Eur J Clin Microbiol Infect Dis (2014) 33:927–932
DOI 10.1007/s10096-013-2028-0

ARTICLE

Fast and simple epidemiological typing of *Pseudomonas aeruginosa* using the double-locus sequence typing (DLST) method

P. Basset · D. S. Blanc

Received: 25 October 2013 / Accepted: 25 November 2013 / Published online: 11 December 2013

Abstract Although the molecular typing of *Pseudomonas aeruginosa* is important to understand the local epidemiology of this opportunistic pathogen, it remains challenging. Our aim was to develop a simple typing method based on the sequencing of two highly variable loci. 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 (N_T)=35–64; index of discrimination (ID)=0.816–0.964] was lower than the combination of two loci (N_T=78–97; ID=0.966–0.971). As each pairwise combination of loci gave similar results, we selected the most robust combination with *ms172* [reverse; R] and *ms217* [R] to constitute the double-locus sequence typing (DLST) scheme for *P. aeruginosa*. 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 makes it a method of choice for local epidemiological analyses of *P. aeruginosa*. Moreover, the possibility to give unambiguous definition of types allowed to develop an Internet database (http://www.dlst.org) accessible by all.

This study has been presented at the 23rd European Congress of Clinical Microbiology and Infectious Diseases (ECCMID), Berlin, Germany, April 2013.

Electronic supplementary material The online version of this article (doi:10.1007/s10096-013-2028-0) contains supplementary material, which is available to authorized users.

P. Basset (✉) · D. S. Blanc
Service of Hospital Preventive Medicine, Lausanne University Hospital, Rue du Bugnon 48, 1011 Lausanne, Switzerland
e-mail: Patrick.Basset@chuv.ch

D. S. Blanc
Institute of Microbiology, Lausanne University Hospital, Lausanne, Switzerland

Introduction

Pseudomonas aeruginosa is a ubiquitous opportunistic pathogen that infects hosts when local or general defense mechanisms are impaired [1, 2]. Because of the ubiquity of *P. aeruginosa* in the environment and in the endogenous flora of hospitalized patients, it is important to use powerful typing methods to understand the local epidemiology of this bacteria (e.g., to identify potential reservoirs and sources of colonization and/or infection) [3].

The epidemiology of *P. aeruginosa* has been analyzed by an array of different molecular typing methods. Among these, pulsed-field gel electrophoresis (PFGE) is generally considered the gold standard for local epidemiological studies because of its high discriminatory power. Yet, this method is labor-intensive and shows low intra- and inter-laboratory reproducibility [4, 5]. In this context, we developed the double digest selective label (DDSL) typing method to offer a simple and fast alternative to PFGE [6]. Nevertheless, this method requires technical skills and it does not yield definitive typing results.

In contrast to band pattern methods, sequence-based methods are portable and definitive methods, offering good intra- and inter-laboratory reproducibility [7]. Recently, several multiple-locus variable-number tandem repeat analysis (MLVA) typing schemes for *P. aeruginosa* have been described [8–10]. Each isolate is characterized by the number of repeats observed at a large number of loci. The occurrence of repeats of intermediate size makes it difficult to unambiguously define the types, yielding inter-laboratory standardization complex. Multilocus sequence typing

(MLST) uses the nucleotide sequence data of several (generally seven) housekeeping genes. Although this method is useful to understand the global population structure of *P. aeruginosa* [11], it is rather expensive and its discriminatory power is too low to investigate local epidemiology.

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 [12]. Similarly 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 [13, 14]. Consequently, this method can be easily incorporated into long-term routine surveillance programs. Our goal in this study was to set up an efficient sequence-based typing scheme similar to the DLST scheme of *S. aureus* to investigate the local epidemiology of *P. aeruginosa*.

Materials and methods

Bacterial isolates

The *P. aeruginosa* collection was composed of 282 isolates recovered at the Lausanne University Hospital between 1994 and 2007 which were previously typed by the DDSL method [6, 15]. This collection included: (i) 128 isolates from 83 intensive care unit (ICU) patients hospitalized in 2007 [15], (ii) 66 duplicate isolates from 23 patients hospitalized between 1994 and 2006, (iii) 34 multi-resistant isolates from patients hospitalized between 2002 and 2007, among which 14 were involved in an outbreak [15], and (iv) 54 isolates from 20 taps between 1998 and 2008.

Selection of loci

A literature review was completed to identify potentially highly variable loci in the *P. aeruginosa* genome. Eleven loci (*ms142*, *ms172*, *ms173*, *ms194*, *ms207*, *ms214*, *ms215*, *ms217*, *ms222*, *ms223*, *oprD*) [8, 9] 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.

Molecular procedure

Bacterial isolates were incubated overnight at 37 °C in 5 ml of lysogeny broth (LB) medium. Crude DNA extracts were prepared by mixing 10 μl of this broth with 90 μl of TE in 96-well plates and heating this solution at 95 °C for 10 min. Primers used for polymerase chain reaction (PCR) amplification for each locus (*ms172*, *ms217*, *oprD*) were those previously described (Table 1) [9, 16]. PCR amplification was carried out in 22-μl reaction volumes containing 2 μl of the DNA extract solution, 1U of *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA), 1× Taq Reaction Buffer, 1.5 mM $MgCl_2$ (supplied with the Taq polymerase), 0.3 μM of each primer, and 0.2 mM of each dNTP (Invitrogen). Cycling conditions were 5 min of initial denaturation at 94 °C; 35 cycles consisting of 30 s at 94 °C, 30 s at 60 °C (*ms172*), or 64 °C (*ms217* and *oprD*), 45 s at 72 °C; and a final extension for 10 min at 72 °C. Standard gel electrophoresis was performed, and stained gels were evaluated for success and specificity, i.e., one clear band visible per PCR. PCRs were then purified with the MultiScreen® PCRμ kit (Millipore, Billerica, MA, USA), according to the manufacturer's instructions. Sequencing reactions were carried out with the Big Dye Terminator kit, version1.1 (Applied Biosystems, Carlsbad, CA, USA) and sequencing reaction cleanup was performed with the BigDye XTerminator kit (Applied Biosystems), according to the manufacturer's instructions. Purified samples were analyzed with the ABI 3130xl sequencer (Applied Biosystems), according to standard protocols. Sequences were analyzed using the BioNumerics software version 6.5 (Applied Math, Sint-Martens-Latem, Belgium). The procedure was repeated a second time when no sequence was obtained or if the sequence quality was too low. If no sequence of good quality was obtained after this second round, the result for the isolate was considered a null allele.

Allele assignment

Variant alleles were taken into account, beginning from the smallest possible genetic difference, i.e., a point mutation. An arbitrary number was given to each distinct allele within the locus. Each isolate was, therefore, given two numbers that represented its DLST type.

Resolution of the markers

Discriminatory power was evaluated by the number of types (variant alleles) and by calculation of the index of discrimination (ID, http://biophp.org/stats/discriminatory_power/demo.php) [17]. An ID value of 1 would indicate that the typing method was able to distinguish each isolate from all others. Conversely, an index of 0 would indicate that all isolates were of an identical type.

Stability of the markers

To evaluate the stability of the loci over time, we analyzed isolates recovered with a delay of at least 7 days from the same

Table 1 Characteristics of the analyzed loci

Locus	Direction	Primers (5′-3′)	Start signature of the trimming pattern (5′-3′)	Length of the analyzed sequence (bp)
ms172	Forward [F]	GGATTCTCTCGCACGAGGT	GGGCCCAGCC	400
	Reverse [R]	TACGTGACCTGACGTTGGTG	ACGATGCTGGADCCA	400
ms217	[F]	TTCTGGCTGTCGCGACTGAT	TTRCCGTTTG	350
	[R]	GAACAGCGTCTTTTCCTCGC	CAVCATGGCGGTGRT	350
oprD	[F]	ATGAAAGTGATGAAGTGGAGC	TCGCCRTRGC	500
	[R]	CAGGATCGACAGCGGATAGT	TCGTGGTGCT	500

patient/tap: the first isolate was compared to subsequent isolates to form pairs of isolates. If these pairs showed DDSL profiles with more than nine bands difference (i.e., different types, as previously defined [6]; DDSL profiles with less than 9 bands difference were considered as subtypes), it was considered as a new colonization/infection event. Those pairs were, therefore, not considered for further evaluation of the markers stability.

Results

Typability

DNA sequences were obtained for the three selected loci, *ms172*, *ms217*, and *oprD*. As expected, the length of DNA sequences was variable among isolates. However, all isolates showed amplification products larger than 350 bp. For each locus and for both directions (forward [F] and reverse [R]), we found a well-conserved region that allowed identifying a start signature (Table 1). Analysis of the length and quality of these sequences lead us to define fragment lengths of 400, 350, and 500 bp from the start signatures for *ms172*, *ms217*, and *oprD*, respectively, for both [F] and [R] sequences. With these parameters, the number of null alleles varied between 0 (*ms217* [F]) and 8 (*ms172* [F]) (Table 2).

Resolution of markers

The performance of each marker was determined for all 282 isolates (Table 2). In addition, to estimate the discriminatory power of each locus in unique isolates (one per patient and one per tap), a subpopulation was formed by selecting only the first isolate recovered in each patient/tap (N =152). In both datasets, DDSL showed the highest performance with the number of types and ID. None of the three loci alone could match this resolution. As *ms172* [R], *oprD* [F], and *ms217* [R] showed the best performances in terms of typability and resolution, we formed and tested combinations of these loci to further increase the resolution. As the three combinations of two loci gave resolution results only slightly lower than the combination of the three loci (Table 2), we proposed to use a combination of only two loci, similarly to *S. aureus*. We selected for this purpose the combination *ms172* [R] and *ms217* [R] to form the DLST scheme, as it generally appeared to be more robust when larger numbers of isolates were typed (data not shown).

Stability of the marker

Multiple isolates (range 2–8) were recovered over time from 44 patients and 11 taps. For each patient/tap, the first isolate was compared to the subsequent isolates to form pairs of isolates. The time period between the first and last isolates varied between 7 days to more than 5 years. Among the 92 possible pairs analyzed, 20 pairs (21.9 %) showed different DDSL clones, suggesting colonization by a different strain. These pairs were removed from further analyses of the stability but, interestingly, all these pairs showed DLST types with different alleles at both loci. Among the remaining 72 pairs, 68 (94.4 %) showed identical DLST types and only 4 (5.6 %) showed a different allele at one DLST locus, suggesting the occurrence of a mutation (Fig. 1).

Epidemiological concordance

The 14 isolates belonging to an outbreak that occurred in the ICU during the period 2002 to 2004 all belonged to the DLST type 20-30. Interestingly, at that time, using PFGE as a molecular typing method, two different genotypes were identified (10 bands differences). DDSL confirmed the relatedness of these two strains (Fig. 2, DDSL 6 and 5).

Discussion

Our data showed that combining partial sequences of two highly variable loci such as *ms172* and *ms217* can be a useful method for the epidemiological typing of *P. aeruginosa*. Adding a third locus such as *oprD* slightly increased the number of types and the discrimination power, and might be useful in situations in which it is important to confirm or refute

Table 2 Resolution of partial sequences of three variable loci and their combination in comparison to double digest selective label (DDSL) typing

	No. of isolates with null alleles	Unique isolates (N=152)		All isolates (N=282)	
		No. of genotypes	ID[a]	No. of genotypes	ID[a]
DDSL all	0	136	0.9979	171	0.9942
ms172[R]*oprD*[F]*ms217*[R]	9	80	0.9829	98	0.9774
ms172[R]*ms217*[R]	6	75	0.9758	86	0.9683
ms172[R]*oprD*[F]	8	67	0.9765	78	0.9707
oprD[F]*ms217*[R]	4	75	0.9777	87	0.9664
ms217[R] 350 bp	1	61	0.9691	64	0.9644
ms217[F] 350 bp	0	56	0.967	62	0.9627
ms172[F] 400 bp	8	48	0.9643	49	0.9575
ms172[R] 400 bp	5	42	0.9597	44	0.9547
oprD[F] 500 bp	3	33	0.8535	41	0.8547
oprD[R] 500 bp	7	25	0.8337	35	0.8156

[a] Index of discrimination

a link between pairs of isolates. However, this gain in performance is probably not large enough to justify the additional costs of systematically sequencing a third locus.

Short sequence length allows single-strand sequencing, reducing the costs of analysis [12]. Moreover, this method can easily be adapted to be performed in batches of 96 isolates, thus further reducing the cost and working time. As a result, it can be easily incorporated into a surveillance program involving the typing of hundreds of isolates. A similar strategy based on the single-strand sequencing of ca. 500 bp of the variable region of the *clfB* and *spa* genes has already shown its value in *S. aureus* for routine surveillance [13, 14, 18].

So far, PFGE has been considered the gold standard for the epidemiological typing of *P. aeruginosa*. However, because of the difficulties in using this method for routine surveillance, a number of other methods have been developed to replace PFGE. For example, the recently developed DDSL typing technique uses a faster and easier procedure while obtaining very similar results to PFGE, as it is based on DNA fragments restricted at one end by the same rare cutting enzyme [6]. This method has, in particular, been shown to efficiently decipher the local epidemiology of *P. aeruginosa* in our hospital [15]. Consequently, we used this method as the reference against which we compared our new DLST scheme. DLST results were generally in good agreement with DDSL typing results. Nevertheless, DDSL often shows a higher discriminatory power, which could lead to differences in clone delimitation. This is illustrated by the examples of two major DLST types (i.e., 1-21 and 20-30) within which different DDSL types were observed (Fig. 2). However, the comparison of banding patterns such as those observed for DDSL is problematic when multiple isolates are analyzed. In such situations, sequence data like the proposed DLST scheme are much more suitable. Yet, DDSL is probably the ideal typing complement to resolve DLST clusters in special situations. Caution must, however, be taken for the interpretation of related patterns.

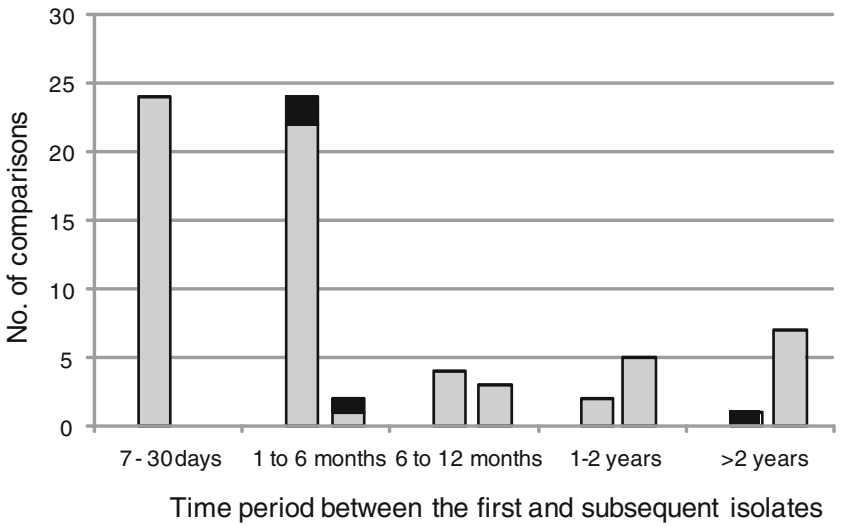

Fig. 1 Comparison of pairs of isolates belonging to the same double digest selective label (DDSL) clone in patients (*first columns*) and taps (*second columns*). Pairs with identical double-locus sequence typing (DLST) types are indicated in *gray* and pairs with one allele difference are indicated in *black*. The duration between the first and subsequent isolates was categorized into five groups

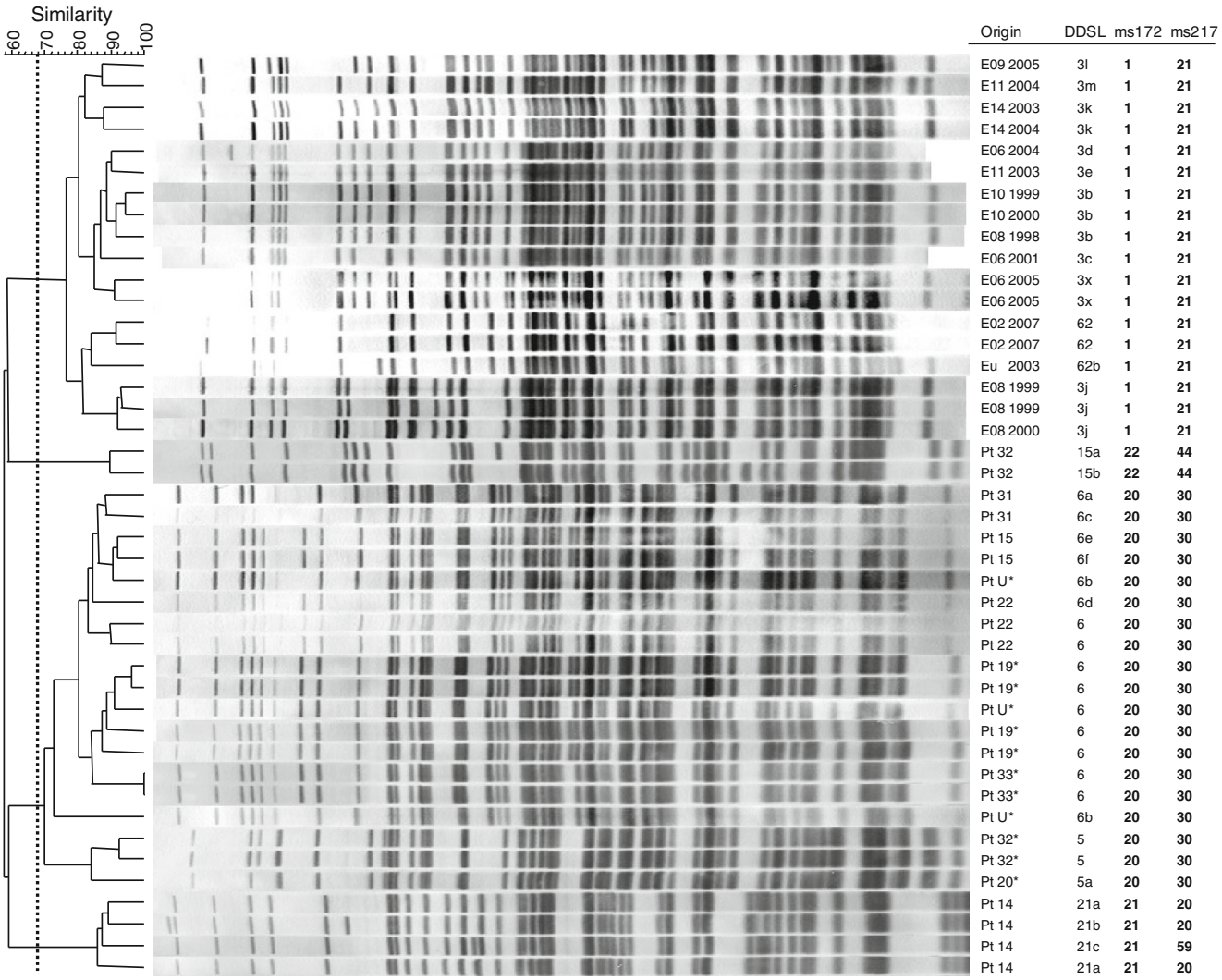

Fig. 2 Example of double digest selective label (DDSL) banding patterns compared to the typing results of double-locus sequence typing (DLST). The origin of isolates are indicated (*E*, environment [taps]; *Pt*, patients), as well as the year of isolation for taps. DDSL types are labeled by a number, whereas subtypes (<9 bands differences) are indicated by a letter; outbreak isolates are marked with an asterisk. The *ms172* and *ms217* allele numbers constitute the DLST type. DDSL patterns were clustered using the Pearson coefficient of similarity and the unweighted pair group method with arithmetic mean (UPGMA) method in BioNumerics software version 6.5

Epidemiological analyses at a local level aim at identifying the transmission of isolates. This requires that the markers are stable during the period of investigation, i.e., during a period of at least several months. To evaluate the stability of our DLST scheme, we compared isolates recovered from the same patient/tap during different periods of time. These results showed that all the pairs showing mutations at the two DLST alleles also showed different DDSL clones, suggesting the acquisition of a different strain. When two isolates from the same DDSL clone were compared, their DLST types were identical in almost 95 % of the cases, and this proportion remained high even for isolates recovered more than 2 years apart (Fig. 1). Consequently, DLST appears to be adapted and stable to compare isolates recovered during investigations with durations of several months or even years. Clustering isolates sharing one allele in common has efficiently been used to discriminate *S. aureus* clones [13]. In contrast to *S. aureus*, the population structure of *P. aeruginosa* has not yet been clearly identified. Contrasting models ranging from panmictic to clonal structure have been proposed [11], and recent data suggested a highly diverse nonclonal epidemic population structure [19, 20]. Therefore, the value of clustering *P. aeruginosa* isolates sharing a DLST allele to reveal the structure of this population might be low and still needs to be assessed.

Useful genetic variation can be captured with an arbitrary approach by considering only 350 or 400 bp of highly variable loci. By defining a given marker length, we deliberately discarded information about repeat organization and considered alleles only at the level of nucleotides. As a result, no

special search algorithms and assembly procedures are needed. Therefore, simple comparisons with international or local databases comprising currently available allele sequences can be used for the interrogation and identification of related strains. In such a context, a web-based database (http://www.dlst.org) has recently been developed to identify allele profiles and DLST types. This could be used as a reference to assess the local and international diversities of the *P. aeruginosa* population. In this context, we are currently investigating the epidemiology of *P. aeruginosa* in ICU patients over 3 years. This represents over 700 isolates from about 350 patients and 260 isolates from the environment. Such a large-scale study would not be possible without an efficient typing method.

Conclusions

Using a local epidemiological collection of isolates, we assessed the value of single-strand sequencing of two highly variable loci for the epidemiological typing of *Pseudomonas aeruginosa*. The method was found to have high typability, reproducibility, and discriminatory power. We showed that these two markers were generally stable over a period of time similar to that of local epidemiological investigations. Its low cost and ease of use enable the typing of large collections of isolates. More importantly, sequence-based typing data allowed an unambiguous and standardized definition of types, which we implemented on an Internet website (http://www.dlst.org).

Acknowledgments We are grateful to Caroline Choulat for the technical assistance and Roland Sahli for the helpful discussions about molecular procedures.

Conflict of interest Nothing to declare.

References

1. Lyczak JB, Cannon CL, Pier GB (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. Microbes Infect 9:1051–1060
2. Wilson R, Dowling RB (1998) Lung infections. 3. *Pseudomonas aeruginosa* and other related species. Thorax 53(3):213–219
3. Blanc DS (2004) The use of molecular typing for epidemiological surveillance and investigation of endemic nosocomial infections. Infect Genet Evol 4:193–197
4. 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(9):2233–2239
5. Botes J, Williamson G, Sinickas V, Gürtler V (2003) Genomic typing of *Pseudomonas aeruginosa* isolates by comparison of Riboprinting and PFGE: correlation of experimental results with those predicted from the complete genome sequence of isolate PAO1. J Microbiol Methods 55(1):231–240
6. Terletskiy V, Kuhn G, Francioli P, Blanc DS (2008) Application and evaluation of double digest selective label (DDSL) typing technique for *Pseudomonas aeruginosa* hospital isolates. J Microbiol Methods 72(3):283–287
7. Aires-de-Sousa M, Boye K, de Lencastre H, Deplano A, Enright MC, Etienne J, Friedrich A, Harmsen D, Holmes A, Huijsdens XW, Kearns AM, Mellmann A, Meugnier H, Rasheed JK, Spalburg E, Strommenger B, Struelens MJ, Tenover FC, Thomas J, Vogel U, Westh H, Xu J, Witte W (2006) High interlaboratory reproducibility of DNA sequence-based typing of bacteria in a multicenter study. J Clin Microbiol 44(2):619–621
8. Onteniente L, Brisse S, Tassios PT, Vergnaud G (2003) Evaluation of the polymorphisms associated with tandem repeats for *Pseudomonas aeruginosa* strain typing. J Clin Microbiol 41(11):4991–4997
9. Vu-Thien H, Corbineau G, Hormigos K, Fauroux B, Corvol H, Clément A, Vergnaud G, Pourcel C (2007) Multiple-locus variable-number tandem-repeat analysis for longitudinal survey of sources of *Pseudomonas aeruginosa* infection in cystic fibrosis patients. J Clin Microbiol 45(10):3175–3183
10. Turton JF, Turton SE, Yearwood L, Yarde S, Kaufmann ME, Pitt TL (2010) Evaluation of a nine-locus variable-number tandem-repeat scheme for typing of *Pseudomonas aeruginosa*. Clin Microbiol Infect 16(8):1111–1116
11. Wiehlmann L, Wagner G, Cramer N, Siebert B, Gudowius P, Morales G, Köhler T, van Delden C, Weinel C, Slickers P, Tümmler B (2007) Population structure of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 104(19):8101–8106
12. Kuhn G, Francioli P, Blanc DS (2007) Double-locus sequence typing using *clfB* and *spa*, a fast and simple method for epidemiological typing of methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 45(1):54–62
13. 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(5):941–947
14. Basset P, Senn L, Prod'hom G, Bille J, Francioli P, Zanetti G, Blanc DS (2010) Usefulness of double locus sequence typing (DLST) for regional and international epidemiological surveillance of methicillin-resistant *Staphylococcus aureus*. Clin Microbiol Infect 16(8):1289–1296
15. Cuttelod M, Senn L, Terletskiy V, Nahimana I, Petignat C, Eggimann P, Bille J, Prod'hom G, Zanetti G, Blanc DS (2011) Molecular epidemiology of *Pseudomonas aeruginosa* in intensive care units over a 10-year period (1998–2007). Clin Microbiol Infect 17(1):57–62
16. Pirnay JP, De Vos D, Cochez C, Bilocq F, Vanderkelen A, Zizi M, Ghysels B, Cornelis P (2002) *Pseudomonas aeruginosa* displays an epidemic population structure. Environ Microbiol 4(12):898–911
17. Hunter PR (1990) Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol 28(9):1903–1905
18. Basset P, Senn L, Vogel V, Zanetti G, Blanc DS (2010) Diversity of staphylococcal cassette chromosome mec elements in predominant methicillin-resistant *Staphylococcus aureus* clones in a small geographic area. Antimicrob Agents Chemother 54(11):4589–4595
19. Pirnay JP, Bilocq F, Pot B, Cornelis P, Zizi M, Van Eldere J, Deschaght P, Vaneechoutte M, Jennes S, Pitt T, De Vos D (2009) *Pseudomonas aeruginosa* population structure revisited. PLoS One 4(11):e7740
20. van Mansfeld R, Jongerden I, Bootsma M, Buiting A, Bonten M, Willems R (2010) The population genetics of *Pseudomonas aeruginosa* isolates from different patient populations exhibits high-level host specificity. PLoS One 5(10):e13482