Repeat)-Cas Array Profiling

MLST has been useful in describing the diversity of the genus and the recognition of pathovars; however, the strong element of clonality may restrict the ability of MLST to distinguish between unrelated strains during microbial source tracking. Previously studies demonstrated that unrelated strains may still be indistinguishable by PFGE and conventional MLST [25, 42]. In order to address this issue, Ogrodzki and Forsythe [94, 95] developed 'clustered regularly interspaced short palindromic repeats' (CRISPRs) and CRISPR-associated gene (*cas*) protein-coding gene (CRISPR-*cas*) array profiling. In brief, CRISPR-*cas* systems may be composed of up to three sections: (a) *cas* genes, (b) an AT-rich leader sequence upstream of the array and (c) a CRISPR array, composed of short (~24–48 nucleotides) direct repeat sequences separated by similarly sized, unique spacers. These spacers are usually derived from mobile genetic elements such as bacteriophages and plasmids. Since CRISPR-*cas* arrays reflect the exposure of strains to phages and plasmids, there should be differences between clonal strains which are unrelated [94, 111–113]. CRISPR arrays may differ between closely related strains due to their different exposure histories to phages and plasmids, leading to differences in their spacer acquisitions.

Therefore, these loci can be used for molecular subtyping, offering greater discrimination between strains than MLST, especially useful for highly clonal species, such as *C. sakazakii*. Ogrodzki and Forsythe [94, 95] analysed the CRISPR-*cas*

arrays of strains from the major pathovars *C. sakazakii* CC1, CC4, CC8 and ST12 and for further comparison of strains from other species across the genus. The CRISPR-*cas* operon architecture is mainly I-E (Ecoli), with a few examples of I-F (Ypseudo). The distribution of the operon types was not phylogenetically related. In general, strains had a large number of CRISPR-*cas* arrays within each genome. Some strains in *C. dublinensis* and *C. muytjensii* lacked *cas* genes. The lack of *cas* genes was also found in the species type strains of *C. universalis* and *C. condimenti*. Strains within the same ST were differentiated according to their CRISPR-*cas* array profiles. An example is reproduced in Fig. 3.9, showing strains in the clonal complex *C. sakazakii* CC4 including those from a neonatal intensive care unit [6].

The CRISPR-*cas* array profiling has been applied to another suspected outbreak in a hospital and showed greater discriminatory power than other genotyping techniques [114]. Zeng et al. [115] genotyped 257 isolates of *C. sakazakii*, *C. malonaticus* and *C. dublinensis* based on CRISPR locus. Results showed that 161 *C. sakazakii* strains could be divided into 129 CRISPR types (CTs), 65 *C. malonaticus* strains were divided into 42 CTs and 31 *C. dublinensis* strains belonged to 31 CTs. There was also a relationship among CT, ST, food types and serotype. Compared to MLST, this method has greater power to distinguish similar strains and had better accordance with WGS.

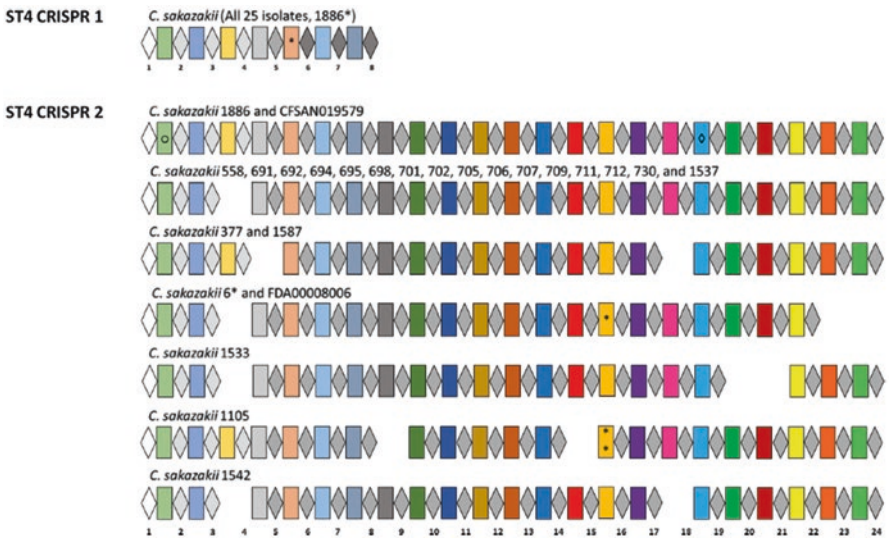

Fig. 3.9 *Cronobacter sakazakii* ST4 CRISPR array profiles. The first array boxes (grey scale) represent the direct repeats. Those with the same shading correspond to the same sequence. Similarly, the second array (coloured) represents the individual spacers between the direct repeats. The same colours correspond to the same sequence within the CRISPR array. Gaps represent the absence of the spacer and its corresponding direct repeat. Ogrodzki and Forsythe [112]

3.6.8 Single-Nucleotide Polymorphism (SNP) Analysis

Typing methods based on whole genome sequencing offer greater discriminatory power than other typing methods, and one means of comparing genomes is by identifying SNPs. These can cause significant changes in phylogenetic distances, but they may not change the PFGE pattern. SNP analysis has been applied by Masood et al. [48] to reanalyse the strains from the French 1994 NICU outbreak reported by Caubilla-Barron et al. [6]. The former study reported there had been three concurrent outbreaks according to the pulsotypes, and the isolates from PIF did not match those from the infants. Masood et al. [48] concurred with the results of PFGE and added better discrimination and relatedness between strains within the same pulsotype. In addition, SNP analysis revealed the *C. sakazakii* ST4 strains differed with those from one baby differing by 300 nt from other ST4 isolates. The source of the *C. sakazakii* could have been from extrinsic contamination of reconstituted PIF from the NICU environment and personnel. This pool of strains would have contributed to the prolonged duration of the outbreak, which was up to 3 months.

For the future, there will be standardization issues for the widespread adoption of SNP analysis for outbreak investigations. These include the need for a comparative reference strain and consistent pipeline for SNP calling and analysis. Currently, a strain (possibly the index strain) can be used within a localized investigation.

To conclude, the advances in next-generation sequencing have greatly contributed to our ability to distinguish *Cronobacter* isolates from closely related genera and also to define genotypes which include pathovars. The common use of MLST has been facilitated by the establishment of a centralized, curated, open-access database for international use.

3.7 Summary

This chapter presents the important characteristics of the bacterium *Cronobacter*, an emergent foodborne pathogen associated with powdered infant formula. The taxonomy history, pathogenicity and virulence of the genus are reviewed, including epidemiological aspects of the seven species according to their clinical relevance. The most common identification and typing methods applied to *Cronobacter* are also discussed.

References

1. Iversen C, Lehner A, Mullane N et al (2007) The taxonomy of *Enterobacter sakazakii*: proposal of a new genus *Cronobacter* gen. nov. and descriptions of *Cronobacter sakazakii* comb. nov. *Cronobacter sakazakii* subsp. *sakazakii*, comb. nov., *Cronobacter sakazakii* subsp. *malonaticus* subsp. nov., *Cronobacter turicensis* sp. nov., *Cronobacter muytjensii* sp. nov., *Cronobacter dublinensis* sp. nov. and *Cronobacter* genomospecies 1. BMC Evol Biol 7:64
2. Iversen C, Mullane N, McCardell B, et al (2008) *Cronobacter* gen., a new genus to accommodate the biogroups of *Enterobacter sakazakii*, and proposal of *Cronobacter sakazakii*

gen. nov., comb. nov., *Cronobacter malonaticus* sp. nov., *Cronobacter turicensis* sp. nov., *Cronobacter muytjensii* sp. nov., *Cronobacter dublinensis* sp. nov., *Cronobacter* genomospecies 1, and of three subspecies, *Cronobacter dublinensis* subsp. *dublinensis* subsp. nov., *Cronobacter dublinensis* subsp. *lausannensis* subsp. nov. and *Cronobacter dublinensis* subsp. *lactaridi* subsp. nov. Int J Syst Evol Microbiol 58(Pt6):1442–1447

3. Joseph S, Cetinkaya E, Drahovska H, et al (2012a) *Cronobacter condiment* sp. nov., isolated from spiced meat and *Cronobacter universalis* sp. nov., a novel species designation for *Cronobacter* sp. genomospecies 1, recovered from a leg infection, water, and food ingredients. Int J Syst Evol Microbiol 62(Pt6):1277–1283
4. Masood N, Moore K, Farbos A et al (2013) Draft genome sequence of the earliest *Cronobacter sakazakii* sequence type 4 strain, NCIMB 8272. Genome Announc 1(5):e00782–e00713
5. Farmer III JJ, Asbury MA, Hickman FW, et al (1980) *Enterobacter sakazakii*: a new species of "Enterobacteriaceae" isolated from clinical specimens. Int J System Bacteriol 30(3):569–584
6. Caubilla-Barron J, Hurrell E, Townsend S et al (2007) Genotypic and phenotypic analysis of *Enterobacter sakazakii* strains from an outbreak resulting in fatalities in a neonatal intensive care unit in France. J Clin Microbiol 45(12):3979–3985
7. Townsend SM, Hurrell E, Caubilla-Barron J et al (2008) Characterization of an extended-spectrum beta-lactamase *Enterobacter hormaechei* nosocomial outbreak, and other *Enterobacter hormaechei* misidentified as *Cronobacter (Enterobacter) sakazakii*. Microbiology 154(Pt12):3659–3667
8. Czerwicka M, Forsythe SJ, Bychowska A et al (2010) Chemical structure of the O-polysaccharide isolated from *Cronobacter sakazakii* 767. Carbohydr Res 345(7):908–913
9. Iversen C, Waddington M, On SL et al (2004) Identification and phylogeny of *Enterobacter sakazakii* relative to *Enterobacter* and *Citrobacter* species. J Clin Microbiol 42(11):5368–5370
10. Iversen C, Waddington M, Farmer JJ III et al (2006) The biochemical differentiation of *Enterobacter sakazakii* genotypes. BMC Microbiol 6:94
11. Joseph S, Hariri S, Forsythe SJ (2013) Lack of continuity between *Cronobacter* biotypes and species as determined using multilocus sequence typing. Mol Cell Probes 27(3–4):137–139
12. Brady C, Cleenwerck I, Venter S et al (2013) Taxonomic evaluation of the genus *Enterobacter* based on multilocus sequence analysis (MLSA): Proposal to reclassify *E. nimipressuralis* and *E. amnigenus* into *Lelliottia* gen. nov. as *Lelliottia nimipressuralis* comb. nov. and *Lelliottia amnigena* comb. nov., respectively, *E. gergoviae* and *E. pyrinus* into *Pluralibacter* gen. nov. as *Pluralibacter gergoviae* comb. nov. and *Pluralibacter pyrinus* comb. nov., respectively, *E. cowanii*, *E. radicincitans*, *E. oryzae* and *E. arachidis* into *Kosakonia* gen. nov. as *Kosakonia cowanii* comb. nov., *Kosakonia radicincitans* comb. nov., *Kosakonia oryzae* comb. nov. and *Kosakonia arachidis* comb. nov., respectively, and *E. turicensis*, *E. helveticus* and *E. pulveris* into *Cronobacter* as *Cronobacter zurichensis* nom. nov., *Cronobacter helveticus* comb. nov. and *Cronobacter pulveris* comb. nov., respectively, and emended description of the genera *Enterobacter* and *Cronobacter*. System Appl Microbiol 36(5):309–319
13. Stephan R, Grim CJ, Gopinath GR et al (2014) Re-examination of the taxonomic status of *Enterobacter helveticus*, *Enterobacter pulveris* and *Enterobacter turicensis* as members of the genus *Cronobacter* and their reclassification in the genera *Franconibacter* gen. nov. and *Siccibacter* gen. nov. as *Franconibacter helveticus* comb. nov., *Franconibacter pulveris* comb. nov. and *Siccibacter turicensis* comb. nov., respectively. Int J Syst Evol Microbiol 64(Pt10):3402–3410
14. Jackson EE, Masood N, Ibrahim K et al (2015) Description of *Siccibacter colletis* sp. nov., a novel species isolated from plant material, and emended description of *Siccibacter turicensis*. Int J Syst Evol Microbiol 65(Pt 4):1335–1341
15. Forsythe SJ (2018) Updates on the *Cronobacter* genus. Annu Rev Food Sci and Technol 9:23–44
16. Schmid M, Iversen C, Gontia I et al (2009) Evidence for a plant-associated natural habitat for *Cronobacter* spp. Res Microbiol 160(8):608–614
17. Osaili T, Forsythe S (2009) Desiccation resistance and persistence of *Cronobacter* species in infant formula. Int J Food Microbiol 136(2):214–220

18. Umeda NS, de Filippis I, Forsythe SJ et al (2017) Phenotypic characterization of *Cronobacter* spp. strains isolated from foods and clinical specimens in Brazil. Food Res Int 102:61–67
19. Forsythe S (2018) Microbial source tracking of Cronobacter spp. Adv Appl Microbiol 103:49–101
20. Joseph S, Sonbol H, Hariri S et al (2012) Diversity of the *Cronobacter* genus as revealed by multi locus sequence typing. J Clin Microbiol 50(9):3031–3039
21. Food and Agriculture Organization/World Health Organization (2004) *Enterobacter sakazakii* and other microorganisms in powdered infant formula, meeting report, MRA series 6. World Health Organization, Geneva. Available at http://www.who.int/foodsafety/publications/micro/mra6/en/index.html
22. Patrick ME, Mahon BE, Greene SA et al (2014) Incidence of *Cronobacter* spp. infections, United States, 2003–2009. Emerg Infect Dis 20(9):1520–1523
23. Holy O, Petrzelova J, Hanulik V et al (2014) Epidemiology of *Cronobacter* spp. isolates from hospital patients from the University Hospital Olomouc (Czech Republic). Epidemiol Mikrobiol Immunol 63(1):69–72
24. Friedemann M (2009) Epidemiology of invasive neonatal *Cronobacter (Enterobacter sakazakii)* infections. Eur J Clin Microbiol Infect Dis 28(11):1297–1304
25. Alsonosi A, Hariri S, Kajsik M et al (2015) The speciation and genotyping of *Cronobacter* isolates from hospitalised patients. Eur J Clin Microbiol Infect Dis 34(10):1979–1988
26. Lepuschitz S, Ruppitsch W, Pekard-Amenitsch S et al (2019) Multicenter study of *Cronobacter sakazakii* infections in humans, Europe, 2017. Emerg Infect Dis 25(3):515–522
27. Bowen AB, Braden CR (2006) Invasive *Enterobacter sakazakii* disease in infants. Emerg Infect Dis 12(8):1185–1189
28. Holy O, Forsythe SJ (2014) *Cronobacter* species as emerging causes of healthcare-associated infection. J Hosp Infect 86(3):169–177.
29. Chaves CEV, Brandão MLL, Lacerda MLGG (2018) Fatal *Cronobacter sakazakii* sequence type 494 meningitis in a Newborn, Brazil. Emerg Infect Dis 24(10):1948–1950
30. Zeng H, Lei T, He W et al (2018) Novel multidrug-resistant *Cronobacter sakazakii* causing meningitis in neonate, China, 2015. Emerg Infect Dis 24(11):2121–2124
31. van Acker J, de Smet F, Muyldermans G et al (2001) Outbreak of necrotizing enterocolitis associated with *Enterobacter sakazakii* in powdered milk formula. J Clin Microbiol 39(1):293–297
32. Himelright I, Harris E, Lorch V, et al (2002) *Enterobacter sakazakii* infections associated with the use of powdered infant formula, Tennessee, 2001. MMWR Morb Mortal Wkly Rep 51(14):297–300
33. Food and Agriculture Organization/World Health Organization (2006) *Enterobacter sakazakii* and *Salmonella* in powdered infant formula. Second risk assessment workshop. Meeting report, MRA series 10. World Health Organization, Geneva. http://www.who.int/foodsafety/publications/micro/mra10/en/index.html
34. Food and Agriculture Organization/World Health Organization (2008) *Enterobacter sakazakii* (*Cronobacter* spp.) in powdered follow-up formulae. MRA series. Available at http://www.who.int/foodsafety/publications/micro/MRA_followup.pdf
35. Codex Alimentarius Commission (2008) Code of hygienic practice for powdered formulae for infants and young children. CAC/RCP 66–2008. Available at http://www.codexalimentarius.net/download/standards/11026/CXP_066e.pdf
36. Bowen A, Wiesenfeld HC, Kloesz JL et al (2017) Notes from the field: *Cronobacter sakazakii* infection associated with feeding extrinsically contaminated expressed human milk to a premature infant – Pennsylvania, 2016. MMWR Morb Mortal Wkly Rep 66(28):761–762
37. McMullan R, Menon V, Beukers AG, et al (2018) *Cronobacter sakazakii* infection from expressed breast milk, Australia. Emerg Infect Dis 24(2):393–394
38. Brandao MLL, Umeda NS, Filippis I (2018) *Cronobacter* spp.: infecções, ocorrência e regulação em alimentos – uma revisão no Brasil. Braz J Food Technol 21:e2017074
39. Ray P, Das A, Gautam V, et al (2007) *Enterobacter sakazakii* in infants: novel phenomenon in India. Indian J Medical Microbiol 25(4):408–410

40. Tsai HY, Liao CH, Huang YT, et al (2013) *Cronobacter* infections not from infant formula, Taiwan. Emerg Infec Dis 19(1):167–169
41. Yong W, Guo B, Shi X et al (2018) An investigation of an acute gastroenteritis outbreak: *Cronobacter sakazakii*, a potential cause of food-borne illness. Front Microbiol 9:2549
42. Forsythe SJ, Dickins B, Jolley KA (2014) *Cronobacter*, the emergent bacterial pathogen *Enterobacter sakazakii* comes of age; MLST and whole genome sequence analysis. BMC Genomics 15:e1121
43. Joseph S, Forsythe SJ (2012) Insights into the emergent bacterial pathogen *Cronobacter* spp., generated by multilocus sequence typing and analysis. Front Microbiol 3:397
44. Henry M, Fouladkhah A (2019) Outbreak history, biofilm formation, and preventive measures for control of *Cronobacter sakazakii* in infant formula and infant care settings. Microorganisms 7(3):pii:E77
45. Gosney MA, Martin MV, Wright AE, et al (2006) *Enterobacter sakazakii* in the mouths of stroke patients and its association with aspiration pneumonia. Eur J Intern Med 17(3):185–188
46. Kim JB, Cho SH, Park YB et al (2008) Surveillance of stool samples for the presence of *Enterobacter sakazakii* among Korean people. Yonsei Med J 49(6):1017–1022
47. Hurrell E, Kucerova E, Loughlin M et al (2009) Neonatal enteral feeding tubes as loci for colonisation by members of the *Enterobacteriaceae*. BMC Infect Dis 9:146
48. Masood N, Moore K, Farbos A et al (2015) Genomic dissection of the 1994 *Cronobacter sakazakii* outbreak in a French neonatal intensive care unit. BMC Genomics 16:750
49. Iversen C, Druggan P, Schumacher S et al (2008) Development of a novel screening method for the isolation of "*Cronobacter*" spp. (*Enterobacter sakazakii*). Appl Environ Microbiol 74(8):2550–2553
50. Cetinkaya E, Joseph S, Ayhan K et al (2013) Comparison of methods for the microbiological identification and profiling of *Cronobacter* species from ingredients used in the preparation of infant formula. Mol Cell Probes 27(1):60–64
51. Brandao MLL, Umeda NS, Jackson E et al (2017) Isolation, molecular and phenotypic characterization, and antibiotic susceptibility of *Cronobacter* spp. from Brazilian retail foods. Food Microbiol 63:129–138
52. Vasconcellos L, Carvalho CT, Tavares RO et al (2018) Isolation, molecular and phenotypic characterization of *Cronobacter* spp. in ready-to-eat salads and foods from Japanese cuisine commercialized in Brazil. Food Res Int 107:353–359
53. Kucerova E, Clifton SW, Xia XQ et al (2010) Genome sequence of *Cronobacter sakazakii* BAA-894 and comparative genomic hybridization analysis with other *Cronobacter* species. PLoS One 5(3):e9556
54. Franco AA, Hu L, Grim CJ et al (2011) Characterization of putative virulence genes on the related RepFIB plasmids harbored by *Cronobacter* spp. Appl Environ Microbiol 77:3255–3267
55. Joseph S, Desai P, Li Y et al (2012) Comparative analysis of genome sequences covering the seven *Cronobacter* species. PLoS One 7(11):e49455
56. Bao X, Yang L, Chen L et al (2017) Analysis on pathogenic and virulent characteristics of the *Cronobacter sakazakii* strain BAA-894 by whole genome sequencing and its demonstration in basic biology science. Microb Pathog 109:280–286
57. Aly MA, Domig KJ, Kneifel W et al (2019) Whole genome sequencing-based comparison of food isolates of *Cronobacter sakazakii*. Front Microbiol 10:1464
58. Townsend S, Hurrell E, Forsythe S (2008) Virulence studies of *Enterobacter sakazakii* isolates associated with a neonatal intensive care unit outbreak. BMC Microbiol 8:64
59. Cruz AL, Rocha-Ramirez LM, Gonzalez-Pedrajo B et al (2011) Flagella from *Cronobacter sakazakii* induced an inflammatory response in human monocytes. Cytokine 56:95
60. Pagotto FJ, Nazarowec-White M, Bidawid S, et al (2003) *Enterobacter sakazakii*: infectivity and enterotoxin production *in vitro* and *in vivo*. J Food Protect 66(3):370–375
61. Kim K, Kim KP, Choi J et al (2010) Outer membrane proteins A (*OmpA*) and X (*OmpX*) are essential for basolateral invasion of *Cronobacter sakazakii*. Appl Environ Microbiol 76(15):5188–5198

62. Singh N, Goel G, Raghav M (2015) Insights into virulence factors determining the pathogenicity of *Cronobacter sakazakii*. Virulence 6(5):433–440
63. International Commission on Microbiological Specification for Foods (2002) Microbiological testing in food safety management, vol 7. Academic/Plenum Publisher, New York
64. Chap J, Jackson P, Siqueira R et al (2009) International survey of *Cronobacter sakazakii* and other *Cronobacter* spp. in follow up formulas and infant foods. Int J Food Microbiol 136(2):185–188
65. Santos RFS, da Silva N, Junqueira VCA et al (2013) Screening for *Cronobacter* species in powdered and reconstituted infant formulas and from equipment used in formula preparation in maternity hospitals. Ann Nutr Metab 63(1–2):62–68
66. Pan Z, Cui J, Lyu G et al (2014) Isolation and molecular typing of *Cronobacter* spp. in commercial powdered infant formula and follow-up formula. Foodborne Pathog Dis 11(6):456–461
67. Pei XY, Yan L, Zhu JH et al (2016) The survey of *Cronobacter* spp. (formerly *Enterobacter sakazakii*) in infant and follow-up powdered formula in China in 2012. Biomed Environ Sci 29(2):99–106
68. Zhang H, Hou P, Lv H et al (2017) Surveillance and molecular typing of *Cronobacter* spp. in commercial powdered infant formula and follow-up formula from 2011 to 2013 in Shandong Province, China. J Sci Food Agric 97(7):2141–2146
69. Parra-Flores J, Cerda-Leal F, Contreras A, et al (2018) *Cronobacter sakazakii* and microbiological parameters in dairy formulas associated with a food alert in Chile. Front Microbiol 9:1708.
70. Chen Y, Lampel K, Hammack K (2012) Bacteriological analytical manual chapter 29. Cronobacter, U.S. Food and Drug Administration
71. International Organization for Standardization (2017). Microbiology of the food chain: Horizontal method for the detection of *Cronobacter* spp. ISO rep. TS22964 Geneva, Switzerland: International Organization for Standards
72. Jackson EE, Forsythe SJ (2016) Comparative study of *Cronobacter* identification according to phenotyping methods. BMC Microbiol 16:146
73. Jackson EE, Sonbol H, Masood N et al (2014) Genotypic and phenotypic characteristics of *Cronobacter* species, with particular attention to the newly reclassified species *C. helveticus*, *C. pulveris*, and *C. zurichensis*. Food Microbiol 44:226–235
74. Malorny B, Wagner M (2005) Detection of *Enterobacter sakazakii* strains by real-time PCR. J Food Protect 68(8):1623–1627
75. Derzelle S, Dilasser F, Maladen V et al (2007) Comparison of three chromogenic media and evaluation of two molecular-based identification systems for the detection of *Enterobacter sakazakii* from environmental samples from infant formulae factories. J Food Prot 70(7):1678–1684
76. Hassan AA, Akineden O, Kress C et al (2007) Characterization of the gene encoding the 16S rRNA of *Enterobacter sakazakii* and development of a species-specific PCR method. Int J Food Microbiol 116(2):214–220
77. Liu Y, Gao Q, Zhang X et al (2006) PCR and oligonucleotide array for detection of *Enterobacter sakazakii* in infant formula. Mol Cell Probes 20(1):11–17
78. Cawthorn DM, Botha S, Witthuhn RC (2008) Evaluation of different methods for the detection and identification of *Enterobacter sakazakii* isolated from South African infant formula milks and the processing environment. Int J Food Microbiol 127(1–2):129–138
79. Iversen C, Lehner A, Mullane N et al (2007) Identification of "*Cronobacter*" spp. (*Enterobacter sakazakii*). J Clin Microbiol 45(11):3814–3816
80. Jaradat ZW, Ababneh QO, Saadoun IM et al (2009) Isolation of *Cronobacter* spp. (formerly *Enterobacter sakazakii*) from infant food, herbs and environmental samples and the subsequent identification and confirmation of the isolates using biochemical, chromogenic and molecular methods. BMC Microbiol 9:225

81. Nair MK, Venkitanarayanan KS (2006) Cloning and sequencing of the *ompA* gene of *Enterobacter sakazakii* and development of an *ompA*-targeted PCR for rapid detection of *Enterobacter sakazakii* in infant formula. Appl Environ Microbiol 72(4):2539–2546
82. Silva JN, Vasconcellos L, Forsythe SJ et al (2019) Molecular and phenotypical characterization of *Cronobacter* species isolated with high occurrence from oats and linseeds. FEMS Microbiol Lett 366(1)
83. Stoop B, Lehner A, Iversen C et al (2009) Development and evaluation of *rpoB* based PCR systems to differentiate the six proposed species within the genus *Cronobacter*. Int J Food Microbiol 136(2):165–168
84. Lehner A, Fricker-Feer C, Stephan R (2012) Identification of the recently described *Cronobacter condimenti* by an rpoB-gene-based PCR system. Eur J Intern Med 61(Pt7):1034–1035
85. Carter L, Lindsey LA, Grim CJ et al (2013) Multiplex PCR assay targeting a diguanylate cyclase-encoding gene, *cgcA*, to differentiate species within the genus *Cronobacter*. Appl Environ Microbiol 79(2):734–737
86. Brengi SP, O'Brien SB, Pichel M et al (2012) Development and validation of a PulseNet standardized protocol for subtyping isolates of *Cronobacter* species. Foodborne Pathog Dis 9(9):861–867
87. Mullane N, Healy B, Meade J et al (2008) Dissemination of *Cronobacter* spp. (*Enterobacter sakazakii*) in a powdered milk protein manufacturing facility. Appl Environ Microbiol 74(19):5913–5917
88. Craven HM, McAuley CM, Duffy LL et al (2010) Distribution, prevalence and persistence of *Cronobacter* (*Enterobacter sakazakii*) in the nonprocessing and processing environments of five milk powder factories. J Appl Microbiol 109(3):1044–1052
89. Akineden Ö, Murata KJ, Gross M et al (2015) Microbiological quality of raw dried pasta from the German market, with special emphasis on *Cronobacter* species. J Food Sci 80(12):M2860–M2867
90. Liu M, Hu G, Shi Y et al (2018) Contamination of *Cronobacter* spp. in Chinese retail spices. Foodborne Pathog Dis 15(10):637–644
91. Lou X, Liu T, Zhang W et al (2019) The occurrence and distribution characteristics of *Cronobacter* in diverse cereal kernels, flour, and flour-based products. Food Microbiol 84:103269
92. Nadon C, Van Walle I, Gerner-Smidt P et al (2017) PulseNet international: vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill 22(23):pii30544
93. Brandao MLL, Umeda NS, Carvalho KR et al (2015) Investigação de um surto causado por *Cronobacter malonaticus* em um hospital maternidade em Teresina, Piauí: caracterização e tipificação por eletroforese em gel de campo pulsado. Vigil Sani Debate 3(3):91–96
94. Ogrodzki P, Forsythe S (2015) Capsular profiling of the *Cronobacter* genus and the association of specific *Cronobacter sakazakii* and *C. malonaticus* capsule types with neonatal meningitis and necrotizing enterocolitis. BMC Genomics 6:758
95. Ogrodzki P, Forsythe SJ (2017) DNA-sequence based typing of the *Cronobacter* genus using MLST, CRISPR-cas array and capsular profiling. Front Microbiol 8:1875
96. Czerwicka M, Marszewska K, Forsythe SJ et al (2013) Chemical structure of the O-polysaccharides isolated from *Cronobacter turicensis* sequence type 5 strains 57, 564 and 566. Carbohydr Res 373:89–92
97. Wang Q, Forsythe SJ, Zhao XJ et al (2019) Species identification and molecular characterization of *Cronobacter* spp. isolated from food imported over nine years into Beijing, China. Food Microbiol 82:11–19
98. Mullane NP, O'Gaora P, Nally JE et al (2008) Molecular analysis of the *Enterobacter sakazakii* O-antigen gene locus. Appl Environ Microbiol 74(12):3783–3794
99. Jarvis KG, Grim CJ, Franco AA et al (2011) Molecular characterization of *Cronobacter* lipopolysaccharide O-antigen gene clusters and development of serotype-specific PCR assays. Appl Environ Microbiol 77(12):4017–4026

100. Sun Y, Wang M, Liu H et al (2011) Development of an O-antigen serotyping scheme for *Cronobacter sakazakii*. Appl Environ Microbiol 77(7):2209–2214
101. Yan Q, Jarvis KG, Chase HR et al (2015) A proposed harmonized LPS molecular-subtyping scheme for *Cronobacter* species. Food Microbiol 50:38–43
102. Blažková M, Javůrková B, Vlach J et al (2015) Diversity of O antigens within the genus *Cronobacter*: from disorder to order. Appl Environ Microbiol 81(16):5574–5582
103. Shashkov AS, Wang M, Turdymuratov EM et al (2015) Structural and genetic relationships of closely related O-antigens of *Cronobacter* spp. and *Escherichia coli*: *C. sakazakii* G2594 (serotype O4)/*E. coli* O103 and *C. malonaticus* G3864 (serotype O1)/*E. coli* O29. Carbohydr Res 404:124–131
104. Maiden MC, van Rensburg MJ, Bray JE et al (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 11(10):728–736
105. Baldwin A, Loughlin M, Caubilla-Barron J et al (2009) Multilocus sequence typing of *Cronobacter sakazakii* and *Cronobacter malonaticus* reveals stable clonal structures with clinical significance which do not correlate with biotypes. BMC Microbiol 9:223
106. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23(2):254–267
107. Zhou Z, Alikhan NF, Sergeant MJ et al (2018) GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res 28(9):1395–1404
108. Feil EJ, Li BC, Aanensen DM et al (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186(5):1518–1530
109. Joseph S, Forsythe SJ (2011) Predominance of *Cronobacter sakazakii* sequence type 4 in neonatal infections. Emerg Infect Dis 17(9):1713–1715
110. Hariri S, Joseph S, Forsythe SJ (2013) *Cronobacter sakazakii* ST4 strains and neonatal meningitis, United States. Emerg Infect Dis 19(1):175–177
111. Jolley KA, Bliss CM, Bennett JS et al (2012) Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology 158(Pt4):1005–1015
112. Ogrodzki P, Forsythe SJ (2016) CRISPR-cas loci profiling of *Cronobacter sakazakii* pathovars. Future Microbiol 11:1507–1519
113. Zeng H, Zhang J, Li C et al (2017) The driving force of prophages and CRISPR-Cas system in the evolution of *Cronobacter sakazakii*. Sci Rep 7:e40206
114. Kadlicekova V, Kajsik M, Soltys K et al (2017) Characterisation of *Cronobacter* strains isolated from hospitalised adult patients. Antonie Van Leeuwenhoek 111(7):1073–1085
115. Zeng H, Li C, He W, et al (2019) *Cronobacter sakazakii*, *Cronobacter malonaticus*, and *Cronobacter dublinensis* genotyping based on CRISPR locus diversity. Front Microbiol 10:1989

Oral and Intestinal *Bacteroidetes*

4

Marina C. Claros, Zaida C. Claros, Sareh Said Yekta-Michael, and Georg Conrads

4.1 Introduction

Most of the obligately anaerobic, Gram-negative, non-spore-forming rods of clinical relevance belong to the phylum *Bacteroidetes*, the order *Bacteroidales*, including families *Bacteroidaceae*, *Porphyromonadaceae*, *Prevotellaceae*, and *Rikenellaceae*, with all of them consisting of several genera. We will focus on the genera *Bacteroides*, *Parabacteroides*, *Porphyromonas*, and *Prevotella*, and all four are discussed here but using and explaining different typing methods, exemplarily. As originally isolated from Bacteroides Bile Esculin (BBE) agar and thus traditionally co-investigated with *Bacteroides*, the urease- and catalase-positive, nitrate-reducing anaerobic Gram-negative species *Bilophila wadsworthia*, even though a member of *Deltaproteobacteria* (*Desulfovibrionaceae*), will also be subjected here.

M. C. Claros
Institute of Medical Microbiology and Infectious Epidemiology, University of Leipzig, Leipzig, Germany

Z. C. Claros
Department of Cardiothoracic and Vascular Surgery, German Heart Center Berlin, Berlin, Germany

S. S. Yekta-Michael
Department of Operative and Preventive Dentistry & Periodontology, RWTH Aachen University Hospital, Aachen, Germany

G. Conrads (✉)
Division of Oral Microbiology and Immunology, Department of Operative and Preventive Dentistry & Periodontology, RWTH Aachen University Hospital, Aachen, Germany
e-mail: gconrads@ukaachen.de

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_4

4.2 The Genus *Bacteroides*

Within the family *Bacteroidaceae*, the genus *Bacteroides* consists of saccharolytic, bile-resistant, and nonpigmented species, mainly isolated from the gut. Since the late 1980s, the genus is limited to species within the *Bacteroides fragilis* group [1, 2], and most of the other clinically relevant species became placed in the genus *Porphyromonas* or *Prevotella* [3]. The genus *Bacteroides* currently includes more than 50 species, of which more than half have been recovered from humans (for an update, see www.bacterio.net). Of these, especially *B. fragilis*, *B. thetaiotaomicron*, and *B. ovatus* are most relevant in human infections. The list of approved species within the genus *Bacteroides* changes frequently, and keeping up with all relevant taxonomic revisions is quite a challenge. For instance, almost half of all "*Bacteroides*" species ever described (which are 100 plus 5 subspecies) have moved to other genera over time. However, these changes are of importance both to clinicians and to clinical microbiologists, since taxonomic placement can be an indicator of virulence potential or antimicrobial resistance. Often by using culture-independent approaches such as 16S rRNA gene sequencing, a variety of new species have added to the total number of *Bacteroides* species which were isolated from various clinical sites by a culturomic approach; examples include *Bacteroides ihuae*, *B. mediterraneensis*, *B. neonati*, and *B. timonensis* [4–6].

As *Bacteroides* (*B. fragilis*) may have both a good and a bad nature, molecular typing aims to differentiate between physiological and pathogenic strains. The pathogenicity of *B. fragilis* is related to the "*B. fragilis* pathogenicity island or BfPAI," producing the enterotoxin, which is a zinc metalloprotease [7]. For historical background, in the mid-1980s, it was recognized that some *B. fragilis* strains produce an enterotoxin (ET) that can cause acute diarrhea in humans, young lambs, calves, pigs, and foals [8]. Later, enterotoxigenic *B. fragilis* (ETBF) strains have also been isolated from the feces of children with diarrhea [9, 10]. Kato et al. [11] showed that *B. fragilis* blood culture isolates were more likely to be ETBF and suggested that ET-positive strains are more virulent than ET-negative strains. The corresponding enterotoxin gene (*bft*) was cloned, sequenced, and identified as producing a zinc metalloproteinase with the size of 44.4 kDa [7]. The *bft* gene is located in a 6 kb genetic element termed the *B. fragilis* pathogenicity island (BfPAI). In our studies [12] and unpublished data, it was determined that the incidence of ETBF in different clinical isolates was 11–23%. The prevalence of ETBF among blood culture isolates (23%) was higher than from other specimens, especially the physiological gut isolates. Appendicitis and peritonitis are typical clinical *Bacteroides*-related cases but which often demonstrate mixed infections with *Enterobacteriaceae* (not subjected here) and *Bilophila* [12].

4.2.1 The Genus *Bilophila*

Bilophila (with a single species, *B. wadsworthia*) was first described by Baron et al. as an asaccharolytic, Gram-negative, bile-resistant, strong catalase-positive bacillus

that is often urease positive (approximately 75% of strains) and able to reduce nitrate to nitrite. The G + C content is 39–40 mol% [13, 14]. Growth is stimulated by taurine, a cysteine derivative and major organic solute in humans, which it uses as a source of sulfite and as a terminal acceptor for electron transport [15]. Phylogenetically, the genus *Bilophila* is located in the *Deltaproteobacteria* (*Desulfovibrionaceae*). Several virulence factors such as abscess formation, endotoxin, cytotoxicity, and adherence as well as outer membrane proteins were determined in *B. wadsworthia* [15, 16].

4.2.2 *Bacteroides*: Methods

4.2.2.1 Phenotypic Identification of Gram-Negative Anaerobic Saccharolytic Rods

Molecular typing can never stand alone but needs state-of-the-art conventional identification as a precondition before being performed. The traditional method for identification and classification of anaerobic bacteria uses carbohydrate fermentation and other biochemical tests in combination with metabolic end-product analysis by gas chromatography and, taken together, still provides the "gold standard" for identification. The biochemical scheme for identification of *Bacteroides* species and *B. wadsworthia* has been described previously and updated in detail [17]. In brief, prereduced, anaerobically sterilized (PRAS) biochemicals are used to test the fermentation of arabinose, rhamnose, trehalose, salicin, sucrose, and xylan, the hydrolysis of esculin, and the production of indole and catalase. Bile resistance is usually determined by growth in PRAS peptone/yeast broth containing 20% bile. In addition, key reactions of the RapID ANA II systems are used. In case of *Bacteroides* species, gas chromatography is not much helpful. In general, differentiation of species within the *B. fragilis* group is not an easy task, as they demonstrate a great deal of similarity in colony and cell morphology as well as biochemical reactions [18]. Recent advances in identification of anaerobes from clinical samples include matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) Biotyper. Comparison with phenotypic identification and 16S rRNA sequencing demonstrated that the ability to differentiate between species relies on the MALDI-TOF database used and enables better characterization of types of infection and anticipates antimicrobial susceptibility [19].

4.2.2.2 Concept of PCR Fingerprinting

Molecular genetic methods, including classic genomic fingerprinting, chromosomal DNA probe hybridization, and species-specific PCR, have been used for identification and characterization of bacterial isolates. For example, new species and changes in nomenclature were increasingly established by using DNA homology studies, especially based on 16S rRNA sequencing and/or 16S–23S rDNA spacer region analysis [20–23]. This latter technique will be explained in more detail with *Porphyromonas* (see Sect. 4.3).

Here we will concentrate on PCR fingerprint techniques. These techniques were broadly used for the characterization and identification of bacteria, fungi, and parasites and have proved a versatile method for detection of polymorphisms for identification, characterization, and typing of all kinds of microorganisms. They were described for typing of aerobic and facultative anaerobic bacteria, primarily with arbitrary primer (AP) PCR [24–26]. However, completely arbitrary priming lies at one end of a spectrum of possible targeting strategies for fingerprinting. The other end of the spectrum uses primers derived from known near-perfect dispersed repeats, for example, tDNA intergenic length polymorphisms. In this spectrum lies a cornucopia of other repeats such as purine-pyrimidine motifs that have been successfully used to produce PCR fingerprints. These mini- and microsatellite repeats are particularly useful because primers directed toward them reveal more polymorphisms between closely related individuals. Primer pairs directed toward rRNA genes are also useful because the rRNA gene clusters evolve more slowly than most of the rest of the genome, which is under less stringent selection pressure. These patterns produced by rDNA-directed primers can be used to compare genomes at a higher taxonomic level than is possible with arbitrarily primed PCR [27, 28].

With the use of PCR fingerprint techniques, DNA polymorphisms have been detected that aid in the differentiation of species. Single nonspecific primers or single tDNA primers were used to both identify and characterize selected clinical

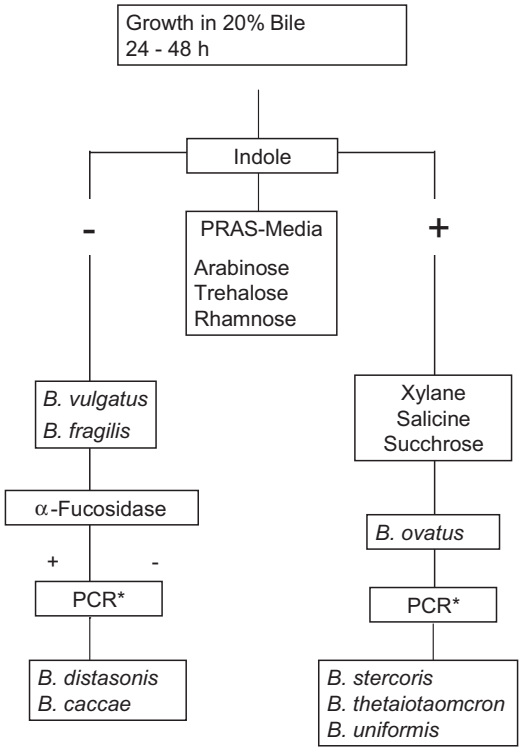

Fig. 4.1 Identification of *Bacteroides* species using molecular and phenotypic methods. *PCR – PCR fingerprinting

isolates of *B. fragilis*, *B. thetaiotaomicron*, or *B. vulgatus* as well as isolates of *Parabacteroides distasonis* (formerly *Bacteroides distasonis*) and *B. caccae* with similar biochemical key reactions (Fig. 4.1).

4.2.3 *Bacteroides*: Detailed Protocols

4.2.3.1 *Bacteroides* Strains, Culture Conditions, and DNA Extraction

Reference strains were obtained from American Type Culture Collection (ATCC), USA; National Collection of Type Cultures (NCTC), GB; Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Germany; and Virginia Polytechnic Institute (VPI), USA. In total, 68 indole-negative and 71 indole-positive *Bacteroides* isolates as well as 101 *Bilophila* isolates from blood and wound cultures that were obtained from different sites in Germany and in the USA were subjected to this study. Isolates were grown for 48 h (*Bacteroides* spp.) on Columbia blood agar or, respectively, 4–6 days (*Bilophila* spp.) on *Bacteroides* bile agar in the anaerobe chamber. Two to ten bacterial colonies were subjected into 100 µl sterile distilled water and incubated for 15 min at 95 °C. After a short centrifugation step (2 min, 11,000 × g), the supernatant was submitted into the PCR master mix. Alternatively, when pure DNA was needed or inhibitors were present (clinical specimens), extraction was performed with the Qiagen Tissue Kit (Qiagen, Germany) using the instructions from the manufacturer.

4.2.3.2 PCR Amplification and Fingerprinting

Primers: The core sequence of the phage M13 (5′-AGGTCGCGGGTTCGAATCC-3′) [24]; the M13universal (also derived from the phage M13) (5′-TTATGAAACGACGGCCAGT-3′) ([25]); the 10mer primer AP3 (5′-TCACGATGCA-3′) [26], and the tDNA primers T3B (5′-AGGTCGCGGGTTCGAATCC-3′), T5A (5′-AGTCCGGTGCTCTAACCAACTGAG-3′), and T3A [25] were used as single primers in the experiments (in detail, *Bacteroides* spp., M13core, M13universal, T3B, T5A, and AP3; *B. wadsworthia*, M13core, T3B, and T3A). Amplification reactions were performed in 50 µl reaction fluid, which contained 2.5 µl DNA extract; 10 x PCR buffer (10 mM Tris-HCl, pH 8,3; 50 mM KCl; 1,5 mM $MgCl_2$; 3 mM magnesium acetate); 200 µM of each dATP, dCTP, dGTP, and dTTP (Pharmacia Biotech, USA); and 2.5 U Taq DNA polymerase (Perkin Elmer Cetus, USA). Negative controls contained PCR-approved water instead of DNA. The primers were submitted in a final concentration of 25 pmol or 50 pmol. Samples were amplified as follows, 1 min at 95 °C and 1 min at 50 °C (universal primers, all other primers) or 30 s at 50 °C (tDNA primer) or 1 min at 36 °C (AP3 primer), followed by an extension cycle of up to 6 min at 72 °C. Reaction tubes were held at 4 °C until further analysis. The samples were concentrated to a volume of 20 µl in a vacuum centrifuge (SpeedVac, Savant, USA) and in relation of 1:10 with gel loading solution (Sigma, Germany) added for gel electrophoresis. All the different PCR assays for an additional group of bacteria were optimized using the Taguchi

scheme [29] for the concentration of chemicals, and with a temperature gradient, the annealing temperature was optimized.

DNA amplicons were separated in a submarine electrophoretic apparatus (Gibco BRL, USA) in 1.2–2.0% agarose gels: 1.2–2.0% agarose (depending on the length of the expected DNA fragments) (Pharmacia Biotech, Germany) in 0.5 × TBE buffer (Tris-borate-EDTA, Sigma, Germany). Electrophoretic separation was performed in a 0.5 × TBE buffer system gel (5 mm × 25 cm × 20 cm) 5–7 h at 3 V/cm. Amplified products were detected by staining with ethidium bromide (2 μg/ml). Gel images were analyzed by direct visual comparison or scanning the banding patterns (ScanJet IIcx Flatbed Scanner, Hewlett Packard, Palo Alto, CA). Absorbance profiles were corrected for gel-to-gel variation on the basis of reference samples run on each gel. Afterward, the patterns were compared by either calculation of the correlation coefficient between absorbance profiles or by using a band position matching coefficient. Natural groupings of similar patterns were found by clustering the matrix and displaying the results as a dendrogram (GelManager, BioSystematica, Prague, Czech Republic). For specific gene detection, the amplification of the *bft* gene was performed using the primers and conditions described by Shetab et al. [30] and Kato et al. [11]. For the detection of the *mpII* gene (metalloprotease gene) as well as the BfPAI (*B. fragilis* pathogenicity island) flanking regions, the primers and method described by Franco et al. were used [31].

4.2.4 *Bacteroides*: Results

4.2.4.1 PCR Fingerprinting

All strains subjected here were pre-identified using several phenotypic tests (see Sect. 4.2.2.1). The *Bacteroides* (including *Parabacteroides*) and *Bilophila* strains were screened using primers of different lengths: M13universal (19mer), M13core (19mer), AP3 (10mer), as well as two different tDNA primers, T3B (19mer) and T5A (24mer). The primers M13universal, M13core, and T3B and T5A produced diverse fragment profiles with species- and strain-specific bands. Nevertheless, amplification products of M13core produced profiles with several main bands. Testing of reference strains of *Bacteroides* species (*B. fragilis* ATCC 25285, *B. thetaiotaomicron* ATCC 29741, *Pa. distasonis* ATCC 8503, *B. ovatus* DSM 1896, *B. vulgatus* ATCC 8482) showed distinct profiles of all the reference strains. The primer M13core was further used for epidemiological testing but also for species identification and characterization among the strains of one or several *B. fragilis* group species. About 15 fragments with a length of 0.3–3 kb were determined. For species and group characteristics, the primers T3B and T3A were also appropriate. Comparing the profiles of all *Bacteroides* reference strains and phenotypically similar strains such as *Prevotella bivia*, the primer T3B produced about three to five main and many more bands with the length of 0.2–4 kb and, therefore, seemed to be suitable for species as well as group identification (Fig. 4.2).

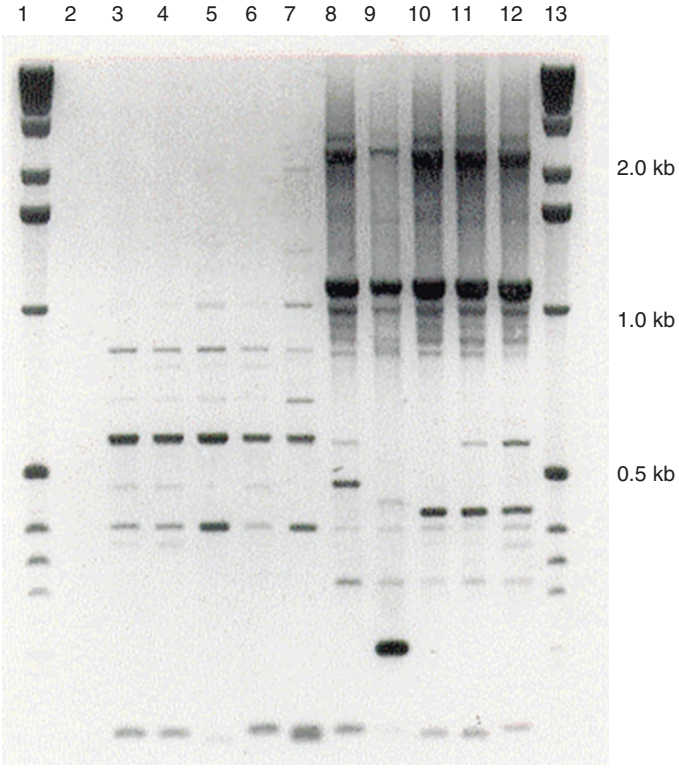

Fig. 4.2 Molecular fingerprinting by using tDNA-directed primers (T3B) of *Bacteroides fragilis* strains separating ten strains into two groups (lanes 1 and 13, length marker; lane 2, negative control; lanes 3–7, VPI 2393-like strains; lanes 8–12, ATCC 25285-type-strain-like strains)

Welsh and McClelland [27] described the tDNA primers as conserved on the species level but also determined genus-specific bands in aerobic bacteria. Using these primers in *Bacteroides*, species-specific as well as genus-specific bands were determined. PCR fingerprinting using the T3B primer confirmed the identity of 34 *B. fragilis* isolates. A species-specific fragment with the length of 530 bp could be determined in all the profiles of these strains, showing the potential of typing. The identification of the 11 isolates phenotypically placed into the species *Pa. distasonis* was also confirmed. Compared to the type strain (ATCC 8503) presenting a main band with a length of 1480 bp, one unusual, indole-negative strain was also identified as *Pa. distasonis*. However, there were a number of discrepancies between the phenotypic and molecular identification of *B. caccae* and *B. vulgatus* isolates. Comparing the species-specific main bands of the type strains of both species, 13 strains were identified as *B. caccae* and 10 strains as *B. vulgatus* (species-specific band with the length of 2.5 kb, data not shown).

4.2.4.2 Characterization of Species and Establishment of Genetic Markers

For *B. fragilis* typing by using the T3B primer, two different unique fingerprint types were established. In total, 30 out of the 34 strains showed PCR fingerprint profiles similar to the type strain ATCC 25285 and established the group I (lanes 8–12 in Fig. 4.2). This group was characterized by three to four main bands. A fragment with the length of 1050 bp was determined as genetic marker for this group. However, 4 of the 34 *B. fragilis* strains demonstrated similarity with the DNA homology group II reference strain (VPI 2393) and, therefore, were put in the PCR group II (lanes 3 [VPI 2393] to 7 in Fig. 4.2). This group showed many different main bands in comparison to group I. A fragment with the length of 370 bp was determined as genetic marker for this group.

4.2.4.3 Development of Specific PCR for Group Resp. Species Detection of *B. fragilis*

For group I, the characteristic 1050 bp fragment and, for group II, the 370 bp fragment were selectively amplified, cloned, and sequenced. From these group-specific sequences, group-specific primer sequences could be determined, showing the potential of typing methods in designing group/species/strain-specific diagnostic oligonucleotides.

4.2.4.4 Amplification of the Enterotoxin Gene in *B. fragilis* Isolates

Two PCR assays were used to detect ETBF strains. Using several sets of primers (see 4.2.3.2), in ten (11%) clinical isolates, the expected 367 bp and 558 bp enhanced virulence gene fragments were amplified [32].

4.2.5 *Bacteroides*: Discussion

In preliminary studies, PCR fingerprinting with single primers was demonstrated to reproducibly produce strain-, species-, and group-specific band patterns. Unique band patterns of unknown strains were compared to suitable reference strains and allowed species and subspecies identification. Using fingerprinting with especially primers M13core and T3B, two *B. fragilis* PCR groups were determined, whereas the biochemical groups – because of limitation in appropriate reactions – did not show major differences. Further comparing ATCC 25285 (type strain, reference strain for DNA homology group I [33]) and VPI 2393 (reference strain for DNA homology group II), the separation of two DNA homology groups was confirmed by testing clinical isolates from different clinical and geographical sites [33]. The majority of strains belonged to the PCR group I, and only a few strains belonged to PCR group II. Performing the amplification reaction with the T3B primer, both groups demonstrated a mixture of specific bands and several group-specific amplicons (Fig. 4.2). This grouping was confirmed using the M13core primer. At the same time, this grouping was confirmed by 16S rRNA sequence analysis, and it was suggested to establish the PCR group II as a second taxon [34]. From our (and the

practical clinical) point of view, it is very important to further determine phenotypic differences between the two groups as the biochemical reactions of all the strains so far did not show relevant differences. In contrast, susceptibility patterns of group II strains demonstrated high resistance against beta-lactam antibiotics, including imipenem (carbapenem) resistance. Appelbaum et al. firstly demonstrated in 1986 changing antibiotic resistance in a few DNA homology group II strains and speculated that this was due to the acquisition of a chromosomally determined metallo-beta-lactamase [35]. These results were confirmed using PCR group II strains as well as the resistance testing using the E test (MICs for imipenem >1 to >32 mg/L). In 1995, the resistance mechanism was described as an endogenous cephalosporinase, encoded by the *cfiA* gene [34]. Referring to the clinical importance and the increasing number of resistant *B. fragilis* isolates, a PCR assay for the differentiation of the PCR groups I and II was developed. Group-specific fragments for groups I and II were chosen, cloned, and sequenced. After sequencing, specific primers for groups I and II were developed, and their specificity was tested and confirmed in PCR assays. Thus, molecular fingerprinting can be a practical approach and precondition to design clinically relevant diagnostic oligonucleotides (for hybridization and PCR). Furthermore, our molecular fingerprinting studies confirmed the finding of *Bilophila wadsworthia* as a rather homogeneous species, since that, using the M13core primer, common bands were found for all but two of the isolates tested (these two isolates were later found to be preliminarily misidentified and belong to other species, unpublished data by Claros M). However, using the T3B primer, at least two distinct PCR fingerprint groups were determined. Interestingly, most of the German strains were found in group I (61 of 78 strains, data not shown). Thus, PCR fingerprinting with the T3B primer seems to detect even small epidemiological differences among strains.

4.3 *Porphyromonas*: A Genus Becoming Diverse

The genus *Porphyromonas* currently includes 20 approved species of asaccharolytic, obligately anaerobic, non-spore-forming, Gram-negative, nonmotile, pleomorphic bacilli. Of human origin are about half of them, namely, *P. asaccharolytica*, *P. bennonis*, "*P. bronchialis*" (a candidate species), *P. catoniae*, *P. endodontalis*, *P. gingivalis*, *P. pasteri*, *P. somerae*, and *P. uenonis* [36]. The other half are of animal origin including the catalase-positive *P. canoris*, *P. cangingivalis*, *P. cansulci*, *P. circumdentaria*, *P. gingivicanis*, *P. loveana* [37], *P. macacae* (which includes the former *P. salivosa*) and *P. pogonae* (with a growing number of human isolates) [38] and the catalase-negative *P. levii*, *P. crevioricanis*, as well as *P. gulae* (*P. gingivalis-like*, [39]). The genus *Parabacteroides*, including currently eight species of human origin with *Pa. distasonis* as a prominent member, is phylogenetically closely related to the genera *Tannerella* and *Barnesiella*. The genus *Tannerella* contains two species, *T. forsythia* (formerly *T. forsythensis*) and the very recently described *T. serpentiformis* [40, 41] and the genus *Barnesiella* with also two species, *B. intestinihominis* from human and *B. viscericola* from chicken feces [42].

The current study was performed to generate ITS data for most of the type strains of *Porphyromonas* spp. along with *T. forsythia*, *Pa. distasonis*, and *Prevotella melaninogenica* (outgroups) and to compare a phylogenetic tree deduced from these data with corresponding 16S rRNA gene data. The ITS sequences were further used to clarify the phylogenetic relationship between *P. gingivalis* and *P. gulae*, as well as between, by molecular typing methods, atypical α-fucosidase-negative and – classical – α-fucosidase-positive isolates of *P. asaccharolytica* (which were indeed later on reclassified as *P. uenonis*). By ITS amplification and sequencing, however, our group published the first hint for this species [23].

4.3.1 *Porphyromonas*: Methods

4.3.1.1 The General Concept of ITS Determination

Searching for "internal transcribed spacer" in September 2021 revealed 3,540,000 hits by Google (www.google.com), 2,219,000 by NCBI Nucleotide, and about 11,100 by NCBI PubMed (for the latter, see www.ncbi.nlm.nih.gov). Clearly this "spacer" does still attract a lot of interest in research. The reason is that the rRNA internal transcribed spacer (ITS) region is a widely used phylogenetic marker. Ribosomal RNAs are integral parts of the protein synthesis apparatus and thus present in all cellular life forms. On the one hand, these molecules and their encoding genes are highly conserved among all prokaryotes (i.e., bacteria and archaea). On the other hand, they contain sufficient sequence variability so that evolutionary relationships between different bacteria can be assessed. In addition, with the development of the PCR and sequence technology and recognition of the 16S rRNA gene

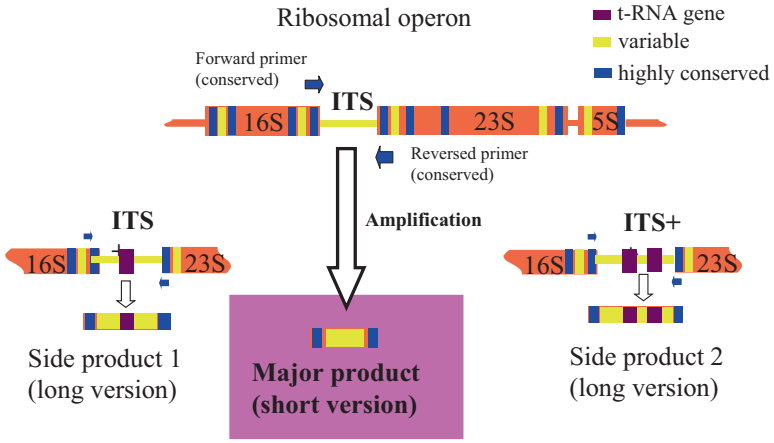

Fig. 4.3 Amplification of the internal transcribed spacer (also known as ribospacer). Depending on the number of ribosomal operons, the distance between 16S rRNA gene end and 23S rRNA start, and the numbers of tRNA genes interspersed, the amplicons can be very different by length and sequence

as outstanding phylogenetic marker gene, specific probes and primers at almost every taxonomic level have been designed and used for detection and phylogenetic characterization of known and novel human pathogens. While the 16S rDNA sequence is a good tool for inferring inter- and intra-generic relationships, the amplification, restriction, and/or sequencing of the 16S–23S rDNA ITS has been suggested to be well suited for typing and identification of bacteria at both the species and the strain level [28], because of marked variation of the ITS in both length and sequence between strains and species (Fig. 4.3). Based on diversities between ITS sequences, it is possible to construct species- and even strain-specific oligonucleotides that can be used to detect or track bacteria in their natural environments including colonized sites in human such as the gut, the vagina, or the oral cavity. Sequence polymorphism and length variation found in the 16S–23S rDNA ITS are increasingly used as tools for the differentiation of bacterial species and subspecies [43–45]. This is because the higher number of variable sites typical for the ITS sequence [46] can overcome the apparent limitations of the phylogenetic resolution of 16S rDNA in some genera as has been described for *Fusobacterium* by our group [45].

4.3.1.2 The Selection of Primers

For any PCR, whether quantitative or conventional, the primer selection is not trivial. In this chapter, we discuss this problem based on broad-range primers that bind at highly conserved regions of the 16S rRNA or 23S rRNA genes, both important for constructing an ITS-directed PCR (Fig. 4.3). When studying the primary literature, the reader will find quite a high number of ITS-directed PCR assays. A critical analysis of any article of interest (with respect to the design of the primers, their validation, and the scientific question that the article aims to address) is important. This is because ITS-directed primers never have the potential to encompass the entire spectrum of bacteria and have to be redesigned for any given taxon of interest. This is simply due to the fact that absolute conserved regions with the 16S/23S rRNA gene – although existing – are generally too short to function as primer-binding regions [47]. While several PCR-based pitfalls due to cell lysis techniques or PCR conditions have well been recognized [48, 49], the lack of universality of "universal" PCR primers and its consequences for routine diagnosis are generally not considered. With ever-increasing public 16S/23S rRNA gene databases, a re-evaluation and possibly redesign of primers are advisable to improve the intended target specificity.

4.3.1.3 The Sample Collection and DNA Extraction

For typing, specimens are taken from pure cultures but, sometimes, e.g., for fast tracking of nosocomial infections, also from the regions associated with infectious disease.

Special care has to be taken in order to avoid contamination during sampling. In the case of abscess puncture or biopsy in the oral cavity or gut, the surrounding area has to be properly isolated and cleaned prior to sampling. Due to the anaerobic lifestyle of all *Bacteroidetes*, rapid transportation into the molecular laboratory is

necessary to avoid or limit cell death and subsequent degradation of released DNA. DNA extraction strongly influences the outcome of any PCR reaction. Basic issues are efficient release of bacterial DNA (consider different cell wall properties of Gram-negative and Gram-positive bacteria), co-extraction of PCR inhibiting substances, and accurate storage of DNA extracts (for long-term storage, freezing at < −70 °C is recommended, and for short durations, storage at 4 °C degrees is possible). Frequent freeze-thawing procedures lead to degradation of genomic DNA and should be avoided. If samples have to be used repeatedly, aliquots should be made prior to freezing. PCR inhibiting substances co-extracted from human samples can be nucleases (critical in *P. gingivalis*-positive oral samples), bile salts, complex polysaccharides in feces, heme, immunoglobulin G, albumin, and lactoferrin in blood [50]. The latter four substances may also be of importance when DNA samples obtained from bleeding oral sites are analyzed. Besides these substances, human DNA itself which is usually co-extracted may interfere with the detection and diagnosis of pathogens [51]. Commercial DNA extraction kits are available for numerous applications including DNA extraction from clinical samples, such as tissue or blood. Note that these kits refer generally to extraction of human DNA. Those that were developed for extracting DNA from bacterial cells are based on the evaluation of selected Gram-positive and Gram-negative pure cultures. However, since for oral diagnostics bacterial DNA extraction has to be performed directly from clinical specimen, the best DNA extraction procedure and extraction kit have still to be tested experimentally.

4.3.2 *Porphyromonas*: Detailed Protocol

4.3.2.1 Bacterial Strains, Culture Conditions, and DNA Extraction

The following bacterial strains were used:

Porphyromonas asaccharolytica ATCC 25260^T and RMA 7115 (sacral wound); 7120 (toe); 7178 (endocervix); 8631 (rectal abscess); 9240 (peritoneal); 9603 (abdominal); 9674 (appendiceal fluid); 10263 (peritoneal); 10884, 10898, 10955, 10966, 10997, 11049, 11138, and 11258 (latter eight from pelvic fluid); 11290 (vaginal cupule); 11582 (endometrial pus); 11690 (endometrium); 11666 (endometrial pus); 11805 (pelvic fluid); and 12959, 12984, and 13273 (latter three from diabetic foot)
P. cangingivalis NCTC 12856^T
P. cansulci NCTC 12858^T
P. circumdentaria NCTC 12469^T
P. endodontalis ATCC 35406^T
P. gingivalis ATCC 33277^T and RMA 3725 (oral, mandible), 4165 (oral, maxilla), and 10371 (peritoneal/abdominal fluid)
P. gulae ATCC 51700^T
P. gingivicanis ATCC 55562^T
P. levii ATCC 29147^T

P. macacae ATCC 33141 and ATCC 49407 ("*P. salivosa*")
Parabacteroides distasonis ATCC 8503^T, *Tannerella forsythia* ATCC 43037^T, and *Prevotella melaninogenica* ATCC 25845^T

The latter three strains were used for contrast. All strains were cultivated at 37 °C on *Brucella* agar (Anaerobe Systems, Morgan Hill, CA) under anaerobic conditions using an anaerobic chamber. Genomic DNA was extracted using the DNeasy Tissue Kit (Qiagen).

4.3.2.2 PCR Amplification and DNA Sequence Analysis

The 16S primer SPFPorph (5′GTA CAC ACC GCC CGT CAA GCC 3′, corresponding to *E. coli* position 1390–1411) and the 23S primer SPRPorph (5′TCG CAG CTT ATC ACG TCC TTC 3′, corresponding to *E. coli* position 62–42) were designed based on the complete genome of *P. gingivalis* W83 (GenBank NC 002950); however, the respective regions among bacterial small and large subunit sequences (RDP) are relatively conserved. PCR was carried out using a Biometra Uno I (Biometra) thermocycler in a volume of 100 µl containing 1 X PCR buffer; 1.5 mM $MgCl_2$; 2 units of Taq-polymerase; 0.2 mM each of dATP, dCTP, dGTP, and dTTP (Boehringer Mannheim); 10 pmol SPFPorph forward primer; 10 pmol SPRPorph reversed primer; and 100 ng template nucleic acids. The amplification was performed using the following temperature profile and 30 cycles: denaturation, 1 min at 94 °C; annealing, 1 min at 52 °C; and elongation, 2.5 min at 72 °C. Amplification products (aliquots of 10 µl) were separated electrophoretically on a 2% macro agarose gel in 1x TPE (80 mM Tris-phosphate, 2 mM EDTA, pH 7.5) for a minimum of 18 h at 30 V.

After purification using the Wizard DNA Clean-Up System (Promega), the spacer DNA was directly sequenced. Sequences were assembled using the program Vector NTI Suite 9.0 (InforMax) and aligned using the program GeneDoc [52]. A phylogenetic tree was constructed by the neighbor-joining method and the programs Clustal W [53], Clustal X [53, 54], and TreeView [55].

4.3.3 *Porphyromonas*: Results

Approximations of ITS lengths were obtained from agarose gels, as demonstrated in Fig. 4.4A. All *Porphyromonas* reference strains showed a single band between 970 bp (*P. gingivalis* ATCC 33277^T) and 710 bp (*P. circumdentaria* NCTC 12469^T). The four strains of *P. gingivalis* analyzed were almost identical by ITS amplicon length (970–960 bp) and sequence (97–99% similarity, data not shown). In contrast, among 24 clinical isolates of *P. asaccharolytica* and the type strain ATCC 25260^T, the length of the ITS amplicons was more variable and ranged from 1044 bp (*P. asaccharolytica* RMA 10263, α-fucosidase-negative strain) to 960 bp (*P. asaccharolytica* ATCC 25260^T, α-fucosidase-positive strain) (Fig. 4.4B). In general, it was not possible to differentiate *Porphyromonas* species by comparing ITS gel electrophoretic profiles alone. Further discrimination without the need of sequencing

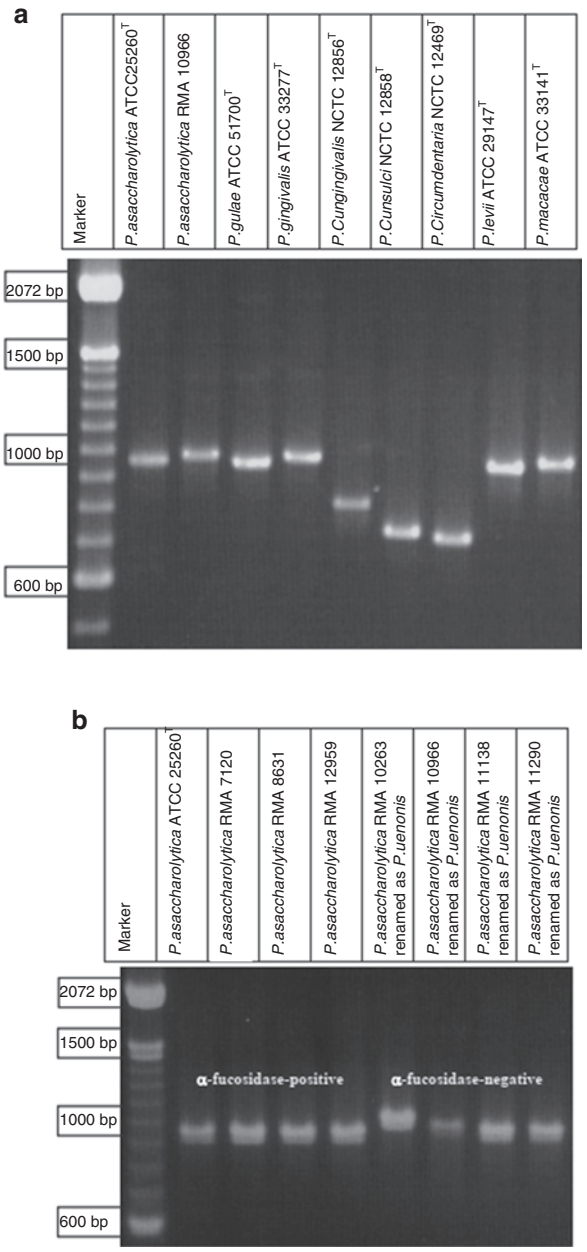

Fig. 4.4 Representative gel electrophoretic ITS amplification patterns of *Porphyromonas* species to demonstrate inter-species (**a**) and in the case of *P. asaccharolytica* also "intra"-species (**b**) heterogeneity – the latter leading to reclassification of α-fucosidase-negative strains as *P. uenonis*

might be possible by ITS restriction digest with endonucleases, since we found considerable variation in restriction sites (e.g., *Ava* I, *Apa*LI, *Cla*I, *Eco* RI, *Hind* III, *Sma* I). Sequencing the purified ITS amplicons of the *Porphyromonas* strains using SPFPorph and SPRPorph as primers led to nearly ambiguity-free sequence determination by comparing both runs and directions. A database search of tRNA

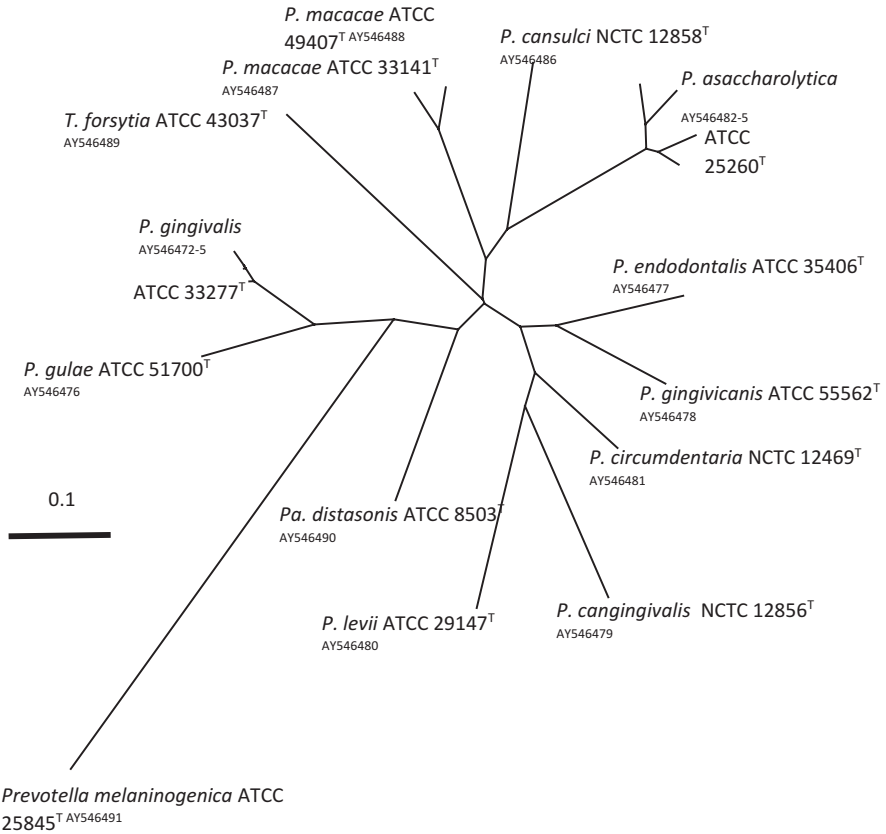

Fig. 4.5 Phylogram (neighbor-joining method) showing the genetic relationships among *Porphyromonas* species based on the DNA sequences of their 16S–23S rDNA spacer regions (GenBank accession numbers are included). *Parabacteroides distasonis* ATCC 8503^T, *Tannerella forsythia* ATCC 43037^T, and *Prevotella melaninogenica* ATCC 25845^T (outgroup, based on the short version of spacer) were included for contrast

consensus sequences (which should always be performed with ITS) and their comparison with our *Porphyromonas* intragenic spacer DNA revealed no matches. Phylogenetic tree reconstruction based on the ITS sequences (short version only in the case of *P. melaninogenica*) is demonstrated in Fig. 4.5**.** The different strains of *P. gingivalis* matched on a 97–99% level, and the two *P. macacae* ATCC strains (ATCC 49407 was formerly referred to as *P. salivosa* and then reclassified) matched on a 94% level; however, *P. asaccharolytica* was more heterogeneous (80–99% range in similarity level). Even more interesting, this latter species, which phenotypically differed in α-fucosidase activity, showed – as expected – two main clusters. Inter-cluster similarity was only 80 to 87%, whereas the intra-cluster similarity was 92–99%. The higher resolution of ITS amplification and sequencing was further used to analyze the relationship between 9 α-fucosidase-positive and 16

α-fucosidase-negative strains of *P. asaccharolytica* and clearly showed that both groups diverged into individual phylogenetic branches [23].

4.3.4 *Porphyromonas*: Discussion

PCR amplification of the ITS region and subsequent gel electrophoresis of 11 different *Porphyromonas* reference strains plus 3 clinical isolates of *P. gingivalis* and 24 of *P. asaccharolytica* showed large heterogeneity in length of amplicons [45]. Furthermore, only one distinct amplification band was produced with *Porphyromonas* species as well as with the relatives *T. forsythia* and *Pa. distasonis*, unlike, e.g., *Fusobacterium* spp. [45] or many other genera analyzed so far [43, 56, 57], which is mainly due to the number of *rrna* operons. Within a species, the length of amplicons and the deduced sequence are relatively constant as we have shown for *P. gingivalis* (four strains) and *P. macacae* (two strains) and for fusobacterial species and subspecies [45]. The high resolution of ITS sequences led to a separation between two clusters of *P. asaccharolytica* strains: one was α-fucosidase-positive as is typical of the type strain and the other was α-fucosidase-negative. Moreover, 11 of the 12 isolates in the larger α-fucosidase-negative group were isolated from endometrial infection specimens. Thus, the heterogeneity found between the 25 *P. asaccharolytica* strains was a first and later confirmed hint for an unrecognized species, *P. uenonis* [36].

The separation between *P. gingivalis* and *P. gulae* as distinct species was supported by our ITS data; thus, *P. gulae* should not be referred to as the "animal strain of *P. gingivalis*" as it is genetically related but not identical with *P. gingivalis*. Fournier and co-authors, describing *P. gulae*, pointed out the paradox that although this species could be distinguished from *P. gingivalis* phenotypically and by DNA-DNA similarity, the differences between genes encoding 16S rRNA appeared tenuous [39]. They also concluded that the recent divergences of ancestral phyla, e.g., after colonizing different mammalian hosts, could not be sufficiently discerned by 16S information. Again, at least in some genera, ITS data give additional information and enhance phylogenetic resolution if discrepancies between DNA-DNA hybridization and 16S sequencing results are observed.

In conclusion, the ITS region is being used increasingly as an important tool for classification and differentiation of bacterial species. Our study was the first to provide this sequence information for most of the *Porphyromonas* species and their relatives. The higher resolution of ITS helped clarify some of the current problems in molecular taxonomy.

4.4 General Discussion and Final Remarks

For *Bacteroidetes*, especially the clinically relevant Bacteroides fragilis and *Porphyromonas* species, PCR-based fingerprinting techniques turned out to be ideal for typing since strain/species/group-specific bands can be found (and further used

for identification and diagnosis) and only a very small DNA amount is needed. The latter is especially important here, since many obligate anaerobic strains are fastidious or often almost nonviable through oxygen contact and grow very slowly in culture. Since the last decade, whole genome shotgun sequencing is increasingly being used as a promising tool for typing *Bacteroidetes* as well as a method for predicting antimicrobial resistance properties. However, with these new technologies and the corresponding bioinformatics pipelines, new challenges such as misassembly might occur and have to be solved [58–60].

4.5 Summary

Bacteroidetes are a phylum of bacteria which consists of several genera including *Bacteroides*, *Porphyromonas*, and *Prevotella*. Whereas the genus *Bacteroides* will be discussed and subjected with infections originating from the intestinal tract, *Porphyromonas* is clearly more associated with oral or vaginal infections. With both genera (and *Bilophila*, *Parabacteroides*, or *Prevotella for contrast*), different methods are described and discussed exemplarily. Critical to this development, however, is a proper understanding and application of the methodologies and knowledge of their limitations. In this chapter, molecular tools based on ITS (internal transcribed spacer) amplification and sequencing as well as PCR fingerprint techniques will be described along with examples showing ways to analyze the datasets. Both methods allow the identification of almost any given bacterial species or strain in pure culture or even directly in clinical samples in a sensitive and reproducible way. This chapter will be complemented by discussing potential pitfalls that should be taken into consideration for producing proper results along with referring the reader to pertinent literature that will allow an individual deepening into the concept of molecular typing in clinical bacteriology.

References

1. Shah HN, Collins MD (1989) Proposal to restrict the genus Bacteroides (Castellani and Chalmers) to Bacteroides fragilis and closely related species. Int J Syst Bacteriol 39:85–87
2. Brook I (2016) Spectrum and treatment of anaerobic infections. J Infect Chemother 22:1–13
3. Shah HN, Collins MD (1988) Proposal for reclassification of Bacteroides asaccharolyticus, Bacteroides gingivalis, and Bacteroides endodontalis in a new genus, Porphyromonas. Int J Syst Bacteriol 38:128–131
4. Andrieu C, Mailhe M, Ricaboni D et al (2018) Noncontiguous finished genome sequences and description of Bacteroides mediterraneensis sp. nov., Bacteroides ihuae sp. nov., Bacteroides togonis sp. nov., Bacteroides ndongoniae sp. nov., Bacteroides ilei sp. nov. and Bacteroides congonensis sp. nov. identified by culturomics. New Microbes New Infect 26:73–88
5. Cassir N, Croce O, Pagnier I et al (2014) Non-contiguous finished genome sequence and description of Bacteroides neonati sp. nov., a new species of anaerobic bacterium. Stand Genomic Sci 9:794–806
6. Ramasamy D, Lagier JC, Rossi-Tamisier M et al (2014) Genome sequence and description of Bacteroides timonensis sp. nov. Stand Genomic Sci 9:1181–1197

7. Kling JJ, Wright RL, Moncrief JS et al (1997) Cloning and characterization of the gene for the metalloprotease enterotoxin of Bacteroides fragilis. FEMS Microbiol Lett 146:279–284
8. Myers LL, Shoop DS (1987) Association of enterotoxigenic Bacteroides fragilis with diarrheal disease in young pigs. Am J Vet Res 48:774–775
9. Sack RB, Albert MJ, Alam K et al (1994) Isolation of enterotoxigenic Bacteroides fragilis from Bangladeshi children with diarrhea: a controlled study. J Clin Microbiol 32:960–963
10. San Joaquin VH, Griffis JC, Lee C et al (1995) Association of Bacteroides fragilis with childhood diarrhea. Scand J Infect Dis 27:211–215
11. Kato N, Kato H, Watanabe K et al (1996) Association of enterotoxigenic Bacteroides fragilis with bacteremia. Clin Infect Dis 23(Suppl 1):S83–S86
12. Claros MC, Claros ZC, Tang YJ et al (2000) Occurrence of Bacteroides fragilis enterotoxin gene-carrying strains in Germany and the United States. J Clin Microbiol 38:1996–1997
13. Baron EJ, Summanen P, Downes J et al (1989) Bilophila wadsworthia, gen. nov. and sp. nov., a unique gram-negative anaerobic rod recovered from appendicitis specimens and human faeces. J Gen Microbiol 135:3405–3411
14. Baron EJ (1997) Bilophila wadsworthia: a unique gram-negative anaerobic rod. Anaerobe 3:83–86
15. Da Silva SM, Venceslau SS, Fernandes CL et al (2008) Hydrogen as an energy source for the human pathogen Bilophila wadsworthia. Antonie Van Leeuwenhoek 93:381–390
16. Wexler HM (2007) Bacteroides: the good, the bad, and the nitty-gritty. Clin Microbiol Rev 20:593–621
17. Jousimies-Somer HR, Summanen P, Citron DM et al (2002) Wadsworth- KTL – anaerobic bacteriology manual. Star Publisher, Belmont
18. Citron DM, Baron EJ, Finegold SM et al (1990) Short prereduced anaerobically sterilized (PRAS) biochemical scheme for identification of clinical isolates of bile-resistant Bacteroides species. J Clin Microbiol 28:2220–2223
19. Coltella L, Mancinelli L, Onori M et al (2013) Advancement in the routine identification of anaerobic bacteria by MALDI-TOF mass spectrometry. Eur J Clin Microbiol Infect Dis 32:1183–1192
20. Smith CJ, Callihan DR (1992) Analysis of rRNA restriction fragment length polymorphisms from Bacteroides spp. and Bacteroides fragilis isolates associated with diarrhea in humans and animals. J Clin Microbiol 30:806–812
21. Rautio M, Eerola E, Vaisanen-Tunkelrott ML et al (2003) Reclassification of Bacteroides putredinis (Weinberg et al., 1937) in a new genus Alistipes gen. nov., as Alistipes putredinis comb. nov., and description of Alistipes finegoldii sp. nov., from human sources. Syst Appl Microbiol 26:182–188
22. Jousimies-Somer H, Summanen P (2002) Recent taxonomic changes and terminology update of clinically significant anaerobic gram-negative bacteria (excluding spirochetes). Clin Infect Dis 35:S17–S21
23. Conrads G, Citron DM, Tyrrell KL et al (2005) 16S-23S rRNA gene internal transcribed spacer sequences for analysis of the phylogenetic relationships among species of the genus Porphyromonas. Int J Syst Evol Microbiol 55:607–613
24. Huey B, Hall J (1989) Hypervariable DNA fingerprinting in Escherichia coli: minisatellite probe from bacteriophage M13. J Bacteriol 171:2528–2532
25. Welsh J, Mcclelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218
26. Williams JG, Kubelik AR, Livak KJ et al (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
27. Welsh J, Mcclelland M (1991) Genomic fingerprints produced by PCR with consensus tRNA gene primers. Nucleic Acids Res 19:861–866
28. Barry T, Colleran G, Glennon M et al (1991) The 16s/23s ribosomal spacer region as a target for DNA probes to identify eubacteria. PCR Methods Appl 1:51–56
29. Cobb BD, Clarkson JM (1994) A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. Nucleic Acids Res 22:3801–3805

30. Shetab R, Cohen SH, Prindiville T et al (1998) Detection of Bacteroides fragilis enterotoxin gene by PCR. J Clin Microbiol 36:1729–1732
31. Franco AA, Cheng RK, Chung GT et al (1999) Molecular evolution of the pathogenicity island of enterotoxigenic Bacteroides fragilis strains. J Bacteriol 181:6623–6633
32. Claros MC, Claros ZC, Hecht DW et al (2006) Characterization of the Bacteroides fragilis pathogenicity island in human blood culture isolates. Anaerobe 12:17–22
33. Johnson J, Ault D (1978) Taxonomy of the Bacteroides. II. Correlation of phenotypic characteristics with deoxyribonucleic acid homology groupings for Bacteroides fragilis and other saccharolytic Bacteroides species. Int J Syst Bacteriol 28:257–268
34. Podglajen I, Breuil J, Casin I et al (1995) Genotypic identification of two groups within the species Bacteroides fragilis by ribotyping and by analysis of PCR-generated fragment patterns and insertion sequence content. J Bacteriol 177:5270–5275
35. Appelbaum PC, Jacobs MR, Spangler SK et al (1986) Comparative activity of beta-lactamase inhibitors YTR 830, clavulanate, and sulbactam combined with beta-lactams against beta-lactamase-producing anaerobes. Antimicrob Agents Chemother 30:789–791
36. Finegold SM, Vaisanen ML, Rautio M et al (2004) Porphyromonas uenonis sp. nov., a pathogen for humans distinct from P. asaccharolytica and P. endodontalis. J Clin Microbiol 42:5298–5301
37. Bird PS, Trott DJ, Mikkelsen D et al (2016) Porphyromonas loveana sp. nov., isolated from the oral cavity of Australian marsupials. Int J Syst Evol Microbiol 66:3771–3778
38. Kawamura Y, Kuwabara S, Kania SA et al (2015) Porphyromonas pogonae sp. nov., an anaerobic but low concentration oxygen adapted coccobacillus isolated from lizards (Pogona vitticeps) or human clinical specimens, and emended description of the genus Porphyromonas Shah and Collins 1988. Syst Appl Microbiol 38:104–109
39. Fournier D, Mouton C, Lapierre P et al (2001) Porphyromonas gulae sp. nov., an anaerobic, gram-negative coccobacillus from the gingival sulcus of various animal hosts. Int J Syst Evol Microbiol 51:1179–1189
40. Maiden MF, Cohee P, Tanner AC (2003) Proposal to conserve the adjectival form of the specific epithet in the reclassification of Bacteroides forsythus Tanner et al. 1986 to the genus Tannerella Sakamoto et al. 2002 as Tannerella forsythia corrig., gen. nov., comb. nov. request for an opinion. Int J Syst Evol Microbiol 53:2111–2112
41. Sakamoto M, Suzuki M, Umeda M et al (2002) Reclassification of Bacteroides forsythus (Tanner et al. 1986) as Tannerella forsythensis corrig., gen. nov., comb. nov. Int J Syst Evol Microbiol 52:841–849
42. Sakamoto M, Lan PT, Benno Y (2007) Barnesiella viscericola gen. Nov., sp. nov., a novel member of the family Porphyromonadaceae isolated from chicken caecum. Int J Syst Evol Microbiol 57:342–346
43. Motoyama Y, Ogata T (2000) 16S-23S rDNA spacer of Pectinatus, Selenomonas and Zymophilus reveal new phylogenetic relationships between these genera. Int J Syst Evol Microbiol 50(Pt 2):883–886
44. Guasp C, Moore ER, Lalucat J et al (2000) Utility of internally transcribed 16S-23S rDNA spacer regions for the definition of Pseudomonas stutzeri genomovars and other Pseudomonas species. Int J Syst Evol Microbiol 50(Pt 4):1629–1639
45. Conrads G, Claros MC, Citron DM et al (2002) 16S-23S rDNA internal transcribed spacer sequences for analysis of the phylogenetic relationships among species of the genus Fusobacterium. Int J Syst Evol Microbiol 52:493–499
46. Soller R, Hirsch P, Blohm D et al (2000) Differentiation of newly described antarctic bacterial isolates related to Roseobacter species based on 16S-23S rDNA internal transcribed spacer sequences. Int J Syst Evol Microbiol 50(Pt 2):909–915
47. Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. J Microbiol Methods 55:541–555
48. Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl Environ Microbiol 62:625–630

49. Von Wintzingerode F, Gobel UB, Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev 21:213–229
50. Kaltenboeck B, Wang C (2005) Advances in real-time PCR: application to clinical laboratory diagnostics. Adv Clin Chem 40:219–259
51. Horz HP, Scheer S, Huenger F et al (2008) Selective isolation of bacterial DNA from human clinical specimens. J Microbiol Methods 72:98–102
52. Nicholas K, Nicholas H (1997) GeneDoc: a tool for editing and annotation multiple sequence alignments. Distributed by the authors https://genedoc.software.informer.com/2.7/ (last access December 2019)
53. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics chapter 2:unit 2 3
54. Jeanmougin F, Thompson JD, Gouy M et al (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405
55. Page RD (2002) Visualizing phylogenetic trees using TreeView. Current protocols in bioinformatics chapter 6:unit 6 2
56. Gurtler V, Rao Y, Pearson SR et al (1999) DNA sequence heterogeneity in the three copies of the long 16S-23S rDNA spacer of Enterococcus faecalis isolates. Microbiology 145:1785–1796
57. Graham TA, Golsteyn-Thomas EJ, Thomas JE et al (1997) Inter- and intraspecies comparison of the 16S-23S rRNA operon intergenic spacer regions of six Listeria spp. Int J Syst Bacteriol 47:863–869
58. Acuna-Amador L, Primot A, Cadieu E et al (2018) Genomic repeats, misassembly and reannotation: a case study with long-read resequencing of Porphyromonas gingivalis reference strains. BMC Genomics 19:54
59. Sydenham TV, Overballe-Petersen S, Hasman H et al (2019) Complete hybrid genome assembly of clinical multidrug-resistant Bacteroides fragilis isolates enables comprehensive identification of antimicrobial-resistance genes and plasmids. Microb Genom 5
60. Sydenham TV, Soki J, Hasman H et al (2015) Identification of antimicrobial resistance genes in multidrug-resistant clinical Bacteroides fragilis isolates by whole genome shotgun sequencing. Anaerobe 31:59–64

Vibrio cholerae

5

Thandavarayan Ramamurthy, Asish K. Mukhopadhyay, Bhabatosh Das, Ranjan K. Nandy, Ankur Mutreja, and G. Balakrish Nair

5.1 Introduction

Cholera caused by toxigenic *Vibrio cholerae* is a major public health problem in many developing countries, where outbreaks and sporadic infections occur at regular intervals. WHO has registered 499,447 cases, including 2990 deaths with case fatality rate of 0.6% in 2018 [1]. The disease is characterized by profuse watery diarrhea that rapidly leads to dehydration, and death occurs in 50–70% of untreated patients.

For more than two centuries, cholera remains one of the great epidemic diseases of the tropical world. Cholera has spread from Asia, where it is endemic to many parts of the world in the form of seven pandemics during the past 200 years [2]. *V. cholerae* serogroup O1, biotype El Tor, has spread from Asia to cause pandemic disease in Africa and South America during the past 50 years. Until 1992, serogroup O1 was considered as the devastating cholera causative agent. A new serogroup, O139, appeared in South Asia in 1992 and changed the whole perception regarding cholera as this was the first non-O1 serogroup related with epidemic cholera. When this serogroup first appeared, it was thought that the next pandemic strain of cholera had emerged, but over the past few years, the prevalence of the O139 serogroup

T. Ramamurthy (✉) · A. K. Mukhopadhyay · R. K. Nandy
Division of Bacteriology, National Institute of Cholera and Enteric Diseases, Kolkata, India
e-mail: ramamurthy.t@icmr.gov.in

B. Das
Infection and Immunity Division, Translational Health Science and Technology Institute, Faridabad, India

A. Mutreja
Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK

G. B. Nair
Rajiv Gandhi Centre for Biotechnology, Trivandrum, India

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_5

abruptly declined. Expansion of the seventh pandemic was accompanied by increased genetic variation among strains of *V. cholerae* O1 and O139, but the relationship of these genetic changes in relation to virulence and in the epidemiology of cholera is not clearly understood.

Advances in molecular genetics have facilitated the development of refined molecular typing techniques, which in turn have assisted in studying the genetic diversity of many bacterial populations. Genotyping studies related to the epidemiology of the disease, otherwise called molecular epidemiology, have revealed clonal diversity among *V. cholerae* strains and emergence of new epidemic clones over the years. Clonal information has the potential to provide information on subtypes of the organism and their source and/or origin of infection and to recognize virulent strains of the organism and monitor vaccination programs. The increasing application of procedures employing several molecular tools has provided new means of discriminating *V. cholerae*. Such studies provide a wealth of information to assist the epidemiologist in tracing and tracking the spread of epidemics and give new insights into the evolution and origin of the newer variants of *V. cholerae*. The complete genome sequences of several strains of *V. cholerae* have provided an important source to begin addressing many questions about the evolution of *V. cholerae* as a human pathogen as well as environmental organism.

Bacterial typing techniques are not always comparable, as each method is related to the specific research question and has its own merit. Our aim of this chapter is to review the impact of their epidemiological applications and inference as evidenced from several studies rather than focusing on the methodology of molecular techniques.

5.2 Background Information on *V. cholerae*

V. cholerae was first described by Filippo Pacini in 1854 from cholera victims followed by its rediscovery in 1883 by Robert Koch in Egypt and India [3]. *V. cholerae* belonging to the serogroups O1 and O139 (synonym "Bengal") are associated with epidemic and pandemics of cholera, and the remaining serogroups (previously referred to as "non-agglutinable" (NAG) vibrios) either were nonpathogenic or cause sporadic infection. The O1 serogroup is subdivided into serotypes, i.e., Ogawa and Inaba, and two biotypes, classical and El Tor. Both the serotypes are found in classical and El Tor biotypes. The first six cholera pandemics were caused by the classical biotype, and the seventh pandemic has been of the El Tor biotype. Hemolysis of sheep erythrocytes, bacteriophage susceptibility with specific classical and El Tor phages, Voges-Proskauer reaction, polymyxin B susceptibility, and hemagglutination of chicken erythrocytes are used for biotyping. Though new phage typing schemes for differentiating between strains of *V. cholerae* O1 and O139 serogroups were established [4, 5], its use is limited to reference laboratories.

5.3 Polymerase Chain Reaction-Based Typing

Polymerase chain reaction (PCR) has generally been used for the rapid detection of infectious agents in clinical samples or bacterial growth and later adapted as one of the molecular typing techniques. PCR typing has an advantage in that the DNA can be amplified (0.1 to >35 kb) even if the template DNA is in minute quantity. Several PCR methods are currently available for typing of *V. cholerae*.

5.3.1 Random Amplification of Polymorphic DNA (RAPD) Profiles

The simplicity and the discriminating capacity of this technique make it useful for detecting genetic diversity among microorganisms from a defined group or for outbreak investigation. During 1993–1994, RAPD-PCR with O139 serogroup indicated that the Asian strains were similar [6]. As shown in this study, the RAPD profiles of the O139 strains resembled to those of E1 Tor strains rather than classical strain. Different RAPD profiles were obtained with *V. cholerae* strains from Malaysia, and there was no correlation with the source of isolation [7]. A collection of *V. cholerae* strains from Brazil showed no correlation with serotype, biotype, or source of the isolates [8]. However, with Iranian strains of *V. cholerae* O1, RAPD was able to discriminate into six distinct groups [9]. *V. cholerae* O1 from a 2006 outbreak in Accra, Ghana, had identical polymorphic DNA profiles, suggesting some genetic dissimilarity was present among the strains [10]. Generally, the discriminatory power of RAPD-PCR is low, and the method is susceptible to technical variations.

5.3.2 Other PCR-Based Typing Methods

The evolutionary relationships and molecular diversity of *V. cholerae* O1 and O139 and non-O1, non-O139 strains were studied using amplified fragment length polymorphism (AFLP) technique. In this, two sets of restriction enzyme-primer combinations were tested in the fingerprinting. Amplification of *Hind*III- and *Taq*I-digested genomic DNA grouped environmental isolates of O1 and non-O1, non-O139 strains and was unable to distinguish between O1 and O139 clinical strains. The AFLP analyses of restriction enzyme *Apa*I- and *Taq*I-digested genomic DNA separated O1 from O139 strains [11]. This study supported that a single clone of pathogenic *V. cholerae* has caused several cholera outbreaks in Asia, Africa, and Latin America during the seventh pandemic. Interestingly, some of the cholera cases were associated with environmental O1 or non-O1, non-O139 strains. Castañeda et al. [12] reported a novel typing method based on *V. cholerae* repeat sequences (VCR) using specific primers. The VCR-PCR of *V. cholerae* O1 supported the hypothesis that a clone with epidemic nature was responsible for the spread of cholera in Latin America.

Rivera et al. [13] have shown the usefulness of enterobacterial repetitive intergenic consensus sequence (ERIC)-PCR to differentiate *V. cholerae* from a cholera epidemic in Brazil from Peruvian, Mexican, and Indian epidemic strains. Majority of the toxigenic *V. cholerae* O1 and O139 strains exhibited a similar fingerprint (FP), whereas the non-toxigenic *V. cholerae* O1 and the non-O1, non-O139 strains belonged to different FP groups. Population diversity of toxigenic *V. cholerae* was analyzed with reference to strain relatedness and their ecological interactions in the human and the aquatic habitats using ERIC-PCR [14]. In this analysis, *V. cholerae* population structure supported the hypothesis that spatial and temporal fluctuations in the aquatic environment can cause shifts in the dynamics of cholera [14]. It has been demonstrated that typing of *V. cholerae* O1 by ERIC-PCR fingerprinting correlated well with ribotyping and was sometimes more discriminating [15]. This PCR assay provides a rapid and simple means of typing strains in epidemiological studies.

5.3.3 Mobile Genetic Elements (MGEs)

It was hypothesized that the composition of MGEs in *V. cholerae* strains would be useful as a phylogenetic typing as it is conserved among *V. cholerae* O1 strains [16]. Three types of MGEs usually account for resistance to antibiotics in *V. cholerae*: (1) plasmids, which for most are large and self-transmissible by conjugation; (2) integrons, which are chromosomal or plasmid-borne gene capture and expression systems; and (3) integrating conjugative elements (ICEs), which are chromosomal self-transmissible MGEs carrying several bacterial adaptive functions including antimicrobial resistance (AMR). SXT/R391 family is one of the ICEs extensively studied in *V. cholerae*.

The genetic characteristics of ICE SXT/R391 in *V. cholerae* are dynamic and region-specific. These ICEs in *V. cholerae* are strongly correlated with resistance to several antibiotics such as tetracycline, streptomycin, and trimethoprim-sulfamethoxazole [17, 18]. In addition, clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated (Cas) protein immune defense system, which protects the bacteria from virus predation, has been predominantly found in MGEs of *V. cholerae* and other vibrios [19]. With the increase of AMR among *V. cholerae*, MGEs have been used in the epidemiological investigations.

5.3.4 Plasmids

Generally, *V. cholerae* strains are devoid of plasmids. Studies conducted for the characterization of plasmids in *V. cholerae* are mostly related to their antimicrobial resistance rather than molecular fingerprinting. Plasmid profile analysis showed the presence of 1.3–4.6 MDa plasmids in non-O1, non-O139 and O1 strains from Malaysia [7]. Conjugative plasmid IncC responsible for multidrug resistance was identified in *V. cholerae* O1/O139 strains responsible for the large cholera outbreaks in Somalia, Haiti, and China [20–22].

5.3.5 Insertion Sequences (IS Elements)

A 628 bp insertion sequence element, IS*1004*, is present in one to eight copies in most of the *V. cholerae* strains [23]. IS*1004*-generated fingerprints discriminated classical and El Tor but not the non-O1, non-O139 strains, which are heterogeneous and unrelated to those of the epidemic *V. cholerae* O1. However, with *V. cholerae* serogroup O37 that was responsible for a large diarrhea outbreak in Sudan, the IS*1004* typing showed that these strains were closely related to classical O1 strains [23]. *V. cholerae* O139 has emerged from the pandemic O1 biotype El Tor through the replacement of a 22 kbp DNA region with a 40 kbp O139-specific DNA fragment. This O139-specific DNA fragment contains an insertion sequence designated IS*1358O139*. Apart from O1 and O139 serogroups, presence of this IS sequence in multiple copies was detected in serogroups O2, O22, and O155 but not in other non-O1, non-O139 serogroups [24]. The nucleotide sequences of IS*1358* in serogroups O22 and O155 are almost identical to that of O1 and O139. The significance of IS elements found in toxigenic strains of *V. cholerae* and their non-toxigenic counterparts is not fully known.

5.3.6 Integrons and ICEs

V. cholerae contains a genetic system called the integron that can integrate and excise DNA elements by site-specific recombination. Evolution of integron arrays can proceed by rearrangements and deletions/insertions of large portions of MGEs in addition to the insertion or excision of single MGE. The extent of resistance patterns and associated MGEs in epidemic *V. cholerae* O1 El Tor was investigated in East Africa with strains isolated from the late 1990s [25]. This study has shown the spread of SXT-related ICEs among *V. cholerae* O1 encoding resistance to chloramphenicol, streptomycin, sulfamethoxazole, and trimethoprim in African isolates. Phenotypically, the reemerged *V. cholerae* O139 in 1995 were susceptible to co-trimoxazole compared to those that appeared in late 1992 [26]. In these reemerged strains from India and Bangladesh, a 3.6 kb region of the SXT element was deleted leading them to become susceptible to co-trimoxazole [17]. More than 30 ICEs have been classified within the SXT/R391 family of clinical and environmental *V. cholerae* strains [27].

5.4 *Vibrio cholerae* Pathogenicity Island and *Vibrio* Seventh Pandemic Islands

Pathogenicity islands (PAI) are present in virulent strains comprising very large genomic regions (10–200 kb). The G+C content of the PIs often differs from the rest of the bacterial genome, the presence of repeat at their ends, and the presence of integrase determinants and other mobility loci support for the generation of PIs by horizontal gene transfer. Such changes are stable and persist for a long time in the bacterial genome and hence used as a marker in bacterial typing.

V. cholerae pathogenicity islands (VPIs) can be detected in epidemic and pandemic strains of *V. cholerae* but are generally absent among nonpathogenic strains [28]. The VPI contains ToxR-regulated genes (*aldA* and *tagA*) and a part of PAI that contains a regulator of virulence (ToxT) and a gene cluster encoding an essential colonization factor and the toxin-coregulated pilus (TCP). Comparative sequence analysis with different strains of *V. cholerae* O1 showed polymorphism in the VPI region. There were differences in several proteins as a result of frameshift mutations [29]. Variations in the VPI region provide preliminary evidence to explain the differences in potential virulent strains appeared between epidemics.

The VPI typing has allowed elucidation of differences in the genetic organization between pre-pandemic and pandemic strains. Osin et al. [30] demonstrated that the genome of pre-seventh pandemic strains of *V. cholerae* O1 isolated during 1910 was devoid of CTX and RS1 prophages, Vibrio pathogenicity islands (VPI-1 and VPI-2), and Vibrio seventh pandemic islands (VSP-1 and VSP-2) that contain key virulence genes. Acquisition of VPI and CTX in *V. cholerae* was shown in cholera outbreak-associated strains that were isolated during 1937. The seventh pandemic strains acquired two additional blocks of genes VSP-1 and VSP-2, which were absent in classical strains [31]. Most *V. cholerae* O1 and O139 strains carried the VSP islands (VSP-1 and VSP-2), whereas the non-O1, non-O139 strains carried several VSP island genes, but not the entire VSP island [32]. Absence of VSP islands in the Australian environmental *V. cholerae* O1 strains indicates their pre-seventh pandemic ancestry [33, 34].

In some non-O1, non-O139 strains, the left end of VPI exhibited extensive DNA rearrangements [35]. This information suggests that potentially pathogenic, non-epidemic, non-O1, non-O139 strains are likely evolved by sequential horizontal acquisition of the VPI and CTX independently rather than by exchange of O-antigen biosynthesis regions in an existing epidemic strain.

5.5 CTX Prophages

Analysis of variations in the cholera toxin (CT) encoding gene (*ctxAB*) and its flanking regions serve one of the essential molecular tools for typing toxigenic *V. cholerae*. The gene *ctxAB* resides in the genome of a lysogenic filamentous phage called CTXΦ [36]. The receptor for CTXΦ is the major colonization factor, TCP. It is evident that the VPI can minimally excise and therefore presumably integrate using a phage-like integrase and attachment site. The CTXΦ genome is composed of several open reading frames (ORFs), located on a 4.5 kb "core region" of the CTX element, which is essential for the morphogenesis of CTXΦ particles. Adjacent to the core is the RS2 region encoding ORFs *rstR*, *rstA2*, and *rstB2*. These genes encode products required for the integration, replication, and regulatory functions of CTXΦ. The *rstR* region is classified into *rstR*Class, *rstR*ET, and *rstR*calc, respectively, for classical, El Tor, and O139 alleles [37].

Based on the structure, organization, and location of the CTX prophages (also known as CTX genetic element), clonal diversity was identified using restriction fragment length polymorphism (RFLP). The clonal nature of the US Gulf Coast *V. cholerae* O1 was identified with 6 and 7 kb *Hin*dIII restriction fragments that contained *ctx* gene, and this pattern was not found in strains from other countries [38, 39]. RFLP analysis was made with several *V. cholerae* O1 strains isolated from different countries [40–43]. The O139 *Vibrio* comprised three or more copies of the *ctxA* gene, and the chromosomal locations of these copies were unlike those of the El Tor or classical vibrios [44]. RFLP of *V. cholerae* O139 Bengal that resurged in Kolkata in 1996 was indistinguishable from the earlier strains by ribotyping, but the structure of the CTX genetic element was different [26, 45–47]. Most *V. cholerae* O139 strains isolated in China from 1993 to 1999 had two or more copies of CTX genetic elements and had extensive restriction patterns, even in strains that belong to the same ribotype [48]. This finding suggests multiple origins of the O139 cholera epidemic or sporadic events. Similarly, *V. cholerae* O1 from Iranian cholera outbreak strains carried either three or two copies of the toxin genes [49].

Among newly emerged strains of *V. cholerae* O1 Inaba from India, the presence of CTX prophage was detected in a single site of the chromosome with at least two RS elements [50]. Incidence of cholera in Mozambique was caused by an El Tor biotype *V. cholerae* O1 strain that carried a classical type (CTX^{class}) prophage [51, 52]. Genomic analysis of CTX prophage together with chromosomal phage integration sites showed that these strains carried two copies of prophages located in the small chromosome in tandem, but the excised phage genome was deficient in replication and did not produce CTX^{class} virion [52]. The possible origin of these strains and the existence of the tandem repeat of the classical prophage in them implicate the presence of the classical CTXΦ [51]. Table 5.1 gives an overall description about the localization and number of prophage alleles of *V. cholerae* O1.

The co-culture of a phage and *V. cholerae* or dilutions of phage-positive cholera stools in nutrient medium supported the emergence of phage-resistant derivatives of the vibrios in vitro by losing their O1 antigen [53]. However, in vivo studies did not permit the selection and persistence of phage-resistant variants and the emerging variants and were thus unable to sustain the ongoing epidemic. This may be the reason why identification of new genetic variants is quite rare at the last phase of an outbreak.

Table 5.1 Localization and type of *ctxB/rstR* alleles of *V. cholerae* O1

	Location		
Biotype	Chromosome I	Chromosome II	Type of *ctxB/rstR*
Classical	Yes	Yes	Classical/classical
El Tor	Yes	No	El Tor/El Tor
Mozambique variant	No	Yes	Classical/classical
El Tor variant	Yes	No	Classical/El Tor
Haitian variant	Yes	No	Haitian/ El Tor

V. cholerae ctxB genotyping scheme was made based on the amino acid substitution at positions 39, 46, and 68. *ctxB1* was conserved in the classical biotype and the US Gulf Coast El Tor, whereas *ctxB2* and *ctxB3* were conserved in the Australian, Latin American, and seventh pandemic El Tor isolates [54]. When the DNA sequencing of *ctxB* from *V. cholerae* O1 strains isolated in 29 countries over a period of 70 years were analyzed, 3 types of CT were identified [54]. The base changes correspond to an amino acid substitution in the B subunit of the CT. Genotype 1 was found in classical biotype worldwide and El Tor biotype strains associated with the US Gulf Coast. Genotype 2 was found in El Tor strains from Australia, and genotype 3 was found in El Tor biotype strains that represented seventh pandemic and the Latin American epidemics. The CT genotype 3, which predominated since the early 1960s, has recently been replaced by CT genotype 1 in Bangladesh and India [55, 56]. The recent El Tor strains belonging to CT genotype 1 are found to be associated with several cholera outbreaks in India [57, 58]. Retrospective analysis with the *V. cholerae* O139 strains isolated during 1998–2005 in Bangladesh indicated the prevalence of new CT genotypes such as 4, 5, and 6 [59].

Currently, 13 different *ctxB* genotypes, and a subtype with an extra 11-amino acid repeat designated *ctxB3b*, have been reported, and among these, *ctxB1*, *ctxB3*, and *ctxB7* have been reported in several cholera outbreaks worldwide [60]. During 2001, *ctxB3* (an altered El Tor) was replaced in Asia and Africa with *ctxB1* [55]. During 2006–2010, further change in the CtxB (histidine to asparagine at the 20th position of CtxB) resulting in substitution of *ctxB1* with *ctxB7* was from India, Bangladesh, Africa, and Haiti [61–63]. However, *V. cholerae* O1 isolates with genotype *ctxB1* have been returned back in Bangladesh during 2013–2014 [64]. Figure 5.1 summarizes many recent genomic changes in *V. cholerae* O1 and O139 strains.

5.6 Multilocus Enzyme Electrophoresis

Multilocus enzyme electrophoresis (MLEE) analysis (also known as zymovar analysis) compares genetic variation among a number of housekeeping genes and on the basis of electrophoretic mobility on starch gels. These variations are used to group the *V. cholerae* into electrophoretic types (ETs) [43, 65]. With the use of 16 enzymes, Wachsmuth et al. [43] found existence of four distinct groups of toxigenic El Tor vibrios, namely, the seventh pandemic, US Gulf coast, Australian, and Latin American clones. With the *V. cholerae* non-O1, non-O139 and classical and El Tor strains from America, Africa, Europe, and Asia, Freitas et al. [66] have shown that the same zymovar may contain more than one serogroup and the South American epidemic strain differs from the seventh pandemic El Tor strain. However, the discriminatory power of MLEE is less but useful in distinguishing strains within a single outbreak.

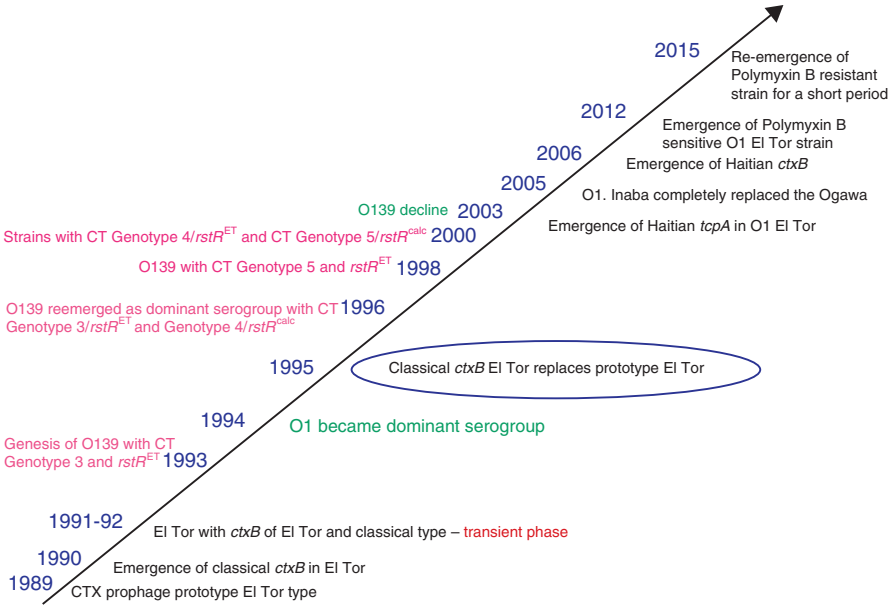

Fig. 5.1 Recorded major events in the evolution of *V. cholerae* O1 El Tor, El Tor variant, and O139 CTX prophages

5.7 Ribotyping

Ribotyping exploits the DNA polymorphism in the chromosomal regions of the highly conserved rRNA genes (*rrn*). *Escherichia coli* RNA is used as a universal probe to screen the restriction patterns of bacterial DNA for the molecular typing purpose. In *V. cholerae*, *Bgl*I is used as a discriminatory enzyme for ribotyping. The *rrn* operons and their flanking regions cause ribotype variation in *V. cholerae* O1 due to recombination in the *rrn* operons [67].

Though the seventh pandemic *V. cholerae* strains from Asia and Africa were clonal and belonged to a single ET, the ribotyping analysis showed that these strains were diverse and belonged to five different types [43]. Based on this observation, it was hypothesized that the observed differences were due to a higher mutation rate in the DNA sequences of flanking rRNA genes than in genes encoding the "housekeeping" enzymes studied using MLEE. A standardized scheme of 27 different *Bgl*I ribotypes and subtypes of *V. cholerae* O1 was developed on the basis of genetic analysis using strains collected over the past 60 years [68]. This analysis revealed 7 and 20 ribotypes among classical and El Tor biotypes, respectively. Six different patterns were found among El Tor vibrios alone. Genetic variation and molecular evolution of sixth and seventh pandemic clone of *V. cholerae* O1 and its relationship with epidemiologically unassociated strains from different countries over 62 years (1931–1993) showed major differences in ribotypes [69]. Majority of seventh

pandemic isolates fell into two groups, the first present from 1961 to the 1993 and found only in Asia and the second rising in 1966 that had spread worldwide.

Ribotype analysis confirmed an association between epidemiologically related clinical isolates and the aquatic environment along with persistence of several clones of the *V. cholerae* O1 from the Australian environment [70]. RFLP of conserved rRNA showed that the El Tor strains isolated before the emergence of *V. cholerae* O139 belonged to four different ribotypes and the one that appeared after the emergence of O139 belonged to a single new ribotype [71]. This finding was fortified by the studies conducted independently in India and Thailand [72, 73]. These results provide evidence that the reemerged El Tor strains represent a new clone of El Tor distinctly different from the earlier clones, which were replaced by the O139 serogroup.

Analysis of *V. cholerae* O139 strains isolated in India and Bangladesh revealed four different ribotypes [17, 74]. Ribotypes I and II were shared by strains isolated from the epidemic outbreak during 1992–1993, ribotype III was represented by a single CT-negative O139 strain from Argentina, and majority of the reemerged strains isolated during 1995–1996 belonged to ribotype IV. These O139 strains may have emerged from similar serotype-specific genetic changes in more than one progenitor. In China, seven different ribotypes were recorded among *V. cholerae* O139 strains isolated between 1993 and 1999, suggesting the diversity of clones in phylogeny [48].

V. cholerae O1 strains isolated in Romania and the Republic of Moldavia (1977–1994) and Somalia (1998–1999) displaced different clones [20, 75]. The RFLP of *Bgl*I-digested DNA probed with five oligonucleotides targeting the conserved regions of 16S and 23S rRNA genes revealed a similar ribotype of *V. cholerae* El Tor strains isolated from outbreaks from different districts of Tehran, Iran [49]. The newly emerged *V. cholerae* O1 Inaba during 2004 and 2005 from different parts of India displayed different ribotypes [50, 76].

Ribotype B27 was identified in the Kenyan cholera epidemic in the 1990s, which has made a rapid spread to West and East Africa during 1994 [77]. *V. cholerae* O1 isolates in 1999–2000 from Thailand disclosed ribotypes D, G, H, and I with the majority of the 2001–2002 isolates showing ribotype G and few belonging to new ribotypes, J and K [78]. In Kolkata, India, *V. cholerae* O1 Ogawa ribotype RIII was replaced by Inaba ribotype RIV in 2005 [50]. Two closely related *V. cholerae* O1 ribotypes, B5a and B8a, were identified in cholera outbreaks that occurred in Somalia from 1998 to 1999 [20]. *Bgl*I rRNA analysis revealed that *V. cholerae* O1 isolated during the 1991–2000 cholera epidemic in Mexico had M5 and M6 ribotype profiles that were identical to those previously designated ribotypes [79]. The other new ribotypes, Mx1, Mx2, and Mx3, were identified among non-toxigenic *V. cholerae* O1 strains isolated between 1998 and 2000 varied from non-toxigenic clones identified in Latin America and on the US Gulf Coast.

5.8 Pulsed-Field Gel Electrophoresis

Pulsed-field gel electrophoresis (PFGE) has proven to be highly effective molecular typing technique for different bacterial species. An international PFGE typing protocol for *V. cholerae* was established for comparison of subtype patterns in the database ([80]; http://pulsenetinternational.org). PFGE was shown to be useful for the identification of spread of specific clones in many cholera outbreak investigations. PFGE results suggested that there was no epidemiological relation among the strains of *V. cholerae* O1 isolated from indigenous cholera in Okinawa during 1994 [81]. With *V. cholerae* O1, 19 subtypes by *Not*I- and *Sfi*I-digested PFGE patterns were identified among Asian strains, suggesting that the pulsotype variation is widely distributed in this region [82]. Based on the PFGE profiling, *V. cholerae* O1 Inaba strains isolated during 1998–1999 in Kolkata, India, were different from the earlier Inaba strains isolated during 1989 but were similar to the prevailing *V. cholerae* O1 Ogawa strains, indicating that the Inaba strains may have the origin of Ogawa strains [83].

V. cholerae O1 strains isolated from a cholera outbreak in Ahmadabad, India, during 2000 resembled a PFGE pattern that was identified in Kolkata many years before, indicating the outbreak was caused by the prevailing clone. However, in the same outbreak, O139 differed in the PFGE patterns with O139 isolates reported during 1992 to 1997 in Kolkata [84]. Clonal analysis using PFGE with non-toxigenic *V. cholerae* O1 strains collected from the Port of Osaka, Japan, during 1987–2001 showed that there is no remarkable change in the fingerprint types [85]. PFGE results obtained from different outbreaks of *V. cholerae* O1 and O139 in Hong Kong and other epidemiologically unrelated strains showed the combination of distinct and similar patterns [86]. In the same study, heterogeneous patterns were also identified among strains that were imported from other parts of Asia. *V. cholerae* O1 strains isolated from two cholera outbreaks (1999–2002) in Thailand demonstrated 17 PFGE banding patterns grouped into four Dice coefficient clusters (PF-I to PF-IV). The prolonged appearance of the clone PF-II, the disappearance of clones PF-I and PF-III, and the emergence of new clones were reported [87]. The patterns of *V. cholerae* O1, El Tor reference strains from Australia, Peru, Romania, and the United States were different from the patterns of reference isolates from Asian countries, such as Bangladesh, India, and Thailand, indicating a close genetic relationship or clonal origin of the isolates in the same geographical region [87].

During 2004–2005, there was a serotype substitution from Ogawa to Inaba in India. New pulsotypes were identified from a cholera outbreak in Delhi during 2004 [76]. The majority of the Inaba isolates belong to "H1" pulsotype, and one isolate is type "H," while the Ogawa isolates were mostly "H" pulsotype [50]. *V. cholerae* O1 Inaba collected during several cholera outbreaks throughout Iran during the summer of 2005 showed an identical ribotype and PFGE patterns in majority of the strains [9]. PFGE analysis of hybrid *V. cholerae* O1 strains isolated during 2004–2005 from cholera patients in Mozambique and Bangladesh showed five closely related patterns and had an El Tor lineage [88]. The restriction patterns grouped the hybrid strains from Mozambique into a separate cluster from Bangladeshi clinical and

environmental strains. This study suggests that hybrid strains differed markedly from classical and El Tor biotypes. In Australia, sporadic cholera was due to indigenous *V. cholerae* O1 El Tor biotype from environmental sources. PFGE analysis revealed that the Australian environmental toxigenic *V. cholerae* O1 strains were more diverse from the non-toxigenic environmental O1 strains [34]. Since there are no reference profiles in the pulsotyping scheme, it is difficult to correlate the enormous data that has been generated through many investigations.

In several countries, the PFGE analysis of latest epidemic strains of *V. cholerae* shown to be related but distinctly different from previous epidemics [89–97]. In Africa, the intercountry spread of cholera is evident as outbreak-associated PFGE profiles of *V. cholerae* O1 found in Ghana during 2014–2015 were found to be similar to DR Congo, Cameroon, Ivory Coast, and Togo isolates [98]. During 2015, a new pulsotype Y15 and a MLVA profile were reported in *V. cholerae* O1 isolates from cholera outbreaks in Myanmar [99]. *V. cholerae* O1 strains isolated in China representing three decades were categorized into three distinct groups in accordance with the epidemiological investigations conducted during the 1980s, 1990s, and 2000s [100].

Analysis of clinical *V. cholerae* isolates from the Haiti outbreak and international travelers returning to the United States from Asia or Africa showed indistinguishable PFGE fingerprints [101]. However, strains isolated in sporadic cholera cases during 2013 displayed PFGE pattern similar to that of *V. cholerae* identified during the epidemic in Haiti in 2010 [102]. In Mexico, *V. cholerae* classical biotype existed during 1991–1997 with endemic cholera, and these isolates differed from Bangladeshi strains indicating its discrete source and evolution [103].

5.9 DNA Sequence-Based Typing Systems

5.9.1 Multilocus Sequence Typing

In multilocus sequence typing (MLST), the genetic variations in multiple housekeeping genes are directly indexed by nucleotide sequencing. This approach is suitable for database storage and software analysis and hence will address long-term epidemiological investigations when the bacterial populations are highly recombinant with large clonal complexes and have substantial time to diversify. An allele number was assigned to each fragment on the basis of its sequence. A sequence type (ST), based on the allelic profile of the several amplicons, was assigned to each strain (http://pubmlst.org/sagalactiae). With the 7 housekeeping genes (*adk*, *gyrB*, *mdh*, *metE*, *pntA*, *purM*, *pyrC*), around 1000 sequence types have been reported in *V. cholerae* (https://pubmlst.org/bigsdb?db=pubmlst_vcholerae_seqdef). An eBURST clonal complex was described as all allelic profiles sharing six identical alleles with at least one other member of the group. Singleton ST refers to an ST that did not cluster into a clonal complex.

MLST with three housekeeping genes, *gyrB*, *pgm*, and *recA*, showed that there was a clustering of epidemic *V. cholerae* O1 and O139 serogroups compared to the

non-epidemic serogroups, and this MLST had the better discriminatory ability than PFGE [104]. The discrimination is more when the number of housekeeping genes increased to nine (*cat*, *chi*, *dnaE*, *gyrB*, *lap*, *pgm*, *recA*, *rstA*, and *gmd*) [100]. With the non-O1, non-O139 strains, MLST revealed that were genetically diverse and clustered in lineages distinct from that of the epidemic strains [35]. The O139 strains were also clustered in several lineages of the dendrogram generated from the matrix of allelic mismatches between the different genotypes [105]. In addition, the application of the Sawyer's test and split decomposition to detect intragenic recombination in the sequenced gene fragments did not indicate the existence of recombination in the tested strains.

Using MLST with 26 housekeeping genes, Salim et al. [106] showed that the US Gulf strains, Australian strains, and some of the strains similar to the El Tor strains belong to the seventh pandemic clone, whereas the sixth pandemic strains were separated in this analysis. An analysis of *V. cholerae* strains isolated between 2004 and 2018 in Public Health England showed that all isolates of *V. cholerae* O1 El Tor and O139 belong to ST69 and the O1 classical variants belong to ST73 [107]. The seventh pandemic strains mostly belonged to ST68 or ST69 [100, 108]. In Taiwan and other Asian countries, ST75 has replaced ST68 since 2009 [108]. From 2002 to 2012, more than 20 genotypes were detected with 2 clonal complexes in Mozambique [109].

5.9.2 Multiple-Locus Variable Number of Tandem Repeat Analysis

In the bacterial genome, repetitive DNA contains monomeric sequences (repeat loci) frequently and is arranged in a head-to-tail configuration. MLVA is used in the genetic analysis of *V. cholerae* that takes advantage of the polymorphism of tandemly repeated DNA sequences. These DNA regions are known as a variable number of tandem repeats (VNTR) that are catalogued on the basis of their repeat unit sizes (ranging from few nucleotides to more than 100 bp). Normally, a VNTR displays a large range of copy numbers, even among highly related strains. For a designated set of tandem repeats, this copy number analysis reveals predictions about the associations at a micro-evolutionary level. In the MLVA, genotype of a strain was expressed as its allelic profile, corresponding to the number of repeats at each VNTR. Many findings use MLVA as a complementary technique to PFGE to generate more detailed differences between bacterial strains that have similar PFGE patterns. The repeat loci in nearly all the VNTR targets are highly conserved, and hence the discrimination power is more comparable to that of MLST.

To explore the relatedness of O139 strains isolated from cholera patients in India, Garg et al. [110] analyzed 9 sequenced loci and found 64 novel alleles distributed among 51 sequence types. Lateral gene transfer (LGT) events produced three times the number of nucleotide changes compared to mutation [110]. In contrast to the traditional concept of epidemic spread of a homogeneous clone, the establishment of variant alleles generated by LGT during the rapid expansion of a clonal bacterial

population may be a paradigm in infections and epidemics. Ghosh et al. [111] evaluated genetic relationships of *V. cholerae* isolates collected between 1992 and 2007 from different states in India by analyzing five VNTR loci. In this study, it was found that each VNTR locus was highly variable, with 5–19 alleles. An eBURST analysis revealed four large groups of genetically related strains. Two groups contained genotypes with O139 serogroup and the other two groups with O1 strains. Using VNTR analysis, it is possible to track the spread of specific genotypes across time and space. It was observed that the minimal overlap in VNTR patterns between the two Bangladeshi communities was consistent, and it was concluded that the small outbreaks of cholera were mainly from local sources [112]. The MLVAType shiny application has been considered to obtain MLVA profiles of *V. cholerae* isolates from WGS data with a longer *k*-mer size [113].

5.10 Fingerprinting of Virulence Genes

It is a general concept that the pathogenic *V. cholerae* is derived from environmental nonpathogenic strains. Potential precursors of new pathogenic strains might require a combination of genes for both ecological fitness and virulence to attain epidemiological predominance. To understand the evolution of pathogenic *V. cholerae* and identify potential precursors of new pathogenic strains, Rahman et al. [32] analyzed environmental or clinical strains for the presence of diverse genes involved in virulence or ecological fitness. This study has shown that 3.9% of the strains carried the pathogenicity island encoding TCP and CT. Few strains carried the TCP island alone and were susceptible to transduction with CTXΦ. Prevalence of putative accessory virulence genes (*mshA*, *hlyA*, and *rtx*) both in toxigenic and non-toxigenic strains of *V. cholerae* supports more recent assumption that these genes impart increased environmental fitness and the epidemiological success of the strains [32].

The evolutionary relationships and multilocus virulence gene profiles of *V. cholerae* isolates indicate that consecutive pandemic strains arose from a common O1 serogroup progenitor through the successive acquisition of new virulence regions. Comparative sequence analysis of malate dehydrogenase gene (*mdh*) revealed that *V. cholerae* O1 and O139 serogroups belonged to the same clonal lineage. Single-strand conformation polymorphism (SSCP) analysis of these O1 and O139 strains at chaperonin (*groE*)-L confirmed the presence of an epidemic clonal complex [31].

5.11 Evidence Showing Intercontinental Spread of *V. cholerae* O1

Several molecular techniques were used for the detection of *V. cholerae* clones that have been spread from one geographical region to the other. With ribotyping and PFGE, spread of a distinct genotype of *V. cholerae* O1 that appeared in Calcutta, India, was detected from cholera cases in Guinea-Bissau from 1993 to 1996 [114]. Molecular epidemiological findings confirm that the epidemic Ukrainian strains are

most closely related to seventh pandemic *V. cholerae* O1 strains from Asia and support a hypothesis that the Ukrainian epidemic during 1994–1995 was caused by toxigenic environmental strains surviving since 1991 [115].

After a decade of absence, *V. cholerae* O1 resurfaced in Italy and Albania during 1994. Ribotype, RAPD, and PFGE patterns indicated that the 1994 isolates belonged to the Asian clone, which emerged in 1990 [116]. The Mozambique *V. cholerae* O1 strains that caused a huge outbreak in 2004 had phenotypic traits of both classical and El Tor biotypes. Interestingly, these strains harbored a CTX prophage in the smaller chromosome similar to that of classical biotype. Retrospective studies conducted in India demonstrated that O1 strains isolated in Kolkata during 1992 were phenotypically identified as El Tor biotype but the *ctxB* was classical genotype [117]. *V. cholerae* O1 Mozambique variant shared most of its genes with the typical El Tor strain N16961 but did not carry the TLC gene cluster and RS1 element adjacent to the CTX prophage [52]. This data further supports the hypothesis that the Mozambique strain has evolved from a progenitor similar to the seventh pandemic strain, involving multiple recombination along with the origination of El Tor strains carrying the classical CTX prophage. Furthermore, the Kolkata strains exhibited an identical ribotype (RI) to that of the Mozambique variant, and the *Not*I pulsotype analysis indicated that the Kolkata O1 strains and the Mozambique variant belonged to closely related clones. Considering the chronological events and the typical identity at the phenotypic and the genotype level, Chatterjee et al. [117] proposed that one of the *V. cholerae* O1 strains from Kolkata in 1992 might have been the progenitor for Mozambique variant O1 strains.

5.12 Quorum-Sensing Systems

During interepidemic periods of cholera, *V. cholerae* survives in aquatic habitats. Recent studies reveal that quorum-sensing systems (QSS) help the vibrios to regulate various cellular functions, pathogenesis, biofilm formation, and protease production [118]. Quorum-sensing systems in geographically diverse *V. cholerae* from epidemic-causing O1 and O139 as well as non-O1, non-O139 and environmental strains revealed an unexpectedly high rate of dysfunctional components and variations in quorum-sensing systems due to environmental selective pressures [118]. The use of QSS as a typing tool may provide several information regarding survival and proliferation of vibrios and subsequent causative agent of cholera in adjacent areas.

5.13 Microarray Analysis

Whole genome sequence of the seventh pandemic El Tor strain N16961 has provided an important tool for addressing questions about the evolution of *V. cholerae* as a human pathogen and environmental organism. To understand *V. cholerae* genome, Dziejman et al. [119] constructed a genomic microarray that displayed

over 93% of the predicted genes of the strain N16961 as spotted features. High degree of positivity was found among the tested strains by DNA hybridization. Genes unique to all pandemic strains as well as genes specific to seventh pandemic El Tor and related O139 serogroup strains were also identified. It was assumed that the odd genes may encode gain-of-function traits specifically associated with displacement of the preexisting classical biotype and might promote the establishment of endemic disease in cholera-free geographical areas.

Pang et al. [120] investigated the genomic diversity of toxigenic and non-toxigenic O1 and O139 strains by comparative genomic microarray hybridization against the genome of El Tor strain N16961. High phylogenetic diversity in non-toxigenic O1 and O139 strains was detected, and most of the genes absent from non-toxigenic strains are clustered together in the N16961 genome. Additionally, sequence variation in virulence-related genes was found in non-toxigenic El Tor strains. The small chromosome of *V. cholerae* is quite conservative outside of the superintegron region (SIR). However, the SIR showed pronounced genetic divergence in both toxigenic and non-toxigenic strains. Comparative genomic microarray analysis of four pathogenic *V. cholerae* non-O1, non-O139 strains indicated that these strains are quite divergent from O1 and O139 strains [121]. In addition, a pathogenic non-O1, non-O139 strain carried a type III secretion system (TTSS). The genes of this *V. cholerae* TTSS appear to be present in many clinical and environmental non-O1, non-O139 strains, including at least one clone that is globally distributed.

5.14 Whole Genome Approach

To understand the origin and relationships of the pandemic clones, Feng et al. [122] did a sequence analysis of genomes of a 1937 pre-pandemic and a sixth pandemic strain and compared them with the published seventh pandemic strain (N16961). Many mutational than recombination events were detected as much as 100-fold higher in seventh pandemic strain compared to the pre-pandemic strains. It was deduced that these pandemic strains have gained pandemic potential, independently with 29 insertions or deletions of genes. There were also substantial changes in the major integron, attributed to gaining of individual cassettes, including replication from within, or loss of blocks of gene cassettes. The genome-based phylogenetic analysis with sequences of *V. cholerae* strains isolated from a variety of sources over the past 98 years revealed 12 distinct lineages, of which, lineage 1 comprises of classical and El Tor biotypes [123]. It was affirmed that transition from sixth to seventh pandemic strains has occurred due to genetic drift with varying genomic composition including laterally trasferred PAIs [123].

Whole genome sequence (WGS) with single-nucleotide polymorphism (SNP) analysis of more than 150 global collections of *V. cholerae* isolates showed that the seventh pandemic has spread from the Bay of Bengal in 3 independent but overlapping waves along with various pancontinental transmission events that influenced the acquisition of the SXT family of antibiotic resistance elements [124]. The *ctxB3*

genotype was found to be associated with wave 1 and *ctxB1* with wave 2 and in early wave 3 strains; and *ctxB7* was found only in the wave 3 strains, including 2010 Haiti cholera outbreak isolates [124].

In West Africa and East/South Africa, the spread of cholera due to single expanded lineage that was introduced at least 11 times since 1970 [125]. Of these, the last five introductions were from Asiatic region. In a subsequent study, cholera epidemics in Uganda showed the spread of type T10 and T13 lineage clones of *V. cholerae* O1 across national borders and East African regions [126]. The cholera epidemics in Latin America during 1991 and 2010 are due to intercontinental introductions of cholera wave 3 with at least seven indigenous lineages [127]. *V. cholerae* O1 isolated during the 1990s from Siberia and the Far East of Russia represented wave 2 and early wave 3 with prophage types of CTX-2a and CTX-3, respectively [128].

Two epidemiological waves have been identified in Yemen during the 2016–2017 outbreak with a single sublineage of the seventh pandemic *V. cholerae* originated from South Asia and that it caused outbreaks in East Africa [129]. These isolates carried susceptibility to many antibiotics, including polymyxin B, and resistance to this drug has been used as a marker of the El Tor biotype. In Tanzania, cholera outbreaks that occurred during 1993–2017 indicated the spread of *V. cholerae* O1 lineage T10 with atypical El Tor (*ctxB1*), which was replaced by the T13 atypical El Tor of the third wave (*ctxB7*) that has caused most cholera outbreaks in East African countries, Yemen, and Lake Victoria. These strains were less drug resistant, as there was a deletion of a 10 Kb region in the SXT element [130].

The major lineages of the seventh pandemic *V. cholerae* in Shanghai during 1961–2011 were found to be closely related to the strains isolated from South or Southeast Asia [131]. Isolates from Papua New Guinea cholera outbreak in 2009–2011 had a tandem repeat of the CTX prophage on chromosome II and without the SXT ICE resembling wave 2 strains that appeared in Southeast and East Asia [132]. Overall, the increasing trend of virulence-related genes in *V. cholerae* genomes suggests the evolutionary advantage of strains in recent years [133].

The WGS of *V. cholerae* has revealed that the core genome is highly conserved, but with several variations in its accessory genome [123, 134]. Changes in the accessory genome are also an important factor in molecular typing. Analysis of publicly available genomes of *Vibrio* spp. provided enormous opportunity to understand the link between genomic contents and continuous evolution of bacterial pathogens (https://www.patricbrc.org). The Pathosystems Resource Integration Center (PATRIC) dendrogram based on the WGS analysis showed similarity between different species of vibrios (Fig. 5.2). The acquisition of virulence factors in *V. cholerae* is implicated with genomic islands (GIs), and hence this information is important in the identification of resurgence and emerging pathogenic strains. GI scanner-based generation of algorithm has been shown to be helpful in the detection of such GIs through comparative genomics and phylogeny-based predictions using *V. cholerae* Genomic Island Database (VCGIDB) [135].

V. cholerae is a naturally competent bacterial species and able to uptake naked DNA from the environment [136, 137]. Beyond the genetic acquisition of novel

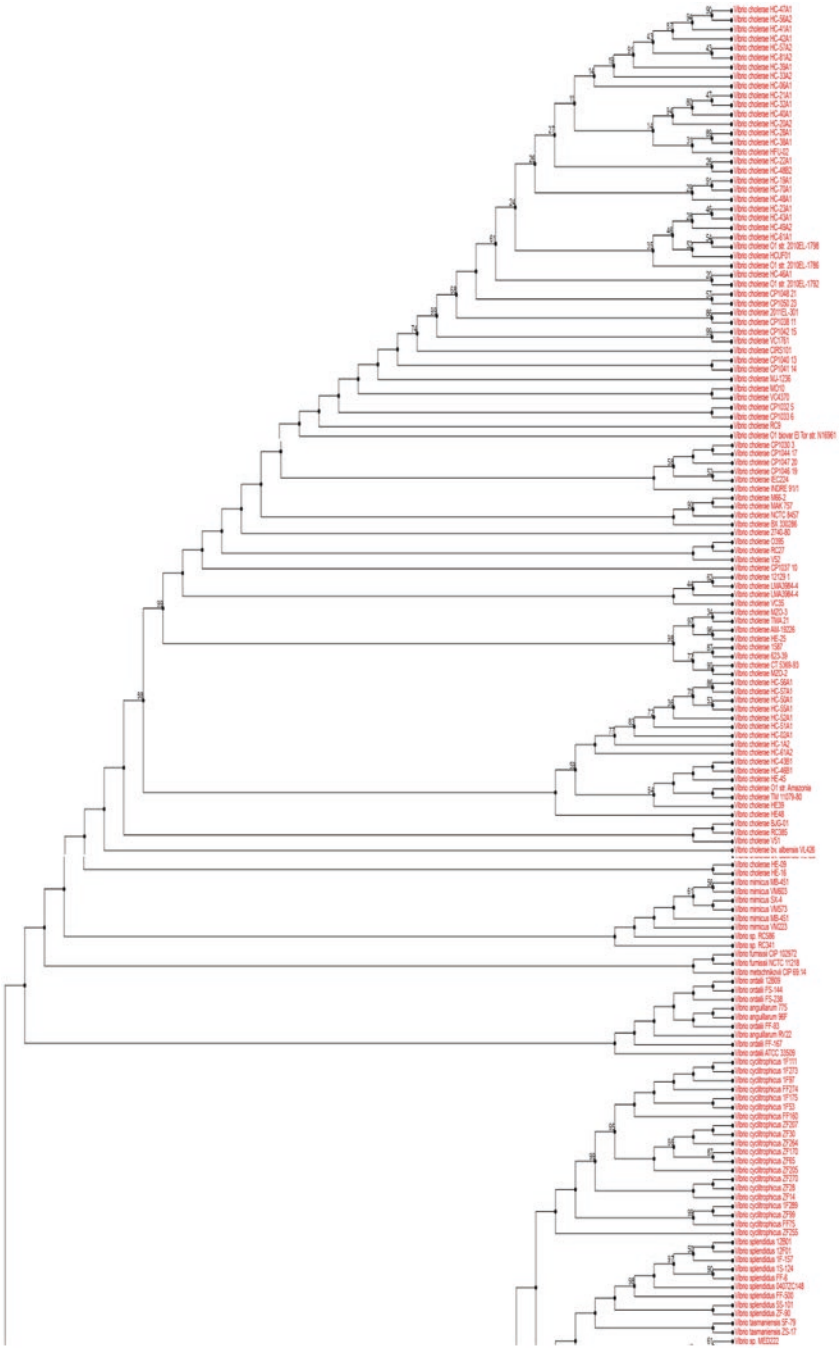

Fig. 5.2 Whole genome-based phylogenetic relationships of vibrios

DNA sequences, natural transformation in *V. cholerae* can also promote the nongenetic inheritance of traits during this conserved mechanism of horizontal gene transfer [138]. Analysis of the genome of *V. cholerae* revealed that the bipartite genome of cholera pathogen has acquired more than 260 accessory functions, which provide virulence (VPI-1, CTXΦ), antibiotic resistance (ICE, integron), metabolic functions (VPI-2, VSP-1, VSP-2), and bactericidal arsenals (GIs) [139]. Both the chromosomes of *V. cholerae* have several hotspots (*dif*, *tmRNA*, *tRNA-Ser*, *tRNA-Met*) that facilitate MGEs, which carry accessory genome to integrate into the chromosome by site-specific recombination using their own DNA recombinases or chromosomally encoded tyrosine recombinases [140]. Some of the accessory genomes are highly stable and transfer to the progeny with the core genome without any detectable event of loss. But the stability of the acquired genome widely varies depending on the nature of MGEs and mode of integration.

5.15 Epilogue

Research on cholera is always inundated with several challenges, especially when the pathogenic clones carry diverse combinations of phenotypic and genotypic properties. Newer concepts are often being proposed for dynamicity of the pathogen supported by different genomic analyses. For example, co-culture of a phage and *V. cholerae* or dilutions of phage-positive cholera stools in nutrient medium supported the emergence of phage-resistant derivatives of the vibrios in vitro by losing their O1 antigen [53]. However, in vivo studies did not permit the selection and persistence of phage-resistant variants, and the emerging variants are thus unable to sustain the ongoing epidemic. This may be the reason why identification of new genetic variants is quite rare during the last phase of an outbreak. The genome of *V. cholerae* is well understood, and several molecular typing methods exist for the detection of subtle changes. Multidisciplinary genomic analysis alone may not be the right direction as we might overlook the simple nucleotide changes that may influence the disease outcome. For example, simple mutational changes such as the one detected in the *ctxB* and the emergence of new CT genotypes of *V. cholerae* were shown to manifest considerable impact on the epidemiology of cholera globally [141, 142]. As mentioned in several sections of this chapter, many molecular typing techniques are now available and are used in different situations and times. In the future, the WGS analysis might cover most of the DNA-based typing techniques and also provide scope for the other novel concepts. However, the use of appropriate method is left to the investigator considering the gravity of the problem, experimental cost, and existing facilities.

5.16 Summary

The disease cholera is caused by toxigenic *Vibrio cholerae* with acute watery diarrhea and severe dehydration. From the known history, cholera has appeared several times in the form of seven pandemics since 1817. Before the age of molecular typing, biochemical and serological characterization of pure cultures are used in identification and classification of bacteria. The DNA-based molecular techniques are not only simple and cost-effective but also less time-consuming and highly precise. Recent advances in molecular studies have facilitated in the development of several highly discriminating techniques, which in turn have assisted not only in investigating the genetic diversity of *V. cholerae* but also changing epidemiology of cholera. Some of the initial techniques such as polymerase chain reaction (PCR)-based assays; analysis of mobile genetic elements, including integrons, pathogenicity islands, and CTXΦ that encode toxin-encoding genes; multilocus enzyme electrophoresis (MLEE); rRNA gene-based ribotyping; pulsed-field gel electrophoresis (PFGE); DNA sequence-based multilocus sequence typing (MLST); and variable number of tandem repeat analysis have helped in the identification of strains that have been involved in international spread. Several cholera toxin (CT) genotypes have been identified in the seventh pandemic *V. cholerae* O1 El Tor. Whole genome sequence (WGS) data have provided an important tool to address many questions about the evolution of *V. cholerae*. Analysis of genetic polymorphisms provides information on variations in phenotypes, antimicrobial resistance, nature of virulence, etc. In the WGS analysis, single-nucleotide polymorphism (SNP) markers are generally identified without any prior genome reduction or a reference genome to assemble large sequencing data that allow the discovery and genetic mapping of thousands of molecular markers. WGS analysis has identified appearance of three waves of cholera during the seventh pandemic. Each wave has specific genetic markers. The large cholera epidemics in Africa, Latin America, and Haiti are well investigated using the WGS. This chapter provides an outline on the usefulness of molecular typing methods of *V. cholerae* and salient features of information generated from these studies.

References

1. World Health Organization (2019) Cholera 2018. Weekly Epidemiol Rec 94:561–580
2. Sack DA, Sack RB, Nair GB, Siddique AK (2004) Cholera. Lancet 363:223–233
3. Pollitzer R (1959) Cholera. World Health Organization, Geneva
4. Chattopadhyay DJ, Sarkar BL, Ansari MQ, Chakrabarti BK, Roy MK, Ghosh AN, Pal SC (1993) New phage typing scheme for *Vibrio cholerae* O1 biotype El Tor strains. J Clin Microbiol 31:1579–1585
5. Chakrabarti AK, Ghosh AN, Nair GB, Niyogi SK, Bhattacharya SK, Sarkar BL (2000) Development and evaluation of a phage typing scheme for *Vibrio cholerae* O139. J Clin Microbiol 38:44–49
6. Makino S, Kurazono T, Okuyama Y, Shimada T, Okada Y, Sasakawa C (1995) Diversity of DNA sequences among *Vibrio cholerae* O139 Bengal detected by PCR-based DNA fingerprinting. FEMS Microbiol Lett 126:43–48

7. Radu S, Ho YK, Lihan S, Yuherman RG, Yasin RM, Khair J, Elhadi N (1999) Molecular characterization of *Vibrio cholerae* O1 and non-O1 from human and environmental sources in Malaysia. Epidemiol Infect 123:225–232
8. Leal NC, Sobreira M, Leal-Balbino TC, de Almeida AM, de Silva MJ, Mello DM, Seki LM, Hofer E (2004) Evaluation of a RAPD-based typing scheme in a molecular epidemiology study of *Vibrio cholerae* O1, Brazil. J Appl Microbiol 96:447–454
9. Pourshafie MR, Bakhshi B, Ranjbar R, Sedaghat M, Sadeghifard N, Zaemi Yazdi J, Parzadeh M, Raesi J (2007) Dissemination of a single *Vibrio cholerae* clone in cholera outbreaks during 2005 in Iran. J Med Microbiol 56:1615–1619
10. Opintan JA, Newman MJ, Nsiah-Poodoh OA, Okeke IN (2008) *Vibrio cholerae* O1 from Accra, Ghana carrying a class 2 integron and the SXT element. J Antimicrob Chemother 62:929–933
11. Jiang SC, Matte M, Matte G, Huq A, Colwell RR (2000) Genetic diversity of clinical and environmental isolates of *Vibrio cholerae* determined by amplified fragment length polymorphism fingerprinting. Appl Environ Microbiol 66:148–153
12. Castañeda NC, Pichel M, Orman B, Binsztein N, Roy PH, Centrón D (2005) Genetic characterization of *Vibrio cholerae* isolates from Argentina by *V. cholerae* repeated sequences-polymerase chain reaction. Diagn Microbiol Infect Dis 53:175–183
13. Rivera IG, Chowdhury MA, Huq A, Jacobs D, Martins MT, Colwell RR (1995) Enterobacterial repetitive intergenic consensus sequences and the PCR to generate fingerprints of genomic DNAs from *Vibrio cholerae* O1, O139, and non-O1 strains. Appl Environ Microbiol 61:2898–2904
14. Zo YG, Rivera IN, Russek-Cohen E, Islam MS, Siddique AK, Yunus M, Sack RB, Huq A, Colwell RR (2002) Genomic profiles of clinical and environmental isolates of *Vibrio cholerae* O1 in cholera-endemic areas of Bangladesh. Proc Natl Acad Sci U S A 99:12409–12414
15. Shangkuan YH, Tsao CM, Lin HC (1997) Comparison of *Vibrio cholerae* O1 isolates by polymerase chain reaction fingerprinting and ribotyping. J Med Microbiol 46:941–948
16. Labbate M, Boucher Y, Joss MJ, Michael CA, Gillings MR, Stokes HW (2007) Use of chromosomal integron arrays as a phylogenetic typing system for *Vibrio cholerae* pandemic strains. Microbiology 153:1488–1498
17. Faruque SM, Siddique AK, Saha MN, Asadulghani RMM, Zaman K, Albert MJ, Sack DA, Sack RB (1999) Molecular characterization of a new ribotype of *Vibrio cholerae* O139 Bengal associated with an outbreak of cholera in Bangladesh. J Clin Microbiol 37:1313–1318
18. Sarkar A, Morita D, Ghosh A, Chowdhury G, Mukhopadhyay AK, Okamoto K, Ramamurthy T (2019) Altered integrative and conjugative elements (ICEs) in recent *Vibrio cholerae* O1 isolated from cholera cases, Kolkata, India. Front Microbiol 10:2072
19. McDonald ND, Regmi A, Morreale DP, Borowski JD, Boyd EF (2019) CRISPR-Cas systems are present predominantly on mobile genetic elements in *Vibrio* species. BMC Genomics 20(1):105
20. Scrascia M, Pugliese N, Maimone F, Mohamud KA, Grimont PA, Materu SF, Pazzani C (2009) Clonal relationship among *Vibrio cholerae* O1 El Tor strains isolated in Somalia. Int J Med Microbiol 299:203–207
21. Folster JP, Katz L, McCullough A, Parsons MB, Knipe K, Sammons SA, Boncy J, Tarr CL, Whichard JM (2014) Multidrug-resistant IncA/C plasmid in *Vibrio cholerae* from Haiti. Emerg Infect Dis 20:1951–1953
22. Wang R, Yu D, Zhu L, Li J, Yue J, Kan B (2015) IncA/C plasmids harboured in serious multidrug-resistant *Vibrio cholerae* serogroup O139 strains in China. Int J Antimicrob Agents 45:249–254
23. Bik EM, Gouw RD, Mooi FR (1996) DNA fingerprinting of *Vibrio cholerae* strains with a novel insertion sequence element: a tool to identify epidemic strains. J Clin Microbiol 34:1453–1461
24. Dumontier S, Trieu-Cuot P, Berche P (1998) Structural and functional characterization of IS*1358* from *Vibrio cholerae*. J Bacteriol 180:6101–6106

25. Pugliese N, Maimone F, Scrascia M, Materu SF, Pazzani C (2009) SXT-related integrating conjugative element and IncC plasmids in *Vibrio cholerae* O1 strains in eastern Africa. J Antimicrob Chemother 63:438–442
26. Mukhopadhyay AK, Basu A, Garg P, Bag PK, Ghosh A, Bhattacharya SK, Takeda Y, Nair GB (1998) Molecular epidemiology of reemergent *Vibrio cholerae* O139 Bengal in India. J Clin Microbiol 36:2149–2152
27. Pande K, Mendiratta DK, Vijayashri D, Thamke DC, Narang P (2012) SXT constin among *Vibrio cholerae* isolates from a tertiary care hospital. Indian J Med Res 135:346–350
28. Karaolis DK, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR (1998) A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. Proc Natl Acad Sci U S A 95:3134–3139
29. Karaolis DK, Lan R, Kaper JB, Reeves PR (2001) Comparison of *Vibrio cholerae* pathogenicity islands in sixth and seventh pandemic strains. Infect Immun 69:1947–1952
30. Osin AV, Nefedov KS, Eroshenko GA, Smirnova NI (2005) Comparative genomic analysis of *Vibrio cholerae* El Tor pre-seventh and seventh pandemic strains isolated in various periods. Genetika 41:53–62. [In Russian]
31. O'Shea YA, Reen FJ, Quirke AM, Boyd EF (2004) Evolutionary genetic analysis of the emergence of epidemic *Vibrio cholerae* isolates on the basis of comparative nucleotide sequence analysis and multilocus virulence gene profiles. J Clin Microbiol 42:4657–4671
32. Rahman MH, Biswas K, Hossain MA, Sack RB, Mekalanos JJ, Faruque SM (2008) Distribution of genes for virulence and ecological fitness among diverse *Vibrio cholerae* population in a cholera endemic area: tracking the evolution of pathogenic strains. DNA Cell Biol 27:347–355
33. Nair GB, Safa A, Bhuiyan NA, Nusrin S, Murphy D, Nicol C, Valcanis M, Iddings S, Kubuabola I, Vally H (2006) Isolation of *Vibrio cholerae* O1 strains similar to pre-seventh pandemic El Tor strains during an outbreak of gastrointestinal disease in an island resort in Fiji. J Med Microbiol 55:1559–1562
34. Safa A, Bhuiyan NA, Murphy D, Bates J, Nusrin S, Kong RY, Chongsanguan M, Chaicumpa W, Nair GB (2009) Multilocus genetic analysis reveals that the Australian strains of *Vibrio cholerae* O1 are similar to the pre-seventh pandemic strains of the El Tor biotype. J Med Microbiol 58:105–111
35. Li M, Kotetishvili M, Chen Y, Sozhamannan S (2003) Comparative genomic analyses of the vibrio pathogenicity island and cholera toxin prophage regions in nonepidemic serogroup strains of *Vibrio cholerae*. Appl Environ Microbiol 69:1728–1738
36. Waldor MK, Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272:1910–1914
37. Bhattacharya T, Chatterjee S, Mait D, Bhadra RK, Takeda Y, Nair GB, Nandy RK (2006) Molecular analysis of the *rstR* and *orfU* genes of the CTX prophages integrated in the small chromosomes of environmental *Vibrio cholerae* non-O1, non-O139 strains. Environ Microbiol 8:526–634
38. Kaper JB, Bradford HB, Roberts NC, Falkow S (1982) Molecular epidemiology of *Vibrio cholerae* in the U.S. Gulf Coast. J Clin Microbiol 16:129–134
39. Kaper JB, Nataro JP, Roberts NC, Siebeling RJ, Bradford HB (1986) Molecular epidemiology of non-O1 *Vibrio cholerae* and *Vibrio mimicus* in the U.S. Gulf Coast region. J Clin Microbiol 23:652–654
40. Finch MJ, Morris JG, Kaviti J, Kagwanja W, Levine MM (1988) Epidemiology of antimicrobial resistant cholera in Kenya and East Africa. Am J Trop Med Hyg 39:484–490
41. Desmarchelier PM, Senn CR (1989) A molecular epidemiological study of *V. cholerae* in Australia. Med J Aust 150:631–634
42. Yam WC, Lung ML, Ng KY, Ng MH (1989) Molecular epidemiology of *Vibrio cholerae* in Hong Kong. J Clin Microbiol 27:1900–1902
43. Wachsmuth IK, Evins GM, Fields PI, Olsvik O, Popovic T, Bopp CA, Wells JG, Carrillo C, Blake PA (1993) The molecular epidemiology of cholera in Latin America. J Infect Dis 167:621–626

44. Faruque SM, Abdul Alim AR, Roy SK, Khan F, Nair GB, Sack RB, Albert MJ (1994) Molecular analysis of rRNA and cholera toxin genes carried by the new epidemic strain of toxigenic *Vibrio cholerae* O139 synonym Bengal. J Clin Microbiol 32:1050–1053
45. Sharma C, Maiti S, Mukhopadhyay AK, Basu A, Basu I, Nair GB, Mukhopadhyaya R, Das B, Kar S, Ghosh RK, Ghosh A (1997) Unique organization of the CTX genetic element in *Vibrio cholerae* O139 strains which reemerged in Calcutta, India, in September 1996. J Clin Microbiol 35:3348–3350
46. Basu A, Mukhopadhyay AK, Sharma C, Jyot J, Gupta N, Ghosh A, Bhattacharya SK, Takeda Y, Faruque AS, Albert MJ, Balakrish Nair G (1998) Heterogeneity in the organization of the CTX genetic element in strains of *Vibrio cholerae* O139 Bengal isolated from Calcutta, India and Dhaka, Bangladesh and its possible link to the dissimilar incidence of O139 cholera in the two locales. Microb Pathog 24:175–183
47. Basu A, Garg P, Datta S, Chakraborty S, Bhattacharya T, Khan A, Ramamurthy T, Bhattacharya SK, Yamasaki S, Takeda Y, Nair GB (2000) *Vibrio cholerae* O139 in Calcutta, 1992–1998: incidence, antibiograms, and genotypes. Emerg Infect Dis 6:139–147
48. Qu M, Xu J, Ding Y, Wang R, Liu P, Kan B, Qi G, Liu Y, Gao S (2003) Molecular epidemiology of *Vibrio cholerae* O139 in China: polymorphism of ribotypes and CTX elements. J Clin Microbiol 41:2306–2310
49. Pourshafie MR, Grimont F, Saifi M, Grimont PA (2000) Molecular epidemiological study of *Vibrio cholerae* isolates from infected patients in Teheran, Iran. J Med Microbiol 49:1085–1090
50. Raychoudhuri A, Chatterjee S, Pazhani GP, Nandy RK, Bhattacharya MK, Bhattacharya SK, Ramamurthy T, Mukhopadhyay AK (2007) Molecular characterization of recent *Vibrio cholerae* O1, El Tor, Inaba strains isolated from hospitalized patients in Kolkata, India. J Infect 55:431–438
51. Lee JH, Han KH, Choi SY, Lucas ME, Mondlane C, Ansaruzzaman M, Nair GB, Sack DA, von Seidlein L, Clemens JD, Song M, Chun J, Kim DW (2006) Mozambique cholera vaccine demonstration project coordination group. Multilocus sequence typing (MLST) analysis of *Vibrio cholerae* O1 El Tor isolates from Mozambique that harbour the classical CTX prophage. J Med Microbiol 55:165–170
52. Faruque SM, Tam VC, Chowdhury N, Diraphat P, Dziejman M, Heidelberg JF, Clemens JD, Mekalanos JJ, Nair GB (2007) Genomic analysis of the Mozambique strain of *Vibrio cholerae* O1 reveals the origin of El Tor strains carrying classical CTX prophage. Proc Natl Acad Sci U S A 104:5151–5156
53. Zahid MS, Udden SM, Faruque AS, Calderwood SB, Mekalanos JJ, Faruque SM (2008) Effect of phage on the infectivity of *Vibrio cholerae* and emergence of genetic variants. Infect Immun 76:5266–5273
54. Olsvik O, Wahlberg J, Petterson B, Uhlén M, Popovic T, Wachsmuth IK, Fields PI (1993) Use of automated sequencing of polymerase chain reaction-generated amplicons to identify three types of cholera toxin subunit B in *Vibrio cholerae* O1 strains. J Clin Microbiol 31:22–25
55. Nair GB, Qadri F, Holmgren J, Svennerholm AM, Safa A, Bhuiyan NA, Ahmad QS, Faruque SM, Faruque AS, Takeda Y, Sack DA (2006) Cholera due to altered El Tor strains of *Vibrio cholerae* O1 in Bangladesh. J Clin Microbiol 44:4211–4213
56. Raychoudhuri A, Patra T, Ghosh K, Ramamurthy T, Nandy RK, Takeda Y, Balakrish Nair G, Mukhopadhyay AK (2009) Classical *ctxB* in *Vibrio cholerae* O1, Kolkata, India. Emerg Infect Dis 15:131–132
57. Kumar P, Jain M, Goel AK et al (2009) A large cholera outbreak due to a new cholera toxin variant of the *Vibrio cholerae* O1 El Tor biotype in Orissa, eastern India. J Med Microbiol 58:234–238
58. Taneja N, Mishra A, Sangar G, Singh G, Sharma M (2009) Outbreak caused by new variants of *Vibrio cholerae* O1 El Tor, India. Emerg Infect Dis 15:352–354
59. Bhuiyan NA, Nusrin S, Alam M, Morita M, Watanabe H, Ramamurthy T, Cravioto A, Nair GB (2009) Changing genotypes of cholera toxin (CT) of *Vibrio cholerae* O139 in Bangladesh and description of three new CT genotypes. FEMS Immunol Med Microbiol 57:136–141

60. Kim EJ, Lee CH, Nair GB, Kim DW (2015) Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. Trends Microbiol 23:479–489
61. Naha A, Pazhani GP, Ganguly M, Ghosh S, Ramamurthy T, Nandy RK, Nair GB, Takeda Y, Mukhopadhyay AK (2012) Development and evaluation of a PCR assay for tracking the emergence and dissemination of Haitian variant *ctxB* in *Vibrio cholerae* O1 strains isolated from Kolkata, India. J Clin Microbiol 50:1733–1736
62. Hasan NA, Choi SY, Eppinger M, Clark PW, Chen A, Alam M, Haley BJ, Taviani E, Hine E, Su Q, Tallon LJ, Prosper JB, Furth K, Hoq MM, Li H, Fraser-Liggett CM, Cravioto A, Huq A, Ravel J, Cebula TA, Colwell RR (2012) Genomic diversity of 2010 Haitian cholera outbreak strains. Proc Natl Acad Sci U S A 109:E2010–E2017
63. Rashed SM, Mannan SB, Johura FT, Islam MT, Sadique A, Watanabe H, Sack RB, Huq A, Colwell RR, Cravioto A, Alam M (2012) Genetic characteristics of drug-resistant *Vibrio cholerae* O1 causing endemic cholera in Dhaka, 2006–2011. J Med Microbiol 61:1736–1745
64. Rashid MU, Rashed SM, Islam T, Johura FT, Watanabe H, Ohnishi M, Alam M (2016) CtxB1 outcompetes CtxB7 in *Vibrio cholerae* O1, Bangladesh. J Med Microbiol 65:101–103
65. Momen H, Salles CA (1985) Enzyme markers for *Vibrio cholerae*: identification of classical, El Tor and environmental strains. Trans R SocTrop Med Hyg 79:773–776
66. Freitas FS, Momen H, Salles CA (2002) The zymovars of *Vibrio cholerae*: multilocus enzyme electrophoresis of *Vibrio cholerae*. Mem Inst Oswaldo Cruz 97:511–516
67. Lan R, Reeves PR (1998) Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of *Vibrio cholerae*. Microbiology 144:1213–1221
68. Popovic T, Bopp C, Olsvik O, Wachsmuth K (1993) Epidemiologic application of a standardized ribotype scheme for *Vibrio cholerae* O1. J Clin Microbiol 31:2474–2482
69. Karaolis DK, Lan R, Reeves PR (1994) Molecular evolution of the seventh-pandemic clone of *Vibrio cholerae* and its relationship to other pandemic and epidemic *V. cholerae* isolates. J Bacteriol 176:6199–6206
70. Desmarchelier PM, Wong FY, Mallard K (1995) An epidemiological study of *Vibrio cholerae* O1 in the Australian environment based on rRNA gene polymorphisms. Epidemiol Infect 115:435–446
71. Faruque SM, Ahmed KM, Abdul Alim AR, Qadri F, Siddique AK, Albert MJ (1997) Emergence of a new clone of toxigenic *Vibrio cholerae* O1 biotype El Tor displacing *V. cholerae* O139 Bengal in Bangladesh. J Clin Microbiol 35:624–630
72. Bag PK, Maiti S, Sharma C, Ghosh A, Basu A, Mitra R, Bhattacharya SK, Nakamura S, Yamasaki S, Takeda Y, Balakrish Nair G (1998) Rapid spread of the new clone of *Vibrio cholerae* O1 biotype El Tor in cholera endemic areas in India. Epidemiol Infect 121:245–251
73. Dalsgaard A, Serichantalergs O, Forslund A, Pitarangsi C, Echeverria P (1998) Phenotypic and molecular characterization of *Vibrio cholerae* O1 isolated in Samutsakorn, Thailand before, during and after the emergence of *V. cholerae* O139. Epidemiol Infect 121:259–268
74. Faruque SM, Ahmed KM, Siddique AK, Zaman K, Alim AR, Albert MJ (1997) Molecular analysis of toxigenic *Vibrio cholerae* O139 Bengal strains isolated in Bangladesh between 1993 and 1996: evidence for emergence of a new clone of the Bengal vibrios. J Clin Microbiol 35:2299–2306
75. Damian M, Koblavi S, Carle I, Nacescu N, Grimont F, Ciufecu C, Grimont PA (1998) Molecular characterization of *Vibrio cholerae* O1 strains isolated in Romania. Res Microbiol 149:745–755
76. Dutta B, Ghosh R, Sharma NC, Pazhani GP, Taneja N, Raychowdhuri A, Sarkar BL, Mondal SK, Mukhopadhyay AK, Nandy RK, Bhattacharya MK, Bhattacharya SK, Ramamurthy T (2006) Spread of cholera with newer clones of *Vibrio cholerae* O1 El Tor, serotype Inaba, in India. J Clin Microbiol 44:3391–3393
77. Scrascia M, Maimone F, Mohamud KA, Materu SF, Grimont F, Grimont PA, Pazzani C (2006) Clonal relationship among *Vibrio cholerae* O1 El Tor strains causing the largest cholera epidemic in Kenya in the late 1990s. J Clin Microbiol 44:3401–3404

78. Tapchaisri P, Na-Ubol M, Jaipaew J, Srimanote P, Chongsa-Nguan M, Yamasaki S, Hayashi H, Nair GB, Kurazono H, Chaicumpa W (2007) Virulence genes of clinical *Vibrio cholerae* O1 isolates in Thailand and their ribotypes. J Infect 55:557–565
79. Lizárraga-Partida ML, Quilici ML (2009) Molecular analyses of *Vibrio cholerae* O1 clinical strains, including new nontoxigenic variants isolated in Mexico during the cholera epidemic years between 1991 and 2000. J Clin Microbiol 47:1364–1371
80. Cooper KL, Luey CK, Bird M, Terajima J, Nair GB, Kam KM, Arakawa E, Safa A, Cheung DT, Law CP, Watanabe H, Kubota K, Swaminathan B, Ribot EM (2006) Development and validation of a PulseNet standardized pulsed-field gel electrophoresis protocol for subtyping of *Vibrio cholerae*. Foodborne Pathog Dis 3:51–58
81. Iwanaga M, Honma Y, Enami M (1997) Molecular epidemiology of *Vibrio cholerae* O1 isolated from sporadic cholera cases in Okinawa, Japan. Microbiol Immunol 41:861–864
82. Arakawa E, Murase T, Matsushita S, Shimada T, Yamai S, Ito T, Watanabe H (2000) Pulsed-field gel electrophoresis-based molecular comparison of *Vibrio cholerae* O1 isolates from domestic and imported cases of cholera in Japan. J Clin Microbiol 38:424–426
83. Garg P, Nandy RK, Chaudhury P, Chowdhury NR, De K, Ramamurthy T, Yamasaki S, Bhattacharya SK, Takeda Y, Nair GB (2000) Emergence of *Vibrio cholerae* O1 biotype El Tor serotype Inaba from the prevailing O1 Ogawa serotype strains in India. J Clin Microbiol 38:4249–4253
84. Chakraborty S, Deokule JS, Garg P, Bhattacharya SK, Nandy RK, Nair GB, Yamasaki S, Takeda Y, Ramamurthy T (2001) Concomitant infection of enterotoxigenic *Escherichia coli* in an outbreak of cholera caused by *Vibrio cholerae* O1 and O139 in Ahmedabad, India. J Clin Microbiol 39:3241–3246
85. Miyagi K, Nakano T, Yagi T, Hanafusa M, Imura S, Honda T, Nakano Y, Sano K (2003) Survey of *Vibrio cholerae* O1 and its survival over the winter in marine water of Port of Osaka. Epidemiol Infect 131:613–619
86. Kam KM, Luey CK, Tsang YM, Law CP, Chu MY, Cheung TL, Chiu AW (2003) Molecular subtyping of *Vibrio cholerae* O1 and O139 by pulsed-field gel electrophoresis in Hong Kong: correlation with epidemiological events from 1994 to 2002. J Clin Microbiol 41:4502–4511
87. Tapchaisri P, Na-Ubol M, Tiyasuttipan W, Chaiyaroj SC, Yamasaki S, Wongsaroj T, Hayashi H, Nair GB, Chongsa-Nguan M, Kurazono H, Chaicumpa W (2008) Molecular typing of *Vibrio cholerae* O1 isolates from Thailand by pulsed-field gel electrophoresis. J Health Popul Nutr 26:79–87
88. Ansaruzzaman M, Bhuiyan NA, Safa A, Sultana M, McUamule A, Mondlane C, Wang XY, Deen JL, von Seidlein L, Clemens JD, Lucas M, Sack DA, Balakrish Nair G (2007) Genetic diversity of El Tor strains of *Vibrio cholerae* O1 with hybrid traits isolated from Bangladesh and Mozambique. Int J Med Microbiol 297:443–449
89. Keddy KH, Nadan S, Govind C, Sturm AW, Group for Enteric, Respiratory and Meningeal Disease Surveillance in South Africa (2007) Evidence for a clonally different origin of the two cholera epidemics of 2001–2002 and 1980–1987 in South Africa. J Med Microbiol 56:1644–1650
90. Chomvarin C, Jumroenjit W, Wongboot W, Kanoktippornchai B, Chaimanee P, Jamjane O, Huttayananont S, Tangkanakul W (2012) Molecular analysis and antimicrobial resistance of *Vibrio cholerae* O1 in northeastern Thailand. Southeast Asian J Trop Med Public Health 43:1437–1446
91. Ismail H, Smith AM, Tau NP, Sooka A, Keddy KH, Group for Enteric, Respiratory and Meningeal Disease Surveillance in South Africa (2013) Cholera outbreak in South Africa, 2008–2009: laboratory analysis of *Vibrio cholerae* O1 strains. J Infect Dis 208(Suppl 1):S39–S45
92. Mahmud ZH, Islam S, Zaman RU, Akter M, Talukder KA, Bardhan PK, Khan AI, Rhodes FC, Kamara A, Wurie IM, Alemu W, Jambai A, Faruque SM, Clemens JD, Islam MS (2014) Phenotypic and genotypic characteristics of *Vibrio cholerae* O1 isolated from the Sierra Leone cholera outbreak in 2012. Trans R Soc Trop Med Hyg 108:715–720

93. Mercy N, Mohamed AA, Zipporah N, Chowdhury G, Pazhani GP, Ramamurthy T, Boga HI, Kariuki SM, Joseph O (2014) Phenotypic and genetic characterization of *Vibrio cholerae* O1 isolated from various regions of Kenya between 2007 and 2010. Pan Afr Med J 19:8
94. Ghosh R, Sharma NC, Halder K, Bhadra RK, Chowdhury G, Pazhani GP, Shinoda S, Mukhopadhyay AK, Nair GB, Ramamurthy T (2016) Phenotypic and genetic heterogeneity in *Vibrio cholerae* O139 isolated from cholera cases in Delhi, India during 2001–2006. Front Microbiol 7:1250
95. Nguyen VH, Pham HT, Diep TT, Phan CD, Nguyen TQ, Nguyen NT, Ngo TC, Nguyen TV, Do QK, Phan HC, Nguyen BM, Ehara M, Ohnishi M, Yamashiro T, Nguyen LT, Izumiya H (2016) *Vibrio cholerae* O1 El Tor from southern Vietnam in 2010 was molecularly distinct from that present from 1999 to 2004. Epidemiol Infect 144:1241–1247
96. Lü H, Yuan Y, Sun N, Bi Z, Guan B, Shao K, Wang T, Bi Z (2017) Characterization of *Vibrio cholerae* isolates from 1976 to 2013 in Shandong Province, China. Braz J Microbiol 48:173–179
97. Hajia M, Sohrabi A (2019) In silico characteristics for reemerging possibility of *Vibrio cholerae* genotypes in Iran. New Microbes New Infect 31:100577
98. Feglo PK, Sewurah M (2018) Characterization of highly virulent multidrug resistant *Vibrio cholerae* isolated from a large cholera outbreak in Ghana. BMC Res Notes 11(1):45
99. Roobthaisong A, Okada K, Htun N, Aung WW, Wongboot W, Kamjumphol W, Han AA, Yi Y, Hamada S (2017) Molecular epidemiology of cholera outbreaks during the rainy season in Mandalay, Myanmar. Am J Trop Med Hyg 97:1323–1328
100. Liao F, Mo Z, Chen M, Pang B, Fu X, Xu W, Jing H, Kan B, Gu W (2018) Comparison and evaluation of the molecular typing methods for toxigenic *Vibrio cholerae* in Southwest China. Front Microbiol 9:905
101. Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, Talkington D, Rowe L, Olsen-Rasmussen M, Frace M, Sammons S, Dahourou GA, Boncy J, Smith AM, Mabon P, Petkau A, Graham M, Gilmour MW, Gerner-Smidt P, V. cholerae Outbreak Genomics Task Force (2011) Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. Emerg Infect Dis 17:2113–2121
102. Díaz-Quiñonez JA (2017) Genome sequence analysis of *Vibrio cholerae* clinical isolates from 2013 in Mexico reveals the presence of the strain responsible for the 2010 Haiti outbreak. Gac Med Mex 153(Supl. 2):S91–S101. [in Spanish]
103. Alam M, Islam MT, Rashed SM, Johura FT, Bhuiyan NA, Delgado G, Morales R, Mendez JL, Navarro A, Watanabe H, Hasan NA, Colwell RR, Cravioto A (2012) *Vibrio cholerae* classical biotype strains reveal distinct signatures in Mexico. J Clin Microbiol 50:2212–2216
104. Kotetishvili M, Stine OC, Chen Y, Kreger A, Sulakvelidze A, Sozhamannan S, Morris JG Jr (2003) Multilocus sequence typing has better discriminatory ability for typing *Vibrio cholerae* than does pulsed-field gel electrophoresis and provides a measure of phylogenetic relatedness. J Clin Microbiol 41:2191–2196
105. Farfán M, Miñana-Galbis D, Fusté MC, Lorén JG (2002) Allelic diversity and population structure in *Vibrio cholerae* O139 Bengal based on nucleotide sequence analysis. J Bacteriol 184:1304–1313
106. Salim A, Lan R, Reeeves PR (2005) *Vibrio cholerae* pathogenic clones. Emerg Infect Dis 11:1758–1760
107. Greig DR, Schaefer U, Octavia S, Hunter E, Chattaway MA, Dallman TJ, Jenkins C (2018) Evaluation of whole-genome sequencing for identification and typing of *Vibrio cholerae*. J Clin Microbiol 56(11):pii: e00831-18
108. Tu YH, Chen BH, Hong YP, Liao YS, Chen YS, Liu YY, Teng RH, Wang YW, Chiou CS (2020) Emergence of *Vibrio cholerae* O1 sequence type 75 in Taiwan. Emerg Infect Dis 26:164–166
109. Garrine M, Mandomando I, Vubil D, Nhampossa T, Acacio S, Li S, Paulson JN, Almeida M, Domman D, Thomson NR, Alonso P, Stine OC (2017) Minimal genetic change in *Vibrio cholerae* in Mozambique over time: multilocus variable number tandem repeat analysis and whole genome sequencing. PLoS Negl Trop Dis 11(6):e0005671

110. Garg P, Aydanian A, Smith DJ, Glenn M Jr, Nair GB, Stine OC (2003) Molecular epidemiology of O139 *Vibrio cholerae:* mutation, lateral gene transfer, and founder flush. Emerg Infect Dis 9:810–814
111. Ghosh R, Nair GB, Tang L, Morris JG, Sharma NC, Ballal M, Garg P, Ramamurthy T, Stine OC (2008) Epidemiological study of *Vibrio cholerae* using variable number of tandem repeats. FEMS Microbiol Lett 288:196–201
112. Stine OC, Alam M, Tang L, Nair GB, Siddique AK, Faruque SM, Huq A, Colwell R, Sack RB, Morris JG Jr (2008) Seasonal cholera from multiple small outbreaks, rural Bangladesh. Emerg Infect Dis 14:831–833
113. Ambroise J, Irenge LM, Durant JF, Bearzatto B, Bwire G, Stine OC, Gala JL (2019) Backward compatibility of whole genome sequencing data with MLVA typing using a new MLVAtype shiny application for *Vibrio cholerae*. PLoS One 14(12):e0225848
114. Sharma C, Ghosh A, Dalsgaard A, Forslund A, Ghosh RK, Bhattacharya SK, Nair GB (1998) Molecular evidence that a distinct *Vibrio cholerae* O1 biotype El Tor strain in Calcutta may have spread to the African continent. J Clin Microbiol 36:843–844
115. Clark CG, Kravetz AN, Dendy C, Wang G, Tyler KD, Johnson WM (1998) Investigation of the 1994–5 Ukrainian *Vibrio cholerae* epidemic using molecular methods. Epidemiol Infect 12:15–29
116. Pazzani C, Scrascia M, Dionisi AM, Maimone F, Luzzi I (2006) Molecular epidemiology and origin of cholera reemergence in Italy and Albania in the 1990s. Res Microbiol 157:508–512
117. Chatterjee S, Patra T, Ghosh K, Raychoudhuri A, Pazhani GP, Das M, Sarkar B, Bhadra RK, Mukhopadhyay AK, Takeda Y, Nair GB, Ramamurthy T, Nandy RK (2009) *Vibrio cholerae* O1 clinical strains isolated in 1992 in Kolkata with progenitor traits of the 2004 Mozambique variant. J Med Microbiol 58:239–247
118. Joelsson A, Liu Z, Zhu J (2006) Genetic and phenotypic diversity of quorum-sensing systems in clinical and environmental isolates of *Vibrio cholerae*. Infect Immun 74:1141–1147
119. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ (2002) Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci U S A 99:1556–1561
120. Pang B, Yan M, Cui Z, Ye X, Diao B, Ren Y, Gao S, Zhang L, Kan B (2007) Genetic diversity of toxigenic and nontoxigenic *Vibrio cholerae* serogroups O1 and O139 revealed by array-based comparative genomic hybridization. J Bacteriol 189:4837–4849
121. Dziejman M, Serruto D, Tam VC, Sturtevant D, Diraphat P, Faruque SM, Rahman MH, Heidelberg JF, Decker J, Li L, Montgomery KT, Grills G, Kucherlapati R, Mekalanos JJ (2005) Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. Proc Natl Acad Sci U S A 102:3465–3470
122. Feng L, Reeves PR, Lan R, Ren Y, Gao C, Zhou Z, Ren Y, Cheng J, Wang W, Wang J, Qian W, Li D, Wang L (2008) A recalibrated molecular clock and independent origins for the cholera pandemic clones. PLoS One 3:e4053
123. Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Lee JH, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci U S A 106:15442–15447
124. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, Croucher NJ, Choi SY, Harris SR, Lebens M, Niyogi SK, Kim EJ, Ramamurthy T, Chun J, Wood JL, Clemens JD, Czerkinsky C, Nair GB, Holmgren J, Parkhill J, Dougan G (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. Nature 477(7365):462–465
125. Weill FX, Domman D, Njamkepo E, Tarr C, Rauzier J, Fawal N, Keddy KH, Salje H, Moore S, Mukhopadhyay AK, Bercion R, Luquero FJ, Ngandjio A, Dosso M, Monakhova E, Garin B, Bouchier C, Pazzani C, Mutreja A, Grunow R, Sidikou F, Bonte L, Breurec S, Damian M, Njanpop-Lafourcade BM, Sapriel G, Page AL, Hamze M, Henkens M, Chowdhury G, Mengel M, Koeck JL, Fournier JM, Dougan G, Grimont PAD, Parkhill J, Holt KE, Piarroux

R, Ramamurthy T, Quilici ML, Thomson NR (2017) Genomic history of the seventh pandemic of cholera in Africa. Science 358(6364):785–789
126. Bwire G, Sack DA, Almeida M, Li S, Voeglein JB, Debes AK, Kagirita A, Buyinza AW, Orach CG, Stine OC (2018) Molecular characterization of *Vibrio cholerae* responsible for cholera epidemics in Uganda by PCR, MLVA and WGS. PLoS Negl Trop Dis 12(6):e0006492
127. Domman D, Quilici ML, Dorman MJ, Njamkepo E, Mutreja A, Mather AE, Delgado G, Morales-Espinosa R, Grimont PAD, Lizárraga-Partida ML, Bouchier C, Aanensen DM, Kuri-Morales P, Tarr CL, Dougan G, Parkhill J, Campos J, Cravioto A, Weill FX, Thomson NR (2017) Integrated view of *Vibrio cholerae* in the Americas. Science 358(6364):789–793
128. Mironova LV, Gladkikh AS, Ponomareva AS, Feranchuk SI, Bochalgin NO, Basov EA, Yu Khunkheeva Z, Balakhonov SV (2018) Comparative genomics of *Vibrio cholerae* El Tor strains isolated at epidemic complications in Siberia and at the Far East. Infect Genet Evol 60:80–88
129. Weill FX, Domman D, Njamkepo E, Almesbahi AA, Naji M, Nasher SS, Rakesh A, Assiri AM, Sharma NC, Kariuki S, Pourshafie MR, Rauzier J, Abubakar A, Carter JY, Wamala JF, Seguin C, Bouchier C, Malliavin T, Bakhshi B, Abulmaali HHN, Kumar D, Njoroge SM, Malik MR, Kiiru J, Luquero FJ, Azman AS, Ramamurthy T, Thomson NR, Quilici ML (2019) Genomic insights into the 2016–2017 cholera epidemic in Yemen. Nature 565(7738):230–233
130. Hounmanou YMG, Leekitcharoenphon P, Kudirkiene E, Mdegela RH, Hendriksen RS, Olsen JE, Dalsgaard A (2019) Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania between 1993 and 2017. PLoS Negl Trop Dis 13(12):e0007934
131. Hu D, Yin Z, Yuan C, Yang P, Qian C, Wei Y, Zhang S, Wang Y, Yuan J, Wang M, Reeves PR, Tu L, Chen M, Huang D, Liu B (2019) Changing molecular epidemiology of *Vibrio cholerae* outbreaks in Shanghai, China. mSystems 4(6):pii: e00561-19
132. Greenhill AR, Mutreja A, Bulach D, Belousoff MJ, Jonduo MH, Collins DA, Kas MP, Wapling J, Seemann T, Lafana A, Dougan G, Brown MV, Horwood PF (2019) Wave 2 strains of atypical *Vibrio cholerae* El Tor caused the 2009–2011 cholera outbreak in Papua New Guinea. Microb Genome 5(3)
133. Li Z, Pang B, Wang D, Li J, Xu J, Fang Y, Lu X, Kan B (2019) Expanding dynamics of the virulence-related gene variations in the toxigenic *Vibrio cholerae* serogroup O1. BMC Genomics 20(1):360
134. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, Gill SR, Nelson KE, Read TD, Tettelin H, Richardson D, Ermolaeva MD, Vamathevan J, Bass S, Qin H, Dragoi I, Sellers P, McDonald L, Utterback T, Fleishmann RD, Nierman WC, White O, Salzberg SL, Smith HO, Colwell RR, Mekalanos JJ, Venter JC, Fraser CM (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature 406(6795):477–483
135. Hur Y, Chalita M, Ha SM, Baek I, Chun J (2019) VCGIDB: a database and web resource for the genomic islands from *Vibrio Cholerae*. Pathogens. 8(4):pii: E261
136. Blokesch M, Schoolnik GK (2008) The extracellular nuclease Dns and its role in natural transformation of *Vibrio cholerae*. J Bacteriol 190:7232–7240
137. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK (2005) Chitin induces natural competence in *Vibrio cholerae*. Science 310(5755):1824–1827
138. Dalia AB, Dalia TN (2019) Spatiotemporal analysis of DNA integration during natural transformation reveals a mode of nongenetic inheritance in bacteria. Cell 179(7):1499–1511.e10
139. Das B, Pazhani GP, Sarkar A, Mukhopadhyay AK, Nair GB, Ramamurthy T (2016) Molecular evolution and functional divergence of *Vibrio cholerae*. Curr Opin Infect Dis 29:520–527
140. Das B (2019) Insights into TLCPhi lysogeny: a twist in the mechanism of IMEX integration. Proc Natl Acad Sci U S A 116:18159–18161
141. Safa A, Sultana J, Dac Cam P, Mwansa JC, Kong RY (2008) *Vibrio cholerae* O1 hybrid El Tor strains, Asia and Africa. Emerg Infect Dis 14:987–988
142. Siddique AK, Nair GB, Alam M, Sack DA, Huq A, Nizam A, Longini IM, Qadri F, Faruque SM, Colwell RR, Ahmed S, Iqbal A, Bhuiyan NA, Sack RB (2009) El Tor cholera with severe disease: a new threat to Asia and beyond. Epidemiol Infect 14:1–6

Part II

Pathogens Causing Healthcare-Associated Infections

Acinetobacter baumannii

6

Dennis Nurjadi and Sébastien Boutin

6.1 Introduction

Bacteria of the genus *Acinetobacter* is abundant and widespread in nature. They can be found in the environment (soil and water samples) and samples from human and animals. Among this genus, *Acinetobacter baumannii* complex, comprising of A. *baumannii, A. nosocomialis, A. pitii* (and *A. calcoaceticus*), are the most clinically relevant. Although most of the *Acinetobacter* spp. are ubiquitous, the emergence of multidrug-resistant (MDR) strains in clinical samples and causing human infections are a major concern. This chapter deals with an overview of the Acinetobacter genus, focusing on the clinically relevant *A. baumannii* complex and various typing methods with particular emphasis on next-generation sequencing (NGS) for molecular typing.

6.2 Microbiology of *Acinetobacter* spp.

The morphology of bacteria of the genus *Acinetobacter* is coccoid, or rod-shaped. In the Gram stain, Acinetobacter can appear quite variable. They are Gram-negative but may appear slightly Gram-positive due to destaining difficulties [1]. They are non-motile, oxidase negative, catalase-positive, and non-fermenting with the optimal growth temperatures between 20 and 37 °C. Some species such as *A. baumannii* and *A. nosocomialis* may tolerate higher temperatures (>40 °C) [2]. Acinetobacter isolated from human sample grow well on solid media, such as Columbia sheep blood agar, tryptic soy agar, and MacConkey agar at an incubation temperature of

D. Nurjadi (✉) · S. Boutin
Department of Infectious Diseases, Medical Microbiology and Hygiene, Heidelberg University Hospital, Heidelberg, Germany
e-mail: dennis.nurjadi@uni-heidelberg.de

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_6

37 °C. *Acinetobacter* spp. can be distinguished from other non-fermenting Gram-negatives by a transformation assay of Juni. This differentiation method is based on the unique property of a naturally transformable tryptophan auxotroph mutant *Acinetobacter* strain *Acinetobacter baylyi* (BD413 trpE27) to be transformed by crude DNA of any *Acinetobacter* spp. and revert to a wild-type phenotype [3, 4].

6.3 Taxonomy of *Acinetobacter* spp.

Acinetobacter belong to the *Gammaproteobacteria*. The first description of *Acinetobacter* dates back to the early twentieth century, when an organism, then named *Micrococcus calcoaceticus*, was isolated from soil by enrichment culture with calcium acetate as a carbon source [5]. Later, similar organisms were described independently and thus assigned as different genera and species. A comprehensive history of this genus is reviewed by Henriksen [6]. The current genus, *Acinetobacter*, was initially proposed by Brisou and Prévot in 1954 [7], but it was not until 1968 that his designation became widely accepted [8]. Due to the lack of phenotypic criteria to distinguish the species within this genus, the different species were categorized together into one genus, Acinetobacter, consisting of a single species, *Acinetobacter calcoaceticus* [9].

It was not until 1986 when DNA-DNA hybridization distinguished 12 groups or genospecies. Some of which were given formal species names known today, including *A. baumannii*, *A. calcoaceticus*. Since then, additional DNA-DNA hybridization groups have been described, and to date, well over 50 species have been validated and given formal species names (https://lpsn.dsmz.de/genus/Acinetobacter). *A. baumannii*, *A. pittii*, *A. nosocomialis* (all three species belonging to *Acinetobacter baumannii* complex), and the environmental species *A. calcoaceticus* are genetically closely related and phenotypically difficult to distinguish so that they are often referred to as the *A. calcoaceticus-A. baumannii* complex (ACB-Complex) (see Fig. 6.1) [10].

In the genomic era, the advancement and access to molecular methods with high discriminatory power, such as 16S rRNA sequencing and NGS, cast doubts on the conservative microbial taxonomy, which relies heavily on phenotypic information and growth properties. However, for the genus *Acinetobacter*, phylogenetic classification was shown to be consistent with classical taxonomic classification [11]. Nevertheless, the taxonomy of *Acinetobacter* remains complex. In particular, the designation of the DNA-DNA hybridization groups is somewhat confusing, since identical hybridization group designation (genomospecies) has been given independently, or several names have been used for the same species. Furthermore, in older literature, several synonyms may have been used interchangeably due to reclassifications, when referring to the same species. Figure 6.1 display the genetic relationship of publicly available sequenced *Acinetobacter sp.* genome.

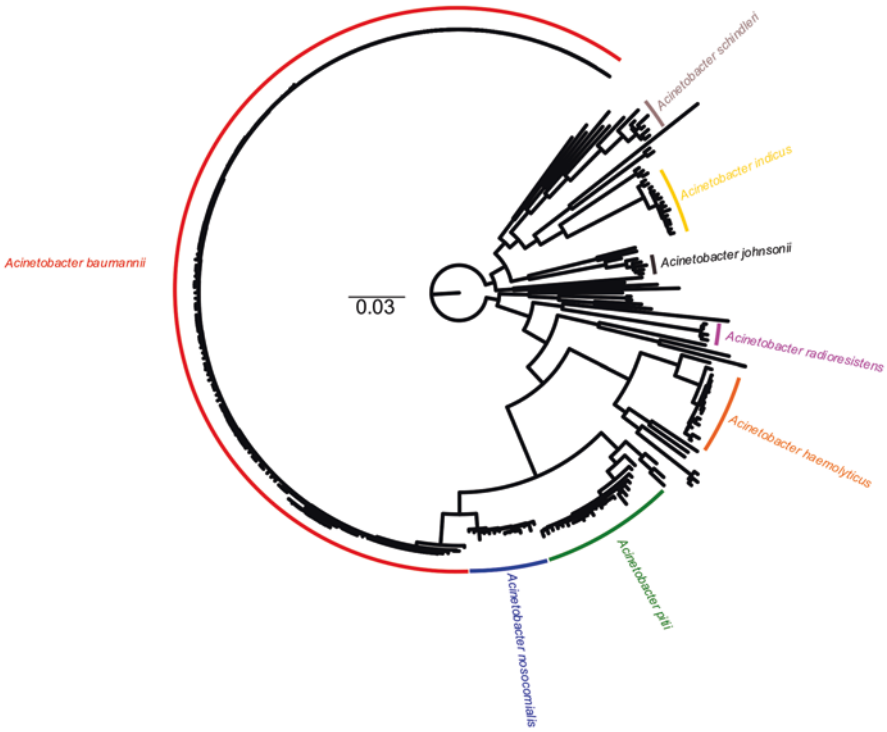

Fig. 6.1 Phylogenetic tree of sequenced and publicly (NCBI Genbank as of June 2020) available *Acinetobacter* spp. 235 complete genomes were downloaded from NCBI Genbank and a core-genome was calculated using Roary. The core-genome was relatively small (36 genes, 4734 polymorphic sites). In total 28 species were represented. 19 genomes were not assigned to a species level. Only eight species were represented by four or more isolates (displayed in the tree). Most of the publicly available sequences are those of *A. baumannii* complex (*A. baumannii* (n = 184), *A. nosocomialis* (n = 14), *A. pitii* (n = 22)). The overrepresentation of *A. baumannii* complex emphasize the clinical importance of this complex

6.4 Clinical Significance of A. baumannii

A. baumannii is associated with healthcare-associated infections with diverse manifestations ranging from wound infections and urinary tract infections to more severe invasive infections such as ventilator-associated pneumonia, bloodstream infections, and meningitis. However, more common than infections is colonization. The distinction of colonization and infection can be challenging due to sensitivity issues and the fact that the natural habitat of *A. baumannii* is not the human body [12, 13]. In fact, it is unclear which body sites should be sampled to obtain reliable colonization status of the patients [12].

A. baumannii infections, especially infections with MDR strains, have been attributed to higher morbidity and mortality rates [14]. However, these rates vary greatly in the available literature and there is an ongoing debate on data reliability

and to what extent the infecting organism contributes to mortality or if the underlying diseases is an independent and decisive factor for the outcome. An increased length of stay associated with MDR *A. baumannii* infection highlights the need for infection control and surveillance strategies to prevent nosocomial transmissions [15].

Some of the known risk factors for *A. baumannii* infections are serious comorbidities (critically ill patients), immunosuppression, ventilation, and previous exposure to antibiotics [16–20]. Although not considered as highly virulent, *A. baumannii* is a difficult pathogen to control as demonstrated by occasional outbreaks and in-hospital cross transmission events [21, 22]. In addition, they are very tenacious and can persist over longer periods on inanimate objects in the hospital, even under sub-optimal conditions, and even maintain their virulence [23, 24].

6.5 Antibiotic Resistance

A. baumannii is included as one of highest priority of clinically relevant antibiotic-resistant bacteria by the WHO Pathogens Priority List Working Group in 2018 [25]. *A. baumannii* is well known for the high burden of disease in healthcare-associated infections and their resistances to numerous antibiotics. Due to multiple intrinsic resistances, the therapy of option for *A. baumannii* is significantly narrowed down to several broad-spectrum, such as carbapenems, and second-line antibiotics. The increasing prevalence of carbapenem resistance in *A. baumannii* is alarming, as carbapenems have been regarded as one of the last resort options to treat *A. baumannii* infections. Thus, infectious disease specialists resort to older, more toxic, or sub-optimal substances to treat carbapenem-resistant *A. baumannii* (CRAB), such as colistin or tigecycline. Even so, resistance to these substances is emerging and pan-drug-resistant *A. baumannii* have been frequently described.

6.6 Epidemiology of Carbapenem-Resistant *A. baumannii*

Most epidemiological studies focus on the aspect of global spread and dissemination of carbapenem-resistant *A. baumannii* (CRAB). The epidemiology of multi-resistant bacteria is highly dynamic, and the increased access to and affordability of high-resolution molecular typing methods constantly contribute new data and insight into the global epidemiological data. This section will focus mainly on the current epidemiology of CRAB based on data published up to 2020.

Over the years, the number of reports reporting CRAB has increased significantly, which is a strong indication of its global dissemination [26, 27]. Most of these reports have been reports on outbreaks due to inter- and intrahospital transfers from various countries across the globe. These outbreaks have been attributed to clones, commonly referred to as global clone 1 or international clone 1 (GC1/IC1) and global clone 2 or international clone 2 (GC2/IC2). In a recent review, Hamidian and Nigro performed an analysis of available *A. baumannii* genomes available in the

GenBank as of April 2019 and demonstrated that around two-thirds of sequenced *A. baumannii* belong to GC1 and GC2, with ST2 (Pasteur MLST) as the predominant type, followed by ST1, ST79, and ST25 as other dominant clones [27].

A. baumannii is considered as naturally resistant to various antibiotics. Partly due to the presence of intrinsic $bla_{\text{OXA-51-like}}$ genes, with $bla_{\text{OXA-69}}$ and $bla_{\text{OXA-66}}$ being the most common. Carbapenem resistance in *A. baumannii* is largely due to acquisition of additional genes that encode carbapenem-hydrolyzing enzymes. To date, variants of $bla_{\text{OXA-23}}$ is the predominant mechanism of carbapenem resistance in *A. baumannii*, followed by variants of $bla_{\text{OXA-24}}$, $bla_{\text{OXA-58}}$ families, which all have been found in isolates involved in nosocomial outbreaks worldwide. Apart from Ambler class D carbapenemase (oxacillinase). Despite quite rare, Ambler class B carbapenemase (metallo-betalactamase), such as bla_{NDM}, may also be found in *A. baumannii*. In addition, overexpression of the intrinsic $bla_{\text{OXA-51}}$ due to the presence of an insertion sequence IS*AbaI* upstream of the OXA-gene has been described as a potential mechanism for carbapenem resistance [28].

Nevertheless, genome-based epidemiological data should be assessed and interpreted carefully, since the publicly available sequences have a skewed geographical distribution, with only five countries (USA, China, Thailand, Australia and Pakistan) contributing to over 60% of uploaded genome data. Hence some regions and continents may be under-represented, and caution is needed before drawing conclusions.

6.7 Typing Methods for *Acinetobacter* spp.

6.7.1 Phenotypic Typing

The classical phenotyping of Acinetobacter spp. based on physiological and nutritional needs as well as enzymatic tests was useful initially to delimitate the species boundaries of the genus [29, 30]. However, due to the advances in automatization, this technique has been replaced mostly by spectroscopic approaches. One of the most used technique is the matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) technology which allow a rapid and accurate identification of the strain [31]. The mass spectra are compared to a reference database of profiles corresponding to known microbes. This technic also offers a scoring system validating the accuracy of the typing. Each strain with a score lower than 2 should be typed with a second method (phenotypic or genotypic). Based on spectroscopy, other methodologies have been developed such as Fourier Transform Infrared Spectroscopy (FTIR) and Raman spectroscopy which can be used to analyze in the same way a spectrum facilitating the identification of the microorganisms to a taxonomic level [32]. The difference relies on the absorbance of infrared or near-infrared radiation by the molecules from whole cells to generate characteristic spectra. This spectrum can be interpreted as a fingerprint to identify microbes with similar resolution to the MALDI-TOF [33]. However, the resolution to the strain and clonal complex level is not achievable for all species with those techniques.

The MALDI-TOF for example has been shown some limitations to give a strain resolution [34]. In the case of *A. baumannii*, strain typing using MALDI has shown a good concordance to repetitive sequence-based PCR [35] but failed to differentiate strains from outbreak defined by pulsed-field gel electrophoresis (PFGE) or multilocus sequence typing (MLST) [36, 37]. On the other hand, Raman and FTIR methods showed promising results regarding the resolution for strain-level identification of *A. baumannii* which indicate that a wider spectrum of analysis is needed to get this level of discrimination [38, 39]. Based on that observation, a proteomic approach based on LC-MS/MS profiling of peptides from the whole bacterial isolates digested with a rapid tryptic digestion showed the ability to classify isolates into groups consistent with their MLST types. This technique is of high interest because it relies on the in silico construction of a species "panpeptidome" to design a set of peptide-markers and can therefore be adapted to any species where the genomic information is available and relatively abundant in the database to design accurate markers [40].

6.7.2 Genotypic Typing

Molecular typing encompasses several methods for typing microorganism to the species or strain level with different resolution based on their genomic content. Genotyping also allows investigating clonal relationship and identifying the source of the original infection.

6.7.2.1 Band Pattern

Ribotyping [41] and amplified fragment length polymorphism (AFLP) [42] were among the first typing techniques used to demonstrate the clonality of *A. baumannii*. These methods however are obsolete and their usage for the typing of *Acinetobacter* spp. has declined over the past years [43] and therefore they are not discussed in detail in this chapter.

In the typing of Acinetobacter, restriction fragment length polymorphism (RFLP) is a standard method used for decades. Shortly, DNA is fragmented with restriction enzymes and the digested DNA is separated according to length by agarose gel electrophoresis, obtaining a restriction pattern allowing the identification of the bacteria. This method was improved by the use of pulsed-field gel electrophoresis (PFGE) where the electric field applied to the gel matrix is periodically changing directions [44]. This method increases the resolution of a large-size DNA fragment (>50 kb) which would migrate together in a normal electrophoresis. Currently, PFGE is still considered as the "gold standard" typing methodology in many microbiology diagnostic departments [45].

Multiple-locus VNTR (variable number of tandem repeats)" analysis (MLVA) is another method which has been used for molecular typing of *Acinetobacter* spp. Briefly, VNTRs in genome sequences are identified and flanking primers were used to amplify these regions by PCR. The fragment sizes, corresponding to the number of repeats, are used for differentiation of organisms [46]. There are multiple schemes

available for *A. baumannii* but MLVA-8 was the most used, which assessed differences of repeat numbers in eight VNTR loci [47], four of which are large repeat with slow evolution, useful for phylogeny, and four small repeats with rapid evolution, useful for outbreak investigations [48]. For each locus, amplicons were separated by agarose electrophoresis, and each strain was assigned a code representing the number of repeats per locus [49]. MLVA data can be stored in a database and exchanged between institutes (http://mlva.u-psud.fr).

Random amplification of polymorphic DNA (RAPD) is another PCR-based typing method using short random primers and low annealing temperature to obtain fingerprint-like band patterns visualized on agarose gel, which has gain popularity due to its simplicity [50, 51]. This method is still in use in laboratory practice because of its reasonable discriminatory power for a low cost and fast time to result. The downside of most of these band-pattern-based typing methods, with the exception of PFGE, is the interlaboratory comparability [43].

6.7.2.2 Partial Sequencing

Sanger sequencing of a particular variable gene or regions, such as the protein A (*spa*) variable repeat region for Staphylococcus aureus (*spa*-typing) or a defined set of housekeeping genes (multi-locus sequence type; MLST), can be used for strain characterization [52]. A number is assigned to each variant of a housekeeping gene and a certain combination of variants in seven housekeeping genes are assigned to a sequence type (ST). Using MLST the population can be stratified into clones and MLST is a useful tool to study and describe the population structure of bacteria (epidemiology) [53]. For *A. baumannii*, two different MLST schemes (Oxford [53] and Pasteur [54]) are available. Both schemes target seven housekeeping genes, of which three are in common. In the pre-WGS era, due to a consensus nomenclature, MLST is considered one of the gold standards to study the population structure as interlaboratory variability is no longer an issue. MLST is still widely used today and is still very useful for typing purposes, although the lack of its discriminatory power may be a major limitation. This can be circumvented by using other methods with high discriminatory power such as WGS [55].

6.7.2.3 Whole Genome Sequencing

Due to the development of high-throughput sequencer and the decreasing price of next-generation sequencing, WGS becomes an affordable all-in-one tool for routine microbiological laboratories. The sequencing of the full genome with a coverage of 35-fold to 50-fold allows us to predict accurately the antimicrobial resistance profile and potential virulence as well as offering a high-resolution typing of the pathogenic strain [56, 57]. The typing with NGS can be considered as the highest resolution to evaluate global and local epidemiology in contrast to other conventional typing methods such as PFGE [58, 59]. While successful to study the relationship between isolates and the population structure of the *Acinetobacter* genus, it has been shown that the conventional methods do not reflect the accurate phylogenetic relationships between isolates as they analyzed only a small part of genomes. In most of the cases, the conventional methods overestimate the relatedness of isolates

due to similar band pattern or high similarity in the defined genetic region of interest. This overestimation of clonal relatedness can affect more than 70% of the isolates in certain cases [60]. However, more problematic issues occur when conventional methods do not indicate relatedness while WGS showed no single nucleotide polymorphisms (SNP) between isolates [61]. This phenomenon can be explained by insertion sequences of small deletion in the region where restriction enzymes or primers are binding creating a different band patterns but have moderate impact on WGS-based typing approaches. Furthermore, WGS is analyzing the full content of the genomes and core genomes often encompass more than 2000 genes, offering a better definition of the clonality than multi-locus sequencing typing (MLST) [62].

6.7.2.4 WGS Application

One of the first objectives of microbiology is to test the microbes for antimicrobial resistance. WGS can be in this field an invaluable tool and is quickly progressing using machine learning algorithm [63–65]. For *Acinetobacter* spp., the prediction focused mostly on beta-lactams and carbapenems resistance (e.g., meropenem and imipenem) because it is used to treat nosocomial infections [66]. The most prevalent resistance genes are mobile, or plasmid carried $bla_{\text{OXA-23}}$, $bla_{\text{OXA-24}}$, and $bla_{\text{OXA-58}}$ and chromosomally encoded *ampC* which confers resistance to narrow-spectrum cephalosporins [67, 68]. Another highly prevalent carbapenemase gene in *Acinetobacter* spp. is the $bla_{\text{OXA-51-like}}$ which can confer resistance to carbapenems when the insertion element IS*Aba1* lies upstream of the gene [69]. Many genes are involved in the resistance to aminoglycoside such as the enzymes encoding *aacC1*, *aphA6*, *aadA1*, and *aadB* or the 16S rRNA methyltransferase *armA* and to a lower extent efflux pump such as AdeABC and AbeM. Macrolides' resistances are associated with the mutation in the *folA* gene or the presence of the *tetM* gene in the genome. Finally, quinolone resistance can occur due to mutations in the genes *parC*, *gyrA*, and *gyrB*. The knowledge in the genomic regions associated with antimicrobial resistance allowed the creation and maintenance of multiple databases such as CARD, ResFinder, ARG-ANNOT used as screening references to infer the resistance profile (resistome) of the isolate. However, this method is still too simplistic to explain the phenotypic resistance and a lack of knowledge in the causing mutations; the influence of the genomic islands needs to be acknowledged. Furthermore, the presence of a gene or a mutation does not necessarily reflect the phenotypic resistance due to potential influences on the expression of the gene [69–71].

The same caveat applies to the virulence gene repertoire, while the advantage of exploring the full genome of the isolates is undeniable and understanding of the virulome of a pathogen is a crucial step to predict the severity of the disease; the gene presence cannot be the sole indicator of virulence. To overcome this issue, RNA sequencing can be performed from the colonies but the differences between the ecological niches (plate versus host) might falsify the results or we can couple the WGS exploration with transposon mutagenesis techniques. This technique allowed identifying essential virulence factors for the persistence and survival of *A. baumannii* in lungs and bloodstream [72, 73].

While the prediction of antibiotic resistance and virulence is important, the main use of WGS is the typing and epidemiology. The access to the full genome (core and accessory) allows us to study the population structure of the Acinetobacter spp. with an evolutionary perspective. The global epidemiology based on WGS allows highlighting the successful mutations and evolutionary trajectory of some emerging successful clone and understanding the genetic background explaining their success and emergence. For example, *A. baumannii* population structure is mostly explained by a recombination event which has driven successful evolutionary trajectories to some highly successful clonal complexes such as the global clone 1 (GC1) [74]. The analysis of the genetic content of GC1 highlights a divergence in two lineages due to recombination events affecting surface structures such as the capsule and lipooligosaccharide loci and the acquisition of antimicrobial resistance [74]. The exploration of the geographic diversity and temporal evolution of the difference clonal lineage is of crucial importance to understand the impact of the selective pressure and the adaptation to the current therapies [62]. Understanding the mechanisms between clonal shift/replacement or adaptation to antibiotic regimen will help us to adapt our antibiotic stewardship [75, 76]. Furthermore, understanding the global evolution, geographical distribution of the clonal group, and the mutation rate will help to set up the threshold to estimate potential transmission in the local settings [77]. *A. baumannii* shows a high mutation rates (24 mutations/year within a patient) and the treatment with antibiotic pressures the strains and allows the appearance of hypermutator phenotype [78, 79]. Local epidemiology and outbreak surveillance using WGS can highlight in real time the emergence or the fast transmission of a successful clone. The genetic distance between clones as well as the acquisition of specific plasmid or mobile element allows us to highlight the origin of the outbreak and the potential reservoir in the hospital. In the fine-scale epidemiology, the accessory genome may play an important role as a rapid acquisition of a plasmid or mobile element by a local clone might be the only characteristic of the outbreak-causing strain [80]. Genome-wide analysis is a powerful tool, and therefore, the scale of the epidemiological settings and the aims should drive the technology and analytic methodology to be as accurate as possible.

6.7.2.5 WGS Technology and Bioinformatics

Three next-generation instruments have dominated the field. The first company Illumina offers several short-read instruments for low-, mid-, and high-throughput sequencing (Miniseq, Miseq, Novaseq). The two other companies offer a long-read sequencing system (Pacific Biosciences and Oxford Nanopore). The advantage of short-read technology is the low error rate in the assembly due to a high coverage for a low cost of sequencing. However, the use of short-read technology limits the completion of a "closed" assembly. The raw reads generated by WGS are filtered to remove low-quality data, then assembled into large segments known as contigs, but it is difficult to close the assembly into one complete chromosome using only short-read technology. Therefore, it is usual to use a combination of both short and long reads to obtain an accurate and "closed" assembly [56].

Two options exist to assemble sequencing data: mapping or de novo assembly. The mapping method is omputationally less demanding than de novo assembly as it consist of aligning and mapping the clean raw reads to close reference genomes. SNPs and gene annotation will then depend highly on the reference used for the mapping. The second method consists of assembling the reads in contigs without the use of a reference. A multitude of de novo assemblers exist currently such as IDBA [81], RAY [82], VELVET [83], and Skesa [84]; but the most used one is Spades [85]. However, it is important to keep in mind that various assembly algorithms will give different results based on biological and technical parameters and it can be hard to define the "best" assembler. Therefore, it is important to evaluate your assembly to check the coverage, the size of the draft genome, and several parameters such as the N50 or the NG50. For both approaches, the pipeline and/or the reference used will modify the results. Therefore, it is important to be consistent in the analysis and to build criteria to validate results.

In the case of typing, two approaches can be used to understand the phylogenetic relationship between isolates: single nucleotide polymorphism (SNP)-based or gene-based approach [86]. SNP is a variation in a single nucleotide at a specific position in the genome. SNP typing is widely used for bacteria, and genome-wide screening of SNPs is a powerful discriminatory technique that enables the identification of strain-specific genetic markers. The power of SNP-based approach is that even if bacterial genomes of the same species share a common group of genes (i.e. core genome), they will vary by point mutations at different genome positions in genes or intergenic regions. Because intergenic regions are less pressured by natural selection, the resolution power will increase among closely related strains and allow accurate local and fine-scale typing. In order to identify SNPs, the reads or contigs can be either aligned to a reference genome or by building a core genome alignment between the conserved sequences of the genomes of interest, followed by SNPs calling [86]. SNPs also need also to be curated to reflect accurately the evolutionary change in the genome and the phylogenetic relationship between strains. This step is crucial to not overestimate the SNP distance due to sequencing errors, artifact alignment on reference or assembly and recombinations [87]. Furthermore, epidemiological studies should use harmonized definitions and nomenclature to ensure comparability in defining the threshold for clonality [56]. SNP thresholds will differ between species due to different generation time and the rate of mutations. For example, a recent study on *Acinetobacter* epidemiology suggested two core SNP thresholds to define first the lineages between *A. calcoaceticus* and *A. baumannii* complex and clonality. The study showed that a threshold of 2500 core SNPs allows differentiation of *A. baumannii* isolates from different clonal lineages (as defined by MLST). Most importantly, core SNP threshold (2.5 core SNPs) can discriminate outbreak strains from sporadic isolates [59]. Another study used pre-defined classification using PFGE to define the thresholds and concluded that strain should be considered "genomically indistinguishable" if they are separated by 3 or fewer variants, "closely related" if they are separated by up to 12 variants, and "unrelated" if they are distinguishable by 13 or more variants [60]. However, those thresholds are influenced by the methods used as well as a clinically or molecularly predefined

outbreak. The evolution rate and inter-individual heterogeneity are not taken into account. Finally, the number of polymorphic sites included in the study will necessarily influence those thresholds as they are using a raw number of SNPs. Therefore, a study using 50,000 polymorphic sites could not be compared to another study using 1,000,000 polymorphic sites and the same problem will apply to the sites of the core genomes. Studies using inter-individual heterogeneity or SNPs filtration gave us more tools to set up strain definition at the genetic levels and offer approaches that still need to be applied on Acinetobacter genus for the normalization or harmonization of the divergence thresholds [88, 89].

The second approach in WGS-typing is the gene-by-gene approach which is an extension of the previous MLST scheme. The idea came to standardize the WGS typing to allow comparable studies by using genes present in a core genome designed for the isolates of the same species (core genome MLST; cgMLST) but can also be used with all genes across the genome (Whole genome MLST; wgMLST) [90]. cgMLST has the advantage of offering a comparable genotyping delineation when it follows a publicly available scheme or allelic patterns which are updated by the users. Each allelic profile is annotated with a complex type (CT) number and can be used all over the world for local and global scale outbreak management. *A. baumannii* complex is the only species from the genus to have a CT scheme due to its clinical relevance. The first scheme was developed in 2016 with an ad hoc typing scheme querying the 3319 genes of the references genome. Of those 2592–2876 genes were actually present in their dataset and the cgMLST typing confirmed the clonal spread of *A. baumannii* in an ICU and give a similar clustering of the isolates compared to the SNP-based typing [91]. Later, an official cgMLST was developed and integrated into commercial software (Ridom SeqSphere+) [92]. This scheme is based on 2390 defined core genes derived from 1339 sequenced *A. baumannii* genomes. This scheme was tested with strains which were previously identified as outbreak using PFGE. With a threshold of eight allelic difference, the CT cluster correlates well with the pulsotype. The difference between pulsotypes ranges from 40 to 2166 alleles, and the highest number of allelic differences (>2000) was between the isolates representing the international clones. The usefulness of the scheme was validated in recent studies with slightly more elevated threshold (9–10 allelic differences) [93–95]. A scheme for wgMLST is also now available for *A. baumannii* containing 5619 loci in the commercial BioNumerics software. This offers of course a higher resolution, but the usage of accessory genome with genes which are not represented in all the isolates might inflate the divergence due to low coverage sequencing. Indeed, some genes might be missing from the assemblies just due to low coverage or high GC content in this specific region. wgMLST or extended cgMLST is a good option to increase the resolution, but it should still be used with the knowledge of the coverage and CG content of your assembly. Some pipelines such as chewBBACA allow building your own scheme based on pangenome (wgMLST) or core genome (cgMLST) but with the disadvantage of losing the comparability with other studies offered by the official cgMLST [96].

6.8 Concluding Remarks and Recommendations

WGS has revolutionized molecular typing of *Acinetobacter* spp. due to the unprecedented discriminatory power. This technology however is not always accessible and is still costly for most users. Moreover, knowledge in bioinformatics is advantageous for data analysis. Publicly available WGS data for *Acinetobacter* spp. is largely dominated by *A. baumannii* complex, reflecting the general interest and the clinical relevance of this pathogen. More importantly, there is an information bias for publicly available WGS data, since most uploaded sequences originated from several countries only and that great care is needed for the interpretation of global epidemiology of *A. baumannii*. Species identification for microbiological diagnostics can be determined reliably using MALDI-TOF, which has replaced biochemical and PCR-based methods. The increase of multi-resistance in *A. baumannii* is worrisome and should be monitored closely. For the investigation of outbreaks, WGS has become a new gold standard. However, many of the conventional typing methods, such as MLST and PCR-based typing, are still widely used today and are still useful for typing of *A. baumannii*, especially in resource-restricted settings.

6.9 Summary

Bacteria of the genus *Acinetobacter* is abundant and widespread in nature. They can be found in the environment (soil and water samples) and samples from human and animals. Among this genus, *Acinetobacter baumannii* complex, comprising *A. baumannii, A. nosocomialis, A. pitii* (and *A. calcoaceticus*), are the most clinically relevant. The emergence of multidrug-resistance (MDR) in *Acinetobacter baumannii* complex is a major public health concern and should be monitored closely. Whole genome sequecing (WGS) has become the new gold standard for molecular typing of *Acinetobacter* spp. allowing to discriminate between outbreak and nonoutbreak strains in clonal populations. WGS data should be interpreted carefully, according to the bioinformatics analysis method and applied parameters.

References

1. Peleg AY, Seifert H, Paterson DL (2008) Acinetobacter baumannii: emergence of a successful pathogen. Clin Microbiol Rev 21:538–582
2. Nemec A, Krizova L, Maixnerova M, van der Reijden TJ, Deschaght P, Passet V et al (2011) Genotypic and phenotypic characterization of the Acinetobacter calcoaceticus-Acinetobacter baumannii complex with the proposal of Acinetobacter pittii sp. nov. (formerly Acinetobacter genomic species 3) and Acinetobacter nosocomialis sp. nov. (formerly Acinetobacter genomic species 13TU). Res Microbiol 162:393–404
3. Juni E (1972) Interspecies transformation of Acinetobacter: genetic evidence for a ubiquitous genus. J Bacteriol 112:917–931
4. Vaneechoutte M, Young DM, Ornston LN, De Baere T, Nemec A, Van Der Reijden T et al (2006) Naturally transformable Acinetobacter sp. strain ADP1 belongs to the newly described species Acinetobacter baylyi. Appl Environ Microbiol 72:932–936

5. Dijkshoorn L, Van Ooyen A, Hop WC, Theuns M, Michel MF (1990) Comparison of clinical Acinetobacter strains using a carbon source growth assay. Epidemiol Infect 104:443–453
6. Henriksen SD (1973) Moraxella, Acinetobacter, and the Mimeae. Bacteriol Rev 37:522–561
7. Brisou J, Prevot AR (1954) Studies on bacterial taxonomy. X. the revision of species under Acromobacter group. Ann Inst Pasteur (Paris) 86:722–728
8. Baumann P, Doudoroff M, Stanier RY (1968) A study of the Moraxella group. II. Oxidative-negative species (genus Acinetobacter). J Bacteriol 95:1520–1541
9. Bergey DH, Buchanan RE, Gibbons NE, American Society for Microbiology (1975) Bergey's manual of determinative bacteriology, 8th edn. Williams & Wilkins Co., Baltimore
10. Gerner-Smidt P, Tjernberg I, Ursing J (1991) Reliability of phenotypic tests for identification of Acinetobacter species. J Clin Microbiol 29:277–282
11. Chan JZ, Halachev MR, Loman NJ, Constantinidou C, Pallen MJ (2012) Defining bacterial species in the genomic era: insights from the genus Acinetobacter. BMC Microbiol 12:302
12. Apisarnthanarak A, Warren DK (2013) Screening for carbapenem-resistant Acinetobacter baumannii colonization sites: an implication for combination of horizontal and vertical approaches. Clin Infect Dis 56:1057–1059
13. Marchaim D, Navon-Venezia S, Schwartz D, Tarabeia J, Fefer I, Schwaber MJ et al (2007) Surveillance cultures and duration of carriage of multidrug-resistant Acinetobacter baumannii. J Clin Microbiol 45:1551–1555
14. Trottier V, Namias N, Pust DG, Nuwayhid Z, Manning R, Marttos AC Jr et al (2007) Outcomes of Acinetobacter baumannii infection in critically ill surgical patients. Surg Infect 8:437–443
15. Sunenshine RH, Wright MO, Maragakis LL, Harris AD, Song X, Hebden J et al (2007) Multidrug-resistant Acinetobacter infection mortality rate and length of hospitalization. Emerg Infect Dis 13:97–103
16. Garcia-Garmendia JL, Ortiz-Leyba C, Garnacho-Montero J, Jimenez-Jimenez FJ, Perez-Paredes C, Barrero-Almodovar AE et al (2001) Risk factors for Acinetobacter baumannii nosocomial bacteremia in critically ill patients: a cohort study. Clin Infect Dis 33:939–946
17. Lortholary O, Fagon JY, Hoi AB, Slama MA, Pierre J, Giral P et al (1995) Nosocomial acquisition of multiresistant Acinetobacter baumannii: risk factors and prognosis. Clin Infect Dis 20:790–796
18. Forster DH, Daschner FD (1998) Acinetobacter species as nosocomial pathogens. Eur J Clin Microbiol Infect Dis 17:73–77
19. Levin AS, Mendes CM, Sinto SI, Sader HS, Scarpitta CR, Rodrigues E et al (1996) An outbreak of multiresistant Acinetobacter baumanii in a university hospital in Sao Paulo, Brazil. Infect Control Hosp Epidemiol 17:366–368
20. Kaul R, Burt JA, Cork L, Dedier H, Garcia M, Kennedy C et al (1996) Investigation of a multiyear multiple critical care unit outbreak due to relatively drug-sensitive Acinetobacter baumannii: risk factors and attributable mortality. J Infect Dis 174:1279–1287
21. D'Agata EM, Thayer V, Schaffner W (2000) An outbreak of Acinetobacter baumannii: the importance of cross-transmission. Infect Control Hosp Epidemiol 21:588–591
22. Eigenbrod T, Reuter S, Gross A, Kocer K, Gunther F, Zimmermann S et al (2019) Molecular characterization of carbapenem-resistant Acinetobacter baumannii using WGS revealed missed transmission events in Germany from 2012–15. J Antimicrob Chemother 74:3473–3480
23. Chapartegui-Gonzalez I, Lazaro-Diez M, Bravo Z, Navas J, Icardo JM, Ramos-Vivas J (2018) Acinetobacter baumannii maintains its virulence after long-time starvation. PLoS One 13:e0201961
24. Wendt C, Dietze B, Dietz E, Ruden H (1997) Survival of Acinetobacter baumannii on dry surfaces. J Clin Microbiol 35:1394–1397
25. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL et al (2018) Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18:318–327
26. Peleg AY, Paterson DL (2006) Multidrug-resistant Acinetobacter: a threat to the antibiotic era. Intern Med J 36:479–482

27. Hamidian M, Nigro SJ (2019) Emergence, molecular mechanisms and global spread of carbapenem-resistant Acinetobacter baumannii. Microb Genom 5:e000306
28. Turton JF, Woodford N, Glover J, Yarde S, Kaufmann ME, Pitt TL (2006) Identification of Acinetobacter baumannii by detection of the blaOXA-51-like carbapenemase gene intrinsic to this species. J Clin Microbiol 44:2974–2976
29. Traub WH (1989) Acinetobacter baumannii serotyping for delineation of outbreaks of nosocomial cross-infection. J Clin Microbiol 27:2713–2716
30. Bouvet PJM, Grimont PAD (1986) Taxonomy of the genus Acinetobacter with the recognition of Acinetobacter baumannii sp. nov., Acinetobacter haemolyticus sp. nov., Acinetobacter johnsonii sp. nov., and Acinetobacter junii sp. nov. and emended descriptions of Acinetobacter calcoaceticus and Acinetobacter lwoffii. Int J Syst Evol Microbiol 36:228–240
31. Singhal N, Kumar M, Kanaujia PK, Virdi JS (2015) MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol 6:791
32. Zarnowiec P, Lechowicz L, Czerwonka G, Kaca W (2015) Fourier transform infrared spectroscopy (FTIR) as a tool for the identification and differentiation of pathogenic bacteria. Curr Med Chem 22:1710–1718
33. Dieckmann R, Hammerl JA, Hahmann H, Wicke A, Kleta S, Dabrowski PW et al (2016) Rapid characterisation of Klebsiella oxytoca isolates from contaminated liquid hand soap using mass spectrometry, FTIR and Raman spectroscopy. Faraday Discuss 187:353–375
34. Spinali S, van Belkum A, Goering RV, Girard V, Welker M, Van Nuenen M et al (2015) Microbial typing by matrix-assisted laser desorption ionization-time of flight mass spectrometry: do we need guidance for data interpretation? J Clin Microbiol 53:760–765
35. Mencacci A, Monari C, Leli C, Merlini L, De Carolis E, Vella A et al (2013) Typing of nosocomial outbreaks of Acinetobacter baumannii by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 51:603–606
36. Rim JH, Lee Y, Hong SK, Park Y, Kim M, D'Souza R et al (2015) Insufficient discriminatory power of matrix-assisted laser desorption ionization time-of-flight mass spectrometry dendrograms to determine the clonality of multi-drug-resistant Acinetobacter baumannii isolates from an intensive care unit. Biomed Res Int 2015:535027
37. Sousa C, Botelho J, Grosso F, Silva L, Lopes J, Peixe L (2015) Unsuitability of MALDI-TOF MS to discriminate Acinetobacter baumannii clones under routine experimental conditions. Front Microbiol 6:481
38. Ghebremedhin M, Heitkamp R, Yesupriya S, Clay B, Crane NJ (2017) Accurate and rapid differentiation of Acinetobacter baumannii strains by Raman spectroscopy: a comparative study. J Clin Microbiol 55:2480–2490
39. Sousa C, Silva L, Grosso F, Lopes J, Peixe L (2014) Development of a FTIR-ATR based model for typing clinically relevant Acinetobacter baumannii clones belonging to ST98, ST103, ST208 and ST218. J Photochem Photobiol B 133:108–114
40. Wang H, Drake SK, Yong C, Gucek M, Tropea M, Rosenberg AZ et al (2016) A novel peptidomic approach to strain typing of clinical Acinetobacter baumannii isolates using mass spectrometry. Clin Chem 62:866–875
41. Gerner-Smidt P (1992) Ribotyping of the Acinetobacter calcoaceticus-Acinetobacter baumannii complex. J Clin Microbiol 30:2680–2685
42. Janssen P, Maquelin K, Coopman R, Tjernberg I, Bouvet P, Kersters K et al (1997) Discrimination of Acinetobacter genomic species by AFLP fingerprinting. Int J Syst Bacteriol 47:1179–1187
43. Rafei R, Osman M, Dabboussi F, Hamze M (2019) Update on the epidemiological typing methods for Acinetobacter baumannii. Future Microbiol 14:1065–1080
44. Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37:67–75
45. Parizad EG, Parizad EG, Valizadeh A (2016) The application of pulsed field gel electrophoresis in clinical studies. J Clin Diagn Res 10:DE01–DE04

46. Turton JF, Matos J, Kaufmann ME, Pitt TL (2009) Variable number tandem repeat loci providing discrimination within widespread genotypes of Acinetobacter baumannii. Eur J Clin Microbiol Infect Dis 28:499–507
47. Rafei R, Kempf M, Eveillard M, Dabboussi F, Hamze M, Joly-Guillou ML (2014) Current molecular methods in epidemiological typing of Acinetobacter baumannii. Future Microbiol 9:1179–1194
48. Rafei R, Dabboussi F, Hamze M, Eveillard M, Lemarie C, Gaultier MP et al (2014) Molecular analysis of Acinetobacter baumannii strains isolated in Lebanon using four different typing methods. PLoS One 9:e115969
49. Pourcel C, Minandri F, Hauck Y, D'Arezzo S, Imperi F, Vergnaud G et al (2011) Identification of variable-number tandem-repeat (VNTR) sequences in Acinetobacter baumannii and interlaboratory validation of an optimized multiple-locus VNTR analysis typing scheme. J Clin Microbiol 49:539–548
50. Piran A, Shahcheraghi F, Solgi H, Rohani M, Badmasti F (2017) A reliable combination method to identification and typing of epidemic and endemic clones among clinical isolates of Acinetobacter baumannii. Infect Genet Evol 54:501–507
51. Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
52. Ecker JA, Massire C, Hall TA, Ranken R, Pennella TT, Agasino Ivy C et al (2006) Identification of Acinetobacter species and genotyping of Acinetobacter baumannii by multilocus PCR and mass spectrometry. J Clin Microbiol 44:2921–2932
53. Bartual SG, Seifert H, Hippler C, Luzon MA, Wisplinghoff H, Rodriguez-Valera F (2005) Development of a multilocus sequence typing scheme for characterization of clinical isolates of Acinetobacter baumannii. J Clin Microbiol 43:4382–4390
54. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S (2010) The population structure of Acinetobacter baumannii: expanding multiresistant clones from an ancestral susceptible genetic pool. PLoS One 5:e10034
55. Castillo-Ramirez S, Grana-Miraglia L (2019) Inaccurate multilocus sequence typing of Acinetobacter baumannii. Emerg Infect Dis 25:186–187
56. Quainoo S, Coolen JPM, van Hijum S, Huynen MA, Melchers WJG, van Schaik W et al (2017) Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. Clin Microbiol Rev 30:1015–1063
57. Desai A, Marwah VS, Yadav A, Jha V, Dhaygude K, Bangar U et al (2013) Identification of optimum sequencing depth especially for de novo genome assembly of small genomes using next generation sequencing data. PLoS One 8:e60204
58. Kanamori H, Parobek CM, Weber DJ, van Duin D, Rutala WA, Cairns BA et al (2015) Next-generation sequencing and comparative analysis of sequential outbreaks caused by multidrug-resistant Acinetobacter baumannii at a large academic burn center. Antimicrob Agents Chemother 60:1249–1257
59. Fitzpatrick MA, Ozer EA, Hauser AR (2016) Utility of whole-genome sequencing in characterizing Acinetobacter epidemiology and Analyzing hospital outbreaks. J Clin Microbiol 54:593–612
60. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT (2015) Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol 53:1072–1079
61. Bogaty C, Mataseje L, Gray A, Lefebvre B, Levesque S, Mulvey M et al (2018) Investigation of a carbapenemase-producing Acinetobacter baumannii outbreak using whole genome sequencing versus a standard epidemiologic investigation. Antimicrob Resist Infect Control 7:140
62. Feng Y, Ruan Z, Shu J, Chen CL, Chiu CH (2016) A glimpse into evolution and dissemination of multidrug-resistant Acinetobacter baumannii isolates in East Asia: a comparative genomics study. Sci Rep 6:24342
63. De Socio GV, Rubbioni P, Botta D, Cenci E, Belati A, Paggi R et al (2019) Measurement and prediction of antimicrobial resistance in bloodstream infections by ESKAPE pathogens and Escherichia coli. J Glob Antimicrob Resist 19:154–160

64. Davis JJ, Boisvert S, Brettin T, Kenyon RW, Mao C, Olson R et al (2016) Antimicrobial resistance prediction in PATRIC and RAST. Sci Rep 6:27930
65. Chowdhury AS, Call DR, Broschat SL (2019) Antimicrobial resistance prediction for gram-negative bacteria via game theory-based feature evaluation. Sci Rep 9:14487
66. Bush K (2018) Past and present perspectives on beta-lactamases. Antimicrob Agents Chemother 62:e01076-18
67. Poirel L, Nordmann P (2006) Carbapenem resistance in Acinetobacter baumannii: mechanisms and epidemiology. Clin Microbiol Infect 12:826–836
68. Liu Y, Liu X (2015) Detection of AmpC beta-lactamases in Acinetobacter baumannii in the Xuzhou region and analysis of drug resistance. Exp Ther Med 10:933–936
69. Turton JF, Ward ME, Woodford N, Kaufmann ME, Pike R, Livermore DM et al (2006) The role of ISAba1 in expression of OXA carbapenemase genes in Acinetobacter baumannii. FEMS Microbiol Lett 258:72–77
70. Ruppe E, Cherkaoui A, Lazarevic V, Emonet S, Schrenzel J (2017) Establishing genotype-to-phenotype relationships in bacteria causing hospital-acquired pneumonia: a prelude to the application of clinical metagenomics. Antibiotics (Basel) 6:30
71. Roe CWC, Vazquez AJ et al (2019) Bacterial genome wide association studies (bGWAS) and transcriptomics identifies cryptic antimicrobial resistance mechanisms in Acinetobacter baumannii. bioRxiv. https://doi.org/10.1101/864462
72. Wang N, Ozer EA, Mandel MJ, Hauser AR (2014) Genome-wide identification of Acinetobacter baumannii genes necessary for persistence in the lung. MBio 5:e01163–e01114
73. Subashchandrabose S, Smith S, DeOrnellas V, Crepin S, Kole M, Zahdeh C et al (2016) Acinetobacter baumannii genes required for bacterial survival during bloodstream infection. mSphere 1:e00013-15
74. Hamidian M, Hawkey J, Wick R, Holt KE, Hall RM (2019) Evolution of a clade of Acinetobacter baumannii global clone 1, lineage 1 via acquisition of carbapenem- and aminoglycoside-resistance genes and dispersion of ISAba1. Microb Genom 5:e000242
75. Adams MD, Wright MS, Karichu JK, Venepally P, Fouts DE, Chan AP et al (2019) Rapid replacement of Acinetobacter baumannii strains accompanied by changes in lipooligosaccharide loci and resistance gene repertoire. MBio 10:e00356-19
76. Lopez M, Rueda A, Florido JP, Blasco L, Fernandez-Garcia L, Trastoy R et al (2018) Evolution of the quorum network and the mobilome (plasmids and bacteriophages) in clinical strains of Acinetobacter baumannii during a decade. Sci Rep 8:2523
77. Duchene S, Holt KE, Weill FX, Le Hello S, Hawkey J, Edwards DJ et al (2016) Genome-scale rates of evolutionary change in bacteria. Microb Genom 2:e000094
78. Hawkey J, Ascher DB, Judd LM, Wick RR, Kostoulias X, Cleland H et al (2018) Evolution of carbapenem resistance in Acinetobacter baumannii during a prolonged infection. Microb Genom 4:e000165
79. Hammerstrom TG, Beabout K, Clements TP, Saxer G, Shamoo Y (2015) Acinetobacter baumannii repeatedly evolves a hypermutator phenotype in response to tigecycline that effectively surveys evolutionary trajectories to resistance. PLoS One 10:e0140489
80. Grana-Miraglia L, Lozano LF, Velazquez C, Volkow-Fernandez P, Perez-Oseguera A, Cevallos MA et al (2017) Rapid gene turnover as a significant source of genetic variation in a recently seeded population of a healthcare-associated pathogen. Front Microbiol 8:1817
81. Peng Y, Leung HC, Yiu SM, Chin FY (2011) Meta-IDBA: a de Novo assembler for metagenomic data. Bioinformatics 27:i94-101
82. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J (2012) Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol 13:R122
83. Zerbino DR (2010) Using the Velvet de novo assembler for short-read sequencing technologies. Curr Protoc Bioinformatics Chapter 11:Unit 11.5
84. Souvorov A, Agarwala R, Lipman DJ (2018) SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol 19:153

85. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477
86. Schurch AC, Arredondo-Alonso S, Willems RJL, Goering RV (2018) Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. Clin Microbiol Infect 24:350–354
87. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD et al (2015) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using gubbins. Nucleic Acids Res 43:e15
88. Hall MD, Holden MT, Srisomang P, Mahavanakul W, Wuthiekanun V, Limmathurotsakul D et al (2019) Improved characterisation of MRSA transmission using within-host bacterial sequence diversity. elife 8:e46402
89. Cremers AJH, Coolen JPM, Bleeker-Rovers CP, van der Geest-Blankert ADJ, Haverkate D, Hendriks H et al (2020) Surveillance-embedded genomic outbreak resolution of methicillin-susceptible Staphylococcus aureus in a neonatal intensive care unit. Sci Rep 10:2619
90. Maiden MC, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA et al (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 11:728–736
91. Willems S, Kampmeier S, Bletz S, Kossow A, Kock R, Kipp F et al (2016) Whole-genome sequencing elucidates epidemiology of nosocomial clusters of Acinetobacter baumannii. J Clin Microbiol 54:2391–2394
92. Higgins PG, Prior K, Harmsen D, Seifert H. (2017) Development and evaluation of a core genome multilocus typing scheme for whole-genome sequence-based typing of Acinetobacter baumannii. PLoS One 12(6):e0179228. https://doi.org/10.1371/journal.pone.0179228. PMID: 28594944; PMCID: PMC5464626
93. Venditti C, Vulcano A, D'Arezzo S, Gruber CEM, Selleri M, Antonini M et al (2019) Epidemiological investigation of an Acinetobacter baumannii outbreak using core genome multilocus sequence typing. J Glob Antimicrob Resist 17:245–249
94. Cerezales M, Xanthopoulou K, Wille J, Bustamante Z, Seifert H, Gallego L et al (2019) Acinetobacter baumannii analysis by core genome multi-locus sequence typing in two hospitals in Bolivia: endemicity of international clone 7 isolates (CC25). Int J Antimicrob Agents 53:844–849
95. Wendel AF, Malecki M, Otchwemah R, Tellez-Castillo CJ, Sakka SG, Mattner F (2018) One-year molecular surveillance of carbapenem-susceptible A. baumannii on a German intensive care unit: diversity or clonality. Antimicrob Resist Infect Control 7:145
96. Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, Santos S et al (2018) chewBBACA: a complete suite for gene-by-gene schema creation and strain identification. Microb Genom 4:e000166

Enterococcus

7

Lúcia M. Teixeira, Adriana R. Faria, Stephanie S. R. Souza, and Vânia Lúcia C. Merquior

7.1 Introduction

The enterococci are a diverse and versatile group of bacteria with several intrinsic characteristics that allow them to survive and grow under a variety of conditions and a remarkable metabolic adaptability in order to fulfill diverse roles as commensals and as opportunistic pathogens. These microorganisms are widely distributed in nature, mainly on the mucosal surfaces of humans and animals, and they are also found in soil, water, dairy products and other foodstuffs, and on plants. In humans, they are predominantly inhabitants of the gastrointestinal tract, and are less commonly found in other body sites, such as in the genitourinary tract, the oral cavity, and skin.

Over the last decades, they have emerged as medically important multiple antibiotic-resistant pathogens, especially in the context of healthcare-associated infections, contributing significantly to patient morbidity and mortality, as well as increasing healthcare costs. A number of intrinsic characteristics of the enterococci

L. M. Teixeira (✉)
Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil
e-mail: lmt2@micro.ufrj.br

A. R. Faria
Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

S. S. R. Souza
Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Departamento de Microbiologia, Imunologia e Parasitologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

V. L. C. Merquior
Departamento de Microbiologia, Imunologia e Parasitologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_7

allow them to persist almost everywhere and present a challenge when controlling of their dissemination is necessary. In addition, they have the capacity to acquire resistance to a wide variety of antimicrobial agents which represents a serious additional threat in the management of patients with enterococcal infections [1–4].

Enterococcal infections were traditionally considered to be acquired endogenously from the patient's own normal flora, and the epidemiology of enterococcal infection attracted little attention. This perspective has dramatically changed, and a major interest has focused on the epidemiology of enterococcal infections, because of the increasing documentation of *Enterococcus* as a leading nosocomial pathogen. Furthermore, the emergence and dissemination of multidrug resistance among enterococcal strains and the evidence supporting the concept of exogenous acquisition of enterococcal infections have generated an additional need for typing the isolates as a means of assisting infection control and epidemiological studies both within and among various medical institutions. Therefore, the investigation of epidemiological aspects of nosocomial outbreaks as well as the dissemination of enterococcal strains harboring antimicrobial resistance markers is of major interest, particularly in the light of the increasing occurrence of vancomycin-resistant enterococci (VRE). Besides outbreak analysis, the methods used for epidemiological investigation of enterococcal isolates must be able to track enterococcal dissemination in different environments and hosts, and the evolution of multidrug-resistant (MDR) strains. In this chapter we comment on molecular techniques that can be used for the identification, rapid detection, and typing of enterococcal isolates.

7.2 Characteristics and Current Classification of the Genus

The genus *Enterococcus* is composed of Gram-positive cocci that occur singly, in pairs, or as short chains. They are nonspore-forming, facultatively anaerobic, catalase- negative bacteria, with a fermentative metabolism resulting in L(+) lactic acid as the major product of glucose fermentation. Characteristics such as growth in broth containing 6.5% NaCl and hydrolysis of esculin in the presence of bile salts (bile-esculin [BE] test) are useful to identify enterococcal strains. Other characteristics presented by most enterococci include hydrolysis of leucine-β-naphthylamide (LAP) and L-pyrrolidonyl-β-naphthylamide (PYR) [1, 4, 5].

The enterococci have a historical connection with the genus *Streptococcus*, and they were earlier considered as a major branch within this genus, distinguished by their higher resistance to chemical and physical agents and accommodating most of the serological group D streptococci. After the introduction of molecular methods for characterization of these microorganisms, they have undergone considerable changes in taxonomy, which started with the recognition of *Enterococcus* as a separate genus [6]. *Streptococcus faecalis* and *Streptococcus faecium* were the first species to be transferred to the new genus as *Enterococcus faecalis* and *Enterococcus faecium*, respectively. Since then, the increasing use of molecular approaches has led to the description of several new enterococcal species and entitled many nomenclatural changes, resulting in the present recognition of 65 species assigned to the

genus *Enterococcus* (as of September 2021; for further details see references [1] and [4], as well as the *List of Prokaryotic Names with Standing in Nomenclature* (http://www.bacterio.net).

Current criteria for inclusion in the genus *Enterococcus* and for the description of new enterococcal species encompass a polyphasic approach resulting from a combination of different molecular techniques and phenotypic testing. Sequencing of the 16S rRNA gene is a practical and powerful tool in aiding the identification of enterococccal species, although limitations for differentiating a few enterococcal species have been documented [4, 5]. Sequencing assays targeting other genetic determinants have increasingly been used as additional tools to assess the phylogenetic relationships among enterococcal species and to formulate the description of new species. Housekeeping genes coding for proteins involved in basic cellular functions [such as the phenylalanyl-tRNA synthase (*pheS* gene), the RNA polymerase α-subunit (*rpoA* gene), the RNA polymerase β-subunit (*rpoB* gene) and the ATP synthase α-subunit (*atpA* gene)] are the most common targets for these sequencing approaches. These assays may be based on the analysis of a single gene or of multiple genes, under the format of multilocus sequence analysis (MLSA) of concatenated sequences [5, 7–9]**.** More recently, comparative analysis of spectra generated by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) as well as analysis of sequences generated by whole genome sequencing (WGS) have also been incorporated into the array of methodologies used for taxonomic studies of the genus *Enterococcus* [10–12]**.**

In diagnostic laboratory settings, identification of enterococcal species is generally accomplished by using a series of conventional physiological tests [1, 4www.cdc.gov/ncidod/biotech/strep/strep-doc/index.htm]. Several miniaturized, manual, semiautomated, and automated identification systems constitute alternatives for the phenotypic identification of enterococcal species in routine diagnostic laboratories [1, 13]. In recent years, MALDI-TOF MS has increasingly been used for diagnostic purposes and was shown to be useful in providing rapid, accurate, and highly convenient identification of a diversity of enterococcal species [13–15]. Future improvements of this technique may also become widely available for the rapid and precise detection of enterococci directly in clinical samples [16–19].

7.3 Clinical Significance and Epidemiology

The enterococci are opportunistic agents of a variety of infections, particularly in elderly patients with serious underlying diseases and other immunocompromised patients who have been hospitalized for prolonged periods, treated with invasive devices, and/or have received broad-spectrum antimicrobial therapy. They most commonly cause infections of the urinary tract, bloodstream, endocardium, wounds (mostly surgical site, decubitus ulcers, and burn wounds), catheters, and other implanted medical devices [2, 20]. They are also frequently associated with intra-abdominal and pelvic infections. Enterococcal infections of the respiratory tract or the central nervous system, as well as otitis, sinusitis, septic arthritis, and

endophthalmitis, may occur but are rare. The ubiquitous presence of enterococci, however, requires caution in establishing the clinical significance of a particular isolate. The gastrointestinal tract is believed to represent the major reservoir for enterococcal strains associated with disease; from this location they may migrate to cause infections and can also disseminate to other hosts and to the environment [3, 21].

The pathogenesis of enterococcal infections is still poorly understood. Several potential virulence factors have been identified, although none has been established as having a major contribution to enterococcal virulence. Nevertheless, epidemiological studies show the existence of clonal relationships among outbreak isolates and support the notion that a subset of virulent lineages are often responsible for the infections of epidemic proportions [22–25].

Although the enterococci can cause human infections in the community and in healthcare settings, these microorganisms began to be recognized with increasing frequency as common causes of healthcare-associated infections (HAIs) in the late 1970s, paralleling the increasing resistance to most currently used antimicrobial agents. As a result, enterococci have emerged as one of the leading therapeutic challenges when associated with serious or life-threatening infections. The ratios of isolation of the different enterococcal species can vary according to each setting and can be affected by a number of aspects, including the dissemination of outbreak-related strains such as VRE. *E. faecalis* and *E. faecium* are usually the most common enterococcal species isolated from human clinical specimens. Historically, *E. faecalis* represented about 80–90% of the clinically significant enterococcal isolates, while *E. faecium* was found in 5–10% of enterococcal infections. However, a trend for a progressive decline in the ratio of *E. faecalis* to *E. faecium* is notable, particularly among isolates from bloodstream infections [23, 26, 27].

VRE infections are difficult to treat and often appear as part of long-lasting outbreaks in healthcare institutions, representing considerable challenges for infection control. As the occurrence and spread of VRE has reached a more global dimension, transmission of these antimicrobial resistant variants became a major public health concern worldwide. Although only a small percentage of colonized patients will develop serious systemic enterococcal infections, intestinal colonization with VRE has been clearly associated with subsequent VRE infections. However, in certain specific clinical situations (i.e., liver transplant recipients, patients on chronic hemodialysis, and patients with hematological malignancies) VRE-colonized patients appear to be at increased risk for developing serious enterococcal infections [2, 20]. This underscores the importance of active surveillance in high-risk patient groups to prevent transmission and outbreaks. Therefore, hospitals are encouraged to implement surveillance programs for VRE detection [2, 20, 28]. In an attempt to overcome the inherent limitations of the culture-based methods of detection, PCR-based methods have been evaluated for direct detection of these microorganisms in clinical and surveillance specimens. Three major systems for molecular detection directly in clinical and surveillance specimens are commercially available. They are: the LightCycler *vanA*/*vanB* detection assay (Roche Molecular Diagnostics, Indianapolis, IN), the GenOhm VanR assay (Becton-Dickinson Microbiology

Systems, Cockeysville, MD), and the Cepheid Xpert *vanA*/*vanB* assay (Cepheid, Sunnyvale, CA) [29–31]. They are based on the detection of the genetic determinants (*vanA* and *vanB*) associated with the two major types of resistance to vancomycin. Improved molecular detection of VRE over that obtained with conventional culture techniques contributes to decreasing the time for detection and potentially reduces the risk of VRE transmission, improving patient care and reducing costs for healthcare facilities.

7.4 Resistance to Antimicrobial Agents

Resistance to several commonly used antimicrobial agents is a remarkable characteristic of most enterococcal species and can be either intrinsic or acquired. The occurrence of acquired traits leading to high-level resistance to aminoglycosides (HLR-A), and resistance to glycopeptides, especially to vancomycin, is of particular clinical significance due to the impact in the treatment of enterococcal infections.

As already mentioned, the isolation of VRE has been continuously reported, indicating epidemic proportions in diverse geographic locations. Taking this phenomenon into account, the Centers for Disease Control and Prevention (CDC) has recognized VRE as a major threat to public health (https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf), and the World Health Organization (WHO) has included vancomycin-resistant *E. faecium* in the list of bacteria with high priority for the development of new control strategies (<http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/>).

VRE strains have been classified according to phenotypic and genotypic features. Nine types of resistance to glycopeptides have already been described among enterococci. Each type is associated with different genetic elements, some of which, in turn, can be divided into subtypes. The *vanA* and *vanB* are considered the most clinically relevant genotypes and are usually associated with *E. faecium* and *E. faecalis* isolates, while the VanC resistance is an intrinsic characteristic of *E. gallinarum* (*vanC1* genotype) and *E. casseliflavus* (*vanC2*–- *vanC4* genotypes). The additional types of glycopeptide resistance, encoded by the *vanD*, *vanE*, *vanG*, and *vanL-vanN* genes, seem to occur rarely among enterococci [4, 32–34].

The development of resistance to several other drugs currently available highlights the difficulties faced by clinicians to control certain enterococcal infections and requires the use of new or innovative therapeutic approaches that involve both old and new antimicrobials [35]. Molecular methods have been designed to detect specific antimicrobial resistance genes and have substantially contributed to the understanding of the spread of acquired resistance among enterococci, especially resistance to vancomycin, as already commented. Considering the high frequency and diversity of antimicrobial resistance traits among enterococcal isolates, determination of the profile of genes associated with resistance to a variety of antimicrobials may be used as an additional valuable tool for epidemiology and typing purposes.

7.5 Typing Methods

7.5.1 Early Typing Methods

Early epidemiological investigations of enterococcal infections were based on phenotypic typing methods, including biotyping and antibiotyping, serotyping, bacteriocin typing, and bacteriophage typing [1, 36]. Although these approaches have occasionally yielded useful information, they frequently fail to adequately discriminate among enterococcal strains, and therefore, they are of limited value for comprehensive epidemiological studies. On the other hand, the use of phenotypic typing methods in conjunction with molecular typing approaches may contribute with valuable information.

7.5.2 Molecular Typing Methods

The introduction of molecular techniques has substantially improved the ability to discriminate enterococcal isolates and has provided critical insights into the epidemiology of the enterococci. By using molecular typing approaches, it was possible to demonstrate the exogenous acquisition of enterococcal strains by direct and indirect contact among patients, breaking the traditional conception that enterococcal infections were endogenous in nature. Intrahospital transmission and interhospital spread have been extensively documented for antimicrobial resistant enterococci, especially VRE [1, 22, 37]. In addition to epidemiological investigations, some of the molecular typing techniques are now used to trace the dissemination of enterococci in different environments and hosts, phylogenetic relationship, and the evolution of multidrug-resistant strains, greatly expanding our understanding of enterococcal epidemiology, population structure, antimicrobial resistance, and virulence. Emergence and global dispersion of certain epidemic enterococcal clonal complexes have been identified [22–25, 37–41].

Several molecular methods have been proposed to type enterococcal isolates as previously reviewed [1, 36]. The first molecular techniques developed for typing of enterococci were the analysis of plasmids profiles (including both plasmid composition and restriction endonuclease analysis of specific plasmids) and the restriction enzyme analysis (REA) of genomic DNA by conventional electrophoresis. These techniques may be helpful in some instances, but problems related to inconsistencies in plasmid yield and to difficulties in accurate interpretation of the electrophoretic profiles have been encountered with the use of these methods. Multilocus enzyme electrophoresis (MLEE), ribotyping, and the polymerase chain reaction (PCR)-based typing methods, such as the random amplified polymorphic DNA (RAPD-PCR) assay, and the repetitive element sequence (REP)-PCR have also been used to investigate the genetic relationship among enterococcal strains. These methods also have limitations, such as poor reproducibility and/or high technical complexity. DNA sequencing of PCR products and restriction fragment length polymorphism (RFLP) analysis of PCR products have been used to trace and

determine differences among specific resistance genes in enterococci, therefore representing additional tools for typing resistant strains.

A remarkable contribution to the ability to discriminate among enterococcal strains was noted with the use of techniques involving the analysis of chromosomal DNA restriction endonuclease profiles by pulsed-field gel electrophoresis (PFGE). This methodology has been extensively evaluated for epidemiological characterization of enterococcal outbreaks, showing improved strain discrimination and allowing the delineation of clonal relationships among enterococcal isolates, identifying clusters of strains that predominate among MDR enterococci, mainly strains with HLR-A and VRE [1, 37, 42–47]. *Sma*I is the restriction enzyme more frequently used to digest enterococcal DNA, and the usefulness of other enzymes, such as *Apa*I and *Sfi*I, has also been documented [1].

Several protocols for performing PFGE typing of enterococcal strains have been published. However, standardized protocols for execution, interpretation, and nomenclature of PFGE profiles, as a result of collaborative studies, are still not available, making difficult to perform inter-laboratory data exchange and comparisons. On the other hand, although PFGE is quite discriminatory, epidemiological interpretation of PFGE profiles is not always clear-cut. The occurrence of genetic events can be associated with substantial changes in the PFGE profiles, leading to problems in clonality assessment [48]. Due to the possibility of such inconsistencies in DNA banding patterns of enterococci, PFGE is recommended mostly for the purpose of evaluating the genetic relatedness and tracing transmission of strains that are associated in time and location, as usefulness for long-term epidemiological studies may be limited. The use of PFGE in conjunction with at least one additional typing technique, or independent PFGE analysis using different restriction enzymes, is highly recommended to help in clarifying epidemiological interpretation. General principles proposed for the interpretation of molecular typing data based on fragment differences are usually applied to interpret PFGE profiles obtained for enterococcal strain. Well-characterized control strains should be evaluated along with unknown isolates. For that purpose, two reference strains, *faecalis* OG1RF (ATCC® 47077™) and *E. faecium* GE1 (ATCC® 51558™), have been proposed [49].

In the context of robust techniques for typing enterococci, multilocus sequence typing (MLST) has emerged as a methodology which exhibits a number of advantages for diversity analyses [50, 51]. This technique circumvents the difficulties in data exchange between different laboratories by generating information that is suitable for the development of web-based databases. MLST is based on identifying alleles after sequencing of internal fragments of a number of selected housekeeping genes, resulting in a numeric allelic profile. Each profile is assigned a sequence type (ST). Internet sites with the possibility for data exchange have been developed, which contain MLST protocols for *E. faecium* (https://pubmlst.org/efaecium/) and *E. faecalis* (https://pubmlst.org/efaecalis/). MLST schemes for these two species are based on sequence analysis of seven loci, each one corresponding to a separate set of different genes [50, 51]. Examples of application of MLST analysis to type enterococcal isolates are shown in Figs. 7.1 and 7.2. Figure 7.1 shows a goeBURST diagram a population snapshot of *E. faecium* sequence types (STs) and clonal

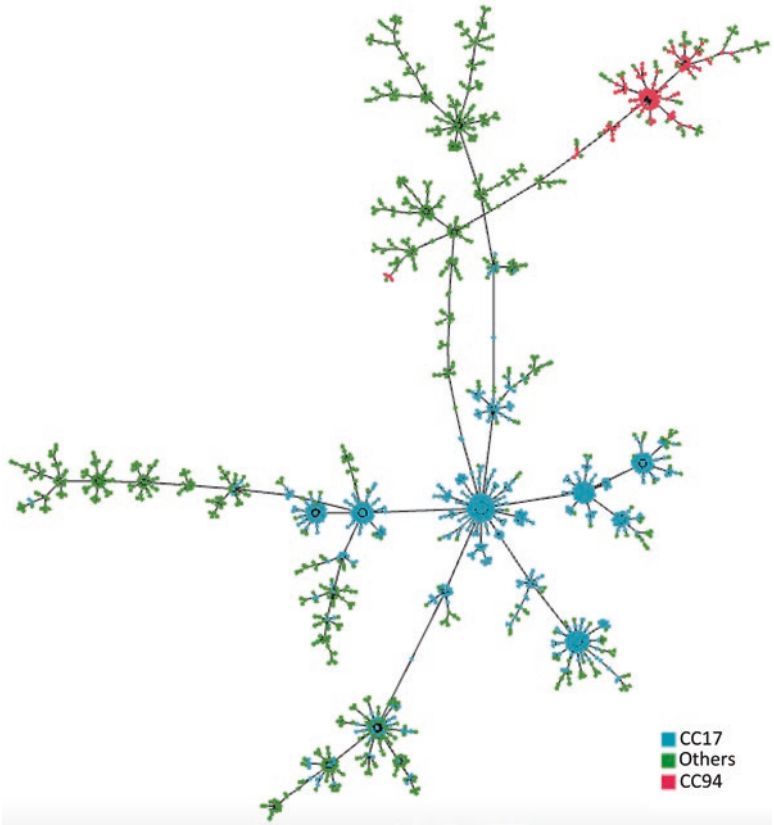

Fig. 7.1 Diagram constructed by using the goeBURST algorithm and displayed on PHYLOViZ software representing a population snapshot of *Enterococcus faecium* sequence types (STs) and clonal complexes (CC) available at the MLST database (https://pubmlst.org/efaecium/ as of November 2019). Each ST is represented as a node colored according to the respective CC. Lines connect single locus variants: STs that differ in only one of the seven housekeeping genes

complexes (CC) available at the MLST database (as of November 2019). Figure 7.2 depicts clonal complexes defined by MLST analysis of *E. faecalis* strains presenting HLR-A, isolated from patients in Brazilian hospitals.

MLST is particularly effective in epidemiological evaluations that include long periods of analysis in retrospective studies. Analysis of population structure of *E. faecalis* and *E. faecium*, by using MLST, has been an important tool in epidemiological investigation of highly specialized enterococcal subpopulations adapted to the hospital environment. Major clonal complexes (CC) were identified and found to be widely distributed, such as CC17 among *E. faecium*, and CC2 and CC9 among *E. faecalis* [25, 45–47, 52, 53].

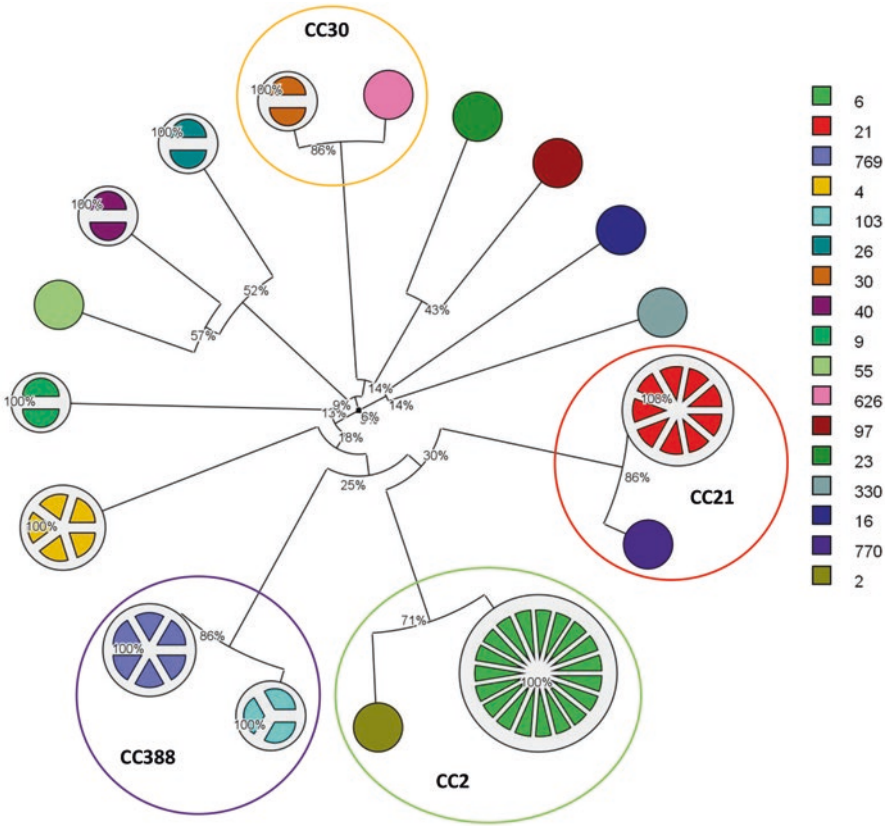

Fig. 7.2 Circular dendrogram representing clonal complexes defined by MLST analysis of *Enterococcus faecalis* isolates showing high-level resistance to aminoglycosides (HLR-A) obtained from healthcare institutions located in Rio de Janeiro, Brazil. The dendrogram was created by using the BioNumerics v.7.6 software. Pie charts represent the set of strains that were typed as a given ST corresponding to the color in the caption, and the slices indicate the number of strains. Each filled circle represent a ST observed in only one strain with the color corresponding to the ST in the caption

However, MLST analyses have shown differences in the distribution of *E. faecalis* and *E. faecium* clonal complexes. Despite the correlation between certain subpopulations and the hospital environment for both species, this evidence is more strongly significant in *E. faecium*. The population structure of *E. faecalis* includes CCs that often group hospital and community strains together, suggesting an evolutionary identity among them. On the other hand, considering that both species are highly recombinogenic, it has been suggested that, depending on the analytical algorithm applied, the definition of CCs may not be precise in indicating the genetic relationships of these microorganisms. Analytical methods as eBURST and global

optimal eBURST (goeBURST) are widely used to indicate strains relationships. The Bayesian Analysis of Population Structure (BAPS), which uses a genetic-statistical model to partition a set of strains into related groups according to genetic recombination and ancestry patterns, has been proposed as a more efficient tool to infer phylogenetic arrangements [25, 46, 54].

Overall, PFGE and MLST are still expensive methodologies and often require specialized personnel, limiting their use particularly in routine laboratories. Thus, the development of methods that allow fast, simple, and reliable characterization of microorganisms is still a major goal. Among these, the MALDI-TOF MS technique has also emerged as a potential method for molecular typing of several bacterial species, including *Enterococcus*. Despite present limitations, it has been endorsed mainly due to the potential to discriminate high-risk VRE strains in the hospital setting [47, 55–57].

Multiple-locus variable-number tandem repeat analysis (MLVA), a method based on differences in variable-number of tandem repeats (VNTR) in multiple loci dispersed over the genome, can also be considered a lower cost alternative for typing enterococci. For each VNTR locus, the number of repeats is determined by PCR using primers based on the conserved flanking regions of the tandem repeats. PCR products are separated on agarose gels and the band size determines the number of repeats. These numbers together result in a MLVA profile and each profile is assigned a MLVA type (MT). The MLVA scheme for *E. faecium* is based on six VNTR loci present in noncoding regions. On the other hand, the MLVA typing scheme for *E. faecalis* is based on seven targets obtained from known genes. Comparative studies indicate that both MLST and MLVA techniques can achieve high degrees of discrimination between isolates and have comparable discriminatory power that appears to be similar to that of PFGE-based typing [37, 45, 51, 58, 59].

More recently, molecular typing based on whole-genome sequencing (WGS) has increased in use for epidemiological analysis, assessing genomic variability of bacterial pathogens, including *Enterococcus* species. WGS analysis of entire genomes provides markedly higher resolution than those of other methods, and full genetic information can be obtained. Investigations of relatedness can be based on entire genome sequences, or, e.g., in silico MLST or various single nucleotide polymorphisms (SNPs) analyses can be performed. Typing by WGS usually evaluates the bacterial strains diversity considering SNPs in core genome or by the identification of unique allelic patterns. Core genome MLST (cgMLST) combines the advantages offered by the MLST technique with the great possibility of information generated by WGS analysis. In this way, SNPs profiling was shown to be cost-effective and a good alternative to conventional MLST in *Enterococcus* typing. Furthermore, if occasionally core genome typing is not enough discriminatory for epidemiological purposes, the inclusion of analysis of accessory genes and plasmid sequences certainly will provide additional data to base molecular typing. Multigenome analyses of *E. faecalis* and *E. faecium* have demonstrated that both species have an open pangenome, indicating that they can efficiently acquire and integrate external DNA

into their genetic pool. The acquisition and exchange of moving elements have been valued as important attributes to be considered in the analyses of the genomic diversity of these species [52, 53, 60–63]. Extensive use of these more recent molecular approachs to type enterococcal isolates from different sources and geographical areas will certainly contribute with new insights on the epidemiology of the enterococci and their roles as important opportunistic infectious agents, as well as components of human, animal and environmental microbiomas.

7.6 Summary

The genus *Enterococcus* is composed by Gram-positive cocci widely distributed in nature, presenting either as harmless commensals or as multifaceted opportunistic pathogens recognized among the leading causes of difficult to treat nosocomial infections. Over the last decades, they have emerged as medically important multiple antibiotic-resistant pathogens, especially in the context of healthcare-associated infections, contributing significantly to patient morbidity and mortality, as well as increasing healthcare costs. Several characteristics of the enterococci allow them to persist almost everywhere and acquire resistance to multiple antimicrobial drugs, thus presenting a considerable challenge for infection control. The introduction of molecular typing techniques, including PFGE, MLST, MLVA, and other PCR based-methodologies, and, more recently, MALDI-TOF MS and WGS, has substantially improved the ability to discriminate enterococcal isolates and has provided critical epidemiology insights. Exogenous acquisition of enterococcal strains by direct and indirect contact among patients has been demonstrated, breaking the traditional conception that enterococcal infections were endogenous in nature. Occurrence and dissemination of successful clonal complexes of antimicrobial-resistant *Enterococcus faecium* and *Enterococcus faecalis* have extensively been documented, particularly among VRE isolates. In addition to epidemiological investigations of human infections, robust molecular typing techniques are now used to trace the dissemination of enterococci in different environments and hosts, and the evolution of multidrug-resistant strains, greatly expanding our understanding of enterococcal epidemiology, population structure, antimicrobial resistance and virulence.

References

1. Facklam RR, Carvalho MGS, Teixeira LM (2002) History, taxonomy, biochemical characteristics, and antibiotic susceptibility testing of enterococci. In: Gilmore MS, Clewell DB, Courvalin P et al (eds) The enterococci: pathogenesis, molecular biology, and antibiotic resistance. ASM, Washington, DC
2. Arias CA, Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. Nat Rev Microbiol 10:266–278

3. Lebreton F, Willems RJL, Gilmore MS (2014) *Enterococcus* diversity, origins in nature, and gut colonization. In: Gilmore MS, Clewell DB, Ike Y et al (eds) Enterococci: from commensals to leading causes of drug resistant infection. Massachusetts Eye and Ear Infirmary, Boston
4. Teixeira LM, Carvalho MG, Facklam RR et al (2019) Enterococcus. In: Carroll KC, Pfaller MA, Landry ML et al (eds) Manual of clinical microbiology, 12th edn. ASM, Washington, DC
5. Švec P, Devriese LA (2009) Genus I. *Enterococcus*. In: De Vos P, Garrity GM, Jones D et al (eds) Bergey's manual of systematic bacteriology, vol 3: the firmicutes, 2nd edn. Springer, New York
6. Schleifer KH, Kilpper-Balz R (1984) Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. Int J Syst Bacteriol 34:31–34
7. Naser SM, Thompson FL, Hoste B et al (2005) Application of multilocus sequence analysis (MLSA) for rapid identification of *Enterococcus* species based on *rpoA* and *pheS* genes. Microbiology 151:2141–2150
8. Naser S, Thompson FL, Hoste B et al (2005) Phylogeny and identification of enterococci by *atpA* gene sequence analysis. J Clin Microbiol 43:2224–2230
9. Li X, Xing J, Li B et al (2012) Use of *tuf* as a target for sequence-based identification of Gram-positive cocci of the genus *Enterococcus*, *Streptococcus*, coagulase-negative *Staphylococcus*, and *Lactococcus*. Ann Clin Microbiol Antimicrob 11:31–36
10. Kadri Z, Spitaels F, Cnockaert M et al (2015) *Enterococcus bulliens* sp. nov., a novel lactic acid bacterium isolated from camel milk. Antonie Van Leeuwenhoek 108:1257–1265
11. Le Page S, Cimmino T, Togo A et al (2016) Noncontiguous finished genome sequence and description of *Enterococcus massiliensis* sp. nov. New Microbes New Infect 12:90–95
12. Sun Y, Li X, Wang G et al (2016) Genome sequence of *Enterococcus pernyi*, a pathogenic bacterium for the Chinese oak silkworm, *Antheraea pernyi*. Genome Announc 4:e01764–e01715
13. Fang H, Ohlsson AK, Ullberg M et al (2012) Evaluation of species-specific PCR, Bruker MS, VITEK MS and the VITEK 2 system for the identification of clinical *Enterococcus* isolates. Eur J Clin Microbiol Infect Dis 31:3073–3977
14. Rychert J, Burnham CA, Bythrow M et al (2013) Multicenter evaluation of the Vitek MS matrix-assisted laser desorption ionization-time of flight mass spectrometry system for identification of Gram-positive aerobic bacteria. J Clin Microbiol 51:2225–2231
15. Christ APG, Ramos SR, Cayô R et al (2017) Characterization of *Enterococcus* species isolated from marine recreational waters by MALDI-TOF MS and Rapid ID API® 20 Strep system. Mar Pollut Bull 118:376–381
16. Ashizawa K, Murata S, Terada T et al (2017) Applications of copolymer for rapid identification of bacteria in blood culture broths using matrix-assisted laser desorption ionization time-of-flight mass spectrometry. J Microbiol Methods 139:54–60
17. Li W, Sun E, Wang Y et al (2019) Rapid identification and antimicrobial susceptibility testing for urinary tract pathogens by direct analysis of urine samples using a MALDI-TOF MS-based combined protocol. Front Microbiol 10:1182
18. Libertucci J, Bassis CM, Cassone M et al (2019) Bacteria detected in both urine and open wounds in nursing home residents: a pilot study. mSphere 4(4):pii: e00463-19
19. Pinault L, Chabrière E, Raoult D et al (2019) Direct identification of pathogens in urine by use of a specific matrix-assisted laser desorption ionization-time of flight spectrum database. J Clin Microbiol 57(4):pii: e01678-18
20. Agudelo Higuita NI, Huycke MM (2014) Enterococcal disease, epidemiology, and implications for treatment. In: Gilmore MS, Clewell DB, Ike Y et al (eds) Enterococci: from commensals to leading causes of drug resistant infection. Massachusetts Eye and Ear Infirmary, Boston
21. Gilmore MS, Lebreton F, van Schaik W (2013) Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. Curr Opin Microbiol 16:10–16
22. Leavis HL, Bonten MJ, Willems RJ (2006) Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. Curr Opin Microbiol 9:454–460

23. Top J, Willems R, Blok H et al (2007) Ecological replacement of *Enterococcus faecalis* by multiresistant clonal complex 17 *Enterococcus faecium*. Clin Microbiol Infect 13:316–319
24. Willems RJ, Bonten MJ (2007) Glycopeptide-resistant enterococci: deciphering virulence, resistance and epidemicity. Curr Opin Infect Dis 20:384–390
25. Lee T, Pang S, Abraham S et al (2019) Antimicrobial-resistant CC17 *Enterococcus faecium*: the past, the present and the future. J Glob Antimicrob Resist 16:36–47
26. Weiner LM, Webb AK, Limbago B et al (2016) Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011–2014. Infect Control Hosp Epidemiol 37:1288–1301
27. European Centre for Disease Prevention and Control (2017) Antimicrobial resistance surveillance in Europe 2016. Annual report of the European antimicrobial resistance surveillance network (EARS-net). ECDC, Stockholm
28. Weber SG, Huang SS, Oriola S et al (2007) Legislative mandates for use of active surveillance cultures to screen for methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci: position statement from the joint SHEA and APIC task force. Infect Control Hosp Epidemiol 28:249–260
29. Usacheva EA, Ginocchio CC, Morgan M et al (2010) Prospective, multicenter evaluation of the BD GeneOhm VanR assay for direct, rapid detection of vancomycin-resistant *Enterococcus* species in perianal and rectal specimens. Am J Clin Pathol 134:219–226
30. Werner G, Serr A, Schutt S et al (2011) Comparison of direct cultivation on a selective solid medium, polymerase chain reaction from an enrichment broth, and the BD GeneOhm VanR assay for identification of vancomycin-resistant enterococci in screening specimens. Diagn Microbiol Infect Dis 70:512–521
31. Holzknect BJ, Hansen DS, Nielsen L et al (2017) Screening for vancomycin-resistant enterococci with Xpert *vanA*/*vanB*; diagnostic accuracy and impact on infection control decision making. New Microbes New Infect 16:54–59
32. Depardieu F, Perichon B, Courvalin P (2004) Detection of the *van* alphabet and identification of enterococci and staphylococci at the species level by multiplex PCR. J Clin Microbiol 42:5857–5860
33. Werner G, Coque TM, Hammerum AM et al (2008) Emergence and spread of vancomycin resistance among enterococci in Europe. Euro Surveill 13:1–11
34. Faron ML, Ledeboer NA, Buchan BW (2016) Resistance mechanisms, epidemiology, and approaches to screening for vancomycin-resistant *Enterococcus* in the health care setting. J Clin Microbiol 54:2436–2447
35. Arias CA, Contreras GA, Murray BE (2010) Management of multidrug-resistant enterococcal infections. Clin Microbiol Infect 16:555–562
36. Domig KJ, Mayer HK, Kneifel W (2003) Methods used for the isolation, enumeration, characterization and identification of *Enterococcus* spp. 2. Pheno- and genotypic criteria. Int J Food Microbiol 88:165–188
37. Freitas AR, Novais C, Ruiz-Garbajosa P et al (2009) Clonal expansion within clonal complex 2 and spread of vancomycin-resistant plasmids among different genetic lineages of *Enterococcus faecalis* from Portugal. J Antimicrob Chemother 63:1104–1111
38. Willems RJ, Top J, van Santen M (2005) Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. Emerg Infect Dis 11:821–828
39. McBride SM, Fischetti VA, Leblanc DJ et al (2007) Genetic diversity among *Enterococcus faecalis*. PLoS One 2:e582
40. Dai D, Wang H, Xu X (2018) The emergence of multi-resistant *Enterococcus faecalis* clonal complex, CC4, causing nosocomial infections. J Med Microbiol 67:1069–1077
41. Kuo AJ, Shu JC, Liu TP et al (2018) Vancomycin-resistant *Enterococcus faecium* at a university hospital in Taiwan, 2002–2015: fluctuation of genetic populations and emergence of a new structure type of the Tn*1546-like* element. J Microbiol Immunol Infect 51:821–828

42. Murray BE, Singh KV, Heath JD et al (1990) Comparison of genomic DNAs of different enterococcal isolates using restriction endonucleases with infrequent recognition sites. J Clin Microbiol 28:2059–2063
43. Mondino SSB, Castro ACD, Mondino PJJ et al (2003) Phenotypic and genotypic characterization of clinical and intestinal enterococci isolated from inpatients and outpatients in two Brazilian hospitals. Microb Drug Resist 9:167–174
44. Top J, Banga NM, Hayes R et al (2008) Comparison of multiple-locus variable-number tandem repeat analysis and pulsed-field gel electrophoresis in a setting of polyclonal endemicity of vancomycin-resistant *Enterococcus faecium*. Clin Microbiol Infect 14:363–369
45. Wardal E, Markowska K, Zabicka D et al (2014) Molecular analysis of *vanA* outbreak of *Enterococcus faecium* in two Warsaw hospitals: the importance of mobile genetic elements. Biomed Res Int 2014:575367
46. Freitas AR, Tedim AP, Francia MV et al (2016) Multilevel population genetic analysis of *vanA* and *vanB Enterococcus faecium* causing nosocomial outbreaks in 27 countries (1986–2012). J Antimicrob Chemother 71:3351–3366
47. Rangberg A, Larsen AL, Kacelnik O, Sæther HS, Bjørland M, Ringstad J, Jonassen CM (2019) Molecular analysis and epidemiological typing of vancomycin-resistant *Enterococcus* outbreak strains. Sci Rep 9:11917
48. Kawalec M, Gniadkowski M, Hryniewicz W (2000) Outbreak of vancomycin-resistant enterococci in a hospital in Gdansk, Poland, due to horizontal transfer of different Tn*1546*-like transposon variants and clonal spread of several strains. J Clin Microbiol 38:3317–3322
49. Tenover FC, Arbeit R, Goering RV et al (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33:2233–2239
50. Homan WL, Tribe D, Poznanski S et al (2002) Multilocus sequence typing scheme for *Enterococcus faecium*. J Clin Microbiol 40:1963–1971
51. Ruiz-Garbajosa P, Bonten MJM, Robinson DA et al (2006) A multilocus sequence typing scheme for *Enterococcus faecalis* reveals hospital-adapted genetic complexes in a background of high rates of recombination. J Clin Microbiol 44:2220–2228
52. Raven KE, Reuter S, Reynolds R et al (2016) A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland. Genome Res 26:1388–1396
53. van Hal SJ, Espedido BA, Coombs GW et al (2017) Polyclonal emergence of *vanA* vancomycin-resistant *Enterococcus faecium* in Australia. J Antimicrob Chemother 72:998–1001
54. Tedim AP, Ruiz-Garbajosa P, Corander J et al (2015) Population biology of intestinal *Enterococcus* isolates from hospitalized and nonhospitalized individuals in different age groups. Appl Environ Microbiol 81:1820–1831
55. Freitas AR, Sousa C, Novais C et al (2017) Rapid detection of high-risk *Enterococcus faecium* clones by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Diagn Microbiol Infect Dis 87:299–307
56. Schlebusch S, Price GR, Gallagher RL et al (2017) MALDI-TOF MS meets WGS in a VRE outbreak investigation. Eur J Clin Microbiol Infect Dis 36:495–499
57. Savas S, Hazirolan G, Karagoz A et al (2018) From days to hours: can MALDI-TOF MS system replace both conventional and molecular typing methods with new cut off level for vancomycin resistant *Enterococcus faecium*. J Microbiol Methods 162:62–68
58. Titze-de-Almeida R, Willems RJ, Top J et al (2004) Multilocus variable-number tandem- repeat polymorphism among Brazilian *Enterococcus faecalis* strains. J Clin Microbiol 42:4879–4881
59. Top J, Schouls LM, Bonten MJ et al (2004) Multiple-locus variable-number tandem repeat analysis, a novel typing scheme to study the genetic relatedness and epidemiology of *Enterococcus faecium* isolates. J Clin Microbiol 42:4503–4511
60. de Been M, Pinholt M, Top J et al (2015) Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. J Clin Microbiol 53:3788–3797

61. Lytsy B, Engstrand L, Gustafsson Å et al (2017) Time to review the gold standard for genotyping vancomycin-resistant enterococci in epidemiology: comparing whole-genome sequencing with PFGE and MLST in three suspected outbreaks in Sweden during 2013–2015. Infect Genet Evol 54:74–80
62. Pinholt M, Gumpert H, Bayliss S et al (2017) Genomic analysis of 495 vancomycin-resistant *Enterococcus faecium* reveals broad dissemination of a *vanA* plasmid in more than 19 clones from Copenhagen, Denmark. J Antimicrob Chemother 72:40–47
63. Neumann B, Prior K, Bender JK et al (2019) A core genome multilocus sequence typing (cgMLST) scheme for *Enterococcus faecalis*. J Clin Microbiol 57(3):pii: e01686-18

Pseudomonas aeruginosa

8

Junyan Liu, Ruirui Xu, Zerong Lu, Guangchao Yu, and Zhenbo Xu

8.1 Phenotyping Methods for *P. aeruginosa*

Serotyping is based on an agglutination reaction between bacterial lipopolysaccharide (LPS) and type-specific rabbit antisera. In 1983, an International Antigenic Typing System (IATS) was developed, establishing the existence of at least 17

J. Liu · R. Xu · Z. Lu
School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

G. Yu
Center of Clinical Laboratory Medicine, First Affiliated Hospital of Jinan University, Guangzhou, China

Z. Xu (✉)
School of Food Science and Engineering, Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Engineering Research Center of Starch and Vegetable Protein Processing Ministry of Education, South China University of Technology, Guangzhou, China

Home Economics Technology, Rajamangala University of Technology Phra Nakhon, Bangkok, Thailand

Research Institute for Food Nutrition and Human Health, Guangzhou, China

Department of Laboratory Medicine, the Second Affiliated Hospital of Shantou University Medical College, Shantou, China

Research Center of Translational Medicine, Second Affiliated Hospital of Shantou University Medical College, Shantou, China

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD, USA

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_8

major heat-stable somatic antigens, and 3 new antigens have been recently added to the existing groups. This serotyping scheme characterizes most strains of P. aeruginosa possessing smooth LPS, although it shows poor discrimination against rough LPS isolates which are common in patients with cystic fibrosis.

Seventy-four isolates from 3 different hospitals and 18 reference strains were studied by Hernández et al. [1] Serotyping provided a good index of discrimination, although 11 isolates could not be serotyped. Eleven strains (12%) could not be assigned to any international serotype and were coded as autoagglutinable (one strain), non-agglutinable (three strains), or polyagglutinable (seven strains). The discrimination index of serotyping analysis referring to clinical isolates was 0.886. Overall, the predominant serotypes were O11 and O1, representing 24% and 19%, respectively, of serotypable strains. Three serotypes (O14, O15, and O17) were not found among the strains isolated from infected patients. The type strain of *P. aeruginosa* (NCTC 10332) was serotype O6. No association was observed between the strain serotype and the hospital.

Among the strains isolated in the hospital, O11 is the most common serotype. Serotyping of 88 non-repetitive clinical isolates is determined by the slide agglutination technique using specific antisera, polyvalent and monovalent. The results of serotyping showed that serotype O11 was the most common ($N = 14$, $P = 16\%$) followed by serotype O7 ($N = 11$, $P = 12.5\%$) and serotype O2 ($N = 10$, $P = 11.36\%$) [2]. In addition, During December 2013–December 2014, 229 *P. aeruginosa* were isolated from infected patients in the clinics of UCCK from a variety of clinical sites. Eighty isolates were studied as primary isolates from the group of infected patients. Serotyping resulted in the detection of eight serogroups. The most prevalent were O11 (65%) and O1 (17.5%). Other serotypes all with a prevalence of less than five were found in 17.5% of isolates: O4, O6, O9 were found in 3 (3.8%) samples each, O12 and O3 in 2 (2.57%) samples each, and O7 in 1(1.2%). No other serotypes were detected. O11 serotype was distributed in almost all clinics, but it was more common in the ICU. A greater diversity of serotypes was observed in the Pulmonology clinic where from six samples received during the study, 2 (O1), 1(O11), 1(O3), 1(O7), and 1(O9) were discovered [3].

Pyocin typing has been used extensively to trace the routes of infections within hospitals. Poor reproducibility has been of some, although lesser, concern in the case of pyocin typing. In the Gillies and Govan method of pyocin typing for *P. Aeruginosa*, a cross-streaking technique was used, and 105 main types and 25 subtypes were identified by the patterns of inhibition observed on 13 indicator strains [3]. The disadvantages of the technique included the need to remove the test strain growth before the application of the indicator strains, the 48-h period needed to obtain a result, and the inability to reliably type mucoid *P. aeruginosa*. The same 13 indicator strains which are already used internationally have been utilized in a revised technique by Janet et al. [4]. Test strains were rapidly applied to the surface of agar plates with a multiple inoculator. After incubation for 6 h and exposure to chloroform, the indicator strains were applied in agar overlays without prior removal of the test strain growth. After 18 h of incubation, the pyocin type was recognized by the inhibition of particular indicator strains. Additionally, the activity of

particulate (R and F) and nonparticulate (S) pyocins could be distinguished on the basis of inhibition zone size, which thus allowed further discrimination. The revised technique allows typing within 24 h, increases the number of identifiable types, and can be used to type mucoid strains.

One hundred and twelve clinical isolates of *P. aeruginosa* were determined by pyocin typing. Four pyocin types (1, 10, 3, and 5) dominated (respective frequencies: 56%, 15%, 12%, and 6%). The reproducibility of pyocin typing was distinctly inferior to both plasmid profiling and serotyping [5]. In addition, the typing method, which determined the pyocin activity of clinical isolates of *P. aeruginosa* on 27 indicator strains, was 43.7% reproducible, but the elimination of 9 indicator strains doubled the reproducibility and yielded more readable pyocin inhibition zones. Seventy-eight of 1084 isolates (7.2%) were untypable [6].

8.2 Restriction-Based Methods

8.2.1 Pulsed-Field Gel Electrophoresis (PFGE)

During the last decade, traditional phenotypic typing methods for epidemiologic and outbreak studies have been replaced by molecular methods. Pulsed-field gel electrophoresis (PFGE) typing is one of the most useful discriminating methods to type Pseudomonas spp. PFGE is considered as the "gold standard" typing method, due to its excellent discriminatory power and high epidemiological concordance [7]. PFGE is a technique used by scientists to generate a DNA fingerprint for a bacterial isolate. As the large DNA particles easily break, also due to their great viscosity, pipetting of them is difficult. For this reason, DNA and microrganisms are embedded in agarose gel and affected with restriction endonuclease enzymes.

Macrorestriction analysis by PFGE of DNA was performed according to USA Centers for Disease Control and Prevention's (CDC) highly standardized PFGE protocols for Gram-negative rods with some minor modifications [8]. Bacterial suspensions were prepared from individual bacterial colonies directly obtained from cultures incubated overnight on Mueller-Hinton agar. The suspensions were adjusted to a concentration of 10^9 CFU/mL, which is equal to 1:1.5 NTU in ethylenediaminetetraacetic acid (EDTA)-saline buffer (75 mmol/L NaCl and 25 mmol/L EDTA, pH 7.5). The cell suspension was mixed with an equal volume of 1% low-melting point seaKem Gold Agarose and was allowed to solidify in a 100 μL plug mold. The agarose plug was incubated for 24 h at 37 °C in 500 μL of lysis buffer (6 mmol/L Tris-HCl (pH 7.6), 0.1 mol/L EDTA, 1 mol/L NaCl, 0.5% Brij®58 (polyoxyethylene (20) cetyl ether; Sigma), 0.4% sodium deoxycholate, 0.5% sodium lauryl sarcosine, and 1 mg/mL lysozyme). Next, the lysis buffer was replaced with 500 μL of proteinase K buffer (1% sodium lauryl sarcosine, 0.5 mol/L EDTA (pH 9), and proteinase K (50 μg/mL; Sigma)), and this solution was incubated with gentle shaking at 50 °C for 20 h. The plugs were then washed four times for 30 min at 37 °C with 10 mL of Tris-EDTA buffer (10 mmol/L Tris-HCl (pH 8) and 1 mmol/L EDTA). One-third of a slice of each plug was cut and incubated for 18–20 h with

30 U of SpeI in the restriction buffer (Promega Buffer). DNA restriction fragments were separated by PFGE by using a CHEF DR III apparatus (Bio-Rad, Richmond, CA, USA) at 14 °C, 6 V/cm, for 20 h, with a time switch of 2–40 s. A Salmonella serotype Branderup strain (H9812) ladder (Bio-Rad Laboratories) restricted with XbaI was used as a universal size marker [9]. The gel was stained with ethidium bromide and visualized with the Gel-Doc system (Bio-Rad Laboratories). According to the criteria by Tenover et al. [10] isolates were considered to be genetically indistinguishable or identical if the restriction fragments had the same number of bands and the corresponding bands were with identical apparent size.

Currently, the SpeI enzyme usually produces 14–25 bands [11, 12] for each strain, and up to 37 [13, 14]. Reproducibility using the Spe I enzyme was 100%, and typeability ranged between 95 and 100% [13, 15]. In different studies, PFGE-Spe I was determined to have DIs between 0.98 and 0.998 [13, 15]. For some strains that cannot be typed, most of them are caused by DNA self-degradation. Studies have shown that this phenomenon can be eliminated by using HEPES buffer instead of Tris or adding 50 mM thiourea to the gel buffer to remove reactive Tris free radicals. An epidemiological typing study on the outbreak of *P. aeruginosa* using PFGE-Spe I showed that there are clones where the spatial and temporal distribution of isolates does not allow more than three band differences to be obtained in the PFGE files of different isolates. The correlation can therefore be suggested for transmission. Variations in the four to six bands seem to rule out direct transmission, but it is believed that clonal variations that may infect the same lineage [16, 17]. Studies have shown that DNA fingerprint variability in *P. aeruginosa* strains derived from macroscopic restriction analysis is usually the result of insertion/deletion rather than single nucleotide polymorphism (SNP). During evolution, large DNA fragments (gene islands) can be excised from the *P. aeruginosa* genome. The differences between the Spe I PFGE patterns were observed between different strains sharing the same SNP profile. This indicates that the core genome of *P. aeruginosa* is highly conserved, and the rate of change in the auxiliary DNA fragments is higher than that of the conserved core genome [18].

Between April 2011 and December 2016, samples of lung, liver, and spleen were collected from mink with this disease on 11 mink farms in 5 Chinese provinces. From these samples, 98 isolates of *P. aeruginosa* were obtained, which belonged to 5 serotypes: G ($n = 58$), I ($n = 15$), C ($n = 8$), M ($n = 5$), and B ($n = 2$); 10 isolates were not typable (10/98). More than 90% of the isolates formed biofilms, and 85% produced slime. All 98 isolates were resistant to 10 antibiotics (oxacillin, ampicillin, penicillin G, amoxicillin, ceftriaxone, cefazolin, cefaclor, tilmicosin, tildipirosin, and sulfonamide). However, almost all were susceptible to gentamicin, polymyxin B, and amikacin. Fifty-six unique genotypes by PFGE were identified. These findings have revealed genetic diversity and high antimicrobial resistance in *P. aeruginosa* isolated from mink with hemorrhagic pneumonia and will facilitate the prevention and control of this disease [19].

PFGE has been used as a molecular typing method in many studies to study multidrug resistance *P. aeruginosa* carrying multiple resistance determinants such as metal b-lactamases (MBL) or broad-spectrum b-lactamases (ESBLs) Clusters. A

molecular epidemiological examination in 2005 showed that hospitals in three different towns in northwestern Hungary were involved in an outbreak caused by multi-drug-resistant O11 VIM-4 MBL *P. aeruginosa* [20]. PFGE is performed according to the method described by Poh et al. After some modifications [21, 22], the genomic DNA insert was digested with 20 U Spe I enzyme at 37 °C for 2.5 h. Electrophoresis was performed in CHEF-DRII equipment (Bio-Rad). The DNA fingerprints were compared by Fingerprinting II Informatix software (Bio-Rad). The criticality of the PFGE genotype identified by using 1% band position tolerance and Dice coefficient was 80%. PFGE-Spe I analysis showed that isolates of the outbreak clones were obtained from three different intensive care units (ICU) in three different towns within 6 months, and showed ≥95% similarity by Dice coefficient. A VIM-4-producing *P. aeruginosa* carrier patient transferred between the two ICUs was also found to provide an epidemiological link between the two. The VIM-4 positive isolate from Huihui is an outbreak clone, indicating that the outbreak may play a depot role in the hospital environment in addition to spreading between patients.

The protocols have been established for clinical isolates, where high typeability and discriminatory power was achieved. The literature about PFGE-typing of environmental isolates is very limited and the method was applied in a very low number of isolates derived mainly from nosocomial environments, where the typeability and discriminatory power of the method could not be evaluated. The importance of typing *P. aeruginosa* isolates deriving from water and wastewater samples is crucial for public health reasons; *P. aeruginosa* in bottled water can be considered a risk to profoundly immunocompromised patients. According to European regulations (C. D. 98/83/EC), *P. aeruginosa* should be absent in potable water. Moreover, the choice of pool and spa waters for medical use is increasing. A number of recent studies emphasize the high prevalence of *P. aeruginosa* in hospital water facilities resulting in outbreaks. Epidemiological investigation to determine the source of an outbreak requires fast and reliable methods [23–25].

Studies have shown that PFGE is very effective in typing *P. aeruginosa* and has a high degree of discrimination. However, this method has the characteristics of high labor intensity, time consumption, and high cost. In addition, due to the lack of a widely accepted standard protocol for *P. aeruginosa* typing, results from different laboratories are not easy to compare [26, 27].

8.2.2 Restriction Fragment Length Polymorphism (RFLP)

This was the first methodto be widely used for typing strains of P. aeruginosa from patients with CF. Genomic DNA is extracted from the bacterium of interest, digested with one of several restriction enzymes, and the DNA fragments are then separated by electrophoresis. A radiolabelled probe, directed to a specific portion of the bacterial genome, is then added, and a hybridization reaction is carried out. The most discriminatory probes are those which react with a hypervariable portion of the bacterial genome. Generally, the genomic DNA of *P. aeruginosa* is digested with

two or three restriction enzymes. The obtained DNA digest is then separated by electrophoresis, transferred to a membrane and then hybridized with a radiolabeled probe (*exoA*). Obtain the size and number of different DNA genomic fragments, and then create fingerprints for different strains. Using Bgl II, Sal I and Xho I restriction enzymes to classify RFLP-*exoA* has a discrimination ability (DI) of 0.97, which can distinguish different strains. However, about 5% of the strains cannot achieve the purpose of typing because they do not contain *exoA*. Iisolates with different LPS serotypes and biotypes showed the same RFLP pattern, indicating that the phenotypic variation of *P. aeruginosa* is not necessarily the result of genetic heterogeneity. On the other hand, RFLP-*exoA* assigns serotypes, biotypes, and anti-bioGram indistinguishable isolates from unrelated patients as different types [26].

Although the type of *exoA* can also indicate the clonal relevance of the strain, the technique is proposed to distinguish infections caused by variants of the same clonal lineage, which persist in different geographic locations [17, 28]. At present, RFLP typing has been replaced by pulsed-field gel electrophoresis and random amplified of polymorphic DNA.

8.2.3 Ribotyping

Ribosomal RNA genes are the most conservative, and there are multiple copies on the genome. Using rRNA gene fragments as probes, the differences in the position and number of rRNA genes were detected, and the strains were typed. Generally, DNA restriction fragments are separated by gel electrophoresis, transferred to a membrane, and the conserved region of the rRNA gene is incubated with the probe. Ribotyping of *P. aeruginosa* uses Pvu II restriction endonucleases and rDNA gene probes with low discrimination. Pitt and his colleagues first studied the multi-resistant *P. aeruginosa* strain 012 from Europe using ribotyping. Their results are consistent with outer membrane protein electrophoresis, LPS analysis, and esterase typing results, indicating that the common origin of these strains implies the common origin of these strains [17, 29, 30].

In addition, when Pvu II enzyme and RiboPrinter are used together, DI values can reach 0.88 and 0.93 [31, 32]. The RiboPrinter™ microbial identification system (E.I. duPont de Nemours and Company) is an automated ribotyping system widely used in the past two decades. This method standardizes the technical and interpretive aspects of the program and also uses computer databases to compare products from a large number of isolates. Pvu II was selected for the analysis of *P. aeruginosa* isolates by automated ribotyping in a molecular surveillance study of European quinolone-resistant clinical isolates of *P. aeruginosa* using the RiboPrinter system [33]. The automated ribotyping method has excellent repeatability, Typeability, and high capacity. This method seems to be a convenient way to quickly identify and compare bacterial clones that are ubiquitous in distant geographic areas and time points.

Automated ribotyping was used to investigate the clonal diversity of the 56 *P. aeruginosa* isolates. The 56 clinical isolates of *P. aeruginosa* from 44 patients

warded in the intensive care unit were obtained from Aga Khan Hospital from March 1998 to March 2001. Automated ribotyping indicated that the clinical isolates were clonally related and could be clustered into four major ribogroups based on their similarity index, with ribogroup II being the dominant one. The *P. aeruginosa* isolates in ribogroup II were correlated with their antibiotic resistance pattern, and interestingly, there seemed to be a gradual acquisition of multiple antibiotic resistance associated with the isolates within this group over time [34].

The ribotyping method has low discrimination. Seventy-eight confirmed nosocomial Pseudomonus aeroginosa out of 1520 different sample (nosocomial infections & environmental samples) were collected during a period of 1 year. Six typing methods were evaluated, utilizing the confirmed 78 Pseudomonas strains, to assess their usefulness as tools to study the bacterial diversity. The methods used were antibiogram, pyocin typing, serotyping, extracellular enzyme typing, automated ribotyping, and PFGE. PFGE yielded 56 distinct types of *P. aeruginosa* with 100% distinction capacity (78/78) as all the strains were typable. Compared to PFGE, the distinctive capacities were 88.5% (69/78) for serotyping, 91% (71/78) for pyocin typing and 100% (78/78) for automated ribotyping analysis. The results obtained in PFGE were the easiest to read and interpret and most discriminating (0.99), followed by the pyocin typing (0.96), whereas ribotyping had (0.90) discriminatory power [35].

8.3 Amplified-Based Methods

8.3.1 Random Amplified of Polymorphic DNA (RAPD)

The random amplified polymorphic DNA genotyping (RAPD) technology generated by the random primer polymerase chain reaction was established in 1990 by Welsh et al. This technology uses short arbitrary sequence primers to perform PCR under non-strict conditions. As a result, many sites of genomic DNA were amplified simultaneously. Because the stringency of the annealing reaction is reduced, a single primer can be used for amplification simultaneously. The same primer can perform extension reactions at many sites on the template DNA. The length polymorphism of multiple sites can be detected by the presence or absence of bands at different migration positions in the gel. Using RAPD technology for nucleic acid analysis of *P. aeruginosa* can not only be used for genetics and epidemiology of *P. aeruginosa*.

Crude bacterial lysates were prepared by suspending a 1 µl loopful of bacteria in 20 µl of 50 mM NaOH-0.25% sodium dodecyl sulfate (SDS) and heating for 15 min at 95 °C. Lysates were diluted with 980 µl of water, and 2.5 µl was used for amplification in a 25 µL PCR mixture. PCR tubes further contained 0.5 U of DNA polymerase, a 400 µM concentration of each deoxynucleoside triphosphate, primer, reaction buffer, and 2.5 mM $MgCl_2$. Amplification was performed with annealing at 52 °C, extension at 72 °C, and melting at 97 °C. The PCR products were mixed with 2.5 µl of gel loading buffer and electrophoresed on 2% agarose gel 1.0XTAE buffer

(TrisAcetate-EDTA) at 100 V for 2 h. Molecular size markers used were a 100 bp ladder. Isolates from each patient typed by RAPD analysis in a single PCR run were analyzed by photographing the gels and marking the position of the bands in order to facilitate the comparison of strains between patients. Strain differentiation was done by observing readily discernible band patterns. In most protocols, at least in the first few cycles of PCR, a lower annealing temperature is used, which allows imperfect hybridization at multiple random locations on the chromosome to amplify its random fragments. The PCR products were separated by agarose gel electrophoresis, stained, and analyzed by visual inspection or by calculation methods. Compared with PFGE and several other molecular typing technologies in many laboratories, this method is simpler, faster, lower cost, and more labor intensive.

RAPD typing was performed on 200 strains of *P. aeruginosa* by using 10 nucleotide primers. The results showed that the typeability was 100% and the intralaboratory reproducibility was 98.5% [36]. However, due to most aspects of the PCR program, including the temperature profile and source, and the small differences in different batches of Taq polymerase, it is possible to affect the repeatability of the band diagram. In different laboratories, the band patterns observed for the same *P. aeruginosa* isolates typed by the same RAPD method may show large differences. Under the same conditions, equipment and the same operator, the measurement is repeated, and the measurement is performed regularly rather than sporadically. The RAPD band pattern has the highest repeatability. Therefore, RAPD can only be used for laboratory comparison of *P. aeruginosa* strain collection [37].

Ten to twenty-five bands of 200–3000 bp for the *P. aeruginosa* isolates were obtained when applying the method evaluated by Campbell and colleagues by the use of the 10-nucleotide primer 208, an Invitrogen Taq polymerase and 40 ng purified genomic DNA. The concentration of genomic DNA will not interfere with the experimental results, but the use of different Taq polymerases has a great influence on the quantity and intensity of PCR products. Among the types of Taq enzymes used, Invitrogen Taq polymerase obtains the best results (i.e., the maximum number and intensity of bands) [36, 38].

The RAPD classification can be assisted by using software. In a typing study on MDR *P. aeruginosa* clinical isolates in Hungary and other European countries, it was found that the epidemiologically related *P. aeruginosa* isolates showed genetic similarity>90%. MLST results show that gifts with more than 80% similarity belong to the same clonal complex. Although MLST type and RAPD type of *P. aeruginosa* are in correspondence, the genetic similarity is less than 80% [39, 40].

In the molecular epidemiology and colonization of *P. aeruginosa* in the burn department of Shahid Motahari Hospital in Tehran, Iran, RFLP and RAPD analysess were used to study 127 clinical and two environmental collected from January 2008 to June 2008. In RFLP, the PCR product of the 16S rRNA gene was restriction enzymes Alu I, Hae III, and Rsa I, and the resulting fragments were analyzed by agarose electrophoresis. Molecular typing by RFLP did not show discrimination against P. aeruginosa isolates, but RAPD-PCR showed eight different genotypes designated RAPD1 to RAPD8 in clinical and environmental isolates. RAPD1

is the main genotype of clinical (n = 64, 50.4%) and environmental isolates (n = 1, 50%). The results show that RAPD may have better typing and discriminating ability than RFLP in studying P. aeruginosa [41].

Due to the simplicity and high speed of RAPD, it can be used as the first screening for *P. aeruginosa* epidemic typing. Through RAPD screening, clonal relevance can be determined at a relatively low cost within 24 h.

8.3.2 Amplified Fragment Length Polymorphisms (AFLP)

AFLP is known as the third-generation molecular labeling technology after RFLP, RAPD, etc. The principle is to cut genomic DNA with two or more restriction enzymes to produce sticky ends to form restriction fragments of different sizes, and then connect the artificial short double junction to the stickiness of these fragments. At the end, a pre-amplification reaction and a selective amplification reaction are performed using the specific fragment with a linker as a template. The linker and several adjacent base sequences serve as the binding sites of the primer. Finally, only the digested fragments paired with the selected base can be amplified to achieve specific amplification [44]. AFLP analysis is a selective restriction fragment amplification technique in which adaptors are ligated to genomic restriction fragments, and then these fragments are PCR amplified using adaptor-specific primers. For AFLP analysis, only a limited amount of purified genomic DNA (50–100 ng) is required. Digested with two restriction enzymes, one has an average cutting frequency (such as Eco RI) and the other has a higher cutting frequency (such as MseI). After adaptor ligation and PCR amplification, a polyacrylamide gel electrophoresis of the PCR product can obtain a pattern of usually 40–200 bands [42].

The genotypic relatedness was assessed by using AFLP fingerprinting. Sixty-six *P. aeruginosa* isolates were obtained from sputa/deep-pharyngeal swabs from 27 CF patients belonging to 17 families. Twenty-three distinct genotypes of *P. aeruginosa* were identified. Eleven families each had one distinct genotype. In the other six families more than one genotype was observed; three families each showed two genotypes, two families each had three genotypes, and one family had four genotypes of *P. aeruginosa*. In several cases, siblings with CF from the same family harbored the same strain of *P. aeruginosa*, which were different from the genotypes in other families. On the other hand, there was an overlap in *P. aeruginosa* between closely related families. Some patients show persistent colonization with the same genotype of *P. aeruginosa* over the longitudinal period. The presence of the same genotypes in the siblings of the same family and closely related families suggests cross-transmission of *P. aeruginosa* or acquisition from common environmental exposure [43].

A comparative study of PFGE and AFLP methods on 22 *P. aeruginosa* isolates shows that AFLP has 100% typability, DI value is 0.97, and PFGE is 0.96. In this study, Eco RI and Mse I restriction enzymes were used in AFLP, and clusters of *P. aeruginosa* isolates related to epidemiology have a similar PFGE pattern, showing>90% identity source. On the other hand, using GelCompar software to assist in

fluorescence-labeled AFLP fingerprint analysis, PFGE with more than six epidemiologically unrelated strains with differences showed less than 90% homology [44].

8.3.3 Multilocus Variable-Number Tandem Repeat (VNTR) Analysis (MLVA)

The MLVA technique has been developed to typing bacteria such as Salmonella, Listeria monocytogenes, and Escherichia coli [45]. The basis of the MLVA technique is the identification of the variable number tandem repeat (VNTRs) in specific locuses on the genome of microorganisms [46]. In the MLVA technique after selecting the desired locus and designing the primer for them and extracting the desired strains DNA, the proliferation of the sequences containing the VNTR is performed by PCR. The product obtained from PCR is sequenced and the number of replicates is calculated. Vergnaud and colleagues developed the MLVA protocol for *P. aeruginosa*, which was subsequently improved by adding new epidemiological information markers. The MLVA scheme involves 15 loci with repeated tandem sequences (VNTR). The MLVA genotype of the *P. aeruginosa* isolate (MLVA15) with 15 VNTRs is expressed as an allele, showing the number of repeats for each analyzed VNTR. When a difference is observed on any of the 15 VNTRs, a new genotype number will be given. Pedigree is defined as a group of isolates with up to two different genotypes (VNTR). The total diversity index of 15 markers calculated from 190 isolates was 0.97. For a simpler and more reliable MLVA scheme, when only microsatellites ms142, ms211, ms212, ms213, ms214, ms215, ms216, ms217, ms222, and ms223 participate in the analysis, the number of characteristic VNTRs can be reduced to 10 [47, 48].

In the genetic typing for the number of bacterial strains isolated from patients using MLVA technique [49]. Seventy samples from different strains of *P. aeruginosa* were isolated from a wide range of patients, including those admitted to the ICU and CCU units to outpatients. Although these strains were similar in nature to the biochemical characteristics and analyzes, 39 strains were classified in the MLVA analysis. This number of strains obtained from *P. aeruginosa* strains indicates the high accuracy of this method in differentiating the differences in these strains. These differences, which led to the creation of different types in the MLVA method, showed that they were not detectable by conventional biochemical methods. Urinary tract infections (UTIs) are one of the three most common causes in the community, especially at the hospital level. Urinary tract infection caused by pathogenic bacteria after respiratory infection is the most common type of infection in patients admitted to different parts of hospitals [50]. Patients with UTI are very diverse, and a wide range of bacterial strains have been identified and reported for different characteristics and pathogens [51, 52]. Due to the high variation in the pathogenic strains that come from patients from different parts of the world, which often carry antibiotic resistance genes, researchers are seeking more practical and precise solutions to differentiate these close to each other strains. As most diagnostic methods require

high volume of clinical specimens, and are time consuming and not cost effective, techniques such as MLVA, have become more common to therapists. In recent years, MLVA-based typing, in which isolates are evaluated by the number of replicates in several genetic regions, for a number of important bacteria such as Bacillus anthraise, Staphylococcus aureus, Enterococcus faecium, Haemophilus influenzae, Bordetella pertosis, and many others has been used [53].

MLVA was used to investigate the source of *P. aeruginosa* infection in a pediatric CF center in Paris, France. Between January 2004 and December 2006, *P. aeruginosa* was detected in 46 children, 17 of whom had primary colonization. A total of 163 strains were recovered. After genotyping 15 VNTRs, a total of 39 lineages were observed, consisting of indistinguishable or closely related isolates. One of them corresponds to "Clone C," which is widely distributed in Europe. This shows that the MLVA genotype of *P. aeruginosa* strains recovered from individual patients proved to be stable over time, except for the occasional insertion of IS elements and the addition or deletion of repeated sequences in a single VNTR [47].

During 2004–2008, researchers collected 81 non-repetitive *P. aeruginosa* in two universities and two hospitals in Bulgaria. Determine the 100% repeatability and 97.5% typability of the MLVA scheme. In order to study the difference between MLVA, MLST, and PFGE in the classification of *P. aeruginosa*, 32 strains from sputum samples from CF patients in the Netherlands were typed using three methods. Only 9 of the first 15 VNTRs were analyzed. The DI of PFGE, MLVA, and MLST are 0.988, 0.980, and 0.952, respectively, with overlapping confidence intervals of 95%. There is a high degree of consistency between the three methods at the cloning cluster level. The authors underscored the advantages of MLVA and MLST in their portability and ease of interpretation, and a further advantage of MLVA over MLST was also highlighted in being more cost effective as it does not require sequencing [54]. However, Johansson et al. analyzed 232 isolates of *P. aeruginosa* isolated from cystic fibrosis patients by both methods. In this comparative study, 91% of the results were similar to each other. However, they emphasized that, despite the expensive and time-consuming PFGE, its accuracy is higher than that of MLVA [55].

Seventy *P. aeruginosa* isolates were collected from different hospitals located in Tehran city in 2018. After the amplification of the genes by PCR, electrophoresis on the agarose gel was performed for the products of each of the eight genes. The bonds produced on the gels representing the size of each of the sequences were analyzed by Gene Tools software by comparison with the 100 bp size marker (ZR 100 bp DNA Marker ™). After sequencing and examining the frequency of repetitions in 70 strains, 39 types were obtained. In the phylogeny tree, the relationship between different strains was based on the similarity in their repetitions in a branch. The length of each branch also shows the difference in the number of repetitions in different branches [56]. The MST (minimum spanning tree) pattern derived from the MLVA analysis for the desired strains. Seventy strains classified into 39 types are presented as a clone based on the number of strains that were categorized. In total, the MST pattern obtained in this study consists of 11 clonal complexes (CC). This concept is based on the relationship between the numbers of repetitions that

have been considered as comparisons in each category. The distance between each CC represents markers that are common in each clone. And, if the similarity between the indices was more, the distance between the clones is closer to each other. The MST pattern usually gives us a better understanding of the typing of strains that cause the creation of a common attribute, because it indicates the differences and common indicators as clonal complex simultaneously [57].

The potential disadvantage of the MLVA method is that even with a fluorescence detection system, it is difficult to accurately determine the size of the fragment because it depends on the mobility and the sequence composition. In addition, the evolution of repeated DNA sequences may be too fast, compromising epidemiological consistency. Although this method has rarely been used for epidemiological studies so far, MLVA may become a more widely used typing technology for *P. aeruginosa* in the future, especially if portable typing data is required between laboratories or intercontinental in regional or international studies [54].

8.4 Sequencing-Based Method

8.4.1 Multilocus Sequence Typing (MLST)

Multilocus sequence typing (MLST) is a strain-typing system that focuses exclusively on conserved housekeeping genes [58]. Though pulsed-field gel electrophoresis (PFGE) possesses higher discriminatory power, the lack of a universal standard and portability makes MLST more ideal for comparative analysis of strain types regardless of region or source. Moreover, in a comparative study of molecular techniques for typing *P. aeruginosa*, MLST had the greatest predictive value (100%) in labeling strains as unique [59]. The standardization of MLST has given rise to databases that enable comparative analysis of allele sequences and identification of unique sequence types.

The seven genes acsA, aroE, guaA, mutL, nuoD, ppsA, and trpE were selected according to the MLST scheme for *P. aeruginosa* created by Curran et al. [60]. The loci were originally chosen based upon biological role (e.g., a range of differing central housekeeping roles including mismatch repair, DNA replication, and amino acid biosynthesis), size (>600 bp), location (i.e., a minimum of 6 kbp upstream or downstream from known virulence factors, lysogenic phage, or insertion sequence elements), and suitability for nested primer design and sequence diversity.

During the PCR amplification of the target sequence, some *P. aeruginosa* isolates may require some modifications to the PCR program. These modifications include the addition of pure dimethyl sulfoxide at a concentration of 5 ml per 100 ml of the PCR master mix and increasing the annealing temperature from 55 °C to 58 °C [37]. In addition, in some cases, in order to generate clear and unambiguous DNA sequence data to search in the MLST database, it is also possible to perform additional sequencing reactions on the purified PCR fragments using amplification primers. The polymerase chain reaction products were sequenced and submitted to the *P. aeruginosa* MLST database (https://pubmlst.org/paeruginosa/) for the

assignment of allelic numbers. Each isolate was then assigned a sequence type (ST) based on the combination of seven allelic numbers. Isolates with the same ST are considered as members of the same clone. Up to now, there are 3391 ST types in the database.

A comparative typing study of 90 *P. aeruginosa* isolates obtained from swab cultures around the rectum in patients in the intensive care unit showed that SpeI digestion of the 90 *P. aeruginosa* isolates from different patients identified 85 unique patterns and MLST revealed 60 different STs. Among the 60 STs, 36 were not previously submitted to the *P. aeruginosa* database. Although both PFGE and MLST have a high discriminating ability (DI values of 0.999 and 0.975, respectively), PFGE has greater discriminating power than MLST [61]. Because the results of MLST are highly reproducible and easy to compare between different laboratories, the interpretation of the data is clear, and it is used to determine the cloning relationship between bacterial strains that differ greatly in time and geography and genetic diversity is particularly useful. Therefore, MLST is a key epidemiological tool for studying regional and global epidemiology of multidrug-resistant *P. aeruginosa*. Studies have shown that *P. aeruginosa*, which is popular in Europe when carrying bla_{VIM} MBL and bla_{PER-1} ESBL, mainly contains two clonal complexes of serotype O11 and O12 isolates [39, 62–64].

In order to identify the major multidrug-resistant hospital clones of *P. aeruginosa*, many molecular typing studies were conducted at the National Epidemiology Center in Budapest, Hungary. The research object was a total of 1500 strains of *P. aeruginosa* between 2003 and 2008. The clinical isolates were screened for MBL and ESBL production. The selected isolates were also used for the overexpression of AmpC b-lactamase and the presence of aminoglycoside resistance determinants carried by integrins. A variety of typing methods have been performed on the representative MDR *P. aeruginosa*, including serotyping, RAPD and MLST. The results of the study indicate the key role of four different *P. aeruginosa* clonal complexes (as determined by MLST) in the emergence of MDR isolates. These four clonal complexes also seem to have a wide geographical distribution outside Hungary, but the acquired resistance determinants may show a high degree of variability among isolates from different geographic sources [39, 40, 64, 65]. The first complex (CC4) is characterized by a serotype of O12 and the founder sequence type ST111, which corresponds to the major multi-resistant P12 clone in Europe. In addition, some isolates of this clonal complex in Hungary are associated with the spread of VIM-4 MBL. The second complex (CC11) is characterized by serotype O11 (the founder sequence type ST235) and contributes to the spread of VIM-4 MBL and PER-1 ESBL. The remaining two clonal complexes distributed throughout the country are characterized by serotypes of O4 and O6, sequence types of ST175 and ST395, and contain isolates that overproduce the chromosomal AmpC b-lactamase and carry Integra aadB (aminoglycoside 2′-O-adenylate transferase) gene. In addition, observing the changes based on the resistance gene complexes CC4 and CC11 carried in different countries and hospitals, multiple independent concepts have proposed the acquisition of these two resistance factors of these two universal clones of *P.*

aeruginosa, which seem to be particularly adept at acquiring resistance determinants that result in an MDR phenotype [37, 39, 66, 67].

One hundred and sixty *P. aeruginosa* strains were isolated from a hospital in China. Multilocus sequence typing analysis demonstrated that these isolates were highly diverse; 68 sequence types were identified, of which 28 were novel sequence types. Polygenic and eBURST analysis demonstrated genetically similar clones with dissimilar resistance profiles. The results showed that 68 STs were segregated into 11 CCs because of the sequence identity among five or more alleles. The largest CC consisted of eight STs (ST2378, ST2372, ST274, ST2383, ST2370, ST2373, ST2405, and ST209, with ST209 being the primary founder). The second CC consisted of four STs (ST244, ST2374, ST2371, and ST597, with ST597 being the primary founder). The nine other groups contained two or three STs. Twenty-six STs were classified as singletons. The MLST tree revealed a high genetic diversity in those isolates. The analysis revealed a weak bootstrapping value, especially in major branches. The phylogenetic tree shows that ST2378, ST2372, ST274, ST2383, ST2370, ST2373, ST2405, and ST209 were clustered together. Several close clusters were identified; these were also previously obtained by the eBURST algorithm. In addition, the results showed that the MLST STs and antimicrobial resistance profiles were not correlated in this study. Isolates with the same ST did not show a unique resistance profile pattern. This demonstrated that there is no definitive link between the ST of isolates and their resistance to these 14 antimicrobial agents. Mapping of resistance profile data onto the eBURST analysis data and the phylogenetic tree revealed the following: (1) the most similar resistance profiles did not cluster together and (2) isolates with the same STs did not share similar resistance phenotypes. Taken together, this shows that these strains displayed a relatively high degree of genetic variability, demonstrating that antibiotic resistance was most likely determined by individual genetic combinations.

A total of 2818 *P. aeruginosa* isolates were collected in 2010 from 65 hospitals in 22 regions of China. Susceptibilities to 16 antimicrobial agents were evaluated by the disk diffusion method. Since carbapenems and ceftazidime are currently the most widely used effective anti-Pseudomonas drugs, we selected 896 imipenem-, meropenem-, or ceftazidime-nonsusceptible isolates for a further MLST study to investigate the clonal relationships among drug-nonsusceptible isolates. Of the 896 isolates, 632 belonged to 116 known STs, 201 demonstrated 104 new STs (new combinations of known alleles), and 63 others belonged to 34 new STs (STs contain novel alleles of certain genes). The regional distribution of the top 10 STs revealed the geographic dispersion of STs. The five of the top 10 STs were distributed in more than 10 regions. ST274, ST244, ST235, ST277, and ST357 were found in 16, 13, 11, 10, and 10 regions, respectively. This suggested that they were more common in China. To investigate the clonal relationships of 896 isolates, BioNumerics was used to create a minimum spanning tree and cluster STs into CCs. Although the whole population was nonclonal, there were several large CCs, which meant that the population was partially clonal, and some of those CCs contained globally spread STs and were related to local outbreaks of *P. aeruginosa* infections, such as ST235,

ST244, ST357, etc. Spain reported two outbreaks of *P. aeruginosa* in 2007 and 2008, one of which was caused by ST235 in a hematology department. South Korea also reported the dissemination of multidrug-resistant *P. aeruginosa* belonging to ST235 [68]. ST244 and ST235 were responsible for an outbreak of infections with *P. aeruginosa* that carried the PER-1 β-lactamase in Poland [69]. ST357 producing the IMP-7 metallo-β-lactamase has been reported in Singapore, and it also spread in the Czech-Polish border region [70]. In addition, the relatedness between STs and numbers of isolates was evaluated by generating a correlation curve. A marked linear relationship between ST categories and numbers of isolates was observed in the correlation curve, which meant that the number of ST categories increased with the addition of isolates.

To investigate the drug resistance and genetic background of *P. aeruginosa* at Shaanxi Provincial People's Hospital between July 2016 and January 2017. Sixty-eight STs were segregated into 11 CCs because of the sequence identity among five or more alleles. The largest CC consisted of eight STs (ST2378, ST2372, ST274, ST2383, ST2370, ST2373, ST2405, and ST209, with ST209 being the primary founder). The second CC consisted of four STs (ST244, ST2374, ST2371, and ST597, with ST597 being the primary founder). The nine other groups contained two or three STs. Twenty-six STs were classified as singletons. The MLST tree revealed a high genetic diversity in those isolates. The analysis revealed a weak bootstrapping value, especially in major branches. The phylogenetic tree shows that ST2378, ST2372, ST274, ST2383, ST2370, ST2373, ST2405, and ST209 were clustered together. Several close clusters were identified; these were also previously obtained by the eBURST algorithm. In addition, the results showed that the MLST STs and antimicrobial resistance profiles were not correlated in this study. Isolates with the same ST did not show a unique resistance profile pattern. This demonstrated that there is no definitive link between the ST of isolates and their resistance to these 14 antimicrobial agents. Mapping of resistance profile data onto the eBURST analysis data and the phylogenetic tree revealed that the most similar resistance profiles did not cluster together and isolates with the same STs did not share similar resistance phenotypes. In conclusion, this shows that these strains displayed a relatively high degree of genetic variability, demonstrating that antibiotic resistance was most likely determined by individual genetic combinations [63, 71].

8.4.2 Double-Locus Sequence Typing (DLST)

It has recently been shown in Staphylococcus aureus that, by sequencing small regions (ca. 500 bp) of only two highly variable loci (double-locus sequence typing, DLST), it is possible to investigate the epidemiology of this pathogen [72–74]. Similar to other sequence-based methods, it gave unambiguous definition of types, allowing inter-laboratory comparisons and high reproducibility. Moreover, the possibility to work with batches of 96 isolates allowed a reduction of costs and working time. Consequently, this method can be easily incorporated into long-term routine

surveillance programs. An efficient sequence-based typing scheme similar to the DLST scheme of *S. aureus* to investigate the local epidemiology of *P. aeruginosa* was set up by Basset et al. in 2013.

To identify potentially highly variable loci in the *P. aeruginosa* genome, eleven loci (ms142, ms172, ms173, ms194, ms207, ms214, ms215, ms217, ms222, ms223, oprD) were tested on a subset of isolates. Among these, only three loci (i.e., ms172, ms217, oprD) showed a product size that was larger than 300 bp for all isolates. Therefore, only these loci were selected for further analyses.

Single-strand sequencing of three highly variable loci (ms172, ms217, and oprD) was performed on a collection of 282 isolates recovered between 1994 and 2007 (from patients and the environment). As expected, the resolution of each locus alone (number of types (NT) = 35–64; index of discrimination (ID) = 0.816–0.964) was lower than the combination of two loci (NT = 78–97; ID = 0.966–0.971). As each pairwise combination of loci gave similar results, the most robust combination with ms172 [reverse; R] and ms217 [R] to constitute the double-locus sequence typing (DLST) scheme for *P. aeruginosa* was selected. This combination gave: (i) a complete genotype for 276/282 isolates (typability of 98%), (ii) 86 different types, and (iii) an ID of 0.968. Analysis of multiple isolates from the same patients or taps showed that DLST genotypes are generally stable over a period of several months. The high typability, discriminatory power, and ease of use of the proposed DLST scheme make it a method of choice for local epidemiological analyses of *P. aeruginosa*. Moreover, an Internet database (http://www.dlst.org) was developed to give an unambiguous definition of DLST types.

Reliable molecular typing methods are necessary to investigate the epidemiology of bacterial pathogens. Reference methods such as MLST and PFGE are costly and time-consuming. Cholley et al. compared the DLST method for *P. aeruginosa* to MLST and PFGE on a collection of 281 isolates. DLST was as discriminatory as MLST and was able to recognize “high-risk” epidemic clones. Both methods were highly congruent [75].

An increase in *P. aeruginosa* incidence was observed in the ICUs of the Lausanne University Hospital between 2010 and 2014 [76]. One hundred fifty three isolates retrieved during this period were typed with double locus sequence typing (DLST), which detected the presence of three major genotypes: DLST 1-18, DLST 1-21, and DLST 6-7. DLST 1-18 (ST1076) isolates were previously associated with an epidemiologically well-described outbreak in the burn unit. Nevertheless, DLST 1-21 (ST253) and DLST 6-7 (ST17) showed sporadic occurrence with only few cases of possible transmission between patients. In addition, the comparison of DLST and MLST showed that all DLST 1-18 isolates belonged to ST1076, DLST 1-21 to ST253, and DLST 6-7 to ST17, except for one DLST 6-7 isolate, which was found to be of ST845, a single-locus variant from ST17 at the nuoD locus. This confirms the previously documented congruence between both methods [77].

8.4.3 Whole Genome Sequencing (WGS)

In order to understand the colonization and infection pathways, a strong typing method is needed to study the correlation between strains. Although the PGGE method currently commonly used is the gold standard, it has low reproducibility between different laboratories and is not suitable for large-scale research. At present, methods based on sequence analysis such as MLST and DLST have shown advantages. MLST has shown high efficiency in studying the overall population structure of *P. aeruginosa*. DLST has been successfully used to study the epidemiology of Staphylococcus aureus and *P. aeruginosa*. At present, with the continuous development of new technologies, in the hospital environment, the latest research in *P. aeruginosa* evolution research and epidemiological research has used whole genome sequencing (WGS) [73, 76].

The emergence of high-throughput methods has promoted WGS, bringing hope for the separation of single base pairs between isolates, making it the ultimate molecular typing method for bacteria. Analysis of single nucleotide polymorphisms (SNPs) in the bacterial genome provides a method to determine the correlation between isolates that are epidemiologically linked and to track the evolution of bacteria over months to years. High-throughput sequencing analysis was performed on the five isolates that broke out in UH-NHST-CC hospital. Although these five strains of *P. aeruginosa* belong to the same PFGE and RAPD lineage, the genomic sequence data obtained by high-throughput sequencing technology shows that the isolate PANOTK11 is an outlier compared to the other four isolates and does not belong to the same outbreak pedigree. PANOTK11 has a 48 kb sequence, which is not present in the other four strains. This area is assembled into a contig and annotated as containing 24 CDS. Furthermore, the single nucleotide polymorphism (SNP) between the genomic sequence data showed the key single-base differences accumulated during the outbreak process, thus providing an in-depth understanding of the evolution of the outbreak strain. Differential SNPs have been found in various genes, including lasR, nrdG, tadZ, and algB. The rate of these generations is estimated to be one SNP every 4–5 months. In conclusion, this study proves that single-base resolution of whole-genome sequencing is one of the powerful tools for the analysis of outbreak isolates. It shows the similarity of strains, and evolve over time passing through the gene sequence adapt to changes [77].

After the incidence of P. aeruginosa in the ICU of Lausanne University Hospital increased, clinical and environmental isolates were typed using DLST. Three main types of DLST were identified (DLST 1-18, DLST 1-21, and DLST 6-7), and the identification capabilities of whole genome sequencing (WGS) were used to further study these three main types of DLST. It is one of the advantages of WGS to identify the ST of isolates by MLST inspection. DLST is both cheap and efficient. It is a routine monitoring method for *P. aeruginosa* in ICU wards and classifies all patients and environmental isolates quarterly. Only when several patients have similar genotypes, on-site epidemiological investigation and WGS are performed. Although the cost of WGS is decreasing, as a routine monitoring method for *P. aeruginosa*, its implementation cost is still higher than the currently used DLST. In addition, the

analysis of WGS data requires a certain level of bioinformatics expertise, but not all laboratories have it [77]. Therefore, using DLST as a first-line molecular typing tool for monitoring and WGS to solve problematic clusters will ultimately be an accurate and cost-effective typing strategy.

8.5 Summary

Pseudomonas aeruginosa is characterized by excellent biochemical and ecological versatility, and the genome size in different strains may vary by as much as 30%. Phenotyping methods for *P. aeruginosa* have been reported, including serotyping, pyocin, and antibacterial susceptibility typing and serotyping. Phenotypic characterization continues to play an important role in the management of *P. aeruginosa* infection. The distinguishing ability of molecular typing technology must be high enough to distinguish unrelated strains, but not high enough to assign isolates of a common lineage (such as epidemic clones) to different genotypes. Currently, a variety of molecular typing methods have been developed. This chapter introduces the application of several molecular typing methods such as restriction-based methods, amplified-based methods, sequencing-based methods in the typing of *P. aeruginosa*.

References

1. Hernandez J, Ferrus MA, Hernandez M, Owen RJ (1997) Arbitrary primed PCR fingerprinting and serotyping of clinical P. aeruginosa strains. FEMS Immunol Med Microbiol 17:37–47
2. Lila G, Mulliqi-Osmani G, Bajrami R, Kurti A, Azizi E, Raka L (2016) Antimicrobial resistance profile and serotyping of P. aeruginosa in university clinical center of Kosovo. ACTA Med Mediterr 32:829–835
3. Gillies RR, Govan JR (1966) Typing of Pseudomonas pyocyanea by pyocine production. J Pathol Bacteriol 91:339–345
4. Fyfe JA, Harris G, Govan JR (1984) Revised pyocin typing method for P. aeruginosa. J Clin Microbiol 20:47–50
5. Poh CL, Yap EH, Tay L, Bergan T (1988) Plasmid profiles compared with serotyping and pyocin typing for epidemiological surveillance of P. aeruginosa. J Med Microbiol 25:109–144
6. Sabat AJ, Budimir A, Nashev D, Sa-Leao R et al (2013) Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Euro Surveill 18:20380
7. Leone I, Chirillo MG, Raso T, Zucca M, Savoia D (2008) Phenotypic and genotypic characterization of *P. aeruginosa* from cystic fibrosis patients. Eur J Clin Microbiol Infect Dis 27:1093
8. Corona-Nakamura AL, Miranda-Novales MG, Leanos-Miranda B, Portillo-Gomez L, Aguilar-Benavides S (2001) Epidemiological study of P. aeruginosa in critical patients and reservoirs. Arch Med Res 32:238–242
9. Muller MP, Gubina M (2000) Serotype antimicrobial susceptibility and clone distribution of P. aeruginosa in a university hospital. Zentralbl Bakteriol 289:857–867
10. Tenover FC, Arbeith RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33(9):2233–2239
11. Tenover FC, Arbeit RD, Goering RV et al (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed- filed gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33:2233–2239

12. Spencker FB, Haupt S, Claros MC et al (2000) Epidemiologic characterization of P. aeruginosa in patients with cystic fibrosis. Clin Microbiol Infect 6:600–607
13. Talon D, Cailleaux V, Thouverez M et al (1996) Discriminatory power and usefulness of pulsed-field gel electrophoresis in epidemiological studies of P. aeruginosa. J Hosp Infect 32:135–145
14. Römling U, Tümmler B (2000) Achieving 100% typeability of P. aeruginosa by pulsed-filed gel electrophoresis. J Clin Microbiol 38:464–4653
15. Vatcheva-Dobrevski R, Ivanov I, Dobreva E, et al (2009) Comparison between multi-locus variable number of tandem repeats and pulsed-field gel electrophoresis for typing of nosocomial P. aeruginosa isolates. Paper presented at the 19th European congress of clinical microbiology and infectious diseases, Helsinki, Finland, vol 22, pp 16–19
16. Renders N, Römling Y, Verbrugh H et al (1996) Comparative typing of P. aeruginosa by random amplification of polymorphic DNA or pulsed- field gel electrophoresis of DNA macrorestriction fragments. J Clin Microbiol 34:3190–3195
17. Grundmann H, Schneider C, Hartung D et al (1995) Discriminatory power of three DNA based typing techniques for P. aeruginosa. J Clin Microbiol 33:528–534
18. Morales G, Wiehlmann L, Gudowius P et al (2004) Structure of P. aeruginosa populations analyzed by single nucleotide polymorphism and pulsed-field gel electrophoresis genotyping. J Bacteriol 186:4228–4237
19. Askarian M, Gooran NR (2003) National nosocomial infection surveillance system-based study in Iran: additional hospital stay attributable to nosocomial infections. Am J Infect Control 31:465–468
20. Libisch B, Muzslay M, Gacs M et al (2006) The epidemiology, pathogenesis and treatment of *P. aeruginosa* infections. Antimicrob Agents Chemother 50:351–368
21. Libisch B, Muzslay M, Gacs M et al (2006) Molecular epidemiology of VIM-4 metallo-β lactamase-producing Pseudomonas spp. isolates in Hungary. Antimicrob Agents Chemother 50:4220–4223
22. Poh CL, Yeo CC, Tay L (1992) Genome fingerprinting by pulsed-field gel electrophoresis and ribotyping to differentiate P. aeruginosa serotype O11 strains. Eur J Clin Microbiol Infect Dis 11:817–822
23. Bradbury RS, Champion AC, Reid DW (2009) Epidemiology of P. aeruginosa in a tertiary referral teaching hospital. J Hosp Infect 73(2):151–156
24. Huhulescu S, Simon M, Lubnow M, Kaase M, Wewalka G, Pietzka AT, Stöger A, Ruppitsch W, Allerberger F (2004) Fatal P. aeruginosa pneumonia in a previously healthy woman was most likely associated with a contaminated hot tub. Infection 39:265–269
25. Lanini S, D'Arezzo S, Puro V, Martini L, Imperi F, Piselli P, Montanaro M, Paoletti S, Visca P, Ippolito G (2010) Molecular epidemiology of a P. aeruginosa hospital outbreak driven by contaminated disinfectant-soap dispensers. PLoS One 6(2):e17064
26. Speert DP (2002) Molecular epidemiology of P. aeruginosa. Front Biosci 7:354–361
27. Singh A, Goering RV, Simjee S et al (2006) Application of molecular techniques to the study of hospital infection. Clin Microbiol Rev 19:512–530
28. Foley SL, Lynne AM, Nayak R (2009) Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. Infect Genet Evol 9:430–440
29. Pitt TL, Livermore DM, Pitcher D et al (1989) Multiresistant serotype O12 P. aeruginosa: evidence for a common strain in Europe. Epidemiol Infect 103:565–576
30. Bingen EH, Denamur E, Elion J (1994) Use of ribotyping in epidemiological surveillance of nosocomial outbreaks. Clin Microbiol Rev 7:311–327
31. Hollis RJ, Bruce JL, Fritschel SJ (1999) Comparative evaluation of an automated ribotyping instrument versus pulsed-field gel electrophoresis for epidemiological investigation of clinical isolates of bacteria. Diagn Microbiol Infect Dis 34:263–268
32. Brisse S, Milatovic D, Fluit AC et al (2000) Molecular surveillance of European quinolone-resistant clinical isolates of P. aeruginosa and Acinetobacter spp. using automated ribotyping. J Clin Microbiol 38:3636–3645

33. Snyder LA, Loman NJ, Faraj LA, Levi K, Weinstock G, Boswell TC, Pallen MJ, Ala'Aldeen DA (2013) Epidemiological investigation of *P. aeruginosa* isolates from a six-year-long hospital outbreak using high-throughput whole genome sequencing. Euro Surveill 18:42
34. Sarwari A, Hasan R, Lim CB, Ng Y, Ng C, Zaman S (2004) PCR identification and automated ribotyping of P. aeruginosa clinical isolates from intensive care patients. Scand J Infect Dis 36:342–349
35. Wiehlmann L, Wagner G, Cramer N, Siebert B, Gudowius P, Morales G et al (2007) Population structure of P. aeruginosa. Proc Natl Acad Sci U S A 104:8101–8106
36. Campbell M, Mahenthiralingam E, Speert DP (2000) Evaluation of random amplified polymorphic DNA typing of P. aeruginosa. J Clin Microbiol 38:4614–4615
37. Giske CG, Libisch B, Colinon C et al (2006) Establishing clonal relationships between VIM1-like metallo-β-lactamase-producing P. aeruginosa strains from four European countries by multilocus sequence typing. J Clin Microbiol 44:4309–4315
38. Mahenthiralingam E, Campbell ME, Foster J et al (1996) Random amplified polymorphic DNA typing of P. aeruginosa isolates recovered from patients with cystic fibrosis. J Clin Microbiol 34:1129–1135
39. Libisch B, Watine J, Balogh B et al (2008) Molecular typing indicates an important role for two international clonal complexes in dissemination of VIM-producing P. aeruginosa clinical isolates in Hungary. Res Microbiol 159:162–168
40. Libisch B, Balogh B, Füzi M (2009) Identification of two multidrug-resistant P. aeruginosa clonal lineages with a countrywide distribution in Hungary. Curr Microbiol 58:111–116
41. Snyder LA, Loman NJ, Faraj LA, Levi K, Weinstock G, Boswell TC et al (2013) Epidemiological investigation of P. aeruginosa isolates from a six-year-long hospital outbreak using high-throughput whole genome sequencing. Euro Surveill 18:20611
42. Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
43. Wahab AA, Taj-Aldeen SJ, Hagen F, Diophode S, Saadoon A et al (2014) Genotypic diversity of P. aeruginosa in cystic fibrosis siblings in Qatar using AFLP. Eur J Clin Microbiol Infect Dis 33:265–271
44. Vu-Thien H, Corbineau G, Hormigos K, Fauroux B, Corvol H, Clément A et al (2007) Multiple-locus variable-number tandem-repeat analysis for longitudinal survey of sources of P. aeruginosa infection in cystic fibrosis patients. J Clin Microbiol 45(10):3175–3183
45. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM et al (2014) Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510
46. Dass SC, Abu-Ghannam N, Antony-Babu S, Cummins EJ (2010) Ecology and molecular typing of L. monocytogenes in a processing plant for cold-smoked salmon in the Republic of Ireland. Food Res Int 43(5):1529–1536
47. Vu-Thien H, Corbineau G, Hormigos K et al (2007) Multiple-locus variable-number tandem-repeat analysis for longitudinal survey of sources of P. aeruginosa infection in cystic fibrosis patients. J Clin Microbiol 45:3175–3183
48. Onteniente L, Brisse S, Tassios PT (2003) Evaluation of the polymorphisms associated with tandem repeats for P. aeruginosa strain typing. J Clin Microbiol 41:4991–4997
49. Estaji M, Tabasi M, Heravi FS, Khezerloo JK (2019) Multiple locus variable number tandem repeat analysis (MLVA) for typing P. aeruginosa isolated from urine samples of different patients. Comp Immunol Microbiol Infect Dis 65:23–28
50. Farrell D, Morrissey I, De Rubeis D, Robbins M, Felmingham D (2003) A UK multicentre study of the antimicrobial susceptibility of bacterial pathogens causing urinary tract infection. J Infect 46(2):94–100
51. Glover M, Moreira CG, Sperandio V, Zimmern P (2014) Recurrent urinary tract infections in healthy and nonpregnant women. Urol Sci 25(1):1–8
52. Leclercq SY, Sullivan MJ, Ipe DS, Smith JP, Cripps AW, Ulett GC (2016) Pathogenesis of Streptococcus urinary tract infection depends on bacterial strain and β-hemolysin/cytolysin that mediates cytotoxicity, cytokine synthesis, inflammation and virulence. Sci Rep 6:29000

53. Elberse KE, Nunes S, Sá-Leão R, van der Heide HG, Schouls LM (2011) Multiple-locus variable number tandem repeat analysis for Streptococcus pneumoniae: comparison with PFGE and MLST. PLoS One 6(5):19668
54. van Mansfeld R, Bonten MJM, Willems R (2009) Assessing clonal relatedness of P. aeruginosa isolates: MLVA should be preferred over MLST and PFGE. Paper presented at the 19th European congress of clinical microbiology and infectious diseases, Helsinki, Finland, pp 16–19
55. Johansson E, Welinder-Olsson C, Gilljam M, Pourcel C, Lindblad A (2015) Genotyping of P. aeruginosa reveals high diversity, stability over time and good outcome of eradication. J Cyst Fibros 14(3):353–360
56. Beranek A, Mikula C, Rabold P, Arnhold D, Berghold C, Lederer I et al (2009) Multiple-locus variable-number tandem repeat analysis for subtyping of Salmonella enterica subsp. enterica serovar Enteritidis. Int J Med Microbiol 299(1):43–51
57. Santos AS, Tilburg JJ, Botelho A, Barahona MJ, Núncio MS, Nabuurs-Franssen MH et al (2012) Genotypic diversity of clinical Coxiella burnetii isolates from Portugal based on MST and MLVA typing. Int J Med Microbiol 302(6):253–256
58. Enright MC, Spratt BG (1999) Multilocus sequence typing. Trends Microbiol 7(12):482–487
59. Waters V, Zlosnik JE, Yau YC, Speert DP, Aaron SD, Guttman DS (2012) Comparison of three typing methods for P. aeruginosa isolates from patients with cystic fibrosis. Eur J Clin Microbiol Infect Dis 31(12):3341–3350
60. Curran B, Jonas D, Grundmann H, Pitt T, Dowson CG (2004) Development of a multilocus sequence typing scheme for the opportunistic pathogen P. aeruginosa. J Clin Microbiol 42(12):5644–5649
61. Johnson JK, Arduino SM, Stine OC (2007) Multilocus sequence typing compared to pulsed field gel electrophoresis for molecular typing of P. aeruginosa. J Clin Microbiol 45:3707–3712
62. Giske CG, Libisch B, Colinon C (2006) A new highly discriminatory multiplex capillary-based MLVA assay as a tool for the epidemiological survey of *P. aeruginosa* in cystic fibrosis patients. J Clin Microbiol 44:4309–4315
63. Feil EJ, Li BC, Aanensen DM et al (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186:1518–1530
64. Libisch B, Poirel L, Lepsanovic Z et al (2008) Identification of PER-1 extended-spectrum b-lactamase producing P. aeruginosa clinical isolates of the international clonal complex CC11 from Hungary and Serbia. FEMS Immunol Med Microbiol 54:330–333
65. Libisch B, Muzslay M, Gacs M et al (2006) Molecular epidemiology of VIM-4 metallo-b-lactamase-producing Pseudomonas spp. isolates in Hungary. Antimicrob Agents Chemother 50:4220–4223
66. Edalucci E, Spinelli R, Dolzani L et al (2008) Acquisition of different carbapenem resistance mechanisms by an epidemic clonal lineage of P. aeruginosa. Clin Microbiol Infect 14:88–90
67. Lepsanovic Z, Libisch B, Tomanovic B et al (2008) Characterisation of the first VIM metallo-b-lactamase-producing P. aeruginosa clinical isolate in Serbia. Acta Microbiol Immunol Hung 55:447–454
68. Empel J, Filczak K, Mrowka A, Hryniewicz W, Livermore DM, Gniadkowski M (2007) Outbreak of P. aeruginosa infections with PER-1extended-spectrumbeta-lactamase in Warsaw, Poland: further evidence for an international clonal complex. J Clin Microbiol 45:2829–2834
69. van Mansfeld R, Irene J, Bootsma M (2010) The population genetics of *P. aeruginosa* isolates from different patient populations exhibits high-level host specificity. PLoS One 5(10):e13482
70. Hrabák J, Cervena D, Izdebski R, Duljasz W, Gniadkowski M, Fridrichova M, Urbaskova P, Zemlickova H (2011) Regional spread of P. aeruginosa ST357 producing IMP-7 metallo-beta-lactamase in Central Europe. J Clin Microbiol 49:474–475
71. Ji J, Wang J, Zhou Z, Wang H, Jiang Y, Yu Y (2013) Multilocus sequence typing reveals genetic diversity of carbapenem- or ceftazidime-nonsusceptible P. aeruginosa in China. Antimicrob Agents Chemother 57(11):5697–5700

72. Basset P, Hammer NB, Kuhn G, Vogel V, Sakwinska O, Blanc DS (2009) Staphylococcus aureus clfB and spa alleles of the repeat regions are segregated into major phylogenetic lineages. Infect Genet Evol 9:941–947
73. Basset P, Senn L, Prod'hom G, Bille J, Francioli P, Zanetti G et al (2010) Usefulness of double locus sequence typing (DLST) for regional and international epidemiological surveillance of methicilin-resistant Staphylococcus aureus. Clin Microbiol Infect 16:1289–1296
74. Sakwinska O, Blanc DS, Lazor-Blanchet C, Moreillon M, Giddey M, Moreillon P (2010) Ecological temporal stability of Staphylococcus aureus nasal carriage. J Clin Microbiol 48:2724–2728
75. Cholley P, Stojanov M, Hocquet D, Thouverez M (2015) Comparison of double-locus sequence typing (DLST) and multilocus sequence typing (MLST) for the investigation of P. aeruginosa populations. Diagn Microbiol Infect Dis 82:274–277
76. Magalhaes B, Valot B, Mohamed MH, Prod'hom G, Greu G, Senn L, Blanc DS (2020) Combining standard molecular typing and whole genome sequencing to investigate P. aeruginosa epidemiology in intensive care units. Front Public Health 8:3
77. Blanc DS (2004) The use of molecular typing for epidemiological surveillance and investigation of endemic nosocomial infections. Infect Genet Evol 4:193–197

Staphylococci

9

JoAnn M. McClure and Kunyan Zhang

9.1 Introduction

Staphylococci are Gram-positive cocci that appear as grape-like clusters. The genus is comprised of more than 40 species, most of which are harmless and exist on the skin and mucous membranes of humans or other animals. Staphylococci are divided into coagulase-negative (CoNS) and coagulase-positive members, based on their ability to produce the free enzyme coagulase, which causes blood clot formation. While the majority of staphylococcal species are CoNS, few CoNS have been implicated in human disease. This, however, has been changing, with an increasing number of CoNS infections identified, boosting their clinical significance [1, 2]. *Staphylococcus aureus* (SA), the most notable member of the genus, is coagulase positive and has been the primary focus of clinical identification as it is commonly associated with human infection. Methicillin-resistant *Staphylococcus aureus* (MRSA), in particular, has garnered much of that attention as it is resistant to all penicillins and most β-lactam drugs and associated with higher morbidity and mortality rates among hospitalized patients and higher patient care costs [3–5].

MRSA has been shown to asymptomatically colonize 20–30% of the human population [6, 7] but is also responsible for a wide variety of infections, ranging

J. M. McClure
Centre for Antimicrobial Resistance (CAR), Alberta Health Services/Alberta Precision Laboratories/University of Calgary, Calgary, Canada

K. Zhang (✉)
Centre for Antimicrobial Resistance (CAR), Alberta Health Services/Alberta Precision Laboratories/University of Calgary, Calgary, Canada

Department of Pathology & Laboratory Medicine, Department of Microbiology, Immunology & Infectious Diseases, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Canada
e-mail: kzhang@ucalgary.ca

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_9

from mild skin and soft tissue infections to life-threatening illnesses such as endocarditis, septicemia, and hemorrhagic pneumonia [8]. MRSA infections were initially associated with hospitals and healthcare settings; however, MRSA has since emerged as a major cause of community-associated infection as well. Adding complexity is the fact that, despite the overwhelming attention given to MRSA, methicillin-sensitive *S. aureus* (MSSA) infections are increasingly being recognized as presenting a significant threat to public health [9, 10]. With the ever-changing prevalence and epidemiology of *S. aureus* infections, reliable methods for characterizing strains are essential for outbreak investigations, for tracking clonal spreading, and for the implementation of effective treatment or control measures. At the local level, typing is useful for identifying clones, which aids in disease management and in predicting prognosis. It also helps identify outbreaks and strain spreading within the geographic locale, guiding infection control strategies. At the international level, strain typing aids in investigation related to the evolution and spread of clonal types, both over large areas and over time. Discussed in this chapter are the various phenotypic and molecular methods used to discriminate *S. aureus* lineages.

9.2 Identification of Staphylococcal Species

Differentiation of *S. aureus* from CoNS is accomplished using standard microbiological methods in clinical diagnostic laboratories. Staphylococci are catalase-positive, facultative anaerobes, capable of growing in the presence of bile salts or 6.5% NaCl solution. Columbia or tryptic soy blood agar, with 5% defibrinated sheep or horse blood, is the primary culture plate used for staphylococcal isolation. On blood, *S. aureus* presents as large, round, golden-yellow colonies that are most often β-hemolytic. CoNS colonies, on the other hand, are typically smaller in size, non-pigmented, smooth, glistening, and opaque, although some species can be gray-yellow to yellow-orange in pigmentation and can also be β-hemolytic. Coagulase tube test or rapid latex and hemagglutination assays allow presumptive identification of *S. aureus*, while commercial systems can differentiate the staphylococcal species using biochemical procedures. Systems such as Vitek 2 (bioMérieux), the BBL Crystal Identification System's Rapid Gram-Positive ID Kit (BD Diagnostic Systems, Sparks, MD), the Pos ID Panel family (Siemens Healthcare Diagnostics, Deerfield, IL), the Phoenix Automated Microbiology System (BD Diagnostic Systems), the Biolog systems (Biolog, Hayward, CA), the RapiDEC Staph (bioMérieux), and the API Staph and ID32 Staph strips (bioMérieux, La Balmeles-Grottes, France) are routinely used in clinical laboratories. Antibiotic susceptibility patterns for the staphylococcal species can be obtained on systems such as Vitek 2 (bioMérieux).

While biochemical identification of *S. aureus* is relatively straightforward, CoNS have proven to be more problematic. Common species such as *S. epidermidis*, *S. saprophyticus*, and *S. haemolyticus* are generally successfully identified by biochemical means, while identification of less common species such as *S. warneri* and *S. hominis* shows more variable rates [11–13]. Nucleic acid amplification and sequencing of universally occurring genomic regions offer an effective alternative for speciating staphylococci and can be accomplished quickly with minimal cost. Sequencing of a portion of the *rpoE* gene has been shown to accurately differentiate staphylococcal species [14]; however, sequencing of the 16S rRNA gene is generally considered the gold standard for identification and taxonomic classification of bacterial species. 16S rRNA is the small component of the prokaryotic ribosome that binds to the Shine-Dalgarno sequence, with its gene undergoing slow rates of evolution, making it useful for phylogenetic analysis. The 16S rRNA gene contains highly conserved primer binding regions, as well as nine hypervariable regions (V1–V9), each ranging from 30 to 100 bp in length [15]. Sequencing of the full 16S rRNA gene can be performed; however, more commonly shorter sequences involving the variable regions are targeted. Regions V1–V3, in particular, have been shown to be the most useful in distinguishing among staphylococcal species [16]. Various 16S ribosomal databases exist for analyzing sequencing data, including public databases such as NCBI and secondary ones such as EzBioCloud, Ribosomal Database Project, SILVA, and Greengenes [17–20]. While the public databases are easily accessible and free, the quality of sequences and taxonomic assignments found on the database are often not validated, making secondary databases that collect and validate 16S rRNA sequences superior choices.

As CoNS are not routinely typed beyond species identification and antibiotic susceptibility, the remainder of this chapter will focus on molecular characterization of *S. aureus*. Discrimination of isolates based on phenotypic and genotypic characteristics is important for determining clonal relationships between strains and furthering our understanding of the epidemiology of infectious diseases. Presently, classification schemes for *Staphylococcus aureus* are based less on phenotypic methods and more so on molecular ones. While many of these methods were initially used for research purposes, they are now commonly used in clinical labs as well.

9.3 MRSA Identification and SCC*mec* Typing

Distinguishing MRSA from MSSA is an important first step in *S. aureus* classification. MRSA have acquired and integrated into their chromosome a mobile genetic element known as staphylococcal cassette chromosome *mec* (SCC*mec*), which carries the methicillin resistance genes *mecA* or *mecC*. *mecA* was the first

methicillin resistance gene identified and encodes an alternative penicillin-binding protein (PBP2a or PBP2'), which has low affinity for semisynthetic penicillins and confers resistance to all β-lactam drugs except ceftaroline and ceftobiprole [21]. *mecA* remained the only methicillin resistance gene identified in *S. aureus* until 2011, when the *mecC* gene was described, sharing 70% identity with *mecA*, and coding a PBP2a/2′ sharing 63% homology at the amino acid level [22]. A third homologue, *mecB*, was first identified in 2009 in closely related bacteria, *Macrococcus caseolyticus* [23]; however, in 2018, it was detected for the first time in *S. aureus* on a plasmid [24]. The *mecB* gene shares 60% homology with *mecA* and confers resistance to methicillin. A fourth homologue, *mecD*, has been reported on a genomic island ($McRI_{mecD}$-1 and $McRI_{mecD}$-2) in *M. caseolyticus* but to date has not been detected in *S. aureus*. The Clinical and Laboratory Standards Institute (CLSI) recommends testing for MRSA using broth microdilution or with cefoxitin disk diffusion or Mueller-Hinton agar plates supplemented with 4% NaCl and 6 μg/ml of oxacillin as alternatives [25]. Chromogenic agars, such as CHROMagar™ MRSA, Oxoid Brilliance™ MRSA, MRSASelect, BBL™ CHROMagar™ MRSA, and ChromID MRSA, are also available for MRSA detection, offering highly sensitive and specific detection [26]. The PBP2a latex agglutination test (Oxoid, Hampshire, UK) is also available as an alternate phenotypic test for detecting PBP2a in *S. aureus* colonies; however, it suffers from a large variability in performance [27, 28]. No optimal phenotypic method exists for MRSA detection, as they generally require specialized conditions and results are affected by factors such as inoculum size, incubation temperature and time, or pH and salt concentration.

Nucleic acid amplification tests represent a more precise and reliable form of MRSA identification and have become the gold standard for MRSA detection. These assays have traditionally relied on detection of the *mecA* gene; however, detection of the *mecC* gene also needs to be considered now. Additionally, while the *mecB* gene has only been described in one instance, its detection may become important if the gene spreads. Murakami et al. [29] were the first to develop a PCR assay for MRSA detection, targeting the *mecA* gene, while the first multiplex PCR assay targeting both the *mecA* and 16S rRNA genes was developed by Geha et al. [30]. Since then, a substantial number of assays have been developed targeting the *mecA*/*mecC* genes alone or in conjunction with other targets, such as PVL, *fem*, *nuc*, or 16S rRNA, and using both standard and real-time PCR platforms. In 2008, Zhang et al. developed a multiplex PCR assay that could discriminate staphylococci from non-staphylococcal species while simultaneously distinguishing *S. aureus* from CoNS, identifying MRSA, identifying the Panton-Valentine leukocidin virulence genes, and presumptively identifying USA300 and USA400 epidemic strains [31]. While this assay has been extensively used, it suffers in that it does not detect the *mecB* or *mecC* genes. In 2012, Stegger et al. developed a multiplex PCR assay capable of simultaneously detecting both the *mecA* and *mecC* genes, along with the PVL genes and the staphylococcal protein A gene (*spa*) [32].

The assay allows rapid and inexpensive detection of MRSA, with the ability to perform downstream *spa* typing of isolates, but does not take into account the *mecB* gene.

As mentioned, the *mecA* and *mecC* genes, which confer resistance to β-lactam antibiotics, are carried on a mobile genetic element termed staphylococcal cassette chromosome *mec*. To date, 13 different SCC*mec* elements have been described in *S. aureus* based on the nature of their *mec* and *ccr* gene complex and are further divided into subtypes based on differences in their joining regions. These differences provide an important means of classifying MRSA isolates, as even closely related strains can differ in the type of SCC*mec* element they carry. Initial SCC*mec* typing schemes involved molecular cloning and sequencing or long-range PCR amplification with multiple sets of primers [33–35]. Typing schemes have since improved to include conventional PCR detection of several type-specific loci [36], RFLP analysis [37, 38], multiplex PCR [39], multiplex real-time PCR [40, 41], and targeted DNA microarrays [42]. Multiplex PCR typing is currently the most widely used method of SCC*mec* typing, with several variations developed. A novel multiplex PCR assay for the characterization and concomitant subtyping of SCC*mec* I–V was developed by Zhang et al. in 2005 and later updated in 2012 to make it more accurate and reliable [43, 44]. Similarly, in 2007, Milheirico et al. updated a previous multiplex PCR assay to detect SCC*mec* I–V. These multiplex assays are by far the most commonly used ones for SCC*mec* typing; however, both are limited to detection of types I–V, requiring other methods for the detection of types VI–XIII. Both are also restricted by their inability to classify newly evolving SCC*mec* types and subtypes. Unfortunately, to date, no single PCR assay is available to identify all SCC*mec* types and subtypes. Targeted DNA microarray offers an alternate option for SCC*mec* typing, simultaneously detecting multiple genes associated with SCC*mec*, including *mecA* and its regulatory genes, and sequences in the J regions [42]. As with PCR, only known SCC*mec* types can be identified with this technique, and it suffers from the added disadvantage that specialized equipment and highly trained personnel are required. As such, multiplex PCR remains the best option for SCC*mec* typing at present.

9.4 Historical Typing Methods

In an attempt to understand and track *S. aureus* (particularly MRSA) infections, numerous typing methods were developed to classify lineages. While these historical methods are rarely used routinely anymore, they still can be of value when typing *S. aureus*.

Phage Typing relies on bacterial susceptibility to a defined set of phages, with a set of 23 internationally accepted phages used for typing human strains of *S. aureus* [45, 46]. While the method was the primary one used for several years, it suffered in

that it often lacked reproducibility and was time-consuming and technically challenging and a large percentage of strains remained untypable with the technique [47–50].

Multilocus Enzyme Electrophoresis (MLEE) involves the extraction of constitutively expressed proteins from the bacteria and their separation on gels using electrophoresis, with the rate of migration being dependent on amino acid composition in the proteins. Generally, 12–20 proteins are assessed, each being assigned allelic types based on variation in their charge, with the similarity between isolates determined by the proportion of loci which show differences. While MLEE generally has good reproducibility and typability for *S. aureus*, it is a labor-intensive procedure, and the results are difficult to compare between laboratories [51, 52].

Random Amplification of Polymorphic DNA (RAPD) and Arbitrarily Primed PCR (AP-PCR) rely on parallel non-stringent amplification of random DNA fragments, resulting in unique gel patterns specific to each bacterial strain [53, 54]. In RAPD, short arbitrary primer sequences and low-temperature, non-stringent annealing conditions allow amplification of multiple PCR products of varying sizes. Amplicons are analyzed either by gel electrophoresis or DNA sequencing, with the number and size of fragments used to define an isolate type [54]. AP-PCR is a variant of RAPD, whereby amplification is done in three parts, each of which has a set stringency and reagent concentrations [53]. While these techniques have been used successfully in outbreaks and are relatively inexpensive and easy, they have lower discriminatory power and lower inter- and intra-laboratory reproducibility [55–57].

Repetitive Element PCR (rep-PCR) employs primers that bind to noncoding repetitive sequences in the bacterial genome, producing fingerprint patterns unique to each isolate [58]. The repetitive palindromic extragenic elements (Rep) are sequences 35–38 bp long that occur in variable positions and numbers. Amplification of the elements creates amplicons of varying lengths, which are separated by electrophoresis, creating fingerprints unique to the strains. For *S. aureus*, RepMP3 and inter-IS256 and Tn916 are commonly used targets, with RepMP3 showing greater reproducibility and stability [59]. Rep-PCR has high discriminatory power, with good correlation to PFGE; however, reproducibility can suffer from variations in reagents and electrophoresis systems [60].

Amplified Fragment Length POLYMORPHISM (AFLP) relies on differences in the amplification of digested genomic DNA fragments [61]. Genomic DNA is digested with restriction enzymes, and double-stranded adaptors are

ligated to the sticky ends, followed by amplification of the fragments using primers complementary to the adaptors. The primers are generally fluorescently labelled; therefore, after separation of the amplicons based on size, they can be detected with an automated DNA sequencer and compared by computer. Analysis of the high-resolution banding patterns is used to determine the relationship between strains [62]. While this technique is portable and highly reproducible and has high discriminatory power, it is time-consuming and expensive [63, 64].

Accessory Gene Regulator (*agr*) Typing is a PCR-based typing method that relies on amplification of hypervariable regions present in the *agr* locus to classify strains. The accessory gene regulator (*agr*) is a bacterial regulatory component containing two divergently transcribed units, which has highly conserved and hypervariable regions [65]. Four genes, *agrA*, *agrC*, *agrD*, and *agrB*, are present in the locus. The C-terminal of *agrB* and *agrD* and the N-terminal of *agrC* are highly divergent and constitute the hypervariable region of the locus, which is used to divide *S. aureus* into four *agr* groups (I–IV) [65]. PCR primers for *agr* group determination were developed by Peacock et al. [66], and a multiplex real-time quantitative PCR assay was developed by Francois et al., targeting the variable region of *agrC* and offering good specificity [67]. While *agr* typing is extremely limited in its discriminatory power and would not be useful for defining *S. aureus* lineages, it does provide additional information about strains that can supplement other typing methods.

9.5 Current Molecular Typing Methods

Current typing schemes for *S. aureus* classification rely predominantly on molecular methods based on DNA sequence variations. A proposal was made that MRSA clones should be defined based on a combination of the genomic type of the strain and the SCC*mec* type, a nomenclature system that was accepted in 2002 by the subcommittee of the International Union of Microbiology Societies in Tokyo [68]. This system, which can be amended to describe both MRSA and MSSA (e.g., ST8-MRSA-IVa or ST8-MSSA), relies solely on multilocus sequence typing and SCC*mec* typing (discussed below) to define the strains. While these two methods are important parts of *S. aureus* classification, the addition of other typing schemes provides more complete information about *S. aureus* lineages, which are discussed below.

9.5.1 Pulsed-Field Gel Electrophoresis (PFGE) Typing

PFGE was first described in 1984 and is based on the digestion of bacterial genomes into large fragments with a restriction enzyme and their subsequent separation by

gel electrophoresis [69, 70]. Because larger fragments of DNA will co-migrate and appear as a large diffuse band with conventional gel electrophoresis, in PFGE, the voltage direction is periodically switched (pulsed), allowing effective separation of larger DNA pieces. Migration of the DNA fragments produces a DNA fingerprint, which can be used to compare the relatedness of strains.

For PFGE, genomic DNA needs to be intact and free from mechanical shearing; therefore, bacterial cells are incorporated into agarose plugs prior to lysis to protect the DNA from damage [71]. DNA, which is immobilized in the agarose plug, is digested with a rare-cutting restriction endonuclease, at which time the plugs are loaded onto an agarose gel and subjected to PFGE. PFGE protocols for *Staphylococcus aureus* have been optimized and, with minor variations, include standard features common to typing this species [72–74] (https://www.cdc.gov/mrsa/pdf/ar_mras_PFGE_s_aureus.pdf). A number of restriction endonucleases have been used in PFGE typing of bacterial species; however, smaI was found to be the most useful for *S. aureus*, allowing nearly all isolates to be typed, with reproducible results following repeated subcultures [75–77]. *S. aureus* belonging to the ST398 lineage are the exception, not typable using smaI due to a DNA methyltransferase that modifies the consensus sequence [78]. The restriction enzyme Cfr9I, a neoschizomer of smaI, is able to cleave these strains within the same recognition sequence as smaI and is used for PFGE typing of the ST389 lineage. *S. aureus* gels are generally run with the contour-clamped homogeneous electric field (CHEF) electrophoresis system, where the current is applied in three directions, offset by 120°, using hexagonally arranged electrodes [52, 79].

PFGE is a popular technique used by laboratories around the world and is effective for providing local epidemiological information, as well as for identifying epidemics. In experienced hands, the method can provide information related to the presence or absence of some mobile genetic elements such as the SCC*mec* cassette or phages. The technique has high discriminatory power, and results can be reproducible at both the intra- and inter-laboratory levels when the method is highly standardized [48, 80]. To aid with standardization, the Centers for Disease Control and Prevention in the USA developed PulseNet (https://www.cdc.gov/pulsenet/index.html). It is a national laboratory network that uses bacterial DNA fingerprints (such as PFGE patterns) to detect foodborne illnesses and outbreaks. Standard protocols are available, and data can be shared nationally or internationally. Also helping with standardization is the fact that *S. aureus* gels are run with the *S. braenderup* H2812 control standard and the data normalized and analyzed using BioNumerics software (Applied Maths, Sint-Martens-Latem, Belgium). Data analysis criteria set out by Tenover et al. are useful for comparing strains and determining their relatedness [81], and *S. aureus* PFGE profiles have been assembled into a national database to assist interpretation [72, 82]. In Fig. 9.1, sample PFGE patterns for Canadian and US epidemic reference strains are shown, along with some other common typing information for each strain. PFGE does suffer from limitations, the main ones being the long turnaround time, the high cost for specialized equipment and software, and the skill level required. Without high standardization, data interpretation can be

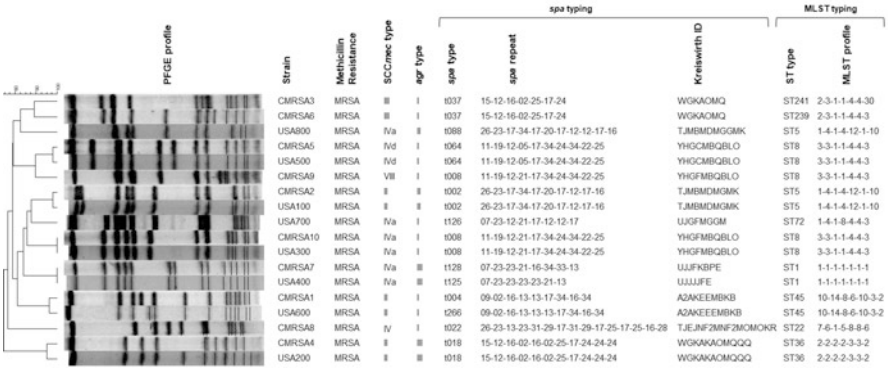

Fig. 9.1 Sample typing results for representative Canadian (CMRSA1-10) and US (USA100-800) epidemic reference strains**.** Different lineages may share the same type when classified using a single typing method but will become distinguishable from each other when multiple typing schemes are used together. Pulsed-field gel electrophoresis (PFGE) profiles, staphylococcal cassette chromosome *mec* (SCC*mec*) type, accessory gene regulator (*agr*) type, staphylococcal protein A (*spa*) type (including Ridom repeat pattern and Kreiswirth ID), and multilocus sequence type (MLST) (including MLST profile) are shown

problematic, as differences in electrophoresis equipment and conditions can affect DNA migration, complicating isolate comparisons within and between laboratories [83, 84]. As well, the technique separates DNA based on size, not sequence, and small changes are enough to affect the fingerprint. For example, the acquisition or loss of mobile genetic elements will alter the banding pattern, as will a point mutation in the smaI recognition sequence.

Despite the limitations, PFGE remains a powerful technique for *S. aureus* typing and classification and is still considered the "gold standard."

9.5.2 Staphylococcal Protein A (*spa*) Typing

The *spa* gene, coding for protein A, is conserved among *S. aureus* and has proven to be an effective target for single-locus sequence typing of this species. The gene is approximately 2 kb in length and contains conserved Fc binding regions, a variable X region, and a conserved C-terminal region. The X region (or repeat region) is comprised of polymorphic variable number of tandem repeats (VNTR), generally consisting of 2–18 repetitive sequences of 21–30 bp (most often 24 bp) in length [85]. Each repeat is given an identifier (numerical or letter code), with the number, order, and sequence of these repeats varying between strains, forming the basis for *spa* typing [86, 87].

Two nomenclature systems, Ridom and Kreiswirth, are used for describing *spa* types and repeats, with Ridom represented by numerical repeat codes and Kreiswirth represented with alpha numeric repeat codes [86, 88]. Conversion between the two

is possible with online tools. The Ridom StaphType software (available for download from www.ridom.de/staphtype/) was developed to ensure uniform assignments of *spa* repeats and types and is useful for MRSA surveillance. The software synchronizes with the Ridom SpaServer (www.SpaServer.ridom.de), which is a freely accessible server developed to collate and harmonize data from around the world, permitting 100% reproducibility between laboratories and providing public access to typing data. Figure 9.1 shows the *spa* type, including the Ridom and Kreiswirth profiles, for Canadian and US epidemic reference strains.

spa typing is a reliable way of assigning lineage and has proven to be effective for both short-term and long-term epidemiological studies [80, 86–89]. The speed and simplicity of targeting a single locus make it favorable for short-term studies, while the stable association of types with lineages over time makes it suitable for long-term studies. Development of the BURP (Based Upon Repeat Pattern) algorithm has provided an automated method to infer clonal relatedness of isolates based on *spa* repeat patterns and was shown to have high concordance with other typing methods [89, 90]. With a high discriminatory power, *spa* typing is a cost-effective, easy-to-use method with excellent reproducibility and portability. The major drawback of *spa* typing is the fact that the method relies on typing a single locus, running the risk that strains can be misclassified due to recombination and/or homoplasy [91]. Strains from different lineages can carry the same *spa* type (Fig. 9.1), and epidemiologically related strains from a lineage may carry different *spa* types, varying in as little as a single repeat. *spa* typing is, consequently, most effective when used in combination with other typing methods.

9.5.3 Multilocus Sequence Typing (MLST)

MLST is similar in principle to MLEE, but variations are examined directly by DNA sequencing. The method relies on sequencing a 402–516 bp fragment from each of seven essential housekeeping genes, present in all *S. aureus* isolates. These genes are crucial to cellular function and, therefore, stable and evolve slowly. Based on point mutations, the genes for each locus are assigned numerical allele designations, with the series of seven numbers (one representing each locus) defining the sequence type (ST type) of a strain. For *S. aureus*, the carbamate kinase (*arcC*), shikimate dehydrogenase (*aroE*), glycerol kinase (*glpF*), guanylate kinase (*gmk*), phosphate acetyltransferase (*pta*), triosephosphate isomerase (*tpi*), and acetyl coenzyme A acetyltransferase (*yqiL*) genes were selected, as they provided the highest number of alleles with the best resolving power for identifying lineages [92]. The genes are arranged in the abovementioned order (i.e., *arcC-aroE-glpF-gmk-pta-tpi-yqiL*) to define the ST type (e.g., ST8 has an MLST profile of 3-3-1-1-4-4-3).

Sequence analysis was initially facilitated by the online server available at MLST.net, a free website which provided the main hub for assigning allele and sequence types, naming new ones, as well as storing other important information related to the clonal types [93]. Now available for analysis is the database at

PubMLST (https://pubmlst.org/saureus/), which contains both sequence definition and epidemiological information [94, 95]. To aid with visualizing and analyzing the evolutionary relationship between isolates, the eBURST (Based Upon Related Sequence Type) algorithm was developed [96, 97]. Strains sharing identical allelic profiles are considered as belonging to the same ST type and lineage, while strains differing by one or two loci (single-locus variants or double-locus variants) are considered to be genetically related, belonging to the same clonal complex (CC). The founding genotype for a clonal complex is the one that differs from the highest number of other genotypes by only one locus, assuming strains emerge as dominant clones and then diversify with time. A representative eBURST image showing the relatedness of MLST types from Canadian and US epidemic strains in the global *Staphylococcus aureus* population is shown in Fig. 9.2.

MLST is a useful tool for assigning lineage and has proven to be effective for studying the origin and evolution of *S. aureus.* The method is unambiguous and portable, making data transfer to, and comparison between, labs around the world simple. The technique is, however, intolerant to sequencing errors, as a single nucleotide change can lead to an incorrect ST assignment. Cost is another drawback to the method, as it requires high-quality sequences for 7 loci, requiring 14 sequencing reads.

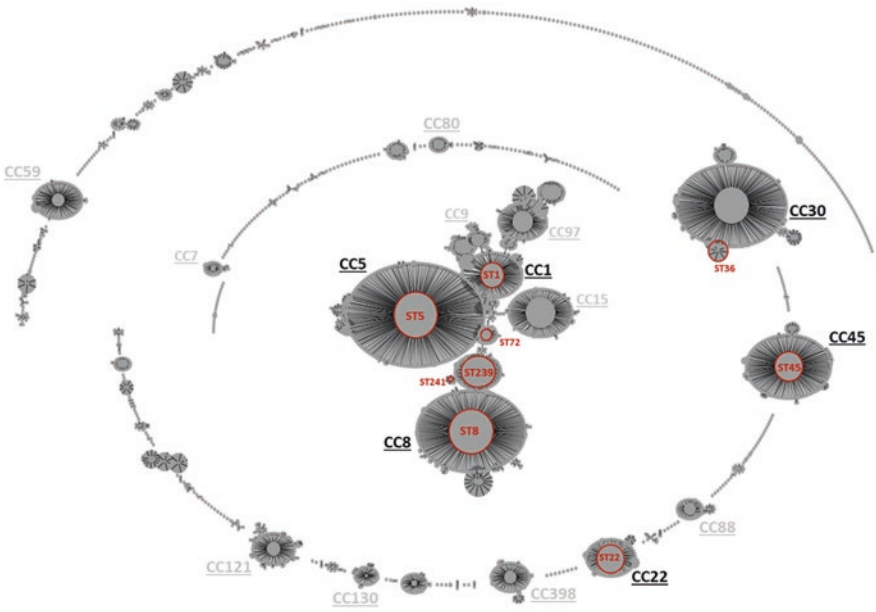

Fig. 9.2 Demonstration of eBURST analysis showing the relatedness of MLST types identified in the Canadian and US epidemic strains CMRSA1-10 and USA100-800 in the global *Staphylococcus aureus* population. Clonal complexes are marked in black font for the strains of interest, while ST types are marked in red (Generated on December 1, 2018)

This makes it less appealing as a tool for studying outbreaks or for use in smaller facilities with limited sequencing capability. Caution also has to be taken when relating MLST types to epidemiology, as strains with significantly different epidemiological significance can share a common MLST type. For example, the major epidemic strain in North America, USA300, belongs to MLST type ST8, a type also found in the infrequently encountered Canadian lineages, CMRSA9 and CMRSA5 (USA500) (Fig. 9.1). Despite the drawbacks, MLST is highly reproducible with high discriminatory power and, in conjunction with SCC*mec* type, remains the gold standard for publishing *S. aureus* epidemiological data.

9.5.4 Microarray

DNA microarrays use DNA probes attached in a known order to a solid surface to type bacterial isolates [98]. The probes can be oligonucleotides or gene segments (PCR amplicons) and can occur in low (100 s) or high (100,000 s) density. Bacterial DNA is labelled and allowed to hybridize to the microarray, such that complementary sequences present in the strain will bind to the probe. The microarray is scanned, and labelled spots are detected and then compared to known strains.

Microarrays are an effective means of typing and, indirectly, assigning lineage for *S. aureus*, simultaneously targeting a large number of strain-specific markers such as genes for antimicrobial resistance, exotoxins, surface components, regulators, and *hsdS* variants [42, 99, 100]. They are also well suited to the detection of complex patterns of virulence genes, mobile genetic elements, and extrachromosomal elements [101, 102] and have been used to understand the molecular mechanisms of pathogenesis, studying regulons such as Agr, Sar, SigB, and Mgr [103–105]. As such, microarrays permit strains to simultaneously be assigned to a lineage while having their resistance and virulence capabilities investigated at the same time.

Numerous microarrays have been designed specifically for *S. aureus* typing, and several companies make it possible to design custom arrays to meet specific needs [106–110]. The Alere StaphType DNA microarray is a commercially available system that covers 334 targets, including 170 genes and their allelic variants [42, 111, 112]. Included are species markers, capsule and *agr* typing markers, toxin and microbial surface components recognizing adhesive matrix molecule (MSCRAMM) genes, resistance gene markers, and SCC*mec* markers. On a larger scale, the Sam-62 microarray was developed based on 62 *S. aureus* whole genome sequencing projects and 153 plasmid sequences. The array targets all open reading frames in the sequences and includes over 29,000 probes, representing 6520 genes and 579 gene variants [113]. Sam-62 has shown potential to identify MRSA, distinguish between extremely similar but non-identical sequences, and be able to identify MRSA transmission events unrecognized using other methods [101].

While DNA microarray is highly accurate, specialized equipment and software are required meaning there is a significant cost associated with their use. Microarrays also suffer in that they cannot directly assign MLST group; strains can only be assigned to a given clonal complex group once the hybridization pattern of a reference strain with known MLST/*spa* types has been defined.

9.6 Whole Genome Sequencing (WGS) and the Future of MRSA Typing

WGS is a powerful tool for *S. aureus* typing, as well as for epidemiological and evolutionary studies, and next-generation sequencing (NGS) has provided a cost-effective means of extracting large amounts of information and identifying genome-wide variations. Today, the most commonly used NGS platform is Illumina (Illumina, Inc., San Diego, CA, USA), which can generate reads up to 300 bp in length. Assembly of a genome can be accomplished via de novo assembly, whereby reads are matched based on overlapping regions, or with reference-guided assembly, where reads are assembled against an existing WGS. De novo assembly in *S. aureus* is challenging, however, because of the small read sizes and the presence of dispersed or tandemly arrayed repeats in the genome. As such, the resulting genome is not continuous, but rather contains numerous contigs with gaps between assembled regions, due in part to the inability to resolve contig order surrounding these repeat elements. Reference-guided assembly can also be challenging because genomic regions, such as mobile genetic elements (MGEs), that are not present in the reference will be assembled poorly, particularly if they contain repeat elements, such as in SCC*mec*. Illumina data is still useful for querying genomic traits and variations, as well as for phylogenetic analysis, but for a complete genome assembly, sequencing platforms that generate longer reads are necessary.

Read lengths of >10 kb (and up to 60 kb) are possible with the "third-generation" PacBio sequencing platform (Pacific Biosciences, Menlo Park, CA, USA), while read lengths in the Mbp range have been achieved using nanopore sequencing technology (Oxford Nanopore, Oxford, UK). These systems suffer in that they can be more expensive and have lower read accuracy than Illumina; however, with tailored assembly methods (such as HGAP for PacBio reads), assemblies with higher accuracy are achieved. Hybrid assemblies, combining Illumina short reads and PacBio or Nanopore long reads, currently offer the most accurate and complete genomes.

A major drawback of WGS is the requirement for significant computer resources and bioinformatics support in order to extract meaningful information from the data. Software such as Lasergene exists for assembly and analysis of the genomes; however, in most cases, more complex pipelines are employed and require trained bioinformaticians. For WGS technology to become useful for

routine typing of *S. aureus*, tools for data analysis that are simple enough for use in clinical settings are required, and a number of web-based and downloadable programs are available to help in this regard. The Center for Genomic Epidemiology (Lyngby, Denmark, available at https://cge.cbs.dtu.dk/services/), for example, has web-based analysis tools that are useful for *S. aureus* WGS analysis and able to extract data from raw reads and assembled or draft genomes generated using Illumina, Ion Torrent, Roche 454, SOLiD, PacBio, or Nanopore platforms. Currently available on the site are MLST, for assigning ST type; spaTyper, for determining *spa* type; and SCCmecFinder, for classifying SCC*mec* type. Also available are ResFinder, for identifying acquired antimicrobial resistance genes and/or chromosomal mutations, VirulenceFinder, and RestrictionModificationFinder. For phylogenetic analysis, CSI Phylogeny will call single-nucleotide polymorphisms (SNPs), filter and validate them, and then infer phylogeny based on the concatenated alignment of the SNPs, generating phylogenetic trees. Also available for phylogenetic analyses are the downloadable software, RAxML (Randomized Axelerated Maximum Likelihood), for sequential and parallel maximum likelihood-based inference of large phylogenetic trees [114], as well as BEAST (Bayesian Evolutionary Analysis Sampling Trees), for inferring rooted, time-measured phylogenies using molecular clock models [115, 116]. Available from the University of Alberta (at http://phaster.ca/) is a web-based tool for rapid identification and annotation of prophage sequences within a bacterial genome, known as PHASTER (PHAge Search Tool – Enhanced Release). The program is able to work on raw DNA sequences as well as annotated GenBank formatted data, providing detailed tables and graphical displays of the phages, with high sensitivity and positive predictive value [117, 118].

WGS is the ultimate tool for the identification of diversity in an organism. In addition to extracting *S. aureus* typing information, WGS data can be used to track transmission events and outbreaks [119–121] and analyze variations between strains within a lineage by SNP analysis [122]. It has shown that related strains have well-conserved core regions but differ in their accessory genetic elements [123] and, likewise, that geographically dispersed isolates of ST239, ST225, and CC30 are stable in their genetic backgrounds, differing by SNPs and MGEs [119, 124, 125]. In the future, we may see the application of extended MLST (eMLST) to *S. aureus* typing, extending typing beyond the seven housekeeping genes to include a subset or all of the genes in the genome. Ribosomal MLST (rMLST) (adding the ribosomal genes), core genome MLST (cgMLST) (including all core genes present in the majority of isolates, and not subject to selection pressure), whole genome MLST (wgMLST) (also including genes subject to selective pressure), and pan-genome approach (including the full complement of genes within the species) would provide the ultimate high-level genomic epidemiology. Available to facilitate eMLST analysis, the Bacterial Isolate Genome Sequence Database (BIGSdb) software stores and analyzes sequence data for bacterial isolates, allowing a large numbers of loci to be defined and allelic profiles for each strain to be determined. BIGSdb is

available within the PubMLST database at https://pubmlst.org/software/database/bigsdb/.

As sequencing costs are reduced and genome analysis tools improve, WGS will almost certainly become the primary tool for *S. aureus* typing and evolutionary and epidemiological studies.

9.7 Conclusions

Each typing scheme for *S. aureus* is met with strengths and limitations, leaving no single method ideal for all situations. PFGE was once considered the gold standard for MRSA typing and remains an effective tool for characterizing outbreaks and understanding *S. aureus* epidemiology, particularly at the local level. With standardization, it can be expanded to the international level; however, lineage cannot be inferred directly from the PFGE pattern. *spa* typing is capable of assigning lineage, is useful for analyzing both outbreaks and long-term molecular evolution, and is rapidly becoming the method of choice for clinical laboratories for epidemiological studies of *S. aureus.* With highly portable and standardized data, it is useful for investigations at both the local and international levels but is not always accurate when assigning lineages. MLST is also an effective tool for assigning lineage and, in combination with SCC*mec* typing, is considered the gold standard for publishing *S. aureus* epidemiological data. Similar to *spa* typing, the data is highly standardized and portable, making it an effective tool for studies at both the local and international levels. However, the cost makes it less appealing for routine use. Microarrays can provide large amounts of strain information within a short timeframe and are well suited for both outbreak investigations and long-term epidemiological studies, particularly at the local level, but suffer in that they cannot directly assign strains to lineages. WGS is the ultimate tool for strain typing and epidemiological studies and will rapidly increase in use as sequencing costs decrease and as easy-to-use data analysis tools are developed.

Ultimately, the technique of choice will depend heavily on the goals and questions that need answering, with a combination of methods offering more detailed information and greater discrimination between isolates. For outbreak situations where speed is important, PCR-based methods may be the better choice, making *spa* typing an effective tool. However, for routine strain typing and epidemiological monitoring at the local level, PFGE and *spa* typing complement well, providing better strain and clone discrimination. For international comparisons, *spa* typing, MLST, and WGS are good for generating highly standardized and portable data, but when detailed strain characterization is desired, a combination of PFGE, *agr* typing, SCC*mec* typing, *spa* typing, and MLST provides a more complete picture. Finally, long-term epidemiological and evolutionary studies benefit from greater detail, making microarrays and WGS attractive options.

9.8 Summary

Staphylococci are Gram-positive bacteria and commonly divided into coagulase-negative staphylococci (CoNS) and coagulase-positive members, based on their ability to produce the free enzyme coagulase. The majority of staphylococcal species are CoNS, with an increasing number of CoNS infections identified, boosting their clinical significance. *Staphylococcus aureus* is coagulase positive and has been the primary focus of clinical identification as it is commonly associated with human infection. Methicillin-resistant *S. aureus* (MRSA), in particular, has garnered much attention as it is resistant to all penicillins and most β-lactam drugs and is associated with higher morbidity and mortality rates and increasingly being recognized as presenting a significant threat to public health. With the ever-changing prevalence and epidemiology of staphylococcal infections, reliable methods for characterizing strains are essential for outbreak investigations, for tracking clonal spreading, and for the implementation of effective treatment or control measures. In this chapter, we discussed various phenotypic and molecular methods used to discriminate staphylococci and *S. aureus* lineages. We first described the methods to identify staphylococcal species and to discriminate MRSA from methicillin-susceptible *S. aureus* (MSSA), including how to characterize different types of staphylococcal cassette chromosome *mec* (SCC*mec*) in MRSA. We then discussed various typing methods applied to study the molecular epidemiology and evolutionary nature of *S. aureus*, starting with the historical methods [phage typing, multilocus enzyme electrophoresis (MLEE), random amplification of polymorphic DNA (RAPD) and arbitrarily primed PCR (AP-PCR), repetitive element PCR (rep-PCR), amplified fragment length polymorphism (AFLP), and accessory gene regulator (agr) typing] and continuing to the current commonly used molecular typing methods [pulsed-field gel electrophoresis (PFGE) typing, staphylococcal protein A (*spa*) typing, multilocus sequence typing (MLST), and microarray] and to the advanced genome approaches (whole genome sequencing). We also discuss the strengths and limitations for each typing scheme and their suitable applications.

References

1. Banerjee SN, Emori TG, Culver DH, Gaynes RP, Jarvis WR, Horan T, Edwards JR, Tolson J, Henderson T, Martone WJ (1991) Secular trends in nosocomial primary bloodstream infections in the United States, 1980–1989. National Nosocomial Infections Surveillance System. Am J Med 91:86S–89S
2. Rupp ME, Archer GL (1994) Coagulase-negative staphylococci: pathogens associated with medical progress. Clin Infect Dis 19:231–243. quiz 244–245
3. Antonanzas F, Lozano C, Torres C (2015) Economic features of antibiotic resistance: the case of methicillin-resistant *Staphylococcus aureus*. Pharmacoeconomics 33:285–325
4. Thampi N, Showler A, Burry L, Bai AD, Steinberg M, Ricciuto DR, Bell CM, Morris AM (2015) Multicenter study of health care cost of patients admitted to hospital with *Staphylococcus aureus* bacteremia: impact of length of stay and intensity of care. Am J Infect Control 43:739–744

5. Whitby M, McLaws ML, Berry G (2001) Risk of death from methicillin-resistant *Staphylococcus aureus* bacteraemia: a meta-analysis. Med J Aust 175:264–267
6. Kluytmans J, van Belkum A, Verbrugh H (1997) Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev 10:505–520
7. Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, Nouwen JL (2005) The role of nasal carriage in *Staphylococcus aureus* infections. Lancet Infect Dis 5:751–762
8. Diekema DJ, Pfaller MA, Schmitz FJ, Smayevsky J, Bell J, Jones RN, Beach M, Group SP (2001) Survey of infections due to *Staphylococcus species*: frequency of occurrence and antimicrobial susceptibility of isolates collected in the United States, Canada, Latin America, Europe, and the Western Pacific region for the SENTRY Antimicrobial Surveillance Program, 1997–1999. Clin Infect Dis 32(Suppl 2):S114–S132
9. David MZ, Boyle-Vavra S, Zychowski DL, Daum RS (2011) Methicillin-susceptible *Staphylococcus aureus* as a predominantly healthcare-associated pathogen: a possible reversal of roles? PLoS One 6:e18217
10. Jackson KA, Gokhale RH, Nadle J, Ray SM, Dumyati G, Schaffner W, Ham DC, Magill SS, Lynfield R, See I (2019) Public health importance of invasive methicillin-sensitive *Staphylococcus aureus* infections: surveillance in 8 US Counties, 2016. Clin Infect Dis doi:5477381 [pii]https://doi.org/10.1093/cid/ciz323
11. Bannerman TL, Kleeman KT, Kloos WE (1993) Evaluation of the Vitek Systems Gram-Positive Identification card for species identification of coagulase-negative staphylococci. J Clin Microbiol 31:1322–1325
12. Ieven M, Verhoeven J, Pattyn SR, Goossens H (1995) Rapid and economical method for species identification of clinically significant coagulase-negative staphylococci. J Clin Microbiol 33:1060–1063
13. Spanu T, Sanguinetti M, Ciccaglione D, D'Inzeo T, Romano L, Leone F, Fadda G (2003) Use of the VITEK 2 system for rapid identification of clinical isolates of Staphylococci from bloodstream infections. J Clin Microbiol 41:4259–4263
14. Mellmann A, Becker K, von Eiff C, Keckevoet U, Schumann P, Harmsen D (2006) Sequencing and staphylococci identification. Emerg Infect Dis 12:333–336
15. Gray MW, Sankoff D, Cedergren RJ (1984) On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. Nucleic Acids Res 12:5837–5852
16. Conlan S, Kong HH, Segre JA (2012) Species-level analysis of DNA sequence data from the NIH Human Microbiome Project. PLoS One 7:e47075
17. Larsen N, Olsen GJ, Maidak BL, McCaughey MJ, Overbeek R, Macke TJ, Marsh TL, Woese CR (1993) The ribosomal database project. Nucleic Acids Res 21:3021–3023
18. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6:610–618
19. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196
20. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617
21. Chambers HF, Deleo FR (2009) Waves of resistance: *Staphylococcus aureus* in the antibiotic era. Nat Rev Microbiol 7:629–641
22. Garcia-Alvarez L, Holden MT, Lindsay H, Webb CR, Brown DF, Curran MD, Walpole E, Brooks K, Pickard DJ, Teale C, Parkhill J, Bentley SD, Edwards GF, Girvan EK, Kearns AM, Pichon B, Hill RL, Larsen AR, Skov RL, Peacock SJ, Maskell DJ, Holmes MA (2011) Methicillin-resistant *Staphylococcus aureus* with a novel *mecA* homologue in human and bovine populations in the UK and Denmark: a descriptive study. Lancet Infect Dis 11:595–603

23. Baba T, Kuwahara-Arai K, Uchiyama I, Takeuchi F, Ito T, Hiramatsu K (2009) Complete genome sequence of *Macrococcus caseolyticus* strain JCSCS5402, [corrected] reflecting the ancestral genome of the human-pathogenic staphylococci. J Bacteriol 191:1180–1190
24. Becker K, van Alen S, Idelevich EA, Schleimer N, Seggewiss J, Mellmann A, Kaspar U, Peters G (2018) Plasmid-Encoded Transferable *mecB*-Mediated Methicillin Resistance in *Staphylococcus aureus*. Emerg Infect Dis 24:242–248
25. CLSI (2018) Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 11th edn, CLSI Standard M07-A11. Clinical and Laboratory Standards Institute, Wayne
26. Xu Z, Hou Y, Peters BM, Chen D, Li B, Li L, Shirtliff ME (2016) Chromogenic media for MRSA diagnostics. Mol Biol Rep 43:1205–1212
27. Sakoulas G, Gold HS, Venkataraman L, DeGirolami PC, Eliopoulos GM, Qian Q (2001) Methicillin-resistant *Staphylococcus aureus*: comparison of susceptibility testing methods and analysis of *mecA*-positive susceptible strains. J Clin Microbiol 39:3946–3951
28. Chapin KC, Musgnug MC (2004) Evaluation of penicillin binding protein 2a latex agglutination assay for identification of methicillin-resistant *Staphylococcus aureus* directly from blood cultures. J Clin Microbiol 42:1283–1284
29. Murakami K, Minamide W, Wada K, Nakamura E, Teraoka H, Watanabe S (1991) Identification of methicillin-resistant strains of staphylococci by polymerase chain reaction. J Clin Microbiol 29:2240–2244
30. Geha DJ, Uhl JR, Gustaferro CA, Persing DH (1994) Multiplex PCR for identification of methicillin-resistant staphylococci in the clinical laboratory. J Clin Microbiol 32:1768–1772
31. Zhang K, McClure JA, Elsayed S, Louie T, Conly JM (2008) Novel multiplex PCR assay for simultaneous identification of community-associated methicillin-resistant *Staphylococcus aureus* strains USA300 and USA400 and detection of *mecA* and Panton-Valentine leukocidin genes, with discrimination of *Staphylococcus aureus* from coagulase-negative staphylococci. J Clin Microbiol 46:1118–1122
32. Stegger M, Andersen PS, Kearns A, Pichon B, Holmes MA, Edwards G, Laurent F, Teale C, Skov R, Larsen AR (2012) Rapid detection, differentiation and typing of methicillin-resistant *Staphylococcus aureus* harbouring either *mecA* or the new *mecA* homologue *mecA*(LGA251). Clin Microbiol Infect 18:395–400
33. Hiramatsu K, Asada K, Suzuki E, Okonogi K, Yokota T (1992) Molecular cloning and nucleotide sequence determination of the regulator region of *mecA* gene in methicillin-resistant *Staphylococcus aureus* (MRSA). FEBS Lett 298:133–136
34. Ito T, Katayama Y, Asada K, Mori N, Tsutsumimoto K, Tiensasitorn C, Hiramatsu K (2001) Structural comparison of three types of staphylococcal cassette chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 45:1323–1336
35. Ito T, Katayama Y, Hiramatsu K (1999) Cloning and nucleotide sequence determination of the entire *mec* DNA of pre-methicillin-resistant *Staphylococcus aureus* N315. Antimicrob Agents Chemother 43:1449–1458
36. Okuma K, Iwakawa K, Turnidge JD, Grubb WB, Bell JM, O'Brien FG, Coombs GW, Pearman JW, Tenover FC, Kapi M, Tiensasitorn C, Ito T, Hiramatsu K (2002) Dissemination of new methicillin-resistant *Staphylococcus aureus* clones in the community. J Clin Microbiol 40:4289–4294
37. van der Zee A, Heck M, Sterks M, Harpal A, Spalburg E, Kazobagora L, Wannet W (2005) Recognition of SCC*mec* types according to typing pattern determined by multienzyme multiplex PCR-amplified fragment length polymorphism analysis of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 43:6042–6047
38. Yang JA, Park DW, Sohn JW, Kim MJ (2006) Novel PCR-restriction fragment length polymorphism analysis for rapid typing of staphylococcal cassette chromosome *mec* elements. J Clin Microbiol 44:236–238

39. Oliveira DC, de Lencastre H (2002) Multiplex PCR strategy for rapid identification of structural types and variants of the *mec* element in methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 46:2155–2161
40. Chen L, Mediavilla JR, Oliveira DC, Willey BM, de Lencastre H, Kreiswirth BN (2009) Multiplex real-time PCR for rapid Staphylococcal cassette chromosome *mec* typing. J Clin Microbiol 47:3692–3706
41. Francois P, Renzi G, Pittet D, Bento M, Lew D, Harbarth S, Vaudaux P, Schrenzel J (2004) A novel multiplex real-time PCR assay for rapid typing of major staphylococcal cassette chromosome *mec* elements. J Clin Microbiol 42:3309–3312
42. Monecke S, Slickers P, Ehricht R (2008) Assignment of *Staphylococcus aureus* isolates to clonal complexes based on microarray analysis and pattern recognition. FEMS Immunol Med Microbiol 53:237–251
43. Zhang K, McClure JA, Conly JM (2012) Enhanced multiplex PCR assay for typing of staphylococcal cassette chromosome *mec* types I to V in methicillin-resistant *Staphylococcus aureus*. Mol Cell Probes 26:218–221
44. Zhang K, McClure JA, Elsayed S, Louie T, Conly JM (2005) Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome *mec* types I to V in methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 43:5026–5033
45. Blair JE, Williams RE (1961) Phage typing of staphylococci. Bull World Health Organ 24:771–784
46. Weller TM (2000) Methicillin-resistant *Staphylococcus aureus* typing methods: which should be the international standard? J Hosp Infect 44:160–172
47. Blair JE (1966) Untypable staphylococci: their identification and possible origin. Health Lab Sci 3:229–234
48. Bannerman TL, Hancock GA, Tenover FC, Miller JM (1995) Pulsed-field gel electrophoresis as a replacement for bacteriophage typing of *Staphylococcus aureus*. J Clin Microbiol 33:551–555
49. Lundholm M, Bergendahl B (1988) Heat treatment to increase phage typability of *Staphylococcus aureus*. Eur J Clin Microbiol Infect Dis 7:300–302
50. Mehndiratta PL, Bhalla P (2012) Typing of Methicillin resistant *Staphylococcus aureus*: a technical review. Indian J Med Microbiol 30:16–23
51. Mulligan ME, Arbeit RD (1991) Epidemiologic and clinical utility of typing systems for differentiating among strains of methicillin-resistant *Staphylococcus aureus*. Infect Control Hosp Epidemiol 12:20–28
52. Tenover FC, Arbeit R, Archer G, Biddle J, Byrne S, Goering R, Hancock G, Hebert GA, Hill B, Hollis R et al (1994) Comparison of traditional and molecular methods of typing isolates of *Staphylococcus aureus*. J Clin Microbiol 32:407–415
53. Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218
54. Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
55. Struelens MJ, Bax R, Deplano A, Quint WG, Van Belkum A (1993) Concordant clonal delineation of methicillin-resistant *Staphylococcus aureus* by macrorestriction analysis and polymerase chain reaction genome fingerprinting. J Clin Microbiol 31:1964–1970
56. Saulnier P, Bourneix C, Prevost G, Andremont A (1993) Random amplified polymorphic DNA assay is less discriminant than pulsed-field gel electrophoresis for typing strains of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 31:982–985
57. van Belkum A, Kluytmans J, van Leeuwen W, Bax R, Quint W, Peters E, Fluit A, Vandenbroucke-Grauls C, van den Brule A, Koeleman H et al (1995) Multicenter evaluation of arbitrarily primed PCR for typing of *Staphylococcus aureus* strains. J Clin Microbiol 33:1537–1547

58. Versalovic J, Koeuth T, Lupski JR (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res 19:6823–6831
59. van der Zee A, Verbakel H, van Zon JC, Frenay I, van Belkum A, Peeters M, Buiting A, Bergmans A (1999) Molecular genotyping of *Staphylococcus aureus* strains: comparison of repetitive element sequence-based PCR with various typing methods and isolation of a novel epidemicity marker. J Clin Microbiol 37:342–349
60. Sabat A, Malachowa N, Miedzobrodzki J, Hryniewicz W (2006) Comparison of PCR-based methods for typing *Staphylococcus aureus* isolates. J Clin Microbiol 44:3804–3807
61. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
62. Mortimer P, Arnold C (2001) FAFLP: last word in microbial genotyping? J Med Microbiol 50:393–395
63. Duim B, Wassenaar TM, Rigter A, Wagenaar J (1999) High-resolution genotyping of *Campylobacter* strains isolated from poultry and humans with amplified fragment length polymorphism fingerprinting. Appl Environ Microbiol 65:2369–2375
64. Zhao S, Mitchell SE, Meng J, Kresovich S, Doyle MP, Dean RE, Casa AM, Weller JW (2000) Genomic typing of *Escherichia coli* O157:H7 by semi-automated fluorescent AFLP analysis. Microbes Infect 2:107–113
65. Novick RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol 48:1429–1449
66. Peacock SJ, Moore CE, Justice A, Kantzanou M, Story L, Mackie K, O'Neill G, Day NP (2002) Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. Infect Immun 70:4987–4996
67. Francois P, Koessler T, Huyghe A, Harbarth S, Bento M, Lew D, Etienne J, Pittet D, Schrenzel J (2006) Rapid *Staphylococcus aureus* agr type determination by a novel multiplex real-time quantitative PCR assay. J Clin Microbiol 44:1892–1895
68. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG (2002) The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci U S A 99:7687–7692
69. Goering RV (2010) Pulsed field gel electrophoresis: a review of application and interpretation in the molecular epidemiology of infectious disease. Infect Genet Evol 10:866–875
70. Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37:67–75
71. Matushek MG, Bonten MJ, Hayden MK (1996) Rapid preparation of bacterial DNA for pulsed-field gel electrophoresis. J Clin Microbiol 34:2598–2600
72. McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC (2003) Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. J Clin Microbiol 41:5113–5120
73. Murchan S, Kaufmann ME, Deplano A, de Ryck R, Struelens M, Zinn CE, Fussing V, Salmenlinna S, Vuopio-Varkila J, El Solh N, Cuny C, Witte W, Tassios PT, Legakis N, van Leeuwen W, van Belkum A, Vindel A, Laconcha I, Garaizar J, Haeggman S, Olsson-Liljequist B, Ransjo U, Coombes G, Cookson B (2003) Harmonization of pulsed-field gel electrophoresis protocols for epidemiological typing of strains of methicillin-resistant *Staphylococcus aureus*: a single approach developed by consensus in 10 European laboratories and its application for tracing the spread of related strains. J Clin Microbiol 41:1574–1585
74. Mulvey MR, Chui L, Ismail J, Louie L, Murphy C, Chang N, Alfa M (2001) Development of a Canadian standardized protocol for subtyping methicillin-resistant *Staphylococcus aureus* using pulsed-field gel electrophoresis. J Clin Microbiol 39:3481–3485
75. Carles-Nurit MJ, Christophle B, Broche S, Gouby A, Bouziges N, Ramuz M (1992) DNA polymorphisms in methicillin-susceptible and methicillin-resistant strains of *Staphylococcus aureus*. J Clin Microbiol 30:2092–2096

76. Ichiyama S, Ohta M, Shimokata K, Kato N, Takeuchi J (1991) Genomic DNA fingerprinting by pulsed-field gel electrophoresis as an epidemiological marker for study of nosocomial infections caused by methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 29:2690–2695
77. Prevost G, Jaulhac B, Piemont Y (1992) DNA fingerprinting by pulsed-field gel electrophoresis is more effective than ribotyping in distinguishing among methicillin-resistant *Staphylococcus aureus* isolates. J Clin Microbiol 30:967–973
78. Argudin MA, Rodicio MR, Guerra B The emerging methicillin-resistant *Staphylococcus aureus* ST398 clone can easily be typed using the Cfr9I SmaI-neoschizomer. Lett Appl Microbiol 50:127–130
79. Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234:1582–1585
80. Cookson BD, Robinson DA, Monk AB, Murchan S, Deplano A, de Ryck R, Struelens MJ, Scheel C, Fussing V, Salmenlinna S, Vuopio-Varkila J, Cuny C, Witte W, Tassios PT, Legakis NJ, van Leeuwen W, van Belkum A, Vindel A, Garaizar J, Haeggman S, Olsson-Liljequist B, Ransjo U, Muller-Premru M, Hryniewicz W, Rossney A, O'Connell B, Short BD, Thomas J, O'Hanlon S, Enright MC (2007) Evaluation of molecular typing methods in characterizing a European collection of epidemic methicillin-resistant *Staphylococcus aureus* strains: the HARMONY collection. J Clin Microbiol 45:1830–1837
81. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33:2233–2239
82. Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, Force CDCPT (2001) PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. Emerg Infect Dis 7:382–389
83. Cookson BD, Aparicio P, Deplano A, Struelens M, Goering R, Marples R (1996) Intercentre comparison of pulsed-field gel electrophoresis for the typing of methicillin-resistant *Staphylococcus aureus*. J Med Microbiol 44:179–184
84. van Belkum A, van Leeuwen W, Kaufmann ME, Cookson B, Forey F, Etienne J, Goering R, Tenover F, Steward C, O'Brien F, Grubb W, Tassios P, Legakis N, Morvan A, El Solh N, de Ryck R, Struelens M, Salmenlinna S, Vuopio-Varkila J, Kooistra M, Talens A, Witte W, Verbrugh H (1998) Assessment of resolution and intercenter reproducibility of results of genotyping *Staphylococcus aureus* by pulsed-field gel electrophoresis of SmaI macrorestriction fragments: a multicenter study. J Clin Microbiol 36:1653–1659
85. Frenay HM, Bunschoten AE, Schouls LM, van Leeuwen WJ, Vandenbroucke-Grauls CM, Verhoef J, Mooi FR (1996) Molecular typing of methicillin-resistant *Staphylococcus aureus* on the basis of protein A gene polymorphism. Eur J Clin Microbiol Infect Dis 15:60–64
86. Koreen L, Ramaswamy SV, Graviss EA, Naidich S, Musser JM, Kreiswirth BN (2004) *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol 42:792–799
87. Shopsin B, Gomez M, Montgomery SO, Smith DH, Waddington M, Dodge DE, Bost DA, Riehman M, Naidich S, Kreiswirth BN (1999) Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus* aureus strains. J Clin Microbiol 37:3556–3563
88. Harmsen D, Claus H, Witte W, Rothganger J, Claus H, Turnwald D, Vogel U (2003) Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. J Clin Microbiol 41:5442–5448
89. Strommenger B, Kettlitz C, Weniger T, Harmsen D, Friedrich AW, Witte W (2006) Assignment of *Staphylococcus* isolates to groups by *spa* typing, SmaI macrorestriction analysis, and multilocus sequence typing. J Clin Microbiol 44:2533–2540

90. Mellmann A, Weniger T, Berssenbrugge C, Rothganger J, Sammeth M, Stoye J, Harmsen D (2007) Based Upon Repeat Pattern (BURP): an algorithm to characterize the long-term evolution of *Staphylococcus aureus* populations based on *spa* polymorphisms. BMC Microbiol 7:98
91. Nubel U, Roumagnac P, Feldkamp M, Song JH, Ko KS, Huang YC, Coombs G, Ip M, Westh H, Skov R, Struelens MJ, Goering RV, Strommenger B, Weller A, Witte W, Achtman M (2008) Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci USA 105:14130–14135
92. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG (2000) Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol 38:1008–1015
93. Aanensen DM, Spratt BG (2005) The multilocus sequence typing network: mlst.net. Nucleic Acids Res 33:W728–W733
94. Jolley KA, Maiden MC (2010) BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595
95. Jolley KA, Bray JE, Maiden MCJ (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124
96. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186:1518–1530
97. Spratt BG, Hanage WP, Li B, Aanensen DM, Feil EJ (2004) Displaying the relatedness among isolates of bacterial species -- the eBURST approach. FEMS Microbiol Lett 241:129–134
98. Ramsay G (1998) DNA chips: state-of-the art. Nat Biotechnol 16:40–44
99. Sanchini A, Campanile F, Monaco M, Cafiso V, Rasigade JP, Laurent F, Etienne J, Stefani S, Pantosti A (2011) DNA microarray-based characterisation of Panton-Valentine leukocidin-positive community-acquired methicillin-resistant *Staphylococcus aureus* from Italy. Eur J Clin Microbiol Infect Dis 30:1399–1408
100. Lindsay JA, Moore CE, Day NP, Peacock SJ, Witney AA, Stabler RA, Husain SE, Butcher PD, Hinds J (2006) Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J Bacteriol 188:669–676
101. McCarthy AJ, Breathnach AS, Lindsay JA (2012) Detection of mobile-genetic-element variation between colonizing and infecting hospital-associated methicillin-resistant *Staphylococcus aureus* isolates. J Clin Microbiol 50:1073–1075
102. McCarthy AJ, Lindsay JA (2012) The distribution of plasmids that carry virulence and resistance genes in *Staphylococcus aureus* is lineage associated. BMC Microbiol 12:104
103. Bischoff M, Dunman P, Kormanec J, Macapagal D, Murphy E, Mounts W, Berger-Bachi B, Projan S (2004) Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. J Bacteriol 186:4085–4099
104. Dunman PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, Wu S, Brown EL, Zagursky RJ, Shlaes D, Projan SJ (2001) Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. J Bacteriol 183:7341–7353
105. Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY (2006) Transcription Profiling of the *mgrA* Regulon in *Staphylococcus aureus*. J Bacteriol 188:1899–1910
106. Dunman PM, Mounts W, McAleese F, Immermann F, Macapagal D, Marsilio E, McDougal L, Tenover FC, Bradford PA, Petersen PJ, Projan SJ, Murphy E (2004) Uses of *Staphylococcus aureus* GeneChips in genotyping and genetic composition analysis. J Clin Microbiol 42:4275–4283
107. Monecke S, Ehricht R (2005) Rapid genotyping of methicillin-resistant *Staphylococcus aureus* (MRSA) isolates using miniaturised oligonucleotide arrays. Clin Microbiol Infect 11:825–833

108. Otsuka J, Kondoh Y, Amemiya T, Kitamura A, Ito T, Baba S, Cui L, Hiramatsu K, Tashiro T, Tashiro H (2008) Development and validation of microarray-based assay for epidemiological study of MRSA. Mol Cell Probes 22:1–13
109. Saunders NA, Underwood A, Kearns AM, Hallas G (2004) A virulence-associated gene microarray: a tool for investigation of the evolution and pathogenic potential of *Staphylococcus aureus*. Microbiology 150:3763–3771
110. Strommenger B, Schmidt C, Werner G, Roessle-Lorch B, Bachmann TT, Witte W (2007) DNA microarray for the detection of therapeutically relevant antibiotic resistance determinants in clinical isolates of *Staphylococcus aureus*. Mol Cell Probes 21:161–170
111. Monecke S, Jatzwauk L, Weber S, Slickers P, Ehricht R (2008) DNA microarray-based genotyping of methicillin-resistant *Staphylococcus aureus* strains from Eastern Saxony. Clin Microbiol Infect 14:534–545
112. Monecke S, Coombs G, Shore AC, Coleman DC, Akpaka P, Borg M, Chow H, Ip M, Jatzwauk L, Jonas D, Kadlec K, Kearns A, Laurent F, O'Brien FG, Pearson J, Ruppelt A, Schwarz S, Scicluna E, Slickers P, Tan HL, Weber S, Ehricht R (2011) A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. PLoS One 6:e17936
113. McCarthy AJ, Witney AA, Gould KA, Moodley A, Guardabassi L, Voss A, Denis O, Broens EM, Hinds J, Lindsay JA (2011) The distribution of mobile genetic elements (MGEs) in MRSA CC398 is associated with both host and country. Genome Biol Evol 3:1164–1174
114. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313
115. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchene S, Fourment M, Gavryushkina A, Heled J, Jones G, Kuhnert D, De Maio N, Matschiner M, Mendes FK, Muller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15:e1006650
116. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4:vey016
117. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21
118. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352
119. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD (2010) Evolution of MRSA during hospital transmission and intercontinental spread. Science 327:469–474
120. Harrison EM, Paterson GK, Holden MT, Larsen J, Stegger M, Larsen AR, Petersen A, Skov RL, Christensen JM, Bak Zeuthen A, Heltberg O, Harris SR, Zadoks RN, Parkhill J, Peacock SJ, Holmes MA (2013) Whole genome sequencing identifies zoonotic transmission of MRSA isolates with the novel *mecA* homologue *mecC*. EMBO Mol Med 5:509–515
121. Price JR, Golubchik T, Cole K, Wilson DJ, Crook DW, Thwaites GE, Bowden R, Walker AS, Peto TE, Paul J, Llewelyn MJ (2014) Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of *Staphylococcus aureus* in an intensive care unit. Clin Infect Dis 58:609–618
122. Koser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, Hsu LY, Chewapreecha C, Croucher NJ, Harris SR, Sanders M, Enright MC, Dougan G, Bentley SD, Parkhill J, Fraser LJ, Betley JR, Schulz-Trieglaff OB, Smith GP, Peacock SJ (2012) Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med 366:2267–2275

123. Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabbinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J (2004) Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci USA 101:9786–9791
124. McAdam PR, Templeton KE, Edwards GF, Holden MT, Feil EJ, Aanensen DM, Bargawi HJ, Spratt BG, Bentley SD, Parkhill J, Enright MC, Holmes A, Girvan EK, Godfrey PA, Feldgarden M, Kearns AM, Rambaut A, Robinson DA, Fitzgerald JR (2012) Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci USA 109:9107–9112
125. Nubel U, Dordel J, Kurt K, Strommenger B, Westh H, Shukla SK, Zemlickova H, Leblois R, Wirth T, Jombart T, Balloux F, Witte W (2010) A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. PLoS Pathog 6:e1000855

Part III

Vector-Borne and Biosafety Level 3 Pathogens

Bartonellaceae

10

Joaquim Ruiz, Cláudia Gomes, and Maria J. Pons

10.1 Introduction

The genus *Bartonella* (in honor of Alberto L. Barton) is the only member of the family *Bartonellaceae*, which is classified on the order *Rhizobiales*, class *Alphaproteobacteria*, phylum *Proteobacteria* (https://lpsn.dsmz.de/family/bartonellaceae).

Currently, the *Bartonella* genus accounts for 37 species with a standing in nomenclature (https://lpsn.dsmz.de/search?word=bartonella), resulting from the fusion of the former *Bartonella*, *Grahamella*, and *Rochalimaea* genera and with increasing and continuous descriptions of new species having been made from the early 2000 onward [1–3]. In addition, there is a large and undefined series of *Candidatus* Bartonella spp. in the literature. Of these, several have been cultured and fully characterized, but mostly remain described from partial DNA sequencing data obtained through molecular or metagenomic studies [4–8].

While *Bartonella apis* has been described as a honey bee symbiont [9], and in a few cases the host or vectors remain to be described [3], almost all *Bartonella* spp. are arthropod-vectorized, with different mammal species, including humans (i.e., *Bartonella bacilliformis*, *Bartonella quintana*, and perhaps *Bartonella ancashensis*) acting as reservoirs [3, 10]. Regarding *B. apis*, despite its symbiont role, it has been recovered from blood samples of apparent healthy dogs [11].

Several *Bartonella* spp. are recognized as human pathogens causing a wide variety of infections such as bacillary angiomatosis, bacteremia, Carrion's disease, cat

J. Ruiz (✉) · M. J. Pons
Laboratorio de Microbiología Molecular y Genómica Bacteriana, Universidad Científica del Sur, Lima, Peru

C. Gomes
Department of Microbiology, New York University Grossman School of Medicine, New York, NY, USA

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_10

scratch disease, chronic fatigue, culture-negative endocarditis, lymphadenitis, meningitis, myocarditis, neuroretinitis, osteomyelitis, peliosis hepatica, pericarditis, and trench fever [3, 10, 12–14]. Of these, *B. bacilliformis*, *Bartonella henselae,* and *B. quintana* account for a majority of cases [3, 10, 14].

10.2 Phenotypic Techniques

Bartonella spp. are hemin-dependent pleomorphic non-fermentative fastidious facultative intracellular slow-growing microorganisms with a G + C DNA content ranging from 37 to 41 mol % [3, 10, 15]. A few exceptions include *B. apis*, with a G + C content of 45.5 mol % [9]. *Bartonella* spp. grow in aerobic or microaerophilic conditions with an optimal culture temperature ranging between 35 °C and 37 °C [10], with exceptions such as *B. bacilliformis*, which is usually cultured at 28–30 °C [3]. *Bartonella* spp. usually grows forming small colonies, which vary from translucent to opaque and from white to tan in color [3, 15]. Furthermore, intra-strain colony morphology variations may be observed during consecutive cultures [3]. The motility of *Bartonella* spp. is variable and is related to the presence or absence of a variable number of flagella. Thus, while most members of the genus lack flagella and are non-motile, at least 9 *Bartonella* species with standing in nomenclature, the most representative including *B. bacilliformis* or *Bartonella clarridgeiae*, present a varying number of flagella conferring motility [3]. In addition, several non-flagellate *Bartonella*, such as *B. henselae* or *B. quintana*, present pili-related twitching motility [16].

While a few biochemical tests, such as L-proline and L-lysine peptidase activity tests, cannot identify the final species, they can contribute to discarding a series of species. In fact, *Bartonella* are typically almost inert to biochemical approaches, which are thereby of no use for identification purposes. Of note, in 1993 Drancourt and Raoult reported that the addition of 100 μg/ml of hemin to different Analytical Profile Index strips enabled the identification of differential biochemical patterns for *B. quintana*, *B. henselae*, and *Bartonella vinsonii* [17]. However, to our knowledge these studies have not been expanded to the vast majority of *Bartonella* spp.

Although *Bartonella* spp. are Gram-negative bacteria, Gram-staining provides poor results, with the most reliable methods being Giemsa, Gimenez, Romanowsky, or Warthin-Starring staining [14, 15]. These techniques may allow direct detection of the presence of *Bartonella* spp. in tissues or blood samples, but they cannot identify *Bartonella* species [10]. It is of note that despite being mostly expertise-dependent, Giemsa staining is the standard diagnostic tool for identifying *B. bacilliformis* in Carrion's disease which is endemic in the rural areas of the South American Andes [3].

Similar to staining techniques, the use of gas chromatography analysis of fatty acids may be used to identify *Bartonella* spp. at a genus level [18]. In addition, fatty acid composition may differ among different *Bartonella* spp. and while it does not allow univocal identification, it may be useful to rule out several species.

In summary, phenotypical techniques have a very limited usefulness for differentiating *Bartonella* spp.

10.3 Molecular Techniques

A series of molecular techniques have been explored in order to identify species or to type isolates at different clonal levels, with most of these techniques being based on different approaches of DNA analysis. While several molecular tools may be used in the absence of bacterial cultures [19, 20], the characteristics of *Bartonella* spp. prevent several classical approaches, such as plasmid typing, because of the scarce number of plasmid descriptions of members of this genus having been made [21, 22]. In addition, most studies have been focused on the main pathogenic species (*B. henselae*, *B. quintana* and *B. bacilliformis*), with a limited number of studies having focused on other *Bartonella* spp. with their subsequent underrepresentation in typing studies.

10.3.1 Species Identification

10.3.1.1 16S-Based Techniques

The *16S ribosomal RNA* (*rRNA)* gene is highly conserved and has been extensively used in phylogenetic studies. Regarding *Bartonella* spp., different typing approaches have been developed based on the use of this gene.

16S sequencing: Sequencing of the *16S rRNA* gene is an important tool for bacterial detection and identification. The following features make this a useful molecular target to study bacterial phylogeny and taxonomy: (i) it is present in almost all bacteria; (ii) its functions have not changed over time; and (iii) the length of this gene (1500 bp) is large enough for informatic purposes [23]. In fact, the first association of *Bartonella* with bacillary angiomatosis was due to the detection of *Bartonella 16S rRNA* gene sequences in the lesions of patients [24]. Similarly, the association of *Bartonella* spp. with cat scratch disease was established by the detection of *B. henselae 16S rRNA* gene sequences in skin test antigen preparations used to diagnose this illness [25]. Later, a rapid, sensitive, and reliable method to generate partial *16S rRNA* sequences of *B. henselae* and *B. quintana* directly from normally sterile clinical specimens was proposed by Goldenberg and colleagues [26]. The amplification of the *16S rRNA* gene has also been proposed for the diagnosis of Carrion's disease in Peru. The authors suggest using both *Bartonella*-specific and universal *16S rRNA* gene primers to detect *B. bacilliformis,* and the use of universal *16S rRNA* gene primers facilitating the detection of other bacterial pathogens [27]. Indeed, in an outbreak of Carrion's disease the identification of *Bartonella rochalimae* was determined by molecular characterization of the *16S rRNA* gene [28].

Some *Bartonella* species, can be clearly differentiated by ribosomal gene sequences [29], but this ability can be problematic for some related species. In fact, the use of the *16S rRNA* gene sequences for differentiating *Bartonella* species has led to unsatisfactory results due to the high degree of conservation of this gene [3, 30]. An example is the study by La Scola and colleagues comparing the ability of 7 gene targets (*16S rRNA*, *glt*A, *gro*El, *rpo*B, *fts*Z, *rib*C, *its*) to discriminate among *Bartonella* species. The authors demonstrated that *16S rRNA* was the least discriminating gene among those tested in the discrimination of *Bartonella* species [31]. Thus, some studies have suggested the use of a combined phylogenetic analysis incorporating other genes to study the phylogenetic relationships among the genus *Bartonella* [32, 33].

16S–23S rRNA intergenic spacer region amplification: This method consists in the polymerase chain reaction (PCR) amplification of *16S–23S rRNA* intergenic spacer region (ITS) sequences. This region is considered to be a hypervariable region with amplicons of a species-specific size resulting in a specific pattern and has, therefore, been proposed as a rapid and reliable method for the detection and subtyping of *Bartonella* species according to their specific patterns. These *16S–23S rRNA* ITS amplicons are usually between 250 and 500 bp in size in prokaryote genomes, being longer in *Bartonella* species, mainly with sizes greater than 1200 bp (for instance, the ITS region of *Bartonella elizabethae* has 1529 bp). There are few exceptions, one being *B. bacilliformis*, which possesses an ITS region of 906 bp in length [34, 35]. Thus, in the 1990s this technique was used alone or combined with a subsequent DNA digestion to discriminate among the species of clinical interest as well as to perform intra-species sub-classifications [35–37]. Furthermore, *Bartonella*-specific primers were designed to amplify variable portions of the ITS region of only small size of around 150–300 bp [34], or medium of around 650–700 bp [36, 38]. The technique has several limitations, such as the continuous increase of new *Bartonella* species, the possible amplification of other members of the order *Rhizobiales* such as *Mesorhizobium* spp. [39], as well as the difficulties inherent to the presence of *Bartonella* spp., such as *B. henselae*, presenting different *16S–23S rRNA* ITS sizes which may lead to misidentification [40]. Thus, with the use of the primers proposed by Maggi and col. [39], Dillon and Iredell reported the presence of up to 5 different *B. henselae 16S–23S rRNA* ITS amplified products with sizes ranging from 648 bp to 693 bp [40]. To solve these problems, the combination of a *16S–23S rRNA* ITS amplification with reverse line blotting was proposed using specific probes for up to 20 *Bartonella* spp. [41]. Although this technique is highly specific, it does not distinguish among *Bartonella birtlesii*, *Bartonella capreoli*, *Bartonella chomelii*, and *Bartonella schoenbuchensis* [41]. All these findings highlight the need for accurate selection of primers and continuous revision of the literature to add new species to the *16S–23S rRNA* ITS schemes.

Despite the above-mentioned limitations, this technique has been largely and successfully used for direct analysis of biological samples, remaining as one of the most used approaches to develop *Bartonella* studies, facilitating the

identification of *Bartonella* spp. as well as the detection of the presence of new *Bartonella* spp. in reservoirs or arthropods [38, 42]. More recent modifications have successfully adapted *16S–23S rRNA* ITS amplification to quantitative PCR (qPCR) techniques for diagnostic purposes [43], which thereafter have also been applied in the detection of known and undescribed species in non-human samples, as in animal reservoirs such as bats [4]. Other authors have included the amplification and sequencing of a portion of the *16S–23S rRNA* ITS on MLST panels as an additional locus to be considered [44].

10.3.1.2 Amplified Ribosomal DNA Restriction Analysis

The use of Amplified Ribosomal DNA Restriction Analysis (ARDRA) (also found in the literature under other denominations such as restriction analysis of the *16S rRNA* amplified gene or restriction fragment length polymorphism – RFLP – of the *16S rRNA* gene) approaches to identify related bacterial species has been largely described, with two main strategies being considered: (a) a series of sequential digestions with different restriction enzymes guided by a decision-tree based on the fragment sizes obtained [45] and (b) a series of parallel digestions resulting in a series of patterns which on reading in combination are specific to each species [46].

Regarding *Bartonella* spp., the first strategy was developed in silico in 2014 to identify the 29 *Bartonella* species (including a few *Candidatus* Bartonella) for which *16S rRNA* data were recorded in GenBank at that time [47]. This approach needed a maximum of 6 consecutive enzymatic digestions to univocally identify 31 *Bartonella* species/subsp. (including 5 *Candidatus* Bartonella spp.), with two of the most relevant species of clinical interest, *B. bacilliformis* and *B. quintana*, being identified in the first digestion step, and *B. henselae* needing 2 additional digestions [47]. While the technique is ductile and, if necessary, may be modified in parallel to new species descriptions by either changing or adding enzymatic digestions, no other attempt to update a fully*16S rRNA* RFLP scheme has been made.

Meanwhile, the second strategy was applied in 1999 to discriminate between *B. elizabethae, B. henselae*, and *B. quintana* [48]. The authors amplified a *Bartonella* genus-specific 296 bp fragment, which was thereafter digested in parallel with *Dde*I and *Mse*I [48]. As above, this approach allows adding other digestions if needed to differentiate among newly added species, but no further attempt to extend the methodology to other *Bartonella* spp. has been found in the literature.

Other authors have used a single *16S rRNA* scheme (alone or combined with the digestion of other genes) for preliminary analysis of isolates recovered in order to select different RFLP patterns to proceed with full sequencing of *16S rRNA* gene (this approach may be extended to other genes analyzed by enzymatic digestion), optimizing economic resources [49].

10.3.1.3 Ribotyping

Ribotyping is based on the use of different restriction enzymes to digest bacterial DNA followed by hybridization with *16S rRNA* and *23S rRNA* probes.

Regarding *Bartonella* spp., this technique has been only used once to type *B. henselae*. Thus, Melter et al. tested the applicability of ribotyping using *Hind*III,

*Bgl*I, and *Pvu*II in 5 cat recovered *B. henselae* [50]. The authors obtained a single ribotype using *Hind*III and were able to differentiate 2 different ribotype patterns using *Bgl*I and *Pvu*II. In all cases, the *B. henselae* ribotype patterns differed from those of *B. quintana, B. clarridgeiae*, and *B. elizabethae,* which were used as controls.

10.3.1.4 DNA-DNA Hybridization

This technique has been largely used as a key technique in the description of new bacterial species. It is based on the reassociation of DNA, with a cut-off of 70% being considered to define a new species [51]. In the case of *Bartonella* spp., this is one of the classical determinations used in the description of new species, being surpassed in recent years by the development of whole genomic sequencing (WGS), with which the inclusion of a percentage of WGS differences has been proposed as a new (or alternative) parameter to define species/subspecies.

10.3.1.5 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS, or usually reported as MALDI-TOF) has emerged as a new protein-based technique for the identification of bacterial species, providing species-specific protein fingerprinting.

MALDI-TOF applications in *Bartonellaceae* are scarce, first approach having been performed in 2009 [52]. The authors obtained unique MALDI-TOF profiles for each of the 17 *Bartonella* species included in the study with no overlap with the bacterial species included in the Bruker database, demonstrating the reliability of the use of this technique for identification purposes.

This technique is also being explored to detect the presence of *Bartonella*-carrying vectors. Studies experimentally infecting different vectors with *B. henselae* or *B. quintana* were able to determine the presence or absence of the infecting *Bartonella* by MALDI-TOF as well as differentiate arthropods infected with *B. henselae* or *B. quintana* [53, 54].

10.3.1.6 Specific PCRs

Several studies in the literature have reported the use of specific PCR approaches to detect and discriminate among *Bartonella* species.

In this sense, the *ialB* gene was considered as specific to *B. bacilliformis* and was proposed for the diagnosis of Carrion's disease [55]. Nevertheless, further studies revealed its presence in other members of the genus, such as *B. henselae* and *B. birtlesii*, making it a useful tool for several *Bartonella* spp. infection [56] but not discriminatory for *B. bacilliformis*.

Most efforts have been addressed to the development of class-PCR targeting either to detect the presence of *Bartonella* spp. (mostly for clinical purposes) or to discriminate at species levels. Furthermore, this approach has also been used in paleomicrobiology, with the *hbp*E, *htr*A, or *gro*EL genes as targets to detect *B. quintana* in ancient samples, such teeth from Napoleon's Grand Army soldiers died on

Vilnius (Lithuania) or from prehistoric (~4000 years) burial [57, 58]. Thus, the riboflavin synthesis genes, such as *rib*C, *rib*D, and *rib*E, are absent in humans, thus being useful for bacterial DNA detection in human samples as well as allowing differentiation of *Bartonella* species. In this sense, species-specific PCR assays based on differences within the *rib*C gene were successfully developed to differentiate *B. henselae*, *B. clarridgeiae*, *B. quintana*, and *B. bacilliformis* [59]. The use of the amplification of the *rib*C as a diagnostic tool was also useful for the recent detection of *B. quintana* DNA from a patient with endocarditis [60]. Other targets has also been explored; thus, in one study by La Scola and colleagues amplification of the *rpo*B and *glt*A genes was one of the most effective PCR approaches for discriminating *Bartonella* at a species level [31].

Several modifications have allowed additional approaches, including the use of nested PCRs or several PCR-RFLP methods. Thus, a nested PCR using the cell division gene *fts*Z to differentiate among *B. bacilliformis*, *B. quintana*, and *B. henselae* has also been described [61]. Meanwhile, RFLP methods have been applied to the amplified products of the RNA polymerase beta subunit (*rpo*B) gene [62], the *fts*Z gene [63], and the citrate synthase (*glt*A) gene [64, 65].

More sensitive tools such as real-time PCR have been used in other studies. A 301 bp region of the *ssr*A gene amplified by real-time PCR was able to discriminate over 30 *Bartonella* species, subspecies, and strains [66]. Furthermore, a two-step protocol combining the amplification of the *glt*A gene by real-time PCR followed by pyrosequencing of the *rpo*B gene has also been described for rapid differentiation of at least 11 medically relevant *Bartonella* spp. [67]. A real-time PCR targeting the NADH dehydrogenase gamma subunit gene (*nuo*G) was found to be sensitive and specific enough to detect diverse *Bartonella* species [68]. More recently, a PCR platform targeting the *glt*A, and *rpo*B genes, was developed for *B. rochalimae,* with the potential to be used for the detection of *B. rochalimae* infections in humans, and it can be used for surveillance studies of vectors and reservoirs [69].

10.3.2 Clonal Determination

10.3.2.1 Multilocus Sequence Typing

Multilocus sequence typing (MLST) was developed in l998 to overcome the difficulty of comparing clonality studies performed by different laboratories [70]. Initial technique validation was performed using *Neisseria meningitidis* [70], having currently been expanded to more than 100 microorganisms. In contrast to PFGE or PCR typing, MLST cannot distinguish between recent epidemiological events, its utility lies in the analysis of long-term epidemiology [70]. The technique is based on the amplification and sequencing of a series (usually around 5–9) of highly conserved housekeeping genes, with each different sequence of a specific gene being numbered, and the combination of all the numbers obtained resulting in a numeric profile which defines a Sequence Type (ST).

Regarding *Bartonella* spp., the first MLST scheme, was designed in 1998 for the epidemiological analysis of *B. henselae* [71]. This scheme was originally based on

the analysis of 9 genes [71]. Further studies modified the MLST scheme using only 7 genes [72], with the current MLST scheme including 8 genes (*16S rRNA*, *bat*R, *rib*C, *gro*EL, *glt*A, *nlp*D, *fts*Z, and *rpo*B) [73]. The MLSTs of *B. henselae* are hosted on the webpage https://pubmlst.org/bhenselae/, with 446 strains classified into 37 *B. henselae* STs.

Subsequent to its first use in *B. henselae*, MLST was expanded to other species, among others, *B. quintana* (9 gene-based scheme: *atp*F, *bqt*R, *fts*Z, *gap*, *glt*A, *gro*EL, *nlp*D, *rib*E, and *rpo*B) [74]. Interestingly, MLST studies of *B. quintana* have demonstrated the ancient divergence among human and macaque isolates which are classified within different and species-specific MLST patterns [75, 76]. These results suggest the possible zoonotic origin of *B. quintana* [75]. Thus, at present at least 22 *B. quintana* STs have been described, with ST1 to ST7 being of human origin, ST8 to ST14 from cynomolgus macaques, ST15 to ST21 from rhesus macaques and ST22 from Japanese macaques [74.76]. Despite the clinical relevance of *B. quintana* [14], no MLST webpage has been developed.

The other two *Bartonella* spp. for which a MLST webpage has been developed are *B. bacilliformis* (https://pubmlst.org/bbacilliformis/) and *Candidatus* Bartonella washoensis (https://pubmlst.org/bwashoensis/). Both schemes were designed in 2011, with *B. bacilliformis* using a 7 gene scheme (*bvr*R, *rnp*B, *fla*A, *rib*C, *fts*Z, *rpo*B and *gro*EL) and *Candidatus* B. washoensis a 6 gene scheme (*16S rRNA*, *fts*Z, *glt*A, *gro*EL, *rib*C,and *rpo*B) [77, 78]. At present, 14 *B. bacilliformis* and 21 *Candidatus* B. washoensis STs have been validated in a bacterial population of 64 and 21 isolates, respectively. Although it cannot be used in vector-samples, MLST analysis of *Bartonella* spp. in infected blood (or possibly other sterile fluids or tissues) may be done. Thus, Pons et al. determined the MLST patterns of 9 *B. bacilliformis* by direct processing of blood samples in the absence of bacterial culture [19].

While no webpage has been developed for other *Bartonella* spp., MLST approaches have also been extended to other *Bartonella* spp. For instance, in *Bartonella bovis*, 27 different ST patterns obtained from 38 isolates analyzed have been described [79–81] and 9 different STs have been detected in 18 isolates of *B. chomelii*, with the presence of more than one infective clone being confirmed in the same biological sample [79]. In both cases the MLST scheme was based on the amplification of 8 different loci (*16S rRNA*, *fts*Z, *gro*EL, *nuo*G, *rib*C, *rpo*B, *ssr*A, and ITS). Of note, the adjudication of the same ST pattern for 2 different allelic combinations highlights the need to organize specific repositories to avoid nomenclature mistakes [79, 81].

In addition to classical analysis of intraspecific clonal relationships, MLST may also be used to establish the phylogenetical relationships among different *Bartonella* spp. For this purpose, the housekeeping genes selected should be amplified, concatenated (either in vitro or in silico), and the results obtained analyzed using a neighbor-joining tree or other similar tools. This approach also allows preliminary screening of possible new species or analysis of the presence of possible *Bartonella* subspecies structures [44, 82].

10.3.2.2 Multispacer Typing

Multispacer typing (MST) has been proposed to increase the discriminatory power of MLST. This technique was conceived as a combination of MLST and *16S–23S rRNA* ITS amplification, in which a series of intergenic spacers placed between two highly conserved genes are amplified. The intergenic regions are thereafter sequenced and each different sequence of a specific intergenic region is numbered, and the combination of all the numbers obtained results in a numeric profile which defines a specific MST. The technique was developed in 2005 [83] and was thereafter expanded to *B. henselae,* with 9 intergenic spacers being established [84]. In addition to *B. henselae* and *B. quintana* this technique has also been used on *B. ancashensis* [85].

Despite its discriminatory power and an attempt to design a webpage database (http://ifr48.timone.univ-mrs.fr/MST_BHenselae/mst), which remained unavailable at the time of writing this chapter, the number of studies using this technique is scarce, and none posterior to 2013 being found searching the terms "MST, *Bartonella*" or "Multispacer Typing, *Bartonella*" in PubMed.

10.3.2.3 Multilocus Variable Copy Numbers of Tandem Repeats Analysis

Multilocus variable copy numbers of tandem repeats analysis (MLVA) is a molecular typing method with high discriminatory power which was developed in 1997 [86, 87] to subtype microbial isolates. This technique reveals insights about the phylogeny within highly homogenous bacterial species subgroups by analyzing the copy number of a selected set of tandem repeats. Even among highly related bacterial strains, the copy number of tandem repeats is usually very diverse. This technique gained importance with the availability of whole genome sequences and free online resources that facilitate the setting-up of new MLVA [88], developed for practically all medically relevant bacterial species. Regarding *Bartonella* spp., the first study developing a MLVA was performed in 2007 for *B. henselae* in which 5 variable number tandem repeats were selected due to the high level of polymorphism provided. In fact, in the same study 31 different profiles were observed in 42 feline and 2 human isolates of *B. henselae* [89]. MLVA has a higher discriminatory power as compared to MSLT and has been proposed for use in epidemiological studies [90, 91]. To date several epidemiological studies analyzing the MLVA profiles of *B. henselae* isolates circulating in reservoir hosts, patients, and vectors have been performed [90–92]. Interestingly, MLVA allowed clear separation between 2 distinct *B. henselae* genotypes; genotype I which is most frequently associated with human infections and genotype II which is most often present in isolates from cats [92]. As far as we know, no MLVA has been developed for other *Bartonella* spp.

10.3.2.4 Pulsed-Field Gel Electrophoresis

Pulsed-field gel electrophoresis (PFGE) consists in electrophoresis of the whole bacterial genome after the application of a restriction enzyme generating a limited number of high-molecular-weight restriction fragments. The analysis of

macrorestriction patterns by PFGE represents the gold standard among band-based typing techniques to establish recent clonal relationships among different bacterial isolates.

While recent reports on the use of PFGE for the typing of *Bartonella* spp. are limited, PFGE has frequently been used for the evaluation of genetic relatedness among *Bartonella* spp. isolates, with the use of different restriction enzymes having been explored (e.g., *Eag*I, *Not*I or *Sma*I) [93, 94]. These reports showed differences in the efficiency to determine the presence of different clones related to higher or lower number of bands obtained, with *Sma*I being considered as one of the best enzymes for typing purposes [93]. Further optimization of the method performed in 2009 confirmed the use of *Sma*I as an elective restriction enzyme, with *Eag*I as a second option [94]. Most studies have been focused on *B. henselae* [72, 95–97], but the PFGE approach has also been validated for other species such as *B. quintana* [93] and *B. vinsonii* [98].

An unexpected problem was described in 2007, when the rapid in vivo emergence of genetic variants of *B. henselae* was observed, leading to the detection of up to three different PFGE patterns (differences ranging from 1 to 3 bands) from primary bacterial cultures [99]. During a study designed to optimize the use of PFGE for typing *Bartonella* spp., Xu et al. also reported this phenomenon in *B. henselae* [94]. Thus, *B. henselae* may mutate frequently, making this a possible limitation for the usefulness of PFGE for typing this microorganism. Notwithstanding, the presence of this high mutation frequency has not been reported in other *Bartonella* spp.

Beyond the classical uses of PFGE, several authors have also applied this technique to validate the presence of genetic differences among isolates of *Bartonella* spp. in the course of studies aimed at describing the presence of *Bartonella* species in different mammal reservoirs [100].

10.4 Whole Genomic Sequencing

WGS is the ultimate typing approach. This technique provides complete information of bacterial genetic material in order to better identify new species and subspecies or to establish phylogenetic relationships among them. The main limitation of this technique (common to several of the techniques described in the present chapter) is the need for a *Bartonella* culture.

The average nucleotide identity (ANI) [101] measures the level of identity to define a species using algorithms modified to the current OrthoANIu with a proposed cut-off value of 95% [102]. This approach has been used in the most recent *Bartonella* spp. descriptions such as those of *Bartonella kosoyi* and *Bartonella krasnovii* [21].

In addition, WGS has been used to analyze intraspecific variability and interspecific relationships by the analysis of single nucleotide polymorphisms (SNPs) and/or orthologous genes [103, 104]. Thus, Guy et al. analyzed 13 *Bartonella* genomes belonging to 16 *Bartonella* spp. [103], while Tay et al. analyzed 27 genomes belonging to 21 different *Bartonella* genomes [104]. These studies allow the *Bartonella*

species analyzed to be classified within different clades. It was of note that the ruminant-adapted bacteria analyzed were consistently classified in a common clade while *B. bacilliformis* and *Candidatus* B. australis were classified within this ruminant clade or as singletons [103, 104].

While highly specific and potent, this approach is currently challenging because of the lack of representative genomes of a high number of *Bartonella* spp., as well as the limited number of sequenced genomes of most of the species. As mentioned above, 37 *Bartonella* species have standing in nomenclature, but the real number of *Bartonella* spp. is predicted to be extremely high. An example of this is that in September 2021 the number of possible *Bartonella* spp. based on DNA sequences (including whole genomes and partial DNA sequences) was 1396 (https://www.ncbi.nlm.nih.gov/taxonomy/?term=txid773[Subtree]; https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=773&lvl=3&keep=1&srchmode=1&unlock). Meanwhile, at present, the genomes of 191 *Bartonella* spp. isolates belonging to 28 species with standing in nomenclature (https://lpsn.dsmz.de/genus/bartonella), 10 *Candidatus Bartonella* spp. and 36 "unclassified *Bartonella*" isolates, are recorded in GenBank (https://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=773). Nonetheless, half of these sequences belong to only 3 species, with 30 *B. henselae*, 18 B. *quintana genomes*, and 17 *B. bacilliformis* genomes. The remaining species are represented by a maximum of 6 genomes, and mainly by a single sequenced genome.

10.5 PCR Typing-Based Techniques

Similar to PFGE, these techniques provide information about recent genetic events, being therefore of special usefulness to differentiate among bacterial strains. Different PCR typing techniques (see next sub-sections) have been used for typing *Bartonella* spp. Although PCR typing is largely absent in the most recent studies because of the use of other more modern molecular approaches, it should be considered, especially due to its cost/effectiveness, which makes it attractive for the typing of *Bartonella*. Furthermore, the use of more than one PCR typing method or in combination with other techniques may enhance epidemiological analysis with more detailed analysis of clonal relationships [105].

10.5.1 Amplified Fragment Length Polymorphism

The amplified fragment length polymorphism (AFLP) technique is based on selective PCR amplification of DNA restriction fragments obtained from a total digestion of genomic DNA. Due to differences in restriction fragment lengths separated by gel electrophoresis, different patterns are created, revealing genetic relationships among various isolates.

While having a high degree of reproducibility and discriminatory power, AFLP has rarely been applied to the typing of *Bartonella* spp. However, it has been applied,

with satisfactory results, in the study of *B. henselae*, digesting the chromosomal DNA with *Hind*III [106], and *B. bacilliformis* using *Pst*I for restriction purposes [107].

10.5.2 Repetitive Extragenic Palindromic PCR

Along the bacterial genome there are multiple interspersed noncoding repetitive sequences that can be used to design primers to amplify the inter-sequence DNA fragments creating a unique DNA profile or fingerprint for individual bacterial strains. This technique is named repetitive extragenic palindromic (REP-PCR) and has been useful for the identification and discrimination of *Bartonella* species [105, 108, 109]. In particular, REP-PCR appears to be useful for subtyping *B. henselae* isolates. Rodriguez-Barradas et al. showed that by combining the results of REP-PCR and enterobacterial repetitive intergenic consensus (ERIC)-PCR, 5 different fingerprint profiles were identified among 17 isolates of *B. henselae*, but only 1 profile was identified among 5 isolates of *B. quintana* [105]. The technique allowed discrimination among *Bartonella* species and among different strains within a species [105]. In a study by Sander and colleagues [109], 17 isolates of *B. henselae* from Germany obtained from blood cultures of domestic cats were subtyped by REP-PCR as well as other typing techniques. Although all the typing methods allowed subtyping of the *B. henselae* isolates, PFGE and ERIC-PCR provided the highest discriminatory power [109]. The results exhibited a high genetic heterogeneity among the *B. henselae* strains studied, showing that these typing techniques can be useful for epidemiological and clinical follow-up studies [109]. Regarding other *Bartonella* spp. only a few reports are available, such as a study on *B. bacilliformis* typing methods in which REP-PCR was concomitantly used with ERIC-PCR [110].

10.5.3 Enterobacterial Repetitive Intergenic Consensus PCR

ERIC sequences are repetitive, short, and highly conserved extragenic palindromic DNA sequences that are widely disseminated among microorganisms. The ERIC-PCR technique consists in the amplification of the DNA region present within two ERIC sequences. The amplification results in a variable number of bands with an indeterminate size, which are used for genomic fingerprinting of bacterial isolates.

The use of this technique was first described in *Bartonella* spp. in 1995, analyzing 17 *B. henselae*, 5 *B. quintana*, and 2 isolates each of *B. bacilliformis*, *B. elizabethae*, and *B. vinsonii*. The study showed good results suggesting better discriminatory power among *B. henselae* compared to *B. quintana* [105]. The presence of a band of 1.3 kb common for *B. henselae* and another of 1.1 kb present in all *B. quintana* was reported with these bands being absent in the remaining *Bartonella* spp. analyzed [105].

Thereafter, ERIC-PCR has been scarcely used. It has been reported in different studies focused on *B. henselae* [111] being considered as discriminatory as PFGE [109]. Similarly, as mentioned above, ERIC-PCR has also been described in *B. bacilliformis*, being used together with REP-PCR to define different profiles [110]. ERIC-PCR has also been used in the characterization of new *Bartonella* species such as *B. schoenbuchensis* [112] and more recently in the analysis of the diversity of *B. bovis* [81].

10.5.4 Randomly Amplified Polymorphic DNA PCR

Randomly amplified polymorphic DNA (RAPD)-PCR (or arbitrarily primed PCR – AP-PCR) is a genotyping method that uses a single primer of arbitrary nucleotide sequence, which requires no previous knowledge of the DNA template and results in the amplification of random strain-specific segments of DNA products. Regarding *Bartonella* spp. the M13 primer (5′-GAGGGTGGCGGTTCT-3′) has been used.

The number of studies using this technique is very scarce, having been explored for the typing of *B. henselae* [109, 113, 114]. The results, however, were not very satisfactory, suggesting that it has limited value as a genotyping tool of *B. henselae* [114]. Meanwhile, Li et al. used this technique to characterize a strain of *Bartonella* isolated from a domestic cat in China, identifying the strain as *B. clarridgeiae* [115]. In addition, Ciervo et al. used RAPD-PCR to differentiate among *B. henselae*, *B. quintana*, and *B. clarridgeiae* [113].

10.5.5 *16S rRNA* Type-Specific Amplification

This technique has been used to differentiate two main clusters of *B. henselae*. This microorganism present two different *16S rRNA* sequences differing in three nucleotides located at positions 172 to 175 [116–118]. In this region the so-called Cluster I (with the Houston I strain being representative) presents the "-TAG" sequence while Cluster II (with the Marseille strain being representative) possesses the "ATTT" sequence (note the presence of an additional nucleotide in Cluster II) [116]. This difference has been used to design an easy PCR approach to detect the presence of these clusters using a universal *16S rRNA* upper primer and the differential 172–175 nucleotide region as 3′-terminal of the lower primer (primer 1 (BH1): 5′-CCGATAAATCTTTCTCCCTAA-3′ and primer 2 (BH2): 5′-CCGATAAATC TTTCTCCAAAT-3′) [119].

While largely surpassed by more discriminative techniques, several studies have showed geographical differences in the distribution of these clusters, with cluster I, for instance, being predominant in Japan [120].

10.5.6 Infrequent Restriction Site PCR

The infrequent restriction site PCR (IRS-PCR) is based on fully genomic DNA digestion with 3 different enzymes, followed by the use of oligonucleotide adapters and subsequent PCR. This technique allows the use of fluorescent markers, and then the products may be analyzed in DNA sequencers for fingerprinting purposes. Its use in *Bartonella* spp. has only been sporadically reported, with the first study validating its usefulness to differentiate 7 *Bartonella* spp. being performed in 2000 [121]. Nonetheless, subsequent analyses suggested its limited value for the analysis of *B. henselae* isolates [114]. Further analysis confirmed the lack of reproducibility of IRS-PCR related to *B. henselae* phase variation, with this phenomenon not affecting the reproducibility of other PCR-based typing methods such as *16S–23S rRNA* ITS amplification, RAPD, ERIC-PCR or that of methods such as *glt*A or*16S rRNA* sequencing [122]. To the best of our knowledge, IRS-PCR has thereafter only been used in a *B. bacilliformis* fingerprinting analysis, describing the presence of geographical clonal differences as well as the presence of the same clone among relatives living in the same household [123].

10.6 Summary

Bartonella spp. is an arthropod-vectorized bacterial genus responsible for a wide array of vector-borne diseases. Although potentially lethal, *Bartonella* spp. has remained for years out of scope of clinicians. While challenged by the fastidious nature of *Bartonella* spp., the development of molecular tools has allowed identifying an increasing number of fully or partially characterized new *Bartonella* species, contributing to the knowledge of its epidemiology and the further understanding of its pathogenicity. The present chapter summarizes the most relevant approaches to type *Bartonella* spp.

Acknowledgments The authors thank Donna Pringle for language correction. JR was supported by Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (FONDECYT – Perú) within the “Proyecto de Mejoramiento y Ampliación de los Servicios del Sistema Nacional de Ciencia, Tecnología e Innovación Tecnológica” [contract number: 08-2019-FONDECYT-BM-INC-INV].

References

1. Birtles RJ, Harrison TG, Saunders NA et al (1995) Proposals to unify the genera descriptions of *Bartonella talpae* comb. nov., and three new *Grahamella* and *Bartonella*, with comb. nov., *Bartonella peromysci* species, *Bartonella grahamii* sp. nov., *Bartonella taylorii* sp. nov, and *Bartonella doshiae* sp. nov. Int J Syst Bacteriol 45:1–8. https://doi.org/10.1099/00207713-45-1-1
2. Brenner DJ, O’Connor SP, Winkler HH et al (1993) Proposals to unify the genera *Bartonella* and *Rochalimaea*, with descriptions of *Bartonella quintana* comb. nov., *Bartonella vinsonii* comb. nov., *Bartonella henselae* comb. nov., and *Bartonella elizabethae* comb nov.

and to remove the family *Bartonellaceae* from the order Rickettsiales. Int J Syst Bacteriol 43:777–786. https://doi.org/10.1099/00207713-43-4-777
3. Gomes C, Ruiz J (2017) Carrion's disease: the sound of silence. Clin Microbiol Rev 31:e00056–e00017. https://doi.org/10.1128/CMR.00056-17
4. André MR, Gutiérrez R, Ikeda P et al (2019) Genetic diversity of *Bartonella* spp. in vampire bats from Brazil. Transbound Emerg Dis 66:2329–2341. https://doi.org/10.1111/tbed.13290
5. Fournier PE, Taylor C, Rolain JM et al (2007) *Bartonella australis* sp. nov. from kangaroos, Australia. Emerg Infect Dis 13:1961–1962. https://doi.org/10.3201/eid1312.060559
6. Laroche M, Berenger JM, Mediannikov O et al (2017) Detection of a potential new *Bartonella* species "*Candidatus* Bartonella rondoniensis" in human biting kissing bugs (Reduviidae; Triatominae). PLoS Negl Trop Dis 11:e0005297
7. Medkour H, Lo CI, Anani H et al (2019) *Bartonella massiliensis* sp. nov., a new bacterial species isolated from an *Ornithodoros sonrai* tick from Senegal. New Microbes New Infect 32:100596. https://doi.org/10.1016/j.nmni.2019.100596
8. Raya AP, Jaffe DA, Chomel BB et al (2018) Detection of *Bartonella* species, including *Candidatus Bartonella ovis* sp. nov, in ruminants from Mexico and lack of evidence of *Bartonella* DNA in saliva of common vampire bats (*Desmodus rotundus*) predating on them. Vet Microbiol 222:69–74. https://doi.org/10.1016/j.vetmic.2018.06.018
9. Kešnerová L, Moritz R, Engel P (2016) *Bartonella apis* sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria. Int J Syst Evol Microbiol 66:414–421. https://doi.org/10.1099/ijsem.0.000736
10. Okaro U, Addisu A, Casanas B et al (2017) *Bartonella* species, an emerging cause of blood-culture-negative endocarditis. Clin Microbiol Rev 30:709–746. https://doi.org/10.1128/CMR.00013-17
11. Álvarez A, Maggi R, Marr HS et al (2019) Serological and molecular study on *Bartonella* infection and other pathogens in dogs from different geographical areas of Spain. In: 9th international conference on *Bartonella* as emerging pathogens Paris (France), 18–20 September 2019
12. Akbari SHA, Averill CE, Roland JL et al (2018) *Bartonella henselae* infection presenting as cervical spine osteomyelitis: case report. J Neurosurg Pediatr 22:189–194. https://doi.org/10.3171/2018.2.PEDS17600
13. Ghidey FY, Igbinosa O, Mills K et al (2016) Case series of *Bartonella quintana* blood culture-negative endocarditis in Washington, DC. JMM Case Rep 3:e005049. https://doi.org/10.1099/jmmcr.0.005049
14. Ruiz J (2018) *Bartonella quintana*, past, present, and future of the scourge of World War I. APMIS 126:831–837. https://doi.org/10.1111/apm.12895
15. Baldani CD, Santos HA, Massard CL (2014) The family *Bartonellaceae*. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F (eds) The prokaryotes – alphaproteobacteria and betaproteobacteria. Springer, Berlin/Heidelberg, pp 82–114. https://doi.org/10.1007/978-3-642-30197-1_251
16. Welch DF, Carroll KC, Hofmeister EK et al (1999) Isolation of a new subspecies, *Bartonella vinsonii* subsp. *arupensis*, from a cattle rancher: identity with isolates found in conjunction with *Borrelia burgdorferi* and *Babesia microti* among naturally infected mice. J Clin Microbiol 37:2598–2601
17. Drancourt M, Raoult D (1993) Proposed tests for the routine identification of *Rochalimaea* species. Eur J Clin Microbiol Infect Dis 12:710–713. https://doi.org/10.1007/bf02009387
18. Minnick MF, Anderson BE (2015) *Bartonella*. In: Tang Y-W, Liu D, Poxton I, Schwartzman J, Sussman M, Williams H, Versteeg-Buschman L (eds) Molecular medical microbiology, 2nd edn. Elsevier, New York, pp 1911–1939. https://doi.org/10.1016/B978-0-12-397169-2.00105-0
19. Pons MJ, Silva W, del Valle-Mendoza J et al (2016) Multi locus sequence typing of *Bartonella bacilliformis* DNA performed directly from blood of patients with Oroya's fever during a Peruvian outbreak. PLoS Negl Trop Dis 10:e0004391. https://doi.org/10.1371/journal.pntd.0004391

20. Ulloa GM, Vásquez-Achaya F, Gomes C et al (2018) Molecular detection of *Bartonella bacilliformis* in *Lutzomyia maranonensis* in Cajamarca, Peru: a new potential vector of Carrion's disease in Peru? Am J Trop Med Hyg 99:1229–1233. https://doi.org/10.4269/ajtmh.18-0520
21. Gutiérrez R, Shalit T, Markus B et al (2020) *Bartonella kosoyi* sp. nov. and *Bartonella krasnovii* sp. nov., two novel species closely related to the zoonotic *Bartonella elizabethae*, isolated from black rats and wild desert rodent-fleas. Int J Syst Evol Microbiol 70:1656–1665
22. Seubert A, Falch C, Birtles RJ et al (2003) Characterization of the cryptic plasmid pBGR1 from *Bartonella grahamii* and construction of a versatile *Escherichia coli-Bartonella* spp. shuttle cloning vector. Plasmid 49:44–52. https://doi.org/10.1016/s0147-619x(02)00103-8
23. Patel JB (2001) *16S rRNA* gene sequencing for bacterial pathogen identification in the clinical laboratory. Mol Diagn 6:313–321
24. Relman DA, Loutit JS, Schmidt TM et al (1990) The agent of bacillary angiomatosis. An approach to the identification of uncultured pathogens. N Engl J Med 323:1573–1580
25. Anderson B, Kelly C, Threlkel R et al (1993) Detection of *Rochalimaea henselae* in cat-scratch disease skin test antigens. J Infect Dis 168:1034–1036
26. Goldenberger D, Schmidheini T, Altwegg M (1997) Detection of *Bartonella henselae* and *Bartonella quintana* by a simple and rapid procedure using broad-range PCR amplification and direct single-strand sequencing of part of the 16S rRNA gene. Clin Microbiol Infect 3:240–245
27. del Valle J, Silva W, Tinco C et al (2014) Diagnosis of Carrion's disease by direct blood PCR in thin blood smear negative samples. PLoS One 9:e92283
28. Mujica GM, León DF, Espinoza-Culupú A. (2014) Identificación de *Bartonella rochalimae* en un brote de enfermedad de Carrión mediante caracterización molecular del gen *16S rRNA*. Paper presented at the VIII Congreso Internacional del Instituto Nacional de Salud, Lima (Peru), 6–8 November 2014, p 73
29. Dauga C, Miras I, Grimont PAD (1996) Identification of *Bartonella henselae* and *B. quintana 16s rDNA* sequences by branch-, genus- and species-specific amplification. J Med Microbiol 45:192–199
30. Kosoy M, Hayman DTS, Chan KS (2012) *Bartonella* bacteria in nature: where does population variability end and a species start? Infect Genet Evol 12:894–904
31. La Scola B, Zeaiter Z, Khamis A, Raoult D (2003) Gene-sequence-based criteria for species definition in bacteriology: the *Bartonella* paradigm. Trends Microbiol 11:318–321
32. Li DM, Hou Y, Song XP et al (2015) High prevalence and genetic heterogeneity of rodent-borne *Bartonella* species on Heixiazi island, China. Appl Environ Microbiol 81:7981–7992
33. Pitulle C, Strehse C, Brown JW et al (2002) Investigation of the phylogenetic relationships within the genus *Bartonella* based on comparative sequence analysis of the *rnpB* gene, *16S rDNA* and *23S rDNA*. Int J Syst Evol Microbiol 52:2075–2080
34. Jensen WA, Fall MZ, Rooney J et al (2000) Rapid identification and differentiation of *Bartonella* species using a single-step PCR assay. J Clin Microbiol 38:1717–1722
35. Roux V, Raoult D (1995) The *16S-23S rRNA* intergenic spacer region of *Bartonella* (*Rochalimaea*) species is longer than usually described in other bacteria. Gene 156:107–111. https://doi.org/10.1016/0378-1119(94)00919-j
36. Houpikian P, Raoult D (2001) *16S/23S rRNA* intergenic spacer regions for phylogenetic analysis, identification, and subtyping of *Bartonella* species. J Clin Microbiol 39:2768–2778. https://doi.org/10.1128/JCM.39.8.2768-2778.2001
37. Matar GM, Swaminathan B, Hunter SB et al (1993) Polymerase chain reaction-based restriction fragment length polymorphism analysis of a fragment of the ribosomal operon from *Rochalimaea* species for subtyping. J Clin Microbiol 30:1730–1734
38. Birtles RJ, Hazel S, Bown K et al (2000) Subtyping of uncultured *Bartonellae* using sequence comparison of *16S/23S rRNA* intergenic spacer regions amplified directly from infected blood. Mol Cell Probes 14:79–87
39. Maggi RG, Breitschwerdt EB (2005) Potential limitations of the *16S-23S rRNA* intergenic region for molecular detection of *Bartonella* species. J Clin Microbiol 43:1171–1176. https://doi.org/10.1128/JCM.43.3.1171-1176.2005

40. Dillon B, Iredell J (2005) Potential limitations of the *16S-23S rRNA* intergenic region for molecular detection of *Bartonella* species. J Clin Microbiol 43:4921–4922. https://doi.org/10.1128/JCM.43.9.4921-4922.2005
41. García-Esteban C, Gil H, Rodríguez-Vargas M et al (2008) Molecular method for *Bartonella* species identification in clinical and environmental samples. J Clin Microbiol 46:776–779. https://doi.org/10.1128/JCM.01720-07
42. Hornok S, Szőke K, Meli ML et al (2019) Molecular detection of vector-borne bacteria in bat ticks (*Acari: Ixodidae, Argasidae*) from eight countries of the Old and New Worlds. Parasit Vectors 12:50. https://doi.org/10.1186/s13071-019-3303-4
43. Gutiérrez R, Morick D, Gross I et al (2013). *Bartonellae* in domestic and stray cats from Israel: comparison of bacterial cultures and high-resolution melt real-time PCR as diagnostic methods. Vector Borne Zoonotic Dis 13:857–864. https://doi.org/10.1089/vbz.2013.1308
44. Bai Y, Hayman DT, McKee CD et al (2015) Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLoS Negl Trop Dis 9:e0003478. https://doi.org/10.1371/journal.pntd.0003478
45. Vila J, Ruiz J, Gallardo F et al (2003) *Aeromonas* spp. and traveler's diarrhea: clinical features and antimicrobial resistance. Emerg Infect Dis 9:552–555. https://doi.org/10.3201/eid0905.020451
46. Salazar de Vegas EZ, Nieves B, Araque M et al (2006) Outbreak of infection with *Acinetobacter* strain RUH 1139 in an intensive care unit. Infect Control Hosp Epidemiol 27:397–403
47. del Valle LJ, Jaramillo M, Talledo M et al (2014) Development of a *16S rRNA* PCR-RFLP assay for *Bartonella* identification: applicability in the identification of species involved in human infections. Univ J Microbiol Res 2:15–22
48. Matar GM, Koehler JE, Malcolm G et al (1999) Identification of *Bartonella* species directly in clinical specimens by PCR-restriction fragment length polymorphism analysis of a *16S rRNA* gene fragment. J Clin Microbiol 37:4045–4047
49. Ziedins AC, Chomel BB, Kasten RW et al (2016) Molecular epidemiology of *Bartonella* species isolated from ground squirrels and other rodents in northern California. Epidemiol Infect 144:1837–1844. https://doi.org/10.1017/S0950268816000108
50. Melter O, Hercík K, Weyant RS et al (2003) Detection and characterization of feline *Bartonella henselae* in the Czech Republic. Vet Microbiol 93:261–273. https://doi.org/10.1016/s0378-1135(03)00032-4
51. Wayne LG, Brenner DJ, Colwell RR (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bact 37:463–464
52. Fournier PE, Couderc C, Buffet S et al (2009) Rapid and cost-effective identification of *Bartonella* species using mass spectrometry. J Med Microbiol 58(Pt 9):1154–1159. https://doi.org/10.1099/jmm.0.009647-0
53. El Hamzaoui B, Laroche M, Almeras L et al (2018) Detection of *Bartonella* spp. in fleas by MALDI-TOF MS. PLoS Negl Trop Dis 12:e0006189. https://doi.org/10.1371/journal.pntd.0006189
54. El Hamzaoui B, Laroche M, Parola P (2019) Detection of *Bartonella* spp. in *Cimex lectularius* by MALDI-TOF MS. Comp Immunol Microbiol Infect Dis 64:130–137. https://doi.org/10.1016/j.cimid.2019.03.001
55. Padilla C, Ventura G (2003) Diseño y estandarización de una prueba PCR para el diagnóstico de Bartonelosis causada por *Bartonella bacilliformis*. Rev Peru Med Exp Salud Publica 20:5–8. https://doi.org/10.17843/rpmesp.2003.201.880
56. Drancourt M, Tran-Hung L, Courtin J et al (2005) *Bartonella quintana* in a 4000-year-old human tooth. J Infect Dis 191:607–611. https://doi.org/10.1086/427041
57. Raoult D, Dutour O, Houhamdi L et al (2006) Evidence for louse-transmitted diseases in soldiers of Napoleon's Grand Army in Vilnius. J Infect Dis 193:112–120. https://doi.org/10.1086/498534
58. Deng H, Pang Q, Xia H et al (2016) Identification and functional analysis of invasion associated locus B (IalB) in *Bartonella* species. Microb Pathog 98:171–177

59. Bereswill S, Hinkelmann S, Kist M et al (1999) Molecular analysis of riboflavin synthesis genes in *Bartonella henselae* and use of the *rib*C gene for differentiation of *Bartonella* species by PCR. J Clin Microbiol 37(10):3159–3166
60. Mohammadian M, Butt S (2019) Endocarditis caused by *Bartonella quintana*, a rare case in the United States. IDCases 17:e00533
61. Kelly TM, Padmalayam I, Baumstark BR (1998) Use of the cell division protein FtsZ as a means of differentiating among *Bartonella* species. Clin Diagn Lab Immunol 5:766–772
62. Renesto P, Gouvernet J, Drancourt M et al (2001) Use of *rpo*B gene analysis for detection and identification of *Bartonella* species. J Clin Microbiol 39:430–437
63. Zeaiter Z, Liang Z, Raoult D (2002) Genetic classification and differentiation of *Bartonella* species based on comparison of partial *fts*Z gene sequences. J Clin Microbiol 40:3641–3647
64. Joblet C, Roux V, Drancourt M et al (1995) Identification of *Bartonella* (*Rochalimaea*) species among fastidious Gram-negative bacteria on the basis of the partial sequence of the citrate synthase gene. J Clin Microbiol 33:1879–1883
65. Norman AF, Regnery R, Jameson P et al (1995) Differentiation of *Bartonella*-like isolates at the species level by PCR restriction fragment length polymorphism in the citrate synthase gene. J Clin Microbiol 33:1797–1803
66. Diaz MH, Bai Y, Malania L et al (2012) Development of a novel Genus-specific real-time PCR assay for detection and differentiation of *Bartonella* species and genotypes. J Clin Microbiol 50:1645–1649
67. Buss SN, Gebhardt LL, Musser KA (2012) Real-time PCR and pyrosequencing for differentiation of medically relevant *Bartonella* species. J Microbiol Methods 91:252–256
68. Colborn JM, Kosoy MY, Motin VL et al (2010) Improved detection of *Bartonella* DNA in mammalian hosts and arthropod vectors by real-time PCR using the NADH dehydrogenase gamma subunit (*nuo*G). J Clin Microbiol 48:4630–4633
69. Chan D, Geiger JA, Vasconcelos EJR et al (2018) *Bartonella rochalimae* detection by a sensitive and specific PCR platform. Am J Trop Med Hyg 99:840–843
70. Maiden MC, Bygraves JA, Feil E et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 95:3140–3145. https://doi.org/10.1073/pnas.95.6.3140
71. Iredell J, Blanckenberg D, Arvand M et al (2003) Characterization of the natural population of *Bartonella henselae* by multilocus sequence typing. J Clin Microbiol 41:5071–5079. https://doi.org/10.1128/jcm.41.11.5071-5079.2003
72. Arvand M, Viezens J (2007) Evaluation of pulsed-field gel electrophoresis and multi-locus sequence typing for the analysis of clonal relatedness among *Bartonella henselae* isolates. Int J Med Microbiol 297:255–262. https://doi.org/10.1016/j.ijmm.2007.02.001
73. Zhao F, Chaloner G, Darby A et al (2012) Optimization of *Bartonella henselae* multilocus sequence typing scheme using single-nucleotide polymorphism analysis of SOLiD sequence data. Chin Med J 125:2284–2288
74. Arvand M, Raoult D, Feil EJ (2010) Multi-locus sequence typing of a geographically and temporally diverse sample of the highly clonal human pathogen *Bartonella quintana*. PLoS One 5:e9765. https://doi.org/10.1371/journal.pone.0009765
75. Li H, Bai JY, Wang LY et al (2013) Genetic diversity of *Bartonella quintana* in macaques suggests zoonotic origin of trench fever. Mol Ecol 22:2118–2127. https://doi.org/10.1111/mec.12261
76. Sato S, Kabeya H, Yoshino A et al (2015) Japanese macaques (*Macaca fuscata*) as Natural reservoir of *Bartonella quintana*. Emerg Infect Dis 21:2168–2170. https://doi.org/10.3201/eid2112.150632
77. Chaloner GL, Ventosilla P, Birtles RJ (2011) Multi-locus sequence analysis reveals profound genetic diversity among isolates of the human pathogen *Bartonella bacilliformis*. PLoS Negl Trop Dis 5:e1248. https://doi.org/10.1371/journal.pntd.0001248
78. Inoue K, Kabeya H, Hagiya K et al (2011) Multi-locus sequence analysis reveals host specific association between *Bartonella washoensis* and squirrels. Vet Microbiol 148:60–65. https://doi.org/10.1016/j.vetmic.2010.08.007

79. Antequera-Gómez ML, Lozano-Almendral L, Barandika JF et al (2015) *Bartonella chomelii* is the most frequent species infecting cattle grazing in communal mountain pastures in Spain. Appl Environ Microbiol 81:623–629. https://doi.org/10.1128/AEM.03159-14
80. Bai Y, Malania L, Alvarez Castillo D et al (2013) Global distribution of *Bartonella* infections in domestic bovine and characterization of *Bartonella bovis* strains using multi-locus sequence typing. PLoS One 8:e80894. https://doi.org/10.1371/journal.pone.0080894
81. Kho KL, Koh FX, Jaafar T et al (2015) Prevalence and molecular heterogeneity of *Bartonella bovis* in cattle and *Haemaphysalis bispinosa* ticks in Peninsular Malaysia. BMC Vet Res 11:153. https://doi.org/10.1186/s12917-015-0470-1
82. Paul S, Minnick MF, Chattopadhyay S (2016) Mutation-driven divergence and convergence indicate adaptive evolution of the intracellular human-restricted pathogen, *Bartonella bacilliformis*. PLoS Negl Trop Dis 10:e0004712. https://doi.org/10.1371/journal.pntd.0004712
83. Foucault C, La Scola B, Lindroos H et al (2005) Multispacer typing technique for sequence-based typing of *Bartonella quintana*. J Clin Microbiol 43:41–48. https://doi.org/10.1128/JCM.43.1.41-48.2005
84. Li W, Chomel BB, Maruyama S et al (2006) Multispacer typing to study the genotypic distribution of *Bartonella henselae* populations. J Clin Microbiol 44:2499–2506. https://doi.org/10.1128/JCM.00498-06
85. Mullins KE, Hang J, Jiang J et al (2013) Molecular typing of "*Candidatus* Bartonella ancashi," a new human pathogen causing verruga peruana. J Clin Microbiol 51:3865–3868. https://doi.org/10.1128/JCM.01226-13
86. Van Belkum A, Riewerts Eriksen NH, Sijmons M et al (1997) Coagulase and protein A polymorphisms do not contribute to persistence of nasal colonisation by *Staphylococcus aureus*. J Med Microbiol 46:222–232
87. Van Belkum A, Scherer S, Van Leeuwen W et al (1997) Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect Immun 65:5017–5027
88. Grissa I, Bouchon P, Pourcel C et al (2008) On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. Biochimie 90:660–668. https://doi.org/10.1016/j.biochi.2007.07.014
89. Monteil M, Durand B, Bouchouicha R et al (2007) Development of discriminatory multiple-locus variable number tandem repeat analysis for *Bartonella henselae*. Microbiology 153:1141–1148. https://doi.org/10.1099/mic.0.2006/001164-0
90. Azzag N, Haddad N, Durand B et al (2012) Population structure of *Bartonella henselae* in Algerian urban stray cats. PLoS One 7:e43621. https://doi.org/10.1371/journal.pone.0043621
91. Gil H, Escudero R, Pons I, Rodríguez-Vargas M et al (2013) Distribution of *Bartonella henselae* variants in patients, reservoir hosts and vectors in Spain. PLoS One 8:e68248. https://doi.org/10.1371/journal.pone.0068248
92. Bouchouicha R, Durand B, Monteil M et al (2009) Molecular epidemiology of feline and human *Bartonella henselae* isolates. Emerg Infect Dis 15:813–816. https://doi.org/10.3201/eid1505.080995
93. Roux V, Raoult D (1995) Inter- and intraspecies identification of *Bartonella* (*Rochalimaea*) species. J Clin Microbiol 33:1573–1579
94. Xu C, Liu Q, Diao B et al (2009) Optimization of pulse-field gel electrophoresis for *Bartonella* subtyping. J Microbiol Methods 76:6–11. https://doi.org/10.1016/j.mimet.2008.09.001
95. Arvand M, Wendt C, Regnath T et al (1998) Characterization of *Bartonella henselae* isolated from bacillary angiomatosis lesions in a human immunodeficiency virus-infected patient in Germany. Clin Infect Dis 26:1296–1299
96. Chang CC, Chomel BB, Kasten RW et al (2002) Molecular epidemiology of *Bartonella henselae* infection in human immunodeficiency virus-infected patients and their cat contacts, using pulsed-field gel electrophoresis and genotyping. J Infect Dis 186:1733–1739. https://doi.org/10.1086/345764
97. Podsiadly E, Haddad N, Berrich M et al (2012) Characterization of Polish feline *B. henselae* isolates by multiple-locus tandem repeat analysis and pulse-field gel electrophoresis. Ann Agric Environ Med 19:39–43

98. Chang CC, Kasten RW, Chomel BB et al (2000) Coyotes (*Canis latrans*) as the reservoir for a human pathogenic *Bartonella* sp.: molecular epidemiology of *Bartonella vinsonii* subsp. *berkhoffii* infection in coyotes from central coastal California. J Clin Microbiol 38:4193–4200
99. Berghoff J, Viezens J, Guptill L et al (2007) *Bartonella henselae* exists as a mosaic of different genetic variants in the infected host. Microbiology 153(Pt 7):2045–2051. https://doi.org/10.1099/mic.0.2007/006379-0
100. Chomel BB, Molia S, Kasten RW et al (2016) Isolation of *Bartonella henselae* and two new *Bartonella* Subspecies, *Bartonella koehlerae* subspecies *boulouisii* subsp. nov. and *Bartonella koehlerae* Subspecies *bothieri* subsp. nov. from free-ranging Californian mountain lions and bobcats. PLoS One 11:e0148299. https://doi.org/10.1371/journal.pone.0148299
101. Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572. https://doi.org/10.1073/pnas.0409727102
102. Goris J, Konstantinidis KT, Klappenbach JA et al (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57(Pt 1):81–91. https://doi.org/10.1099/ijs.0.64483-0
103. Guy L, Nystedt B, Toft C et al (2013) A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen *Bartonella*. PLoS Genet 9:e1003393. https://doi.org/10.1371/journal.pgen.1003393
104. Tay ST, Kho KL, Lye SF et al (2018) Phylogeny and putative virulence gene analysis of *Bartonella bovis*. J Vet Med Sci 80:653–661. https://doi.org/10.1292/jvms.17-0448
105. Rodriguez-Barradas MC, Hamill RJ, Houston ED et al (1995) Genomic fingerprinting of *Bartonella* species by repetitive element PCR for distinguishing species and isolates. J Clin Microbiol 33:1089–1093
106. Mietze A, Morick D, Köhler H et al (2011) Combined MLST and AFLP typing of *Bartonella henselae* isolated from cats reveals new sequence types and suggests clonal evolution. Vet Microbiol 148:238–245. https://doi.org/10.1016/j.vetmic.2010.08.012
107. Birtles RJ, Fry NK, Ventosilla P et al (2002) Identification of *Bartonella bacilliformis* genotypes and their relevance to epidemiological investigations of human bartonellosis. J Clin Microbiol 40:3606–3612. https://doi.org/10.1128/jcm.40.10.3606-3612.2002
108. Clarridge JE 3rd, Raich TJ, Pirwani et al (1995) Strategy to detect and identify *Bartonella* species in routine clinical laboratory yields *Bartonella henselae* from human immunodeficiency virus-positive patient and unique *Bartonella* strain from his cat. J Clin Microbiol 33:2107–2113
109. Sander A, Ruess M, Bereswill S et al (1998) Comparison of different DNA fingerprinting techniques for molecular typing of *Bartonella henselae* isolates. J Clin Microbiol 36:2973–2981
110. Padilla C, Ventura G (2003) Genotipificación de *Bartonella bacilliformis* por amplificación de elementos repetitivos mediante el uso de REP-PCR y ERIC-PCR. Rev Peru Med Exp Salud Publica 20:128–131
111. Sander A, Bühler C, Pelz K et al (1997) Detection and identification of two *Bartonella henselae* variants in domestic cats in Germany. J Clin Microbiol 35:584–587
112. Dehio C, Lanz C, Pohl R et al (2001) *Bartonella schoenbuchii* sp. nov., isolated from the blood of wild roe deer. Int J Syst Evol Microbiol 51:1557–1565
113. Ciervo A, Petrucca A, Ciarrocchi S et al (2001) Molecular characterization of first human *Bartonella* strain isolated in Italy. J Clin Microbiol 39:4554–4557. https://doi.org/10.1128/JCM.39.12.4554-4557.2001
114. Dillon B, Valenzuela J, Don R et al (2002) Limited diversity among human isolates of *Bartonella henselae*. J Clin Microbiol 40:4691–4699. https://doi.org/10.1128/jcm.40.12.4691-4699.2002
115. Li D, Liu Q, Song X et al (2009) Biological and molecular characteristics of a cat-borne *Bartonella clarridgeiae*. Wei Sheng Wu Xue Bao 49:429–437

116. Bergmans AM, Groothedde JW, Schellekens JF et al (1995) Etiology of cat scratch disease: comparison of polymerase chain reaction detection of *Bartonella* (formerly *Rochalimaea*) and *Afipia felis* DNA with serology and skin tests. J Infect Dis 171:916–923
117. Regnery RL, Anderson BE, Clarridge JE 3rd. et al (1992) Characterization of a novel *Rochalimaea* species. *R. henselae* sp. nov., isolated from blood of a febrile. Human immunodeficiency virus-positive patient. J Clin Microbiol 30:265–274
118. Relman DA, Lepp PW, Sadler KN et al (1992) Phylogenetic relationships among the agent of bacillary angiomatosis. *Bartonella bacilliformis*, and other alpha-proteobacteria. Mol Microbiol 6:1801–1807. https://doi.org/10.1111/j.1365-2958.1992.tb01352.x
119. Bergmans AM, Schellekens JF, van Embden JD et al (1996) Predominance of two *Bartonella henselae* variants among cat-scratch disease patients in the Netherlands. J Clin Microbiol 34:254–260
120. Maruyama S, Nakamura Y, Kabeya H et al (2000) Prevalence of *Bartonella henselae*, *Bartonella clarridgeiae* and the *16S rRNA* gene types of *Bartonella henselae* among pet cats in Japan. J Vet Med Sci 62:273–279. https://doi.org/10.1292/jvms.62.273
121. Handley SA, Regnery RL (2000) Differentiation of pathogenic *Bartonella* species by infrequent restriction site PCR. J Clin Microbiol 38:3010–3015
122. Kyme P, Dillon B, Iredell J (2003) Phase variation in *Bartonella henselae*. Microbiology 149:621–629. https://doi.org/10.1099/mic.0.26014-0
123. Hambuch TM, Handley SA, Ellis B et al (2004) Population genetic analysis of *Bartonella bacilliformis* isolates from areas of Peru where Carrion's disease is endemic and epidemic. J Clin Microbiol 42:3675–3680. https://doi.org/10.1128/JCM.42.8.3675-3680.2004

Brucella

11

Roland T. Ashford and Adrian M. Whatmore

11.1 Introduction

11.1.1 Brucellosis

Bacteria of the genus *Brucella* are responsible for causing brucellosis infections in humans, livestock and wildlife. Brucellosis remains one of the most significant zoonotic infections globally, which is endemic across a wide geographic range [1]. In many low- and middle-income countries, brucellosis represents a dual burden, causing significant human morbidity [2, 3], as well as substantial impacts on livelihoods due to reduced livestock productivity (e.g. Ref. [4]). Furthermore, in those industrialised nations where it has been eradicated from livestock populations, brucellosis remains an infrequent but significant human disease, associated primarily with travel to, or migration from, endemic regions (e.g. Ref. [5, 6]).

In humans, brucellosis is characterised by a severe febrile illness, with chronic debilitating conditions such as arthralgia, myalgia and back pain affecting a significant proportion of individuals [7]. In livestock, the disease is characterised by abortions and infertility [8]. Transmission of brucellosis to humans occurs primarily through occupational exposure to material from infected animals (e.g. when assisting with parturition) or through consumption of unpasteurised dairy products [9]. Hence, the most effective mechanism for controlling human brucellosis is to control the infection in livestock populations [2, 10]. This may be achieved by a variety of means, including the application of test and slaughter policies, and livestock vaccination.

R. T. Ashford · A. M. Whatmore (✉)
Department of Bacteriology, Animal and Plant Health Agency (APHA), Addlestone, Surrey, UK
e-mail: Adrian.Whatmore@apha.gov.uk

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_11

11.1.2 The Genus *Brucella*

The genus *Brucella* belongs to the family *Brucellaceae* (order *Rhizobiales*), within the alphaproteobacteria class of Gram-negative bacteria. The *Brucellaceae* are a diverse bacterial family, containing species with a wide range of habitat or host preferences, encompassing obligate intracellular pathogens of animals, opportunistic pathogens, plant-associated pathogens and symbionts, and environmental organisms.

For several decades, the genus *Brucella* was described as consisting of six "classical" species (*B. abortus*, *B. melitensis*, *B. ovis*, *B. suis*, *B. canis* and *B. neotomae*), with well-characterised mammalian host preferences [2]. Of these, three species (*B. melitensis*; *B. abortus* and *B. suis*) are responsible for the vast majority of human illness and loss of domestic livestock productivity [11]. Latterly, six additional species have been described, with the genus expanding to include *B. pinnipedialis* and *B. ceti* isolated from marine mammals [12], *B. microti* from voles [13], *B. inopinata* isolated from a human infection [14], *B. papionis* from baboons [15] and *B. vulpis*, isolated from foxes [16]. Both *B. inopinata* and *B. vulpis* have been described as genomically atypical of the "core" *Brucella* species, whilst *B. microti* exhibits a number of atypical phenotypic characteristics. Furthermore, there is an expanding body of literature describing the isolation of *Brucella* from amphibian hosts (e.g. Ref. [17]). These species and strains reflect an ongoing expansion of the known host range and genetic diversity of the genus (Table 11.1).

The genome of *Brucella melitensis* is 3.2 Mb in length, and is organised in two circular chromosomes, of approximately 2.1 Mb and 1.2 Mb, respectively, encoding ~3,200 coding sequences [18]. The majority of species and strains sequenced to date share a similar genomic arrangement, though some exceptions do exist [19]. Comparative genomics has demonstrated that *Brucella* genomes are characterised by high levels of synteny, with large insertions, inversions or deletions occurring relatively rarely [19]. Similarly, horizontal gene transfer has been shown to be relatively rare within the core *Brucella* species, which are considered to be largely clonal [20]. However, recently described atypical strains incorporate genomic regions exhibiting sequence identity to various soil living bacteria [17, 21]. At the sequence level, the genus *Brucella* is characterised by very low levels of genetic diversity. Whole-genome analyses demonstrate that average nucleotide identity (ANI) values amongst all members of the genus exceed the commonly recognised threshold for distinguishing between species [22, 23].

11.1.3 Significance of Molecular Typing Approaches to *Brucella* Control and Eradication

A well-established biotyping system exists for classical *Brucella* spp., which applies a suite of phenotypic assays to identify species, and further distinguish a number of "biovars" within the three species of greatest zoonotic significance [24]. Prior to the advent of molecular typing methods, these biovars represented the only sub-species

Table 11.1 Currently recognised *Brucella* species and biovars, highlighting natural host preferences and pathogenicity for humans, where known

Species	Biovar	Strain[a]	Year described	Host species	Zoonotic potential?
B. abortus	1	544^T	1920	Cattle	High
	2	86/8/59			
	3	Tulya			
	4	292			
	5	B3196			
	6	870			
	9	C68			
B. melitensis	1	16M^T	1920	Sheep and goats	High
	2	63/9			
	3	Ether			
B. suis	1	1330^T	1929	Pigs	High
	2	Thomsen		Wild boar, hare	No
	3	686		Pigs	High
	4	40		Reindeer, caribou	High
	5	513		Rodents	Unknown
B. ovis		63/290^T	1956	Sheep	No
B. neotomae		5K33^T	1957	Desert woodrat	No[b]
B. canis		RM6/66^T	1968	Dogs	Moderate
B. ceti		B1/94^T	2007	Cetaceans	Unknown[c]
B. pinnipedialis		B2/94^T	2007	Pinnipeds	Unknown
B. microti		CCM4915^T	2008	Vole, fox	Unknown
B. inopinata		BO1^T	2010	Unknown	Unknown[d]
B. papionis		F8/08-60^T	2014	Baboon	Unknown
B. vulpis		F60^T	2016	Fox	Unknown

[a]Strain identifier of species type, or biovar reference, strain
[b]Previously considered non-zoonotic but human cases recently reported
[c]MLST genotype ST27 has been associated with a number of human infections
[d]Species is currently represented by a single isolate, which was recovered from a human

categorisation available for use in epidemiological investigations [25]. However, whilst biotyping has been widely applied in brucellosis research and control, the methods involved are labour-intensive and a potential biological safety hazard to laboratory personnel. In addition, the suite of biotyping assays requires considerable experience in their application, and their interpretation can be somewhat subjective [25]. Furthermore, the biotyping scheme was initially developed on the basis of characteristics observed amongst isolates of the six classical *Brucella* species and is therefore of limited value for more recently described strains [25]. Similarly, there are numerous examples of strains of the classical species, which do not adhere well to the biotyping scheme, particularly those arising from geographical regions poorly represented in established strain collections (e.g. Ref. [26]). The

epidemiological value of biotyping has been further questioned as it has become increasingly clear that biovars do not necessarily reflect genetic groupings [27].

Molecular typing methods provide a more reliable approach to the identification and sub-species characterisation of *Brucella*, and can provide a far greater degree of resolution than previously available. Additionally, molecular typing methods can provide a direct link to the phylogeny of a given isolate, and hence are of greater value in informing epidemiological investigations. This chapter summarises the most significant methods currently employed for the molecular detection of *Brucella* and for discrimination between existing species. Widely applied molecular typing methods for characterising below the species level are then described in further detail.

11.2 PCR-Based Molecular Detection Methods

Whilst not strictly molecular typing, PCR-based methods for confirming the identity of *Brucella* spp. play a critical role in the application of molecular methods to surveillance and research activities for brucellosis. The application of PCR-based detection enables the presence of *Brucella* spp. to be confirmed without the need for extensive phenotyping, and the associated practical and biological safety implications of these procedures [28]. Initial applications of these methods utilised conventional end-point PCR approaches, in which the amplified product is visualised by agarose gel electrophoresis. Subsequently, real-time (or quantitative) PCR approaches have been developed, which have been widely adopted for diagnostic purposes.

Real-time PCR provides a number of benefits over conventional PCR approaches in a diagnostic laboratory environment [29]. The method is more rapid, as post-amplification electrophoresis is not required. Additionally, as amplification and detection are performed simultaneously, the reaction tube remains closed. This substantially reduces the risk of amplicon contamination occurring, which presents a significant hazard to the reliable application of PCR in a diagnostic environment. Furthermore, the sensitivity of real-time PCR is generally greater than conventional PCR assays for the same target (e.g. Ref. [30]).

11.2.1 *Brucella* Genus-Specific PCR Targets: bcsp31

A number of different genomic targets have been evaluated and applied for the PCR-based detection of *Brucella* organisms. Perhaps the most widely used of these is the gene encoding the 31 kDa *Brucella* cell surface protein, *bcsp*31. An initial study using this genomic target applied a conventional PCR assay and focused on only *B. abortus* and *B. melitensis*, with specificity assessed using a small panel of Gram-negative pathogens [31]. Subsequent studies demonstrated that the specificity of primer pairs amplifying this genomic target was not always perfect, with cross-reactions observed with some closely related members of the genus *Ochrobactrum*

[32, 33]. However, subsequently developed probe-based real-time PCR assays targeting the same locus appear to exhibit greater specificity. Al Dahouk et al. [34], for example, developed a dual FRET hybridisation probe (LightCycler) assay targeting *bcsp*31. This was evaluated with DNA from all classical *Brucella* species and biovars, and a large panel of non-*Brucella* organisms, and exhibited no cross-reaction in non-target organisms. Similarly, Probert et al. [35] reported a probe-based real-time PCR assay for confirmation of presumptive *Brucella* spp. isolates, which they evaluated using an extensive panel of non-*Brucella* isolates without reporting any non-specific amplification.

11.2.2 *Brucella* Genus-Specific PCR Targets: IS*711*

A second genomic locus that has been widely employed as a target for *Brucella* spp.-specific PCR assays is the insertion sequence IS*711* [36]. This insertion sequence has been identified in all *Brucella* species and strains investigated to date and is considered to be specific to the genus. Initial conventional PCR assays targeting IS*711* made use of species-specific insertion locations to identify *Brucella* to the species level (see section 4.2 for further details). Subsequently, the IS*711* sequence has been used as a target for a number of genus-specific real-time PCR assays. For example, Hinić et al. [37] describe an IS*711* sequence-based assay, which they evaluated with *Brucella* reference and field strains, as well as a panel of non-*Brucella* isolates. Similarly, Matero et al. [38] developed an IS*711*-based qPCR assay, which they validated using *Brucella* sp. strains covering six classical species, plus clinically relevant or phylogenetically related non-*Brucella* bacterial isolates.

11.2.3 *Brucella* Genus-Specific PCR Targets: 16S rRNA, *omp*, *per* and Others

Several other genomic targets for *Brucella* PCR confirmation have been applied. Romero et al. [39] performed an early evaluation of the use of a 16S rRNA-based PCR assay, using a primer pair derived from the 16S rRNA sequence of *B. abortus*. This detected all *Brucella* isolates tested, but also amplified a product in a closely related *Ochrobactrum* sp. strain. Other authors have employed primers targeting the 16S–23S intergenic spacer region [40]. Outer membrane protein (*omp*) genes have also been used in a number of studies, as targets for genus-specific PCR assays. Several different *omp* loci have been used for this purpose, including *omp*2 [41] and *omp*28, also known as bp26 [42]. Finally, the perosamine synthetase (*per*) gene has also been used as a target for genus-specific PCR assays. This gene is involved in the synthesis of the lipopolysaccharide O-chain and has been reported to be conserved amongst *Brucella* strains [43]. Lübeck et al. [44] reported the development of a conventional PCR assay targeting the *per* gene, which Bogdanovich et al. [45] adapted by incorporating a TaqMan probe. This assay detected all *Brucella* strains included in the evaluation, but none of the non-*Brucella* organisms, including a

number of serologically cross-reactive organisms known to possess a *per* gene homolog (e.g. *Escherichia coli* O157 and *Yersinia enterocolitica* 0:9).

11.2.4 Application of PCR to Clinical Samples

A very large number of studies have applied PCR-based detection methods directly to clinical samples arising from both humans and livestock. In the case of human brucellosis, for example, a large body of literature exists regarding the performance of PCR detection of *Brucella* DNA in blood or serum samples. In livestock species, other sample types, which may be more readily collected (e.g. milk) or present a greater likelihood of the target pathogen being present (e.g. foetal material), have also been the focus of considerable research. The diagnostic performance of PCR in clinical samples, relative to other methods, is considered to be outside the scope of this chapter; however, a number of previous reviews have addressed this subject (e.g. Ref. [28]). It should be noted that the sample matrix, and its processing prior to PCR (storage, sample volume, DNA extraction method, etc.), is likely to play a very significant role in the diagnostic performance of PCR-based approaches in clinical samples [28].

11.3 Early Approaches to Molecular Typing

Early molecular typing of *Brucella* spp. organisms relied on approaches which did not require extensive prior knowledge of the organism's genome sequence, but rather made use of restriction enzyme targets (e.g. pulsed-field gel electrophoresis [PFGE]; restriction fragment length polymorphism [RFLP]; amplified fragment length polymorphism [AFLP]) or the presence of short target sequences for hybridisation probes (e.g. insertion sequence [IS] typing) for discrimination between isolates. PFGE makes use of infrequently cutting restriction enzymes to generate large genomic fragments, which are then separated by gel electrophoresis, periodically switching the gel's orientation to optimise separation of large fragments. Unlike a number of other bacterial pathogens, however, PFGE never found widespread use as a routine typing tool for *Brucella*, reflecting the very limited diversity identified at the sub-species level [25]. RFLP approaches have been more widely adopted for the study of *Brucella* diversity. In particular, outer membrane protein (*omp*) genes have been used as targets for RFLP analysis, and have been demonstrated to be capable of distinguishing *Brucella* species, and biovars in a number of cases [46]. AFLP is a whole-genome fingerprinting approach, based on selective PCR amplification of restriction fragments. Application of the technique to *Brucella* was able to distinguish all of the classical species (with the exception *B. suis* and *B. canis*), but with only very limited diversity identified at the sub-species level [47]. A technique that has been more widely applied in *Brucella* research is insertion sequence (IS)-based typing, which relies on the multiple genomic locations of IS*711* elements, which can be used to discriminate between isolates [36].

Application of these "DNA banding pattern-based" approaches emphasised the high degree of genetic homogeneity observed within the genus, such that whilst several methods were capable of differentiating *Brucella* species they provided very limited resolution at the sub-species level [25]. Furthermore, these approaches highlighted a number of taxonomic ambiguities within the genus, such as the frequent inability to reliably distinguish *B. suis* and *B. canis*. However, many of these early band-based methods were methodologically demanding, and therefore difficult to reliably replicate in different locations [25]. More recently developed methods generate data, which can be more easily compared between laboratories, thereby lending themselves to the development of publically accessible international databases [25].

11.4 Molecular Typing Assays Discriminating Between *Brucella* Species

As described above, the identification of *Brucella* species has formerly relied on the application of a suite of phenotypic tests, which can be time-consuming, subjective and a potential biosafety hazard. Consequently, a wide range of molecular typing assays have been developed, which are able to reliably distinguish *Brucella* species. Broadly, these methods have either adopted a comprehensive approach, attempting to cover the diversity of the genus *Brucella* at the time of development, or a targeted approach, focusing on specific species of most significance in the context for which they were intended. Hence, several techniques have been developed, which distinguish *B. abortus* and *B. melitensis* (and often *B. suis* and *B. canis*), as the most common causes of human and livestock brucellosis, but which cannot distinguish between other members of the genus.

11.4.1 Single Loci Discriminating Between Species

A number of single loci have been investigated as potentially discriminatory targets for distinguishing between *Brucella* species, though these typically lack sufficient resolution. The classical *Brucella* species exhibit identical 16S ribosomal RNA gene sequences [48]. Atypical mammal-associated species and amphibian-derived strains exhibit only a small degree of 16S ribosomal RNA gene sequence diversity, differing from the classical *Brucella* by just a few bases (e.g. Ref. [16, 49]). Consequently, 16S rRNA sequencing approaches are of very limited use for species typing of *Brucella*. Similarly, the *recA* locus has been used for investigating relationships between *Brucella* and closely related bacteria, but exhibits no diversity amongst classical *Brucella* species [50]. This target has, however, been shown to provide a degree of resolution amongst more recently described atypical *Brucella* species and strains [17, 21, 51]. Other authors have suggested that the *rpoB* gene, encoding the beta-subunit of the DNA-dependent RNA polymerase, may represent a valuable target for discriminating between *Brucella* species, and in some cases

biovars [52]. Whilst the robustness of this approach has not been extensively evaluated, a number of authors have analysed *rpoB* gene sequences to investigate diversity in collections of field samples (e.g. Ref. [53]).

11.4.2 Multiplex Conventional PCR Approaches Discriminating Between Species

One of the first such techniques to be developed and widely applied was the AMOS multiplex PCR, named after the *Brucella* species it identified (i.e. *B. abortus*, *B. melitensis*, *B. ovis* and *B. suis*: [54]). This assay made use of unique insertions of IS*711* in the genomes of these species, by locating one primer in the insertion sequence and the accompanying primer in the unique chromosomal DNA adjacent to the insertion. However, this assay was developed specifically for the disease context in the USA, and consequently, only some biovars of these species were identified (namely *B. abortus* biovars 1, 2 and 4 and *B. suis* biovar 1). A subsequent modification of the assay incorporated additional primers, which were able to distinguish *B. abortus* vaccine strains RB51 and S19 [55]. Ocampo-Sosa et al. [56] incorporated a further primer pair, based on the *ery* locus (AMOS-ERY), which amplified an additional product in *Brucella abortus* biovars 5, 6 and 9. Nonetheless, the AMOS assay retained the disadvantage that it failed to detect certain species and/or biovars. This study also provided early evidence for incongruity between phenotypic biovars and genetic data, as *B. abortus* biovar 3 was demonstrated to form two distinct groups based on amplification of the *ery* locus.

A similar, but more comprehensive, approach was adopted by García-Yoldi et al. [57] to develop a multiplex conventional PCR assay, which could differentiate in a single reaction all the classical *Brucella* species, including those found in marine mammals, and the S19, RB51 and Rev.1 vaccine strains. Subsequently, López-Goñi et al. [58] reported the results of an inter-laboratory ring trial of the multiplex PCR assay, which largely confirmed the consistency of the results with the identification of species by classical phenotyping. These authors also coined the commonly adopted term “Bruce ladder” referring to the banding pattern of multiplexed amplification products observed following agarose gel electrophoresis. However, López-Goñi et al. [58] noted that the original assay was unable to distinguish some *B. canis* strains from *B. suis*. Mayer-Scholl et al. [59] published an advancement of the Bruce ladder multiplex PCR assay, incorporating an additional pair of primers targeting *B. microti*, which then allowed the differentiation of all recognised *Brucella* species. Most recently, López-Goñi et al. [60] proposed a further advancement to the Bruce ladder assay, to overcome the previously observed erroneous identification of some *B. canis* strains as *B. suis*. At around the same time, Kang et al. [61] proposed a similar amendment of the originally described Bruce ladder assay.

Previous Bruce ladder assay iterations did not seek to differentiate between biovars. However, López-Goñi et al. [60] additionally proposed a *B. suis*-specific multiplex assay (referred to as the Suis ladder), to distinguish *B. suis* at the biovar level and differentiate it from *B. canis* and *B. microti*. In this case, the distinction

between biovars is significant, as whilst biovars 1, 2 and 3 may cause disease in domestic pigs, biovar 2 is considered to pose a low risk of zoonotic transmission, and is endemic in wildlife across much of Europe [62].

The Bruce ladder is perhaps the most widely applied "molecular typing" tool for the study of *Brucella*, as demonstrated by the large number of published studies in which it is used. As such, it has been applied to strains isolated from a very wide geographic range and has been shown to be robust and reliable. One of the strengths of the method lies in its relative simplicity, requiring only commonly available laboratory equipment and reagents. However, as a multiplex PCR, amplifying between four (*B. abortus* S19) and eight (*B. microti*) separate fragments, it requires relatively high concentrations of DNA for reliable results to be obtained. Additionally, the Bruce ladder assay was developed prior to recently described *Brucella* species and atypical strains, and consequently, it may fail to correctly identify these. For example, Whatmore et al. [15] report that the profile of bands obtained for *B. papionis* is identical to that of *B. melitensis* and that the assay cannot reliably differentiate *B. ceti* from *B. pinnipedialis* [63].

11.4.3 Targeted Real-Time Assays

An alternative and widely applied approach to identifying *Brucella* species involves the use of real-time PCR assays specific to particular species. In several cases, such assays have been developed as multiplex assays targeting the *Brucella* species of greatest livestock significance and zoonotic potential. Similarly, several discriminatory assays have made use of species-specific insertions of the IS*711* element.

One such assay was described by Redkar et al. [64], which used three species-specific dual FRET hybridisation probe (LightCycler) assays to differentiate *B. abortus*, *B. melitensis* and *B. suis*. In this case, the forward primer for each assay was located at the 3′ end of the IS*711* element, whilst the reverse primers were located in species-specific loci for each species¸ as previously identified by Bricker and Halling [54]. Probert et al. [35] subsequently described a multiplexed real-time PCR assay capable of identifying organisms belonging to the genus *Brucella*, and differentiating between *B. abortus* and *B. melitensis*. Identification at the genus level was based on the detection of the *bcsp*31 genomic target (described above), whilst differentiation between species was based on the species-specific insertion of an IS*711* element downstream of the *alkB* gene in *B. abortus* or of locus BMEI1162 in *B. melitensis*. Al Dahouk et al. [34], however, reported that neither of these assays performed with perfect sensitivity, failing to detect a significant proportion of *B. abortus* isolates, particularly those belonging to biovars 3 and 6.

Hinić et al. [37] reported a more comprehensive panel of PCR assays, targeting all six classical *Brucella* species, which the authors suggest would be suitable for use as either conventional or real-time assays. This panel of assays was based on the identification of species-specific loci in the case of *B. melitensis*, *B. abortus*, *B. suis* and *B. ovis*, whilst assays differentiating *B. neotomae* and *B. canis* targeted unique deletions. Results presented by Hinić et al. [37] indicated that these assays were

able to distinguish the majority of biovars and field strains of these six species, although *B. suis* biovar 4 and *B. canis* could not be differentiated.

11.4.4 Single-Species Assays

Finally, numerous authors have developed real-time PCR-based assays targeting single *Brucella* species of most relevance in the specific disease context of their research. Newby et al. [65], for example, reported a stand-alone *B. abortus*-specific real-time PCR assay, utilising the IS*711* insertion element and the species-specific chromosomal locus, *alkB* as a target. However, Al Dahouk [34] subsequently demonstrated that this failed to identify a significant number of *B. abortus* isolates. Hänsel et al. [66] applied a bioinformatic approach based on alignment of available genomes, to identify a 17 bp repeat within the 3' end of the BS1330_II0657 locus, which was used for the design of a *B. suis* real-time assay. This assay was identified as specific for *B. suis* bv 1–4, with all other *Brucella* spp., including the closely related *B. canis*, lacking the insertion. However, it was also shown to be absent in *B. suis* biovar 5, meaning the assay was unable to detect all members of the species, as currently defined. Kaden et al. [67] developed a species-specific assay targeting a two base pair deletion in the acetyl-CoA acetyltransferase gene, which they describe as highly specific for *B. melitensis*.

11.4.5 Real-Time PCR Assays Based on Canonical SNPs

The availability of expanding databases of genomic sequence data from diverse and comprehensive panels of *Brucella* spp. isolates facilitated the development of a more phylogenetically robust approach to discriminating between *Brucella* isolates [25]. This approach utilised unique single nucleotide polymorphisms (SNPs) to define *Brucella* species or clades. Such "canonical" SNPs have been shown to be reliable for accurately describing phylogenetic lineages of bacteria with clonal population structures [68, 69].

Scott et al. [70], Foster et al. [71] and Gopaul et al. [72] developed similar approaches for distinguishing *Brucella* species, based on the presence of canonical SNPs. In all three cases, the assays were intended to provide a rapid and robust method for distinguishing between the six classical *Brucella* species recognised at the time of publication, plus the marine mammal *Brucella* clade subsequently formally described as two species (*B. ceti* and *B. pinnipedialis*).

Scott et al. [70] developed a multiplex SNP detection assay, based on a primer extension method targeting six SNPs in just three different loci (*omp25*, *glk* and *trpE*). The primer extension method involved the PCR amplification of the target loci, containing the canonical SNPs, followed by a reaction using "SNP interrogation primers", annealing one base short of the target SNP, to insert a dye terminator complementary to the SNP site. The extended product was then sequenced using capillary electrophoresis. The targeted SNPs were evaluated in 330 *Brucella* strains,

representing all species and biovars. The assay was able to reliably differentiate the six classical *Brucella* species plus the marine mammal clade. It was not, however, possible to identify SNP markers specific for *B. suis*, and therefore, members of this species were identified on the basis that they lacked any of the species-specific markers, and therefore had a profile distinct from other species.

Foster et al. [71] and Gopaul et al. [72] independently developed very similar approaches for distinguishing *Brucella* species using canonical SNPs. Both studies developed methods based on TaqMan minor groove-binding (MGB) probes, targeting SNP loci. In this approach, a TaqMan probe is designed for each allele state at a given SNP locus, targeting either the species-defining SNP, or the allele state found in all other members of the genus. Real-time PCR assays for each target locus incorporate probes for both allele states; the probe complimentary to the allele present in a given sample will amplify preferentially, thereby identifying the sample as belonging to a given species or not. By combining several such assays, both Foster et al. [71] and Gopaul et al. [72] were able to differentiate classical species and *Brucella* isolated from marine mammals. The panel of assays developed by Foster et al. [71], however, did not distinguish between *B. canis* and *B. suis*. Gopaul et al. [72] were able to differentiate these two species, though a species-defining SNP encapsulating all *B. suis* biovars could not be identified (with biovar 5 identified solely by the absence of other species-defining SNPs).

In order to address issues with the identification and differentiation of *B. suis* biovars, highlighted above, Fretin et al. [73] developed a panel of real-time PCR-based allelic discrimination assays. Using assays targeting four loci (*ptsP*, *pyrH*, *dnaK* and *rpoB*), allelic profiles unique for each *B. suis* biovar were defined. However, the closely related species *B. canis* remained indistinguishable from *B. suis* biovar 4 in this typing scheme.

11.4.6 High-Resolution Melt (HRM) Analysis Assays Based on Canonical SNPs

High-resolution melt (HRM) analysis provides an alternative approach to distinguishing between species-defining SNPs, which have been applied to differentiate bacterial strains and species. HRM employs an intercalating dye (typically SYBR Green) for real-time PCR amplification, which fluoresces when bound to double-stranded DNA. Amplification is followed by a melt curve, during which double-stranded DNA dissociates, with fluorescence levels recorded at fine-scale intervals across the temperature range. Differences in the nucleotide sequence of the product amplified from different strains influence the dissociation temperature, and hence fluorescence levels detected during high-resolution melt analysis.

Winchell et al. [74] developed a combined real-time PCR and HRM analysis assay, which was able to detect and differentiate six classical *Brucella* species plus marine mammal *Brucella*. However, some biovars of *B. abortus* were poorly represented, or absent, in the panel used for evaluation of this assay. Subsequent work has demonstrated that the SNP in the *glk* locus identified as *B. abortus*-specific is absent

in a subset of isolates of African origin, including *B. abortus* biovar 3 type strain (Tulya) and field isolates of biovars 3 and 6 [75]. Winchell et al. [74] highlighted that a strength of the HRM approach they developed was the "plasticity" of the assay, such that novel strains could be identified by the unique melting profiles they generate. This was demonstrated by incorporating two atypical *Brucella* strains (BO1 and BO2, the former of which was subsequently formally described as the type species, and sole representative, of *B. inopinata*). Both strains exhibited unique profiles in the panel of species-defining HRM assays. Recently, Guzmán-Verri et al. [76] have further demonstrated this principle, by applying the same HRM approach, alongside a range of other molecular tools, to characterise an atypical *Brucella* sp. isolate from a dog suffering epididymitis, which they propose as a potential novel species.

Gopaul et al. [75] also adopted a HRM analysis approach to design a multiplex assay differentiating the five terrestrial *Brucella* species of most importance to human and livestock health (*B. abortus*, *B. canis*, *B. melitensis*, *B. ovis* and *B. suis*). This assay used five separate targets, optimised to allow amplification and melt curve analysis in a single tube. Whilst this approach provided potential savings in reagent costs, it presented a number of technical difficulties, such as requiring a quantified and standardised concentration of DNA template [75].

11.4.7 Vaccine Strain-Specific Assays

Globally, three *Brucella* vaccine strains have been most widely applied (*B. abortus* RB51 and S19; *B melitensis* Rev1) and are currently recommended by the World Organization for Animal Health (Office International des Epizooties [OIE]) for brucellosis control. These vaccine strains retain the capacity to cause abortion in livestock (depending on when administered) and infection in humans. Consequently, it is important to be able to clearly distinguish vaccine strains from wild-type strains of the same species circulating in the population. Whilst vaccine strains can be distinguished by their growth characteristics in culture, the issues relating to biotyping, described above, remain.

A number of conventional PCR assays have been developed, targeting a deletion within the *eryCD* locus for *B. abortus* S19 [77], an IS*711* insertion in the *wboA* gene of *B. abortus* RB51 [78] and a SNP within the *rpsL* locus in *B. melitensis* Rev1 [79]. Furthermore, the multiplex conventional PCR approach distinguishing between species [57–60], described above, incorporates targets which discriminate these three vaccine strains. Additionally, Gopaul et al. [80] adopted the SNP-based approach previously applied to discriminating *Brucella* species in order to develop assays specific for *Brucella* vaccine strains S19, RB51 and Rev1.

11.5 Approaches to Sub-Species Molecular Typing

Two approaches, in particular, have emerged as providing powerful tools for molecular typing of *Brucella* spp., namely multi-locus sequence typing (MLST) and multi-locus variable number tandem repeat analysis (MLVA). Both of these molecular typing approaches provide resolution below the species level, and therefore have been applied in numerous phylogenetic and epidemiological studies. Additionally, both approaches benefit from standardised methods, which can be readily applied in different locations to generate directly comparable data [28]. In both cases, publically accessible databases exist, allowing results to be widely accessed and compared.

11.5.1 Multi-Locus Sequence Typing (MLST)

Multi-locus sequence typing (MLST) has become one of the most widely employed methods for characterising bacterial relationships, typically at the intra-species level [81]. MLST relies on sequencing a short section of a defined number of housekeeping genes (typically seven). Unique sequences at each locus are assigned an allele number, in order of discovery. Alleles described for each of the selected loci are combined into an allelic profile, which is assigned a sequence type (ST) designation [81].

In the case of *Brucella,* the first published MLST scheme incorporated nine loci, equating to 4,396 bp of sequence [20]. In addition to seven housekeeping genes, the *Brucella* scheme also included an outer membrane protein (*omp25*) and a partially intergenic region (int-*hyp*). This scheme was initially evaluated on 160 isolates representing the known diversity of the genus, including all biovars and vaccine strains. At that point, the genus was represented by 27 distinct sequence types (STs). However, subsequent application of the scheme to further isolates has expanded the described diversity of the genus, which is now represented by 99 STs in the 9-locus scheme (sequence database last updated 2020-07-28). Data generated using the *Brucella* MLST scheme can be submitted, along with relevant metadata (e.g. host species and country of isolation), to a publically accessible online database (https://pubmlst.org/brucella/), which is curated to avoid the inclusion of sequence errors that could generate spurious alleles and STs [81, 82]. This database currently contains data for 882 individual isolates, of which 63% belong to the three species of greatest zoonotic significance (isolate database last updated 2020-07-28).

Later work extended the 9-locus *Brucella* MLST scheme to incorporate a further 12 loci, examining diversity at 21 independent genetic sequences (10,257 bp of sequence), in order to increase discriminatory power [27]. Originally evaluated on a geographically diverse collection of over 500 isolates, representing all recognised *Brucella* species, the 21-locus scheme identified 101 STs (Fig. 11.1). Subsequent application of the scheme to further isolates has expanded this diversity, which is now represented by 150 STs (sequence database last updated 2020-07-28). This work highlighted the poor association evident between ST and biovar for both

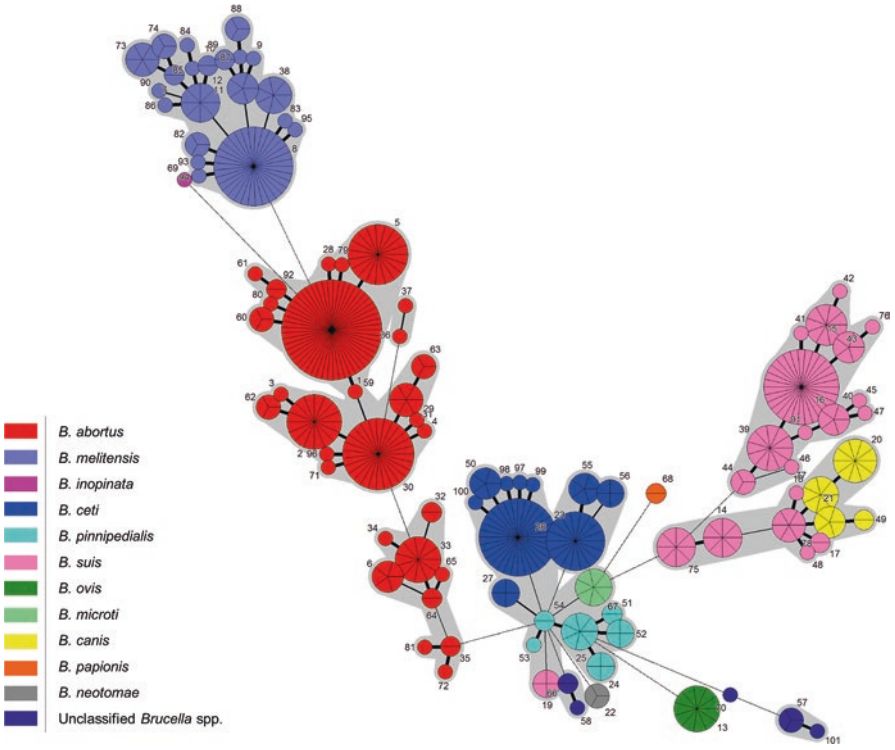

Fig. 11.1 Minimum spanning tree of 21-locus *Brucella* MLST profiles [27]. Each circle denotes a particular ST type with the size of the circle illustrating the number of isolates of that particular type. The colouring inside the circles indicates the *Brucella* species. Thick solid lines joining two types denote types differing at a single locus, thinner solid line types differing at two or three loci and the thinnest solid line types differing at four or five loci. Dashed lines indicate types differing at >5 loci. The grey halos surrounding groupings represent clusters created if neighbours differed in no more than 5 of 21 loci. (Reproduced from Ref. [27])

B. abortus and *B. melitensis*. In the case of *B. melitensis*, for example, many sequence types contained isolates belonging to several biovars, with no clear association between MLST clades and specific biovars. Furthermore, application of the extended MLST scheme also demonstrated that specific STs of some *Brucella* species exhibit a degree of geographical linkage. In the case of *B. abortus*, two clades, each containing several STs, were identified as comprised of isolates predominantly of African origin. Similarly, three broad geographical clusters of *B. melitensis* were identified based on 21-locus MLST data, approximately corresponding to the East and West Mediterranean and American clades originally identified by MLVA [83] (though isolates in the latter cluster in the MLST study were predominantly of African origin). However, there remains insufficient resolution to use MLST for epidemiological investigations at a regional, national or local geographical scale.

One of the strengths of a multi-locus sequencing approach is that it is inherently phylogenetic, therefore particularly useful for placing novel isolates in the context

of the known diversity of the genus. For example, Whatmore et al. [15] applied multi-locus sequence analysis, alongside a suite of other molecular approaches, to describe the novel species *Brucella papionis*, isolated from baboons, and to demonstrate its relationship with existing species and biovars. This approach has more recently been used to characterise novel atypical *Brucella* sp. strains from amphibians [21, 51].

11.5.2 Multi-Locus Variable Number Tandem Repeat Analysis (MLVA)

The very low rates of genetic heterogeneity observed within the *Brucella*, even when using multi-locus sequencing methods, have limited the application of such approaches for investigating local epidemiology. Consequently, multi-locus variable number tandem repeat analysis (MLVA) has emerged as the most widely adopted molecular typing tool for such studies. Variable number tandem repeat (VNTR) loci accumulate variation at a much greater rate than SNP loci, due to their mechanism of mutation [68]. Furthermore, the intrinsic rate of mutation differs between VNTR loci, generating different levels of diversity for different markers, and therefore making them suitable for investigating relationships at different evolutionary scales [68].

Several MLVA schemes have been independently developed for *Brucella* [84–87], though there is considerable overlap in the loci used in these. However, the scheme developed by Le Fleche et al. [86] has become the most widely adopted [88]. This scheme typically employs 16 VNTR loci (MLVA-16), which are divided into two panels. The first panel consists of eight loci characterised by larger, more stable, repeat units ("minisatellites"), whilst the second panel (further divided into 2A and 2B) contains loci with smaller repeat units ("microsatellites"), and generally greater levels of diversity. The adoption of a consistent MLVA scheme, using the same loci, enables data generated in different locations to be easily compared. This is further facilitated by the existence of a publically accessible online database (http://microbesgenotyping.i2bc.paris-saclay.fr/databases/public). The most recent update of this database (Brucella_4_5; created 19/12/2019) contains 6987 entries from taxonomically and geographically diverse *Brucella* isolates.

Alleles of VNTR loci may be visualised by agarose gel electrophoresis, as originally described [86], which permits high-resolution molecular typing data to be generated with laboratory equipment available in most locations. However, it can be challenging to accurately differentiate neighbouring alleles of loci with small repeat units using this method. For this reason, a number of studies have attempted to employ automated electrophoresis systems (e.g. Ref. [89]). Alternatively, PCR amplification can be undertaken using fluorescently labelled primers, allowing the size of the amplified fragment to be estimated by capillary electrophoresis [90, 91]. The use of different fluorophores additionally permits amplification and/or fragment sizing to be performed in multiplex, thereby providing cost savings and increasing efficiency.

VNTR loci are generally not considered appropriate genetic markers for phylogenetic analyses, as they may exhibit high levels of homoplasy (i.e. a shared allele state arising independently in separate lineages). However, in the case of genetically homogeneous *Brucella*, the inclusion of both moderately variable and highly discriminatory loci in the MLVA-16 scheme enables meaningful analyses at a broad geographical and temporal scale [91]. This can be achieved either by performing analyses using only the more stable loci (MLVA-8 or MLVA-11 [92–94]) or by placing greater "weighting" on stable markers during clustering analyses (e.g. Ref. [83, 95]).

This approach has recently been adopted to describe the global population structure of *Brucella*, based on the analysis of MLVA-11 data from more than 4,900 isolates [91]. This work demonstrated that analysis of these VNTR loci was able to delineate a number of major clusters within the main *Brucella* species, which exhibited substantial geographical association and were broadly consistent with the population structure described by MLST [27]. In the case of *B. melitensis*, for example, three major clusters were identified (Fig. 11.2a), broadly characterised as "East Mediterranean" (comprising strains from Europe, the Middle-East and Asia), "West Mediterranean" (strains from Mediterranean Europe and to a lesser extent Africa) and "Americas" (mainly composed of strains from the Americas and Europe). Clusters with broadly the same geographical linkages were also identified in analysis of MLST data, though in this case the "Americas" cluster contained many African isolates [27]. Also consistent with MLST, MLVA data revealed no clear relationship between genotype and biovar in the case of *B. melitensis* (Fig. 11.2b).

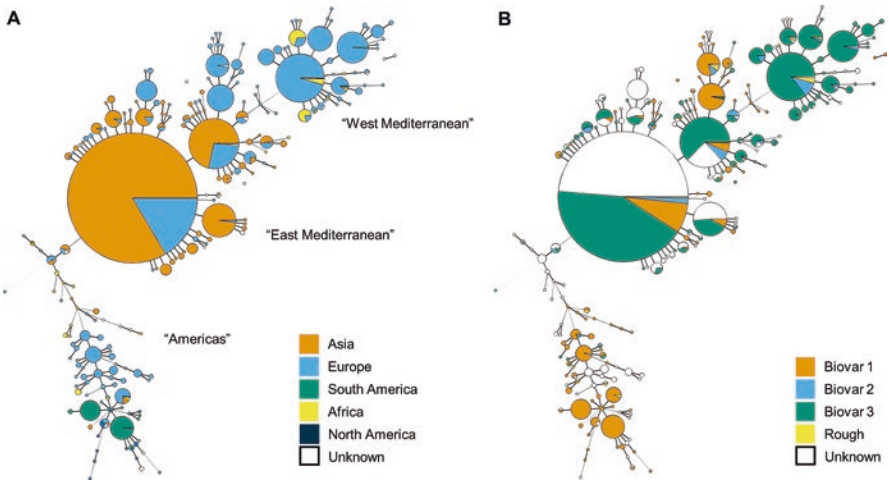

Fig. 11.2 Minimum spanning tree based on MLVA-11 data for *B. melitensis* isolates (n = 3903) retrieved from the publically accessible online *Brucella* MLVA database (reproduced after Ref. [91]). The same minimum spanning tree is colour-coded according to (**a**) continent of origin of isolate and (**b**) biovar of isolate. Grouping descriptions referred to in Figure **a** are taken from Ref. [91]. The size of each circle reflects the number of isolates of that MLVA genotype

In the case of *B. abortus*, three main clusters were identified by MLVA, which corresponded to clades identified by MLST analysis (clades B, C1 and C2 from Ref. [27]). Again, clear geographic associations were identified, with MLVA clade B comprised primarily of isolates from sub-Saharan Africa, mostly belonging to a single MLST sequence type, and clade C exhibiting a more widespread distribution from the continents of America, Asia and Europe.

The greatest value of the MLVA method, however, lies in its ability to provide fine-scale resolution between *Brucella* isolates. A very large number of studies have employed the high discriminatory capacity of MLVA to investigate *Brucella* epidemiology or population structure at a regional, national or local scale. For example, De Massis et al. [96] used MLVA-16 to investigate a *B. melitensis* outbreak in a region of Italy officially free from brucellosis in sheep and goats. MLVA-16 showed that the isolates belonged to a distinct genotype, absent from the rest of Italy, indicating that they were likely to have been introduced from Spain, where this lineage was endemic. Similarly, Hanot Mambres et al. [97] employed MLVA of *B. melitensis* strains isolated from patients in Belgium, in order to identify the likely source of infection, where such information was missing.

A further strength of the MLVA approach lies in its capacity to be adapted for a specific epidemiological context or research question. For example, Munoz et al. [62] applied MLVA-16 to investigate *B. suis* biovar 2 in European wildlife and domestic swine, selecting 11 informative loci from the broader panel of 16, which they referred to as MLVA-11_{suis2}. They were able to identify four major clades with a strong association with their geographic distribution across the continent. These clades were further broken down into clonal complexes (CCs), with more restricted geographical ranges, typically restricted to a single country.

11.6 Whole-Genome Sequencing-Based Approaches

11.6.1 Early Whole-Genome Sequencing Studies

Early applications of whole-genome sequencing to *Brucella* focused primarily on phylogenetic analyses (e.g. Ref. [98]) and comparative genomics (e.g. Ref. [99, 100]). Such studies highlighted the genetic homogeneity characteristic of the genus, in particular the classical species and more recently described core *Brucella* (e.g. Ref. [101]). Atypical strains exhibited greater diversity, and formed an early diverging basal group, though they remained distinct from other closely related bacterial species [101].

Subsequently, sequencing costs have decreased significantly, in large part due to developments in sequencing technologies [102]. Alongside the reduced costs, there have also been significant developments in bioinformatic methods available for analysing the arising sequence data. Such developments have encouraged the widespread application of whole-genome sequencing (WGS) to bacterial strain typing [102]. The increasing application of WGS approaches has expanded the number of strains for which whole-genome data are available in publically accessible

databases (e.g. European Nucleotide Archive [ENA]; National Centre for Biotechnology Information [NCBI]; DNA Data Bank of Japan [DDBJ] or Pathosystems Resource Integration Center [PATRIC]). This has, in turn, facilitated more detailed phylogenetic and epidemiological analyses.

11.6.2 High-Resolution Strain Typing Using Whole-Genome Sequencing

Broadly, two approaches are most commonly applied for high-resolution bacterial strain typing using WGS data [103–105]. The first is an extension of the MLST approach, which describes the diversity of isolates based on alleles identified for defined loci of the species, or genus, of interest [106]. Such a "gene-by-gene" approach can either be applied to all loci common to a selected panel of isolates (including both the core and accessory genome), referred to as whole-genome MLST (wgMLST), or to those loci that are conserved amongst all (or nearly all), members of the taxon, referred to as core genome MLST (cgMLST) [103].

In the case of cgMLST, a scheme is defined for the taxon of interest, typically comprising >1000 targets, and a curated database of known alleles at all loci is maintained. Sequenced isolates are compared against the database, and novel alleles are assigned a unique identifier. Hence, the strength of the approach lies in the scope to develop standardised typing schemes accessible through online databases [107]. As character-type data, wgMLST and cgMLST results are frequently represented as minimum spanning trees (MST), which represent connections linking groups of isolates by the shortest possible distance. Such analyses can be used to define the maximum number of allele differences defining epidemiologically linked isolates within a cluster.

The second approach to high-resolution bacterial strain typing using WGS data involves using high-quality SNPs observed amongst the genomes of interest to infer phylogenetic relationships. Most often, this is achieved by using a reference genome for aligning either short sequencing reads or previously *de novo* assembled genomes (typically as contigs), with SNP variants identified in samples of interest relative to the selected reference [108]. However, alignment- and reference-free approaches to identifying high-quality SNPs amongst panels of isolates have also been developed (e.g. Ref. [109]).

Whole-genome SNP analyses have the potential to provide a high level of resolution, as they differentiate isolates based on SNPs across the entire genome, including intergenic regions. However, the approach lacks the standardised methodology and transferrable nomenclature that MLST-based approaches offer [107]. The set of SNPs identified in a given analysis is dependent on the specific isolates included, and therefore may not be directly comparable between different sample sets. Additionally, SNP panels identified in analyses are typically filtered, to ensure only high-quality variants are utilised, with differing criteria and levels of stringency applied by different pipelines. Furthermore, where a reference-based SNP calling pipeline is used, the results of the analysis will be highly dependent on the selection

of the reference genome, which should ideally be as closely related to the strains of interest as possible, in order to facilitate accurate and efficient read mapping [105]. These factors make whole-genome SNP analyses attractive for high-resolution typing of closely related isolates, but currently challenging to apply as a unified typing approach [103, 104].

11.6.3 Genome-Scale MLST Approaches (cgMLST and wgMLST)

A number of studies have attempted to apply genome-wide MLST schemes for *Brucella*, or for individual species within the genus. Janowicz et al. [110] developed a cgMLST scheme for *B. melitensis* and compared its performance to both whole-genome SNP and MLVA-16 approaches. This scheme was designed using the commercially available SeqSphere+ software (Ridom GmbH, Münster, Germany), based on reference genomes from all three *B. melitensis* biovars, to identify 2,704 targets. These authors concluded that both cgMLST and whole-genome SNP analysis provided greater resolution than MLVA-16, particularly for strains belonging to the same lineage [110]. The nomenclature for this scheme is available online via a public database (https://www.cgmlst.org/ncs/schema/6398355/). However, the submission of new alleles is presently only possible using the SeqSphere+ proprietary software.

Janowicz et al. used data from epidemiologically related isolates to define the maximum number of cgMLST allele differences (or SNPs) for inclusion of a given isolate into a brucellosis outbreak cluster. Putative thresholds of 6 cgMLST genes (or 7 SNPs) were identified, though the authors caution that these values should be interpreted in the context of available epidemiological information. This *B. melitensis* cgMLST scheme has been applied in a number of subsequent studies (e.g. Ref. [6]).

An alternative *Brucella* genome-wide MLST scheme has been developed by the commercial software company BioNumerics (Applied Maths, Sint-Martens-Latem, Belgium). This scheme differs from those described above in that it employs a genus-wide "trunk" scheme of 3246 loci, with eleven "branch" sub-schemas, containing locus panels for separate *Brucella* species. Hence, these schemes can be applied to investigate the position of unknown isolates relative to the diversity of existing genomes, and to provide greater resolution within *Brucella* species. However, the scheme is currently maintained by a commercial company, and therefore, access requires the use of proprietary software.

Finally, a *B. melitensis*-specific wgMLST scheme has been developed using open-source tools [111]. In this scheme, 2656 target loci were identified using the *B. melitensis* 16M reference genome and all complete and draft *B. melitensis* genomes available within the NCBI database at the time. Strains isolated in Portugal were used to identify clusters of genetically related strains with epidemiological support. A threshold corresponding to between 6 and 11 allelic differences was identified as providing the maximum discriminatory power. Additionally, the wgMLST scheme was compared to a previously published 164 locus *Brucella* spp.

cgMLST scheme [112], and exhibited significantly greater resolution. This scheme has been made publically available (https://doi.org/10.5281/zenodo.3575026).

11.6.4 Whole-Genome SNP Approaches

The studies of both Georgi et al. [5] and Pelerito et al. [111] provide examples of approaches to selecting reference genomes for alignment-based SNP calling methods. In the former case, Georgi et al. [5] selected five *de novo* assemblies of their sequenced strains and compared these to complete *B. melitensis* genomes available in the NCBI database, in order to identify a suitable reference genome. The reference genome was then used for short-read alignment of all sequenced strains, and extracting informative SNP positions for phylogenetic analysis. An alternative approach was adopted by Pelerito et al. [111] who utilised representative draft genome assemblies of their own strains for reference-based mapping. Furthermore, these authors applied clade-specific reference genomes for mapping and variant calling of isolates belonging to two distinct clusters, previously identified through wgMLST analysis.

A number of studies have applied whole-genome SNP analysis to undertake high-resolution strain typing of *Brucella*, for purposes such as source attribution and phylo-geographic analyses. One such study employed a whole-genome SNP approach to identify the likely origin of a human *B. suis* infection in the USA [113]. The infected patient had both occupational exposure to commercial swine in the USA and epidemiological links to Tonga. A detailed whole-genome SNP phylogenetic analysis was performed, using field isolates from the USA and human isolates from patients from Tonga. This analysis demonstrated that the patient was likely to have been infected in Tonga rather than the USA, with only 7 SNPs distinguishing the index case isolate from the most closely related Tongan strain. These authors emphasised the value of international databases of WGS data that can be used in such epidemiologic investigations.

Allen et al. [114] presented a high-resolution whole-genome SNP analysis of a *B. abortus* outbreak in Northern Ireland, which occurred between 1997 and 2012, and which had previously been the subject of a detailed epidemiological investigation using MLVA [115]. The previous study employed a customised MLVA panel of hyper-variable VNTR loci, to maximise genetic discrimination of the *B. abortus* population. In the whole-genome sequencing study, a total of 345 SNPs were identified across 76 *B. abortus* isolates. Phylogenetic analysis of SNP data indicated that the outbreak was caused by two distinct lineages. The first of these contained a small number of isolates from the disease outbreak, as well as a single earlier isolate, consistent with the 1997–2012 outbreak being linked to previous endemic infection in Northern Ireland. The second lineage contained the majority of the isolates from the recrudescent outbreak, but exhibited very little population substructure, with a pairwise average of only 2.6 SNPs differentiating all isolates. This was incompatible with findings of the earlier MLVA-based analysis, which indicated that seven distinct clonal complexes were present. The authors of the later

study concluded that the selection of hyper-variable VNTR loci may have resulted in the overestimation of pathogen diversity.

11.6.5 *In Silico* Extraction of Data for Existing Typing Schemes

A major benefit of WGS approaches to molecular typing is that data for existing typing schemes can often be extracted simultaneously, allowing the wealth of information previously generated to be further exploited [116]. In the case of *Brucella*, numerous studies have employed WGS data to retrieve MLST results consistent with the existing 9 and 21 locus schemes. Jaý et al. [51], for example, determined 21-locus MLST profiles *in silico* from WGS data, in order to characterise atypical *B. microti* like *Brucella* sp. isolates from frogs. Similarly, Sacchini et al. [6] retrieved 9-locus MLST data from whole-genome sequencing results and compared these to allele profiles of *B. melitensis* strains available in the public database and containing metadata regarding their geographic origin.

Similarly, an increasing number of published studies have used whole-genome sequencing data to generate MLVA profiles *in silico*. Reconstructing VNTR loci is more bioinformatically challenging than retrieving coding sequences, however, due to their repetitive nature [107]. Nonetheless, a number of software solutions have been developed for this task, though these rely to a significant degree on the quality of the genome assembly. Georgi et al. [5] employed in-house Python scripts to extract MLVA-16 allele calls for *B. melitensis* strains isolated in Germany, which they then compared to genotypes retrieved from the *Brucella* MLVA database. In a similar study of Swedish *B. melitensis* strains [6], a tool designed for *in silico* MLVA typing of *Salmonella* strains was adapted for use with *Brucella*. Finally, Vergnaud et al. [91] employed a tool developed specifically for *in silico* MLVA typing *Brucella* from WGS data and made publically available (https://github.com/dpchris/MLVA). Although not yet extensively evaluated, existing studies have demonstrated that the short-read sequencing protocol employed can be critical in the success of retrieving MLVA-16 profiles from WGS data [91]. In particular, where Illumina paired-end reads are employed, longer read lengths (250–300bp) may be more accurate than shorter (e.g. 150bp) reads.

11.7 Conclusions and Future Directions

Modern molecular typing approaches have contributed hugely to understanding of the agents causing brucellosis, in particular to understanding the relationships within and between groups and in offering tools that can potentially aid control of the disease by better understanding epidemiology from global to local levels. This chapter charts the move from early methods, requiring no detailed knowledge of sequences, to methods based on detailed sequence knowledge that now dominate this area. Multi-locus-based approaches such as MLST, and particularly MLVA, have now become mainstays of molecular typing applied to *Brucella*.

Whole-genome sequencing will continue to become increasingly accessible, and the use of genome data, either in its entirety, or to extract MLST or MLVA data directly and more efficiently, will become increasingly frontline. Various standalone assays will likely continue to be used as rapid tools to categorise isolates into species or other subtypes. However, it needs to be remembered that the performance of these assays is only as good as the panel of isolates against which they were developed and validated. Historically, such panels have been biased towards North America and Western Europe, as the first areas to establish control programmes for brucellosis in livestock. Molecular typing is increasingly revealing previously unindexed diversity in under-sampled areas of the globe, such as Africa, that may compromise the utility of existing assays. Similarly, many of these assays were developed and evaluated prior to the recent expansion of the genus and as such may not be comprehensive. An ongoing challenge remains in developing molecular approaches directly applicable to field material—whilst this is possible for material where bacterial load is high (such as birthing/abortion materials), it remains challenging for more routinely accessible samples such as blood and milk. Further developments in this area would increase the utility of molecular typing for the group and help reduce the potential hazard and need for specialist facilities associated with handling this pathogen.

References

1. Pappas G, Papadimitriou P, Akritidis N, Christou L, Tsianos EV (2006) The new global map of human brucellosis. Lancet Infect Dis 6:91–99
2. Corbel MJ (2006) Brucellosis in humans and animals. World Health Organization, Geneva
3. Rubach MP, Halliday JE, Cleaveland S, Crump JA (2013) Brucellosis in low-income and middle-income countries. Curr Opin Infect Dis 26:404–412
4. Lokamar PN, Kutwah MA, Atieli H, Gumo S, Ouma C (2020) Socio-economic impacts of brucellosis on livestock production and reproduction performance in Koibatek and Marigat regions, Baringo County, Kenya. BMC Vet Res 16:61
5. Georgi E, Walter MC, Pfalzgraf M-T, Northoff BH, Holdt LM, Scholz HC, Zoeller L, Zange S, Antwerpen MH (2017) Whole genome sequencing of *Brucella melitensis* isolated from 57 patients in Germany reveals high diversity in strains from Middle East. PLoS One 12:e0175425
6. Sacchini L, Wahab T, Di Giannatale E, Zilli K, Abass A, Garofolo G, Janowicz A (2019) Whole genome sequencing for tracing geographical origin of imported cases of human brucellosis in Sweden. Microorganisms 7:398
7. Dean AS, Crump L, Greter H, Hattendorf J, Schelling E, Zinsstag J (2012) Clinical manifestations of human brucellosis: A systematic review and meta-analysis. PLoS Negl Trop Dis 6:e1929
8. Roop RM II, Gaines J, Anderson E, Caswell C, Martin D (2009) Survival of the fittest: how *Brucella* strains adapt to their intracellular niche in the host. Med Microbiol Immunol 198:221–238
9. Dadar M, Shahali Y, Whatmore AM (2019) Human brucellosis caused by raw dairy products: A review on the occurrence, major risk factors and prevention. Int J Food Microbiol 292:39–47

10. Roth F, Zinsstag J, Orkhon D, Chimed-Ochir G, Hutton G, Cosivi O, Carrin G, Otte J (2003) Human health benefits from livestock vaccination for brucellosis: case study. Bull World Health Organ 81:867–876
11. Godfroid J, Scholz HC, Barbier T, Nicolas C, Wattiau P, Fretin D, Whatmore AM, Cloeckaert A, Blasco JM, Moriyon I, Saegerman C, Muma JB, Al Dahouk S, Neubauer H, Letesson JJ (2011) Brucellosis at the animal/ecosystem/human interface at the beginning of the 21st century. Prev Vet Med 102:118–131
12. Foster G, Osterman BS, Godfroid J, Jacques I, Cloeckaert A (2007) *Brucella ceti* sp. nov. and *Brucella pinnipedialis* sp. nov. for *Brucella* strains with cetaceans and seals as their preferred hosts. Int J Syst Evol Microbiol 57:2688–2693
13. Scholz HC, Hubalek Z, Sedláček I, Vergnaud G, Tomaso H, Al Dahouk S, Melzer F, Kämpfer P, Neubauer H, Cloeckaert A, Maquart M, Zygmunt MS, Whatmore AM, Falsen E, Bahn P, Göllner C, Pfeffer M, Huber B, Busse H-J, Nöckler K (2008) *Brucella microti* sp. nov., isolated from the common vole *Microtus arvalis*. Int J Syst Evol Microbiol 58:375–382
14. Scholz HC, Nöckler K, Göllner C, Bahn P, Vergnaud G, Tomaso H, Al Dahouk S, Kämpfer P, Cloeckaert A, Maquart M, Zygmunt MS, Whatmore AM, Pfeffer M, Huber B, Busse H-J, De BK (2010) *Brucella inopinata* sp. nov., isolated from a breast implant infection. Int J Syst Evol Microbiol 60:801–808
15. Whatmore AM, Davison N, Cloeckaert A, Al Dahouk S, Zygmunt MS, Brew SD, Perrett LL, Koylass MS, Vergnaud G, Quance C, Scholz HC, Dick EJ, Hubbard G, Schlabritz-Loutsevitch NE (2014) *Brucella papionis* sp. nov., isolated from baboons (*Papio* spp.). Int J Syst Evol Microbiol 64:4120–4128
16. Scholz HC, Revilla-Fernandez S, Al Dahouk S, Hammerl JA, Zygmunt MS, Cloeckaert A, Koylass M, Whatmore AM, Blom J, Vergnaud G, Witte A, Aistleitner K, Hofer E (2016) *Brucella vulpis* sp nov., isolated from mandibular lymph nodes of red foxes (*Vulpes vulpes*). Int J Syst Evol Microbiol 66:2090–2098
17. Scholz HC, Muehldorter K, Shilton C, Benedict S, Whatmore AM, Blom J, Eisenberg T (2016) The change of a medically important genus: worldwide occurrence of genetically diverse novel *Brucella* species in exotic frogs. PLoS One 11:e0168872
18. DelVecchio VG, Kapatral V, Redkar RJ, Patra G, Mujer C, Los T, Ivanova N, Anderson I, Bhattacharyya A, Lykidis A, Reznik G, Jablonski L, Larsen N, D'Souza M, Bernal A, Mazur M, Goltsman E, Selkov E, Elzer PH, Hagius S, O'Callaghan D, Letesson JJ, Haselkorn R, Kyrpides N, Overbeek R (2002) The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proc Natl Acad Sci USA 99:443–448
19. O'Callaghan D, Whatmore AM (2011) *Brucella* genomics as we enter the multi-genome era. Brief Funct Genomics 10:334–341
20. Whatmore AM, Perrett LL, MacMillan AP (2007) Characterisation of the genetic diversity of *Brucella* by multilocus sequencing. BMC Microbiol 7:34
21. Al Dahouk S, Köhler S, Occhialini A, Jiménez de Bagüés MP, Hammerl JA, Eisenberg T, Vergnaud G, Cloeckaert A, Zygmunt MS, Whatmore AM, Melzer F, Drees KP, Foster JT, Wattam AR, Scholz HC (2017) *Brucella* spp. of amphibians comprise genomically diverse motile strains competent for replication in macrophages and survival in mammalian hosts. Sc Rep 7:44420
22. Ashford RT, Muchowski J, Koylass M, Scholz HC, Whatmore AM (2020) Application of whole genome sequencing and pan-family multi-locus sequence analysis to characterize relationships within the family *Brucellaceae*. Front Microbiol 11:1329
23. Leclercq SO, Cloeckaert A, Zygmunt MS (2019) Taxonomic organization of the family *Brucellaceae* based on a phylogenomic approach. Front Microbiol 10:3083
24. Alton GG, Jones LM, Pietz DE, World Health Organisation, & Food and Agriculture Organization of the United Nations (1975) Laboratory techniques in brucellosis / Alton GG, Jones LM, Pietz DE (eds) 2nd edn. World Health Organization, Geneva
25. Whatmore AM (2009) Current understanding of the genetic diversity of *Brucella*, an expanding genus of zoonotic pathogens. Infect Genet and Evol 9:1168–1184

26. Sanogo M, Abatih E, Thys E, Fretin D, Berkvens D, Saegerman C (2013) Importance of identification and typing of *Brucellae* from West African cattle: A review. Vet Microbiol 164:202–211
27. Whatmore AM, Koylass MS, Muchowski J, Edwards-Smallbone J, Gopaul KK, Perrett LL (2016) Extended multilocus sequence analysis to describe the global population structure of the genus *Brucella*: Phylogeography and relationship to biovars. Front Microbiol 7:2049
28. Whatmore AM, Gopaul KK (2012) Recent advances in molecular approaches to *Brucella* diagnostics and epidemiology. In: López-Goñi I, O'Callaghan D (eds) *Brucella*: molecular microbiology and genomics. Caister Academic Press, Wymondham, pp 57–88
29. Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, Vetter EA, Yao JDC, Wengenack NL, Rosenblatt JE, Cockerill FR, Smith TF (2006) Real-time PCR in clinical microbiology: Applications for routine laboratory testing. Clin Microbiol Rev 19:165
30. Bounaadja L, Albert D, Chénais B, Hénault S, Zygmunt MS, Poliak S, Garin-Bastuji B (2009) Real-time PCR for identification of *Brucella* spp.: A comparative study of IS*711*, *bcsp*31 and *per* target genes. Vet Microbiol 137:156–164
31. Baily GG, Krahn JB, Drasar BS, Stoker NG (1992) Detection of *Brucella melitensis* and *Brucella abortus* by DNA amplification. J Trop Med Hyg 95:271–275
32. Casañas MC, Queipo-Ortuño MI, Rodriguez-Torres A, Orduña A, Colmenero JD, Morata P (2001) Specificity of a polymerase chain reaction assay of a target sequence on the 31-kilodalton *Brucella* antigen DNA used to diagnose human brucellosis. Eur J Clin Microbiol Infect Dis 20:127–131
33. Da Costa M, Guillou J-P, Garin-Bastuji B, Thiébaud M, Dubray G (1996) Specificity of six gene sequences for the detection of the genus *Brucella* by DNA amplification. J Appl Bacteriol 81:267–275
34. Al Dahouk S, Nockler K, Scholz HC, Pfeffer M, Neubauer H, Tomaso H (2007) Evaluation of genus-specific and species-specific real-time PCR assays for the identification of *Brucella* spp. Clin Chem Lab Med 45:1464–1470
35. Probert WS, Schrader KN, Khuong NY, Bystrom SL, Graves MH (2004) Real-time multiplex PCR assay for detection of *Brucella* spp., *B. abortus*, and *B. melitensis*. J Clin Microbiol 42:1290–1293
36. Halling SM, Tatum FM, Bricker BJ (1993) Sequence and characterization of an insertion sequence, IS*711*, from *Brucella ovis*. Gene 133:123–127
37. Hinić V, Brodard I, Thomann A, Cvetnić Ž, Makaya PV, Frey J, Abril C (2008) Novel identification and differentiation of *Brucella melitensis*, *B. abortus*, *B. suis*, *B. ovis*, *B. canis*, and *B. neotomae* suitable for both conventional and real-time PCR systems. J Microbiol Methods 75:375–378
38. Matero P, Hemmilä H, Tomaso H, Piiparinen H, Rantakokko-Jalava K, Nuotio L, Nikkari S (2011) Rapid field detection assays for *Bacillus anthracis*, *Brucella* spp., *Francisella tularensis* and *Yersinia pestis*. Clin Microbiol Infect 17:34–43
39. Romero C, Gamazo C, Pardo M, López-Goñi I (1995) Specific detection of *Brucella* DNA by PCR. J Clin Microbiol 33:615–617
40. Rijpens NP, Jannes G, Van Asbroeck M, Rossau R, Herman LM (1996) Direct detection of *Brucella* spp. in raw milk by PCR and reverse hybridization with 16S-23S rRNA spacer probes. Appl Environ Microbiol 62:1683–1688
41. Mukherjee F, Jain J, Patel V, Nair M (2007) Multiple genus-specific markers in PCR assays improve the specificity and sensitivity of diagnosis of brucellosis in field animals. J Med Microbiol 56:1309–1316
42. Mitka S, Anetakis C, Souliou E, Diza E, Kansouzidou A (2007) Evaluation of different PCR assays for early detection of acute and relapsing brucellosis in humans in comparison with conventional methods. J Clin Microbiol 45:1211–1218
43. Cloeckaert A, Grayon M, Verger J-M, Letesson J-J, Godfroid F (2000) Conservation of seven genes involved in the biosynthesis of the lipopolysaccharide O-side chain in *Brucella* spp. Res Microbiol 151:209–216

44. Lübeck PS, Skurnik M, Ahrens P, Hoorfar J (2003) A multiplex PCR-detection assay for *Yersinia enterocolitica* serotype O:9 and *Brucella* spp. based on the Perosamine synthetase gene. In: Skurnik M, Bengoechea JA, Granfors K (eds) The genus *Yersinia*: entering the functional genomic era. Springer US, Boston, pp 451–454
45. Bogdanovich T, Skurnik M, Lübeck PS, Ahrens P, Hoorfar J (2004) Validated 5′ nuclease PCR assay for rapid identification of the genus *Brucella*. J Clin Microbiol 42:2261–2263
46. Al Dahouk S, Tomaso H, Prenger-Berninghoff E, Splettstoesser WD, Scholz HC, Neubauer H (2005) Identification of *Brucella* species and biotypes using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). Crit Rev Microbiol 31:191–196
47. Whatmore AM, Murphy TJ, Shankster S, Young E, Cutler SJ, Macmillan AP (2005) Use of amplified fragment length polymorphism to identify and type *Brucella* isolates of medical and veterinary interest. J Clin Microbiol 43:761–769
48. Gee JE, De BK, Levett PN, Whitney AM, Novak RT, Popovic T (2004) Use of 16S rRNA gene sequencing for rapid confirmatory identification of *Brucella* isolates. J Clin Microbiol 42:3649–3654
49. De BK, Stauffer L, Koylass MS, Sharp SE, Gee JE, Helsel LO, Steigerwalt AG, Vega R, Clark TA, Daneshvar MI, Wilkins PP, Whatmore AM (2008) Novel *Brucella* strain (BO1) associated with a prosthetic breast implant infection. J Clin Microbiol 46:43–49
50. Scholz HC, Al Dahouk S, Tomaso H, Neubauer H, Witte A, Schloter M, Kampfer P, Falsen E, Pfeffer M, Engel M (2008) Genetic diversity and phylogenetic relationships of bacteria belonging to the *Ochrobactrum-Brucella* group by *recA* and 16S rRNA gene-based comparative sequence analysis. Syst Appl Microbiol 31:1–16
51. Jaý M, Girault G, Perrot L, Taunay B, Vuilmet T, Rossignol F, Pitel P-H, Picard E, Ponsart C, Mick V (2018) Phenotypic and molecular characterization of *Brucella microti*-like bacteria from a domestic marsh frog (*Pelophylax ridibundus*). Front Vet Sci 5:283–283
52. Marianelli C, Ciuchini F, Tarantino M, Pasquali P, Adone R (2006) Molecular characterization of the rpoB gene in *Brucella* species: new potential molecular markers for genotyping. Microbes Infect 8:860–865
53. Valdezate S, Navarro A, Villalón P, Carrasco G, Saéz-Nieto JA (2010) Epidemiological and phylogenetic analysis of Spanish human *Brucella melitensis* strains by multiple-locus variable-number tandem-repeat typing, hypervariable octameric oligonucleotide fingerprinting, and *rpoB* typing. J Clin Microbiol 48:2734–2740
54. Bricker BJ, Halling SM (1994) Differentiation of *Brucella abortus* bv. 1, 2, and 4, *Brucella melitensis*, *Brucella ovis*, and *Brucella suis* bv. 1 by PCR. J Clin Microbiol 32:2660–2666
55. Bricker BJ, Halling SM (1995) Enhancement of the *Brucella* AMOS PCR assay for differentiation of *Brucella abortus* vaccine strains S19 and RB51. J Clin Microbiol 33:1640–1642
56. Ocampo-Sosa AA, Agüero-Balbín J, García-Lobo JM (2005) Development of a new PCR assay to identify *Brucella abortus* biovars 5, 6 and 9 and the new subgroup 3b of biovar 3. Vet Microbiol 110:41–51
57. García-Yoldi D, Marín CM, de Miguel MJ, Muñoz PM, Vizmanos JL, López-Goñi I (2006) Multiplex PCR assay for the identification and differentiation of all *Brucella* species and the vaccine strains *Brucella abortus* S19 and RB51 and *Brucella melitensis* Rev1. Clin Chem 52:779–781
58. López-Goñi I, García-Yoldi D, Marín CM, de Miguel MJ, Muñoz PM, Blasco JM, Jacques I, Grayon M, Cloeckaert A, Ferreira AC, Cardoso R, Corrêa de Sá MI, Walravens K, Albert D, Garin-Bastuji B (2008) Evaluation of a multiplex PCR assay (Bruce-ladder) for molecular typing of all *Brucella* species, including the vaccine strains. J Clin Microbiol 46:3484–3487
59. Mayer-Scholl A, Draeger A, Göllner C, Scholz HC, Nöckler K (2010) Advancement of a multiplex PCR for the differentiation of all currently described *Brucella* species. J Microbiol Methods 80:112–114
60. López-Goñi I, García-Yoldi D, Marín CM, de Miguel MJ, Barquero-Calvo E, Guzmán-Verri C, Albert D, Garin-Bastuji B (2011) New Bruce-ladder multiplex PCR assay for the biovar typing of *Brucella suis* and the discrimination of *Brucella suis* and *Brucella canis*. Vet Microbiol 154:152–155

61. Kang S-I, Her M, Kim JW, Kim J-Y, Ko KY, Ha Y-M, Jung SC (2011) Advanced multiplex PCR assay for differentiation of *Brucella* species. Appl Environ Microbiol 77:6726–6728
62. Muñoz PM, Mick V, Sacchini L, Janowicz A, de Miguel MJ, Cherfa M-A, Nevado CR, Girault G, Andrés-Barranco S, Jay M, Di Giannatale E, Zilli K, Ancora M, Dondo A, Zoppi S, Arnal MC, Tittarelli M, De Massis F, Garin-Bastuji B, Blasco JM, Garofolo G (2019) Phylogeography and epidemiology of *Brucella suis* biovar 2 in wildlife and domestic swine. Vet Microbiol 233:68–77
63. Whatmore AM, Dawson C, Muchowski J, Perrett LL, Stubberfield E, Koylass M, Foster G, Davison NJ, Quance C, Sidor IF, Field CL, St. Leger J (2017) Characterisation of North American *Brucella* isolates from marine mammals. PLoS One 12:e0184758
64. Redkar R, Rose S, Bricker B, DelVecchio V (2001) Real-time detection of *Brucella abortus*, *Brucella melitensis* and *Brucella suis*. Mol and Cell Probes 15:43–52
65. Newby DT, Hadfield TL, Roberto FF (2003) Real-time PCR detection of *Brucella abortus*: a comparative study of SYBR green I, 5′-exonuclease, and hybridization probe assays. Appl Environ Microbiol 69:4753–4759
66. Hänsel C, Mertens K, Elschner MC, Melzer F (2015) Novel real-time PCR detection assay for *Brucella suis*. Vet Rec Open 2:e000084
67. Kaden R, Ferrari S, Alm E, Wahab T (2017) A novel real-time PCR assay for specific detection of *Brucella melitensis*. BMC Infect Dis 17:230
68. Keim P, Van Ert MN, Pearson T, Vogler AJ, Huynh LY, Wagner DM (2004) Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. Infect Genet Evol 4:205–213
69. Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, Busch JD, Huynh LY, Dukerich M, Trim CB, Beaudry J, Welty-Bernard A, Read T, Fraser CM, Ravel J, Keim P (2007) Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. J Clin Microbiol 45:47–53
70. Scott JC, Koylass MS, Stubberfield MR, Whatmore AM (2007) Multiplex assay based on single-nucleotide polymorphisms for rapid identification of *Brucella* isolates at the species level. Appl Environ Microbiol 73:7331–7337
71. Foster JT, Okinaka RT, Svensson R, Shaw K, De BK, Robison RA, Probert WS, Kenefic LJ, Brown WD, Keim P (2008) Real-time PCR assays of single-nucleotide polymorphisms defining the major *Brucella* clades. J Clin Microbiol 46:296–301
72. Gopaul KK, Koylass MS, Smith CJ, Whatmore AM (2008) Rapid identification of *Brucella* isolates to the species level by real time PCR based single nucleotide polymorphism (SNP) analysis. BMC Microbiol 8:86
73. Fretin D, Whatmore AM, Al Dahouk S, Neubauer H, Garin-Bastuji B, Albert D, Van Hessche M, Ménart M, Godfroid J, Walravens K, Wattiau P (2008) *Brucella suis* identification and biovar typing by real-time PCR. Vet Microbiol 131:376–385
74. Winchell JM, Wolff BJ, Tiller R, Bowen MD, Hoffmaster AR (2010) Rapid identification and discrimination of *Brucella* isolates by use of real-time PCR and high-resolution melt analysis. J Clin Microbiol 48:697–702
75. Gopaul KK, Sells J, Lee R, Beckstrom-Sternberg SM, Foster JT, Whatmore AM (2014) Development and assessment of multiplex high resolution melting assay as a tool for rapid single-tube identification of five *Brucella* species. BMC Res Notes 7:903
76. Guzmán-Verri C, Suárez-Esquivel M, Ruíz-Villalobos N, Zygmunt MS, Gonnet M, Campos E, Víquez-Ruiz E, Chacón-Díaz C, Aragón-Aranda B, Conde-Álvarez R, Moriyón I, Blasco JM, Muñoz PM, Baker KS, Thomson NR, Cloeckaert A, Moreno E (2019) Genetic and phenotypic characterization of the etiological agent of canine orchiepididymitis smooth *Brucella* sp. BCCN84.3. Front Vet Sci 6:175.
77. Sangari FJ, García-Lobo JM, Agüero J (1994) The *Brucella abortus* vaccine strain B19 carries a deletion in the erythritol catabolic genes. FEMS Microbiol Lett 121:337–342
78. Vemulapalli R, McQuiston JR, Schurig GG, Sriranganathan N, Halling SM, Boyle SM (1999) Identification of an IS*711* element interrupting the *wboA* gene of *Brucella abortus*

vaccine strain RB51 and a PCR assay to distinguish strain RB51 from other *Brucella* species and strains. Clin Diagn Lab Immunol 6:760–764
79. Cloeckaert A, Grayon M, Grépinet O (2002) Identification of *Brucella melitensis* vaccine strain Rev.1 by PCR-RFLP based on a mutation in the *rpsL* gene. Vaccine 20:2546–2550
80. Gopaul KK, Sells J, Bricker BJ, Crasta OR, Whatmore AM (2010) Rapid and reliable single nucleotide polymorphism-based differentiation of *Brucella* live vaccine strains from field strains. J Clin Microbiol 48:1461–1464
81. Urwin R, Maiden MCJ (2003) Multi-locus sequence typing: a tool for global epidemiology. Trends Microbiol 11:479–487
82. Jolley KA, Maiden MCJ (2010) BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinform 11:595–595
83. Al Dahouk S, Flèche PL, Nöckler K, Jacques I, Grayon M, Scholz HC, Tomaso H, Vergnaud G, Neubauer H (2007) Evaluation of *Brucella* MLVA typing for human brucellosis. J Microbiol Methods 69:137–145
84. Bricker B, Ewalt D, Halling S (2003) *Brucella* 'HOOF-Prints': strain typing by multi-locus analysis of variable number tandem repeats (VNTRs). BMC Microbiol 3:15
85. Huynh LY, Ert MNV, Hadfield T, Probert WS, Bellaire BH, Dobson M, Burgess RJ, Weyant RS, Popovic T, Zanecki S, Wagner DM, Keim P (2008) Multiple locus variable number tandem repeat (VNTR) analysis (MLVA) of *Brucella* spp. identifies species-specific markers and insights into phylogenetic relationships. In Totowa NJ (ed). National Institute of Allergy and Infectious Diseases, NIH. Humana Press, pp 47–54
86. Le Fleche P, Jacques I, Grayon M, Al Dahouk S, Bouchon P, Denoeud F, Nockler K, Neubauer H, Guilloteau L, Vergnaud G (2006) Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. BMC Microbiol 6:9
87. Whatmore AM, Shankster SJ, Perrett LL, Murphy TJ, Brew SD, Thirlwall RE, Cutler SJ, MacMillan AP (2006) Identification and characterization of variable-number tandem-repeat markers for typing of *Brucella* spp. J Clin Microbiol 44:1982–1993
88. Scholz HC, Vergnaud G (2013) Molecular characterisation of *Brucella* species. Rev Sci Tech 32:149–162
89. De Santis R, Ciammaruconi A, Faggioni G, D'Amelio R, Marianelli C, Lista F (2009) Lab on a chip genotyping for *Brucella* spp. based on 15-loci multi locus VNTR analysis. BMC Microbiol 9:66
90. Garofolo G, Ancora M, Di Giannatale E (2013) MLVA-16 loci panel on *Brucella* spp. using multiplex PCR and multicolor capillary electrophoresis. J Microbiol Methods 92:103–107
91. Vergnaud G, Hauck Y, Christiany D, Daoud B, Pourcel C, Jacques I, Cloeckaert A, Zygmunt MS (2018) Genotypic expansion within the population structure of classical *Brucella* species revealed by MLVA16 typing of 1404 *Brucella* isolates from different animal and geographic origins, 1974–2006. Front Microbiol 9:1545
92. Ferreira AC, Corrêa de Sá MI, Dias R, Tenreiro R (2017) MLVA-16 typing of *Brucella suis* biovar 2 strains circulating in Europe. Vet Microbiol 210(Suppl C):77–82
93. Gyuranecz M, Wernery U, Kreizinger Z, Juhász J, Felde O, Nagy P (2016) Genotyping of *Brucella melitensis* strains from dromedary camels (*Camelus dromedarius*) from the United Arab Emirates with multiple-locus variable-number tandem repeat analysis. Vet Microbiol 186:8–12
94. Shevtsova E, Vergnaud G, Shevtsov A, Shustov A, Berdimuratova K, Mukanov K, Syzdykov M, Kuznetsov A, Lukhnova L, Izbanova U, Filipenko M, Ramankulov Y (2019) Genetic diversity of *Brucella melitensis* in Kazakhstan in relation to world-wide diversity. Front Microbiol 10:1897
95. Ferreira AC, Chambel L, Tenreiro T, Cardoso R, Flor L, Dias IT, Pacheco T, Garin-Bastuji B, Le Flèche P, Vergnaud G, Tenreiro R, de Sá MIC (2012) MLVA16 typing of Portuguese human and animal *Brucella melitensis* and *Brucella abortus* isolates. PLoS One 7:e42514
96. De Massis F, Ancora M, Atzeni M, Rolesu S, Bandino E, Danzetta ML, Zilli K, Di Giannatale E, Scacchia M (2015) MLVA as an epidemiological tool to trace back *Brucella melitensis* biovar 1 re-emergence in Italy. Trans Emerg Dis 62:463–469

97. Hanot Mambres D, Boarbi S, Michel P, Bouker N, Escobar-Calle L, Desqueper D, Fancello T, Van Esbroeck M, Godfroid J, Fretin D, Mori M (2017) Imported human brucellosis in Belgium: bio and molecular typing of bacterial isolates, 1996–2015. PLoS One 12:e0174756
98. Audic S, Lescot M, Claverie J-M, Cloeckaert A, Zygmunt MS (2011) The genome sequence of *Brucella pinnipedialis* B2/94 sheds light on the evolutionary history of the genus *Brucella*. BMC Evol Biol 11:200
99. Chain PSG, Comerci DJ, Tolmasky ME, Larimer FW, Malfatti SA, Vergez LM, Aguero F, Land ML, Ugalde RA, Garcia E (2005) Whole-genome analyses of speciation events in pathogenic *Brucellae*. Infect Immun 73:8353–8361
100. Halling SM, Peterson-Burch BD, Bricker BJ, Zuerner RL, Qing Z, Li L-L, Kapur V, Alt DP, Olsen SC (2005) Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. J Bacteriol 187:2715–2726
101. Wattam AR, Foster JT, Mane SP, Beckstrom-Sternberg SM, Beckstrom-Sternberg JM, Dickerman AW, Keim P, Pearson T, Shukla M, Ward DV, Williams KP, Sobral BW, Tsolis RM, Whatmore AM, O'Callaghan D (2014) Comparative phylogenomics and evolution of the *Brucellae* reveal a path to virulence. J Bacteriol 196:920–930
102. Balloux F, Brønstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, Wang H, Eldholm V (2018) From theory to practice: Translating whole-genome sequencing (WGS) into the clinic. Trends Microbiol 26:1035–1048
103. European Centre for Disease Prevention and Control (2015) Expert opinion on the introduction of next-generation typing methods for food- and waterborne diseases in the EU and EEA. ECDC, Stockholm
104. Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E (2018) Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clin Microbiol Infect 24:335–341
105. Schürch AC, Arredondo-Alonso S, Willems RJL, Goering RV (2018) Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. Clin Microbiol Infect 24:350–354
106. Maiden MCJ, van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Micro 11:728–736
107. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Kam KM, Perez E, Trees E, Kubota K, Takkinen J, Nielsen EM, Carleton H, Panel F-NE (2017) PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Eurosurveill 22:30544
108. Olson ND, Lund SP, Colman RE, Foster JT, Sahl JW, Schupp JM, Keim P, Morrow JB, Salit ML, Zook JM (2015) Best practices for evaluating single nucleotide variant calling methods for microbial genomics. Front Genet 6:235
109. Gardner SN, Slezak T, Hall BG (2015) kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics 31:2877–2878
110. Janowicz A, De Massis F, Ancora M, Cammà C, Patavino C, Battisti A, Prior K, Harmsen D, Scholz H, Zilli K, Sacchini L, Di Giannatale E, Garofolo G (2018) Core genome multilocus sequence typing and single nucleotide polymorphism analysis in the epidemiology of *Brucella melitensis* infections. J Clin Microbiol 56:e00517–e00518
111. Pelerito A, Nunes A, Núncio MS, Gomes JP (2020) Genome–scale approach to study the genetic relatedness among *Brucella melitensis* strains. PLoS One 15:e0229863
112. Sankarasubramanian J, Vishnu US, Gunasekaran P, Rajendhran J (2019) Development and evaluation of a core genome multilocus sequence typing (cgMLST) scheme for *Brucella* spp. Infect Genet Evol 67:38–43
113. Quance C, Robbe-Austerman S, Stuber T, Brignole T, DeBess EE, Boyd L, LeaMaster B, Tiller R, Draper J, Humphrey S, Erdman MM (2016) Identification of source of *Brucella suis* infection in human by whole-genome sequencing, United States and Tonga. Emerg Infect Dis 22:79–82

114. Allen AR, Milne G, Drees K, Presho E, Graham J, McAdam P, Jones K, Wright L, Skuce R, Whatmore AM, Graham J, Foster JT (2020) Genomic epizootiology of a *Brucella abortus* outbreak in Northern Ireland (1997–2012). Infect Genet Evol 81:104235
115. Allen A, Breadon E, Byrne A, Mallon T, Skuce R, Groussaud P, Dainty A, Graham J, Jones K, Pollock L, Whatmore A (2015) Molecular epidemiology of *Brucella abortus* in Northern Ireland—1991 to 2012. PLoS One 10:e0136721
116. Bletz S, Mellmann A, Rothgänger J, Harmsen D (2015) Ensuring backwards compatibility: traditional genotyping efforts in the era of whole genome sequencing. Clin Microbiol Infect 21:347.e1-347.e4

Coxiella burnetii

12

Dimitrios Frangoulidis, Mathias C. Walter,
Akinyemi M. Fasemore, and Sally J. Cutler

12.1 Introduction

In general, molecular epidemiology and surveillance based on methods which can discriminate between strains is an important and somehow mandatory technique when tracing back infections and outbreaks to their source. In addition we must not forget the biocrime and bioterrorism aspect due to the fact that *Coxiella burnetii* was a pathogen used in bioweapon programs of the cold war era and remains classified on the CDC list as a category B potential bioterrorism agent [1].

Since the first PCR-studies with *C. burnetii* in 1991 [2], several sophisticated molecular-based typing methods have been developed and applied [3–9] – mostly in research studies, sometimes in veterinary medicine and epidemiological investigations, seldomly in human medicine. The reason for this is that this "query" pathogen has a challenging history given its diverse interactions with different hosts. Looking to the animal reservoirs, especially small ruminants (goats and sheep) and cattle, we realized that regular small and sometimes larger outbreaks occur frequently worldwide. Due to an established veterinarian surveillance and notification system,

D. Frangoulidis (✉)
Bundeswehr Medical Service Headquarters VI-2, Medical Intelligence & Information (MI2), Munich, Germany
e-mail: DimitriosFrangoulidis@Bundeswehr.org

M. C. Walter
Bundeswehr Institute of Microbiology, Munich, Germany

A. M. Fasemore
University of Würzburg, Würzburg, Germany

S. J. Cutler
School of Health, Sport & Bioscience, University of East London, London, UK

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9_12

C. burnetii is often identified in animal reservoirs, frequently manifesting through abortions, low birth weight offspring, and so on. Typically, it is these events that are associated with the liberation of huge amounts of *C. burnetii* being released into the environment, creating opportunity for human infection and outbreaks [10–13].

Despite access to heavily contaminated infectious material, isolation, strain attribution, and sequencing including molecular typing are not routine procedures in veterinary medicine. On the other hand, human diagnostics for Q fever are reliant upon serology. Once specific antibodies against *C. burnetii* become detectable, *C. burnetii* DNA is rarely detectable in human samples. *Coxiella burnetii*-specific DNA can be detected by PCR, but generally only during the very early stages of human infection, before the production of an elevated antibody response. Human molecular diagnostics for Q fever are further challenged by the low amount of DNA obtained compared with animal samples, thus preventing the application of advanced sequencing techniques like typing or whole genome sequencing [14–18]. In consequence, human diagnostics should rely on a very early suspicion of acute Q fever, combining both PCR and serology. For those suspected of having chronic Q fever, mainly presenting as a subacute bacterial endocarditis, DNA detection can be recovered from heart valves [19–21].

Management of *C. burnetii* infection fully aligns with the "one health" concept whereby veterinary and human medicine must be tackled collectively for the prevention of zoonotic diseases. This was strikingly demonstrated between 2007 and 2010 when the world's largest Q fever outbreak was seen in the Netherlands. Starting from infected and aborting goats, a vast amounts of *C. burnetii* were spread, contaminating the environment and resulting in the infection of more than 4,000 people [22]. This was also the first time that molecular typing methods (mainly MLVA) were applied regularly to trace back to the source and location of infection and also to identify regional clonal clusters in an extensive manner [23, 24]. Realizing that for a small country like the Netherlands, such a large-scale outbreak posed a threat not only for health, but also for the nationwide economy, a national strategy coupled with investment resulted in the establishment of a long-lasting control and monitoring program. But this model was an exception compared to other countries that have only experienced much smaller infection outbreaks.

Collectively *C. burnetii* molecular typing and sequencing still presents real challenges, not only in veterinary medicine, but especially in human medicine, and is far beyond becoming a routine diagnostic tool for molecular epidemiology and surveillance. In this chapter we give a short summary and utility of established typing methods, showing efforts to overcome the inherent problems of low target DNA-amounts, and providing insights into current database networks facilitating typing and sequencing information applicable to those from a broad range of stakeholders from researchers, veterinarians and those working in human medical disciplines.

12.2 The Pathogen

Q fever is a zoonosis caused by the Gram-negative coccobacillus *C. burnetii* belonging to the class Gammaproteobacteria. Despite the worldwide distribution of the agent, Q fever is rarely reported, probably as a result of both clinical and diagnostic challenges presented by *C. burnetii* infection compounded by the lack of requirement for reporting cases in many countries [25]. Despite this, it persists with a large reservoir among multiple species, regularly leading to outbreaks.

The organism is highly infectious, with experimental estimates suggesting an infectious dose of less than 10 organisms. Furthermore, they are highly resistant against both heat and desiccation, have a ubiquitous distribution, are relatively easy to cultivate, and remain infectious over several kilometers when aerosolized. These characteristics resulted in the inclusion of *C. burnetii* among agents tested in the old biological weapon programs of the United States and the former Soviet Union. Similarly, this was also the justification for its inclusion on the CDC-list as a category B potential bioterrorism agent [1].

Against this background, recent attention focused upon *C. burnetii* has highlighted our limited understanding of its epidemiology, population diversity, ecology, and basic biology. Huge improvements in diagnostic capabilities have been achieved with the introduction of molecular techniques, although serological methods still play a vital role in diagnosing human infections and for population screening. Whole genome sequence data has demonstrated the genetic diversity among strains of *C. burnetii*, and this diversity has enabled the design and application of high-resolution molecular typing systems that can be useful for epidemiologic investigations of infectious episodes.

In addition, these methods have also the potential to differentiate between natural and deliberate release of the pathogen. Another important aspect is the appreciation of the background level of this pathogen in the environment whereby it might circulate through natural ecological interactions. Due to the ubiquitous occurrence of *Coxiella* throughout the world, relevant "traces" of this pathogen could complicate the interpretation of bio-forensic investigations.

Molecular epidemiological investigations can be hindered by diagnostic delays resulting from the very low amounts of DNA often seen in clinical human material; consequently trace back analysis of cases is rarely performed. In recent years some promising publications presented data that could make it possible to do such epidemiological analysis as seen in other diseases (e.g., tuberculosis, salmonellosis) in the future (Driscoll 2009; Malorny et al. 2008).

12.3 Molecular Typing Methods

Due to the zoonotic nature of the pathogen veterinary-based strains are mostly isolated from various animal species. Beyond hen eggs, animal inoculation and cell culture techniques have played a valuable role, complemented in recent years by successful new axenic culture media providing the ability to isolate *Coxiella* strains

with less cumbersome laboratory work [26, 27]. These techniques are very important to gain enough material for genomic typing, overcoming the limitations of restricted amounts of DNA from acute and chronic human disease samples.

12.3.1 History – Typing Is an "Old" Science in Q Fever

Shortly after the publication of the first specific *C. burnetii* PCR in 1991, different genomic targets were used for differentiation issues of strains. Some of them are still in current use. The *determination of the plasmid type* was the first attempt to identify markers synonymic for acute or chronic disease [42]. Although this hope has not been confirmed and the real impact of the four different plasmid types to the course and outcome of the Q fever disease is still not known [28], differentiation into plasmid types became an accepted method to categorize strains, isolates, and their prepared DNA [8]. Most of the *Coxiella* pathogens worldwide harbor the QpH1-plasmid. But QpRS [29] is seen on the Australian continent, and clusters of QpDV-positive strains have been reported from France [30]. A very limited number of *Coxiella* possess QpDG, with some isolates showing a chromosomal integrated plasmid [31].

Other parts of the genome that have been targeted for typing include *mucZ* [32], *com1* [33], *icd* [34], and *16S/23S rRNA* [35, 36], and also restriction fragment length polymorphisms (*RFLP*) subsequently analyzed with SDS-PAGE and pulsed-field gel electrophoresis (PFGE) [8, 37–39]. These techniques are no longer used due to their minimal or variable discriminatory power or lack of standardization to achieve reproducible results (inter- and intralaboratory reproducibility).

12.3.2 Today – Reduced to the Max – A Lot of Epidemiological Studies – No Use in Routine Diagnostics

A reliable and more reproducible typing method for *C. burnetii* was published 2005 by Glazunova, *Multispacer-Sequence-Typing* (*MST*) [7]. For the first time discrimination to geographical regions and/or disease-associated patterns were possible, not in a perfect manner, but for orientation issues acceptable. Another important resource was introduced from this French group enhancing molecular epidemiology and surveillance of Q fever: a web-based database containing typing results and also with the ability to update the information by the database hosting institute in Marseille. Shortly after this publication the second technique still today widely used was introduced and published 2006 – *the multiple loci variable number of tandem repeat analysis* (*VNTR or MLVA*) [3, 40] – and also establishing a web-based database, which we will discuss later.

The MLVA *method* exploits the variability in copy numbers of tandem repeats that exist within a chromosome. In *C. burnetii*, up to 17 different target regions have been used for strain characterization [18]. Recent work from Frangoulidis et al. [6],

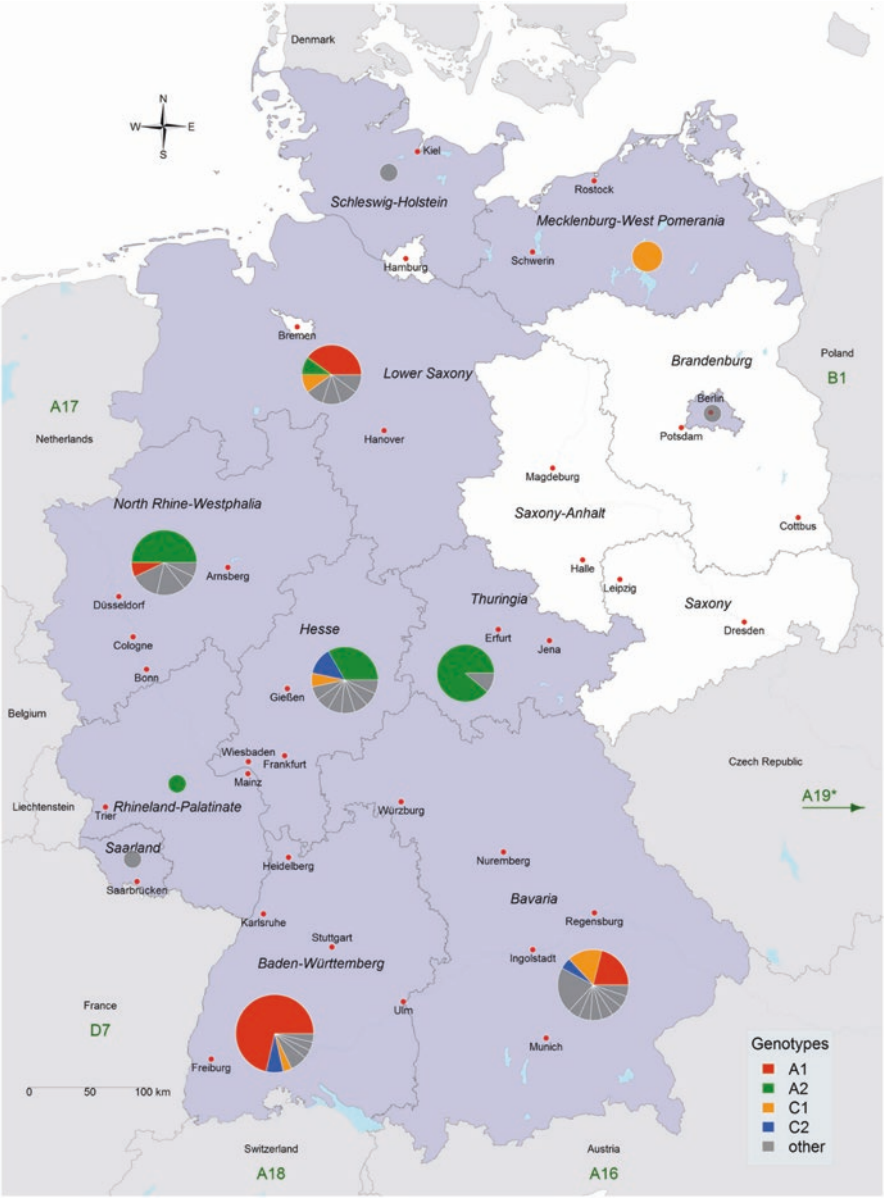

Fig. 12.1 Distribution of genotypes in Germany and representative genotypes of neighboring countries [6]

where they exploited 14 MLVA markers revealed that the MLVA method is highly discriminative enough to even characterize strains from similar geographic locations (Fig. 12.1). Another study by Svraka et al. also revealed that the MLVA method has higher discriminatory power than MST methods [40].

Clinical application of this method for diagnostics was exploited with success during the Q fever outbreak in Netherlands between 2007 and 2010. Klassen et al. [41] used a three-marker panel to analyze clinical samples from both acute and chronic Q fever cases; they collected swabs, sputum, urine, and plasma samples from both humans and animals. They reported that they could characterize four different genotypes. This was a great milestone as this was the first time that a non-serological method was applied in a clinical diagnostics setting in human medicine.

Nevertheless several challenges for this method remain, one of which is the lack of consensus for individual markers, essentially making harmonization of results error prone and not compatible between laboratories [18], also genomic instability of repeat markers means that they can theoretically fluctuate although the selected MLVA markers in *C. burnetii* have been deemed stable [40], future fluctuations cannot be completely ruled out. Nevertheless, through selecting a subset of MLVA-markers good discriminatory power providing a suitable and fast way for strain classification was demonstrated (Fig. 12.2) [6].

Another PCR-based method that has been used to characterize *C. burnetii* strains utilizes *adaA* gene detection and variation. Identified in 2005 [9] it was thought to be a specific marker for acute Q fever disease (*a*cute *di*sease *a*ntigen A = adaA). In a more detailed analysis of the gene region, a much more variability was identified (Fig. 12.3) [5]. Today this method gives additional information to discrimination patterns in animal and human derived materials.

A variation and enhancement upon the MST-techniques described above is provisioned through the use of a *SNP* (*S*ingle *N*ucleotide *P*olymorphism)-based method applied not only to the intergenic regions like the MST but additionally using SNP-variations in the whole genome sequence (Hornstra et al. 2014). It shows a good correlation with other typing methods, particularly providing good concordance with the existing plasmid type distributions, but discriminatory power is variable. As we discuss later, *SNP*-based pattern analysis is especially interesting when

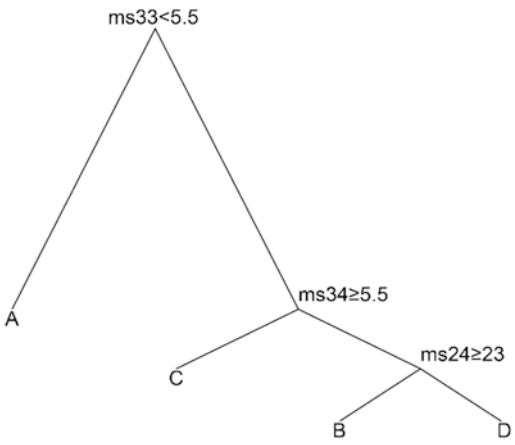

Fig. 12.2 Classification tree grouping the genotypes into four clusters according to the alleles of the *C. burnetii* VNTR markers ms33, ms34 and ms24 [6]

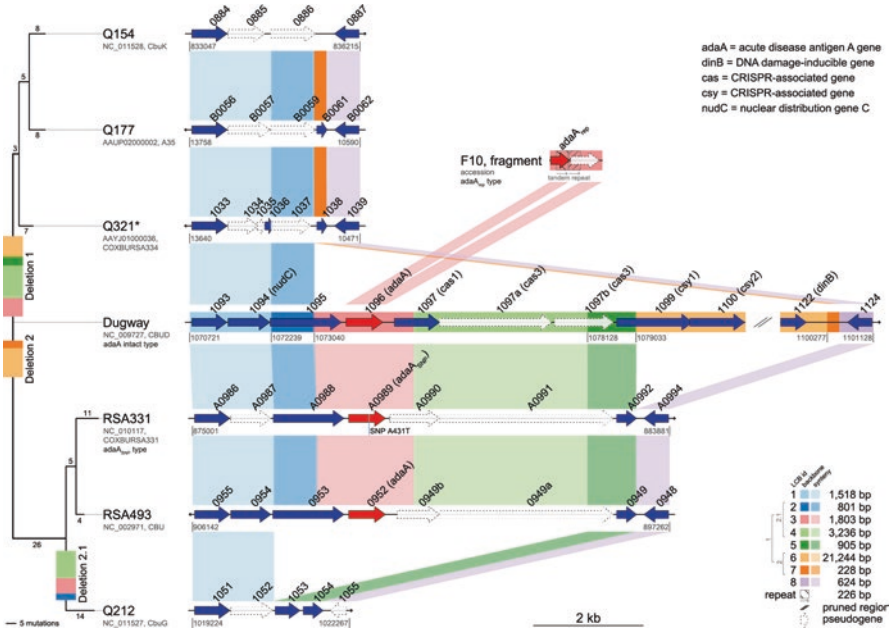

Fig. 12.3 Polymorphisms of the adaA genetic region demonstrated using seven *C. burnetii* reference genomes [5]

analyzing whole complete genomes and has the potential to provide the most robust comparative molecular epidemiological method. Beside an existing bioinformatic bottleneck, there are considerable other challenges that need to be tackled before whole genome sequencing techniques are suitable for diagnostic and routine use.

Finally, the *IS1111* typing method, based on a paper from 2007 [4], is worthy of discussion. Using 10 different upstream regions from the repetitive *IS1111* element in *C. burnetii*, strain discrimination was possible and produced reliable results. Though less popular when compared to MST and MLVA, this method has the potential to be a typing standard in a large number of diagnostic laboratories. Recent data from our group (unpublished) suggests over 80 genomic target points that could be explored, suggesting a high discriminative power. Remembering that this is a PCR-product-based method, it would be technologically feasible for most diagnostic laboratories to implement, providing promising potential for the future application in clinical diagnostic settings.

12.3.3 On the Edge to Tomorrow – Whole Genome Sequencing and Enrichment Techniques

The genomics of *C. burnetii* dated 35 years back to 1985 where the QpRS and QpH1 plasmids were analyzed by Samuel et al. using restriction enzymes [42]. The first DNA sequences of the heat-shock operon and the 16S rRNA were published in 1988 [43, 44]. Five years later, in 1993, the first complete sequence of the plasmid

QpH1 was published [45]. Then it took almost 10 years to sequence the complete genome of *C. burnetii* reference strain Nine Mile RSA 493 [46]. In 2005, Seshadri and Samuel proposed the sequencing of six additional isolates [47] which were finally sequenced in 2008 by a whole genome random shotgun-sequencing approach [48] using the Sanger capillary sequencing technology. In 2009 the first comparative genomics study of these five *Coxiella* genomes was published [49].

Thanks to the development of the axenic medium (Acidified Citrate Cysteine Medium [ACCM]) in 2009 [27] and its several improvements [50–53], *C. burnetii* can now grow outside of host cells. Despite the introduction and improvements of the axenic media, there are still some strains of *C. burnetii* more refractory, grow with lower efficiency, or sometimes show no growth in ACCM2 [26, 54]. Nevertheless, the number of sequenced isolates has now increased tenfold by using next-generation sequencing (NGS).

A downside of the simpler NGS approach is that the sequenced fragments are much shorter compared to Sanger sequencing reads and no complete genomes containing several identical regions can be reconstructed from them. Therefore, a large part of the available *C. burnetii* genomes are fragmented, existing in the form of several contigs or few scaffolds with several gaps (Fig. 12.4). Also, many of them were aligned against the reference genome Nine Mile RSA 493 and artificially concatenated afterward, rather than aligning them to a genome graph constructed from all reference genomes. That leads to incorrectly assembled "whole" genomes and makes it difficult to analyze genome structure dynamics.

Obtaining sufficient amounts of DNA for sequencing remains a major constraint. With upstream cultivation in cell culture or axenic medium, only a limited amount of starting material with a maximum concentration of 10^9 copies/µl can be produced, even with repeated applications. Much more problematic, however, is the significant damage or loss of DNA during inactivation. Hence, only very new *C. burnetii* strains were sequenced [55–57] with the third-generation sequencing technology such as single-molecule real-time (SMRT) sequencing [58] introduced by Pacific Biosciences. Currently, only the attenuated strain Nine Mile RSA 439 has

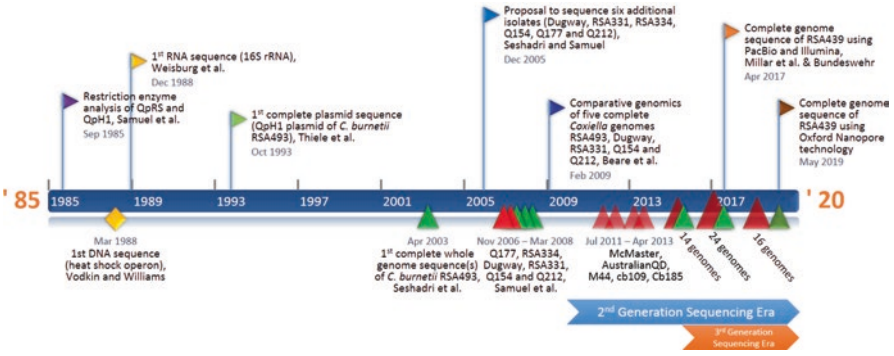

Fig. 12.4 History of *C. burnetii* genome sequencing (red triangles = incomplete genomes, green triangles = complete genomes)

also been sequenced using nanopores [59] introduced by Oxford Nanopore and uploaded to the NCBI BioProject (accession PRJNA541166).

Another approach to obtain higher amounts of DNA is whole genome amplification (WGA) using either rolling circle amplification [60] or multiple displacement amplification [61]. Technologies such as WGA can produce sufficient amounts of DNA from very low starting material even from single cells [62], but conversely introduces several problems like coverage bias or chimera formation [63], causing artificial rearrangements and inversions [64] and copy-number alterations [65] leading to mis-assemblies.

The first whole genome amplification of *Coxiella* DNA was performed by Kumar et al. and evaluated by the comparison of un-amplified and amplified DNA templates at 20 selected loci with a fold-amplification in the range of 3×10^2 to 192×10^2 [66]. Reis investigated the influence of three different WGA kits on the *Coxiella* genome and specific loci used for MLVA and *IS1111* typing. If the DNA is highly fragmented (<2 kb) like after heat inactivation, GenomiPhi® can be used if the starting concentration is at least 1000 genome copies/µl. GenomePlex® also performed well but introduced many errors at the MLVA loci just by not amplifying them. The best-performing kit introducing the least errors and producing the highest yield of DNA was REPLI-g® but requires less fragmented DNA only [67]. Based on these studies, the genome sequence of *C. burnetii* strain Namibia was assembled using a hybrid approach by combining amplified DNA sequences with *IS1111* typing information. Therefore, any chimeric *IS1111* element introduced by MDA could be corrected [68].

One drawback of WGA is that not only the DNA of interest gets amplified; any background DNA (i.e., host DNA in case of clinical samples) gets amplified as well. In many cases, the starting material contains much less *Coxiella* DNA (target) than background DNA. Because the higher amount of background DNA is more accessible during amplification, the ratio of target to background DNA can become worse, even the total amount of *Coxiella* DNA is increased after amplification. That makes it harder to sequence and a higher sequence depth is necessary. To overcome this problem, one can try to deplete host DNA for instance with the NEBNext® Microbiome DNA Enrichment Kit, which depletes CpG-methylated host DNA, but works well only for human DNA. More promising is the application of selective whole genome amplification (SWGA) which amplifies target genomes using nucleotide sequence motifs that are common in the target microbe genome, but rare in the background genomes [69]. This was already tested to sequence *C. burnetii* directly from environmental samples resulting in a 68- to 147-fold enrichment but is limited to non-degraded DNA [70] and likely produce an uneven breath of genome coverage.

Even more promising is the application of target enrichment strategies such as a custom hybridization-based capture process where an adapter-ligated DNA sequencing library is hybridized to single-stranded biotinylated oligonucleotides (baits) complementary to the region of interest. Once the hybridization is complete, the baits are captured using streptavidin magnetic beads and washed to remove nonspecific fragments. The captured fragments are eluted and PCR amplified to obtain enough material for sequencing [71]. Several manufactures offer custom

enrichment kits based on baits designed to target a specific region or whole genome. A study comparing the different kits for *Coxiella* genome sequencing is ongoing. Currently, this target enrichment approach is limited to next-generation sequencing libraries and cannot be used with third-generation sequencing platforms. To sequence long target regions of interest, a Cas9-guided adapter ligation approach using nanopore sequencing was developed [72]. This technique still requires a significant amount of input DNA and is currently not applicable to whole bacterial genomes but likely can be used to sequence longer repetitive elements like the IS elements of *C. burnetii*.

Every typing method described above can also be applied in silico to the steadily increasing number of available (near-) complete *Coxiella* genomes fitting them back into the established typing schemes or even used to rapidly classify them during outbreak investigation. So, when solving the above-described problems according to the quality and quantity of specific DNA, this will maybe the solution for typing issues of *C. burnetii* in the future.

12.3.4 Database Issues

Molecular genotyping of pathogenic species is a data-generating process with public health implications; hence, it is essential that a platform for storage, retrieval, analysis, and sharing of generated data should accompany this process. This is often in the form of a publicly available database where historical strains as well as their typing results are collated alongside meaningful metadata. This resource often serves as the warehouse for future typing inquiries as well as strain discovery. Before 2005, no such resource was available for *C. burnetii* genotyping data, whereas public database resources were already available for other pathogens like *Staphylococcus aureus* [73].

Molecular genotyping methods for *C. burnetii* are varied, differing in their output data and often incompatible with other methods; hence, there is a challenge to collate all the available methods within a single resource, which can serve as a repository for future isolates and for data interrogation during outbreaks.

The first open *Coxiella* genotyping resource was made available in 2005 by Raoult's group in France (URL https://ifr48.timone.univ-mrs.fr/mst/coxiella_burnetii). It focused solely on the MST typing method developed by the same group. This resource provided comprehensive data regarding the MST method that included a list of all available spacer allele sequences in FASTA format, strain lists of historical isolates and their MST genotyped groups, a static UPGMA tree for all known MST groups, a MST in silico sequence typing method, and instruction on how to submit novel spacers. This resource failed to address the challenge to provide further genotyping resources aside the MST methods. Also to highlight some other challenges, there was no facility to query MST groups based on spacer information. Although a comprehensive list is provided, scanning this table visually to locate the MST group of multiple spacers would be a time-consuming task. Also the only

provided metadata for the strain list is the origin of the strain and the geographic source with no filter feature to exclude certain strains for investigative purposes.

The second resource is MLVA bank from Vergnaud's group, also from France, published in 2008 (https://microbesgenotyping.i2bc.paris-saclay.fr). This database is a resource that houses MLVA data for several bacterial species including *C. burnetii*. It includes a private space for users to analyse data they do not want to show to the public. The MLVA method has been developed across several laboratories using different methods; this resource caters for the different methods and offers a query method based on MLVA repeats.

Dynamic UPGMA trees can also be created based on MLVA profiles, and it was the first resource that extended *Coxiella* genotyping data with map graphics, making it easier for researchers and interested users to gather insights based on geographical locations. It also failed to address the challenge of providing other genotyping resources for *C. burnetii* essentially only focusing on the MLVA typing method.

The third and most recent *Coxiella* genotyping resource is CoxBase which is from the Frangoulidis group based in Germany [74]. It was made available to the public in 2019 and it is available at URL https://coxbase.q-gaps.de. It is a comprehensive resource that attempts to solve the challenge to provide multiple *Coxiella* typing resources under the same platform alongside comprehensive metadata for annotation purposes. This provides the opportunity to compare different genotyping methods with as little interruptions as possible. As of October 2020, the CoxBase platform can be used for the analysis of six different *Coxiella* genotyping methods: MLVA [6], MST [7], *adaA* [5], SNP [75], *IS1111* [4], and plasmid typing. This approach allows a researcher to cover all areas of *Coxiella* typing within a few minutes. Aside from the novel approach of combining the varied typing methods, it is also the first platform which provides a retrieval function for *Coxiella* genotyping data, in which completed genotyping analysis can be retrieved at a future date from any location across the globe for reproducibility purposes. It houses genotyping data from more than 400 *Coxiella* isolates and provides several means of isolate discovery and comparison such as query aggregation as well as search features. It also provides visualization features for the distribution of typing data across several countries (Fig. 12.5).

The CoxBase platform also extends the generation of dynamic phylogenetic trees from genotyping analysis, enabling researchers and other stakeholders to compare new isolates to historical isolates based on their MLVA profiles and also metadata analysis using features such as genotype group, MLVA, MST, *IS1111*, *adaA*, plasmid type, source of the isolate, as well as year of isolation to generate hypothesis for epidemiological research. CoxBase also provides submission features for MLVA and MST profiles via web interface; it would be promising to extend the submissions for *IS1111* as well as SNP data.

Database management for *Coxiella* typing data is still challenged by the total number of available *Coxiella* isolates as well as a comparison bias as a result of most of the isolates being of European origin. Active surveillance as well as mandatory reporting and sharing of data especially in geographic regions with little known

Fig. 12.5 Distribution of *Coxiella* isolates in the CoxBase database visualized on the world map. A cluster of isolates is apparent in the European continent; this can be attributed to several factors, chiefly among them will be due to the legislation of mandatory reporting of *C. burnetii* outbreaks as well as data from previous outbreaks

C. burnetii data might help to tackle this challenge since *C. burnetii* is globally distributed.

12.4 Conclusions

Reflecting the different issues, we presented and discussed in our chapter, molecular typing of the Q fever pathogen *C. burnetii* remains far away becoming a routine method in daily diagnostic work, widely accepted from other pathogenic microbes, e.g., MRSA. But realizing the impact of this zoonotic disease in nearly all the countries of the world, molecular, sequencing data is an important resource, which must be created and updated. Starting with veterinary specialized laboratories, the sequence information of strains and variations must be the backbone of a universal, global open database. Reference laboratories in countries spanning veterinary and human medicine should facilitate and offer specialized typing power to the stakeholders to enlarge existing sequencing data. Furthermore, they should be encouraged to conduct whole genome sequencing and share the resulting data with the wider community through data repositories. In addition, all these data should be deposited in centralized databases, like NCBI, but also a more pathogen focused modern and intuitive way to handle sequencing and typing data is mandatory in the future to overcome the barriers between basic research, applied science and stakeholders in biology, veterinary and human medicine. This will overcome most of the existing problems in *Coxiella* typing. Data, information, and finally knowledge management will be the key feature for future work in a one health approach against infectious, zoonotic diseases like the Q fever disease of *C. burnetii.*

References

1. Rotz LD, Khan AS, Lillibridge SR, Ostroff SM, Hughes JM (2002) Public health assessment of potential biological terrorism agents. Emerg Infect Dis 8(2):225–230
2. Hendrix LR, Samuel JE, Mallavia LP (1991) Differentiation of *Coxiella burnetii* isolates by analysis of restriction-endonuclease-digested DNA separated by SDS-PAGE. J Gen Microbiol 137(2):269–276
3. Arricau-Bouvery N, Hauck Y, Bejaoui A, Frangoulidis D, Bodier CC, Souriau A et al (2006) Molecular characterization of *Coxiella burnetii* isolates by infrequent restriction site-PCR and MLVA typing. BMC Microbiol 14
4. Denison AM, Thompson HA, Massung RF (2007) IS1111 insertion sequences of *Coxiella burnetii*: characterization and use for repetitive element PCR-based differentiation of *Coxiella burnetii* isolates. BMC Microbiol 7:91
5. Frangoulidis D, Splettstoesser WD, Landt O, Dehnhardt J, Henning K, Hilbert A et al (2013) Microevolution of the chromosomal region of acute disease antigen a (*adaA*) in the query (Q) fever agent *Coxiella burnetii*. PLoS One 8(1):e53440
6. Frangoulidis D, Walter MC, Antwerpen M, Zimmermann P, Janowetz B, Alex M et al (2014) Molecular analysis of *Coxiella burnetii* in Germany reveals evolution of unique clonal clusters. Int J Med Microbiol IJMM 304(7):868–876
7. Glazunova O, Roux V, Freylikman O, Sekeyova Z, Fournous G, Tyczka J et al (2005) *Coxiella burnetii* genotyping. Emerg Infect Dis 11(8):1211–1217
8. Jäger C, Willems H, Thiele D, Baljer G (1998) Molecular characterization of *Coxiella burnetii* isolates. Epidemiol Infect 120(2):157–164
9. Zhang G, To H, Russell KE, Hendrix LR, Yamaguchi T, Fukushi H et al (2005) Identification and characterization of an Immunodominant 28-Kilodalton *Coxiella burnetii* outer membrane protein specific to isolates associated with acute disease. Infect Immun 73(3):1561–1567
10. Álvarez-Alonso R, Basterretxea M, Barandika JF, Hurtado A, Idiazabal J, Jado I et al (2018) A Q fever outbreak with a high rate of abortions at a dairy goat farm: *Coxiella burnetii* shedding, environmental contamination, and viability. In: Dozois CM (ed) Appl Environ Microbiol [Internet]. 84(20). Available from: http://aem.asm.org/lookup/doi/10.1128/AEM.01650-18
11. Bjork A, Marsden-Haug N, Nett RJ, Kersh GJ, Nicholson W, Gibson D et al (2014) First reported multistate human Q fever outbreak in the United States, 2011. Vector-Borne Zoonotic Dis 14(2):111–117
12. Gilsdorf A, Kroh C, Grimm S, Jensen E, Wagner-Wiening C, Alpers K Large Q fever outbreak due to sheep farming near residential areas, Germany. Epidemiol Infect 136(08) Available from: http://www.journals.cambridge.org/abstract_S0950268807009533
13. Porten K, Rissland J, Tigges A, Broll S, Hopp W, Lunemann M et al (2006) A super-spreading ewe infects hundreds with Q fever at a farmers' market in Germany. BMC Infect Dis 6(1) Available from: http://bmcinfectdis.biomedcentral.com/articles/10.1186/1471-2334-6-147
14. Bae M, Jin CE, Park JH, Kim MJ, Chong YP, Lee S-O et al (2019) Diagnostic usefulness of molecular detection of *Coxiella burnetii* from blood of patients with suspected acute Q fever. Medicine (Baltimore) 98(23):e15724
15. Fournier P-E, Raoult D (2003) Comparison of PCR and serology assays for early diagnosis of acute Q fever. J Clin Microbiol 41(11):5094–5098
16. Hartzell JD, Wood-Morris RN, Martinez LJ, Trotta RF (2008) Q fever: epidemiology, diagnosis, and treatment. Mayo Clin Proc 83(5):574–579
17. Keijmel SP, Krijger E, Delsing CE, Sprong T, Nabuurs-Franssen MH, Bleeker-Rovers CP (2015) Differentiation of acute Q fever from other infections in patients presenting to hospitals, the Netherlands1. Emerg Infect Dis 21(8):1348–1356
18. Massung RF, Cutler SJ, Frangoulidis D (2018) Molecular typing of *Coxiella burnetii* (Q Fever). In: Toman R, Heinzen RA, Samuel JE, Mege J-L (eds) Coxiella burnetii: recent advances and new perspectives in research of the Q fever bacterium [Internet].

Springer, Dordrecht, pp 381–396. Available from: http://www.springerlink.com/index/10.1007/978-94-007-4315-1_19
19. Gauduchon V, Chalabreysse L, Etienne J, Célard M, Benito Y, Lepidi H et al (2003) Molecular diagnosis of infective endocarditis by PCR amplification and direct sequencing of DNA from valve tissue. J Clin Microbiol 41(2):763–766
20. Jang Y-R, Song JS, Jin CE, Ryu B-H, Park SY, Lee S-O et al (2018) Molecular detection of *Coxiella burnetii* in heart valve tissue from patients with culture-negative infective endocarditis. Medicine (Baltimore) 97(34):e11881
21. Lepidi H, Houpikian P, Liang Z, Raoult D (2003) Cardiac valves in patients with Q fever endocarditis: microbiological, molecular, and histologic studies. J Infect Dis 187(7):1097–1106
22. Delsing CE, Kullberg BJ, Bleeker-Rovers CP (2010) Q fever in the Netherlands from 2007 to 2010. Neth J Med 68(12):382–387
23. Roest HIJ, Tilburg JJHC, Hoek WVD, Vellema P, Zijderveld FGV, Klaassen CHW et al (2011) The Q fever epidemic in The Netherlands: history, onset, response and reflection. Epidemiol Infect 139(1):1–12
24. Roest HIJ, Ruuls RC, Tilburg JJHC, Nabuurs-Franssen MH, Klaassen CHW, Vellema P et al (2011) Molecular epidemiology of *Coxiella burnetii* from ruminants in Q fever outbreak, the Netherlands. Emerg Infect Dis 17(4):668–675
25. Fournier PE, Marrie TJ, Raoult D (1998) Diagnosis of Q fever. J Clin Microbiol 36(7):1823–1834
26. Boden K, Wolf K, Hermann B, Frangoulidis D (2015) First isolation of *Coxiella burnetii* from clinical material by cell-free medium (ACCM2). Eur J Clin Microbiol Infect Dis 34(5):1017–1022
27. Omsland A, Cockrell DC, Howe D, Fischer ER, Virtaneva K, Sturdevant DE et al (2009) Host cell-free growth of the Q fever bacterium *Coxiella burnetii*. Proc Natl Acad Sci 106(11):4430–4434
28. Thiele D, Willems H (1994) Is plasmid based differentiation of *Coxiella burnetii* in 'acute' and 'chronic' isolates still valid? Eur J Epidemiol 10(4):427–434
29. Lautenschläger S, Willems H, Jäger C, Baljer G (2000) Sequencing and characterization of the cryptic plasmid QpRS from *Coxiella burnetii*. Plasmid 44(1):85–88
30. Valková D, Ksár J (1995) A new plasmid (QpDV) common to *Coxiella burnetii* isolates associated with acute and chronic Q fever. FEMS Microbiol Lett 125(2–3):275–280
31. Willems H, Ritter M, Jäger C, Thiele D (1997) Plasmid-homologous sequences in the chromosome of plasmidless *Coxiella burnetii* Scurry Q217. J Bacteriol 179(10):3293–3297
32. Sekeyová Z, Roux V, Raoult D (1999) Intraspecies diversity of *Coxiella burnetii* as revealed by *com1* and *mucZ* sequence comparison. FEMS Microbiol Lett 180(1):61–67
33. Zhang G, To H, Yamaguchi T, Fukushi H, Hirai K (1997) Differentiation of *Coxiella burnetii* by sequence analysis of the gene (*com1*) encoding a 27-kDa outer membrane protein. Microbiol Immunol 41(11):871–877
34. Nguyen SV, Hirai K (1999) Differentiation of *Coxiella burnetii* isolates by sequence determination and PCR-restriction fragment length polymorphism analysis of isocitrate dehydrogenase gene. FEMS Microbiol Lett 180(2):249–254
35. McLaughlin HP, Cherney B, Hakovirta JR, Priestley RA, Conley A, Carter A et al (2017) Phylogenetic inference of *Coxiella burnetii* by 16S rRNA gene sequencing. PLoS One 12(12):e0189910
36. Stein A, Kruszewska D, Gouvernet J, Raoult D (1997) Study of the 16S-23S ribosomal DNA internal spacer of *Coxiella burnetii*. Eur J Epidemiol 13(4):471–475
37. Heinzen R, Stiegler GL, Whiting LL, Schmitt SA, Mallavia LP, Frazier ME (1990) Use of pulsed field gel electrophoresis to differentiate *Coxiella burnetii* strains. Ann N Y Acad Sci 590:504–513
38. Hendrix LR, Samuel JE, Mallavia LP (1991) Differentiation of *Coxiella burnetii* isolates by analysis of restriction-endonuclease-digested DNA separated by SDS-PAGE. J Gen Microbiol 137(2):269–276

39. Thiele D, Willems H, Köpf G, Krauss H (1993) Polymorphism in DNA restriction patterns of *Coxiella burnetii* isolates investigated by pulsed field gel electrophoresis and image analysis. Eur J Epidemiol 9(4):419–425
40. Svraka S, Toman R, Skultety L, Slaba K, Homan WL (2006) Establishment of a genotyping scheme for *Coxiella burnetii*. FEMS Microbiol Lett 254(2):268–274
41. Klaassen CHW, Nabuurs-Franssen MH, Tilburg JJHC, Hamans MAWM, Horrevorts AM (2009) Multigenotype Q fever outbreak, the Netherlands. Emerg Infect Dis 15(4):613–614
42. Samuel JE, Frazier ME, Mallavia LP (1985) Correlation of plasmid type and disease caused by *Coxiella burnetii*. Infect Immun 49:5
43. Vodkin MH, Williams JC (1988) A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein in both mycobacteria and Escherichia coli. J Bacteriol 170(3):1227–1234
44. Weisburg WG, Dobson ME, Samuel JE, Dasch GA, Mallavia LP, Baca O et al (1989) Phylogenetic diversity of the *Rickettsiae*. J Bacteriol 171(8):4202–4206
45. Thiele D, Willems H, Haas M, Krauss H (1994) Analysis of the entire nucleotide sequence of the cryptic plasmid QpH1 from *Coxiella burnetii*. Eur J Epidemiol 10(4):413–420
46. Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC et al (2003) Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. Proc Natl Acad Sci 100(9):5455–5460
47. Seshadri R, Samuel J (2005) Genome analysis of *Coxiella burnetii* species: insights into pathogenesis and evolution and implications for biodefense. Ann N Y Acad Sci 1063(1):442–450
48. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269(5223):496–512
49. Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD et al (2009) Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus *Coxiella*. Infect Immun 77(2):642–656
50. Omsland A, Beare PA, Hill J, Cockrell DC, Howe D, Hansen B et al (2011) Isolation from animal tissue and genetic transformation of *Coxiella burnetii* are facilitated by an improved axenic growth medium. Appl Environ Microbiol 77(11):3720–3725
51. Sanchez SE, Vallejo-Esquerra E, Omsland A (2018) Use of axenic culture tools to study *Coxiella burnetii*. Curr Protoc Microbiol 50(1):e52
52. Sandoz KM, Beare PA, Cockrell DC, Heinzen RA (2016) Complementation of arginine Auxotrophy for genetic transformation of *Coxiella burnetii* by use of a defined axenic medium. Appl Environ Microbiol 82(10):3042–3051
53. Vallejo Esquerra E, Yang H, Sanchez SE, Omsland A (2017) Physicochemical and nutritional requirements for axenic replication suggest physiological basis for *Coxiella burnetii* niche restriction. Front Cell Infect Microbiol 7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5449765/
54. Kersh GJ, Oliver LD, Self JS, Fitzpatrick KA, Massung RF (2011) Virulence of pathogenic *Coxiella burnetii* strains after growth in the absence of host cells. Vector-Borne Zoonotic Dis 11(11):1433–1438
55. Hammerl JA, Mertens K, Sprague LD, Hackert VH, Buijs J, Hoebe CJ et al (2015) First draft genome sequence of a human *Coxiella burnetii* isolate, originating from the largest Q fever outbreak ever reported, the Netherlands, 2007 to 2010. Genome Announc 3(3):e00445–e00415
56. Kuley R, Smith HE, Janse I, Harders FL, Baas F, Schijlen E et al (2016) First complete genome sequence of the Dutch veterinary *Coxiella burnetii*Strain NL3262, originating from the largest global Q fever outbreak, and draft genome sequence of its epidemiologically linked chronic human isolate NLhu3345937. Genome Announc 4(2):e00245–e00216
57. Millar JA, Beare PA, Moses AS, Martens CA, Heinzen RA, Raghavan R (2017) Whole-genome sequence of *Coxiella burnetii* Nine Mile RSA439 (Phase II, Clone 4), a laboratory workhorse strain. Genome Announc 5(23) Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5465618/

58. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G et al (2009) Real-time DNA sequencing from single polymerase molecules. Science 323(5910):133–138
59. Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S, Bayley H (2009) Continuous base identification for single-molecule nanopore DNA sequencing. Nat Nanotechnol 4(4):265–270
60. Dean FB, Nelson JR, Giesler TL, Lasken RS (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11(6):1095–1099
61. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P et al (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A 99(8):5261–5266
62. Rodrigue S, Malmstrom RR, Berlin AM, Birren BW, Henn MR, Chisholm SW (2009) Whole genome amplification and de novo assembly of single bacterial cells. Ahmed N (ed). PLoS One 4(9):e6864
63. Lasken RS, Stockwell TB (2007) Mechanism of chimera formation during the multiple displacement amplification reaction. BMC Biotechnol 7(1):19
64. Jiao X, Rosenlund M, Hooper SD, Tellgren-Roth C, He L, Fu Y et al (2011) Structural alterations from multiple displacement amplification of a human genome revealed by mate-pair sequencing. PLoS One 6(7):e22250
65. Pugh TJ, Delaney AD, Farnoud N, Flibotte S, Griffith M, Li HI et al (2008) Impact of whole genome amplification on analysis of copy number variants. Nucleic Acids Res 36(13):e80
66. Kumar S, Gangoliya SR, Berri M, Rodolakis A, Alam SI (2013) Whole genome amplification of the obligate intracellular pathogen *Coxiella burnetii* using multiple displacement amplification. J Microbiol Methods 95(3):368–372
67. Reis S (2014) Der Einfluss von WGA-Techniken auf das Coxiella-Genom und genotypische Dynamik von *Coxiella burnetii* in einer isolierten Herde kleiner Wiederkäuer [Internet]. PhD thesis, Ludwig-Maximilians-Universität München. Available from: https://edoc.ub.uni-muenchen.de/16718/
68. Walter MC, Öhrman C, Myrtennäs K, Sjödin A, Byström M, Larsson P et al (2014) Genome sequence of *Coxiella burnetii* strain Namibia. Stand Genomic Sci 9:22
69. Leichty AR, Brisson D (2014) Selective whole genome amplification for resequencing target microbial species from complex natural samples. Genetics 198(2):473–481
70. Cocking JH, Deberg M, Schupp J, Sahl J, Wiggins K, Porty A et al (2020) Selective whole genome amplification and sequencing of *Coxiella burnetii* directly from environmental samples. Genomics 112(2):1872–1878
71. Kozarewa I, Armisen J, Gardner AF, Slatko BE, Hendrickson CL (2015) Overview of target enrichment strategies. Curr Protoc Mol Biol 112:7.21.1–7.2123
72. Gilpatrick T, Lee I, Graham JE, Raimondeau E, Bowen R, Heron A et al (2020) Targeted nanopore sequencing with Cas9-guided adapter ligation. Nat Biotechnol 38(4):433–438
73. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 95(6):3140–3145
74. Fasemore AM, Helbich A, Walter MC, Dandekar T, Vergnaud G, Förstner KU et al (2020) CoxBase: an online platform for epidemiological surveillance, visualization, analysis and typing of Coxiella burnetii genomic sequence. bioRxiv:2020.11.29.402362
75. Hornstra HM, Priestley RA, Georgia SM, Kachur S, Birdsell DN, Hilsabeck R et al (2011) Rapid typing of *Coxiella burnetii*. PLoS One 6(11):e26201

Index

I. de Filippis (ed.), *Molecular Typing in Bacterial Infections, Volume II*,
https://doi.org/10.1007/978-3-030-83217-9