DNA–DNA hybridization study of *Burkholderia* species using genomic DNA macro-array analysis coupled to reverse genome probing

Vincent Ramisse,[1] Jacques Balandreau,[2] François Thibault,[3] Dominique Vidal,[3] Gilles Vergnaud[1] and Philippe Normand[2]

[1]Centre d'Etudes du Bouchet (CEB), BP3, 91710 Vert le Petit, France

[2]Écologie Microbienne, UMR 5557 CNRS, Bât. 741, Université Claude Bernard Lyon I, 69622 Villeurbanne cedex, France

[3]Centre de Recherches du Service de Santé des Armées, BP87, 38702 La Tronche cedex, France

Correspondence
Vincent Ramisse
vramisse@ceb.etca.fr

The present study was aimed at simplifying procedures to delineate species and identify isolates based on DNA–DNA reassociation. DNA macro-arrays harbouring genomic DNA of reference strains of several *Burkholderia* species were produced. Labelled genomic DNA, hybridized to such an array, allowed multiple relative pairwise comparisons. Based on the relative DNA–DNA relatedness values, a complete data matrix was constructed and the ability of the method to discriminate strains belonging to different species was assessed. This simple approach led successfully to the discrimination of *Burkholderia mallei* from *Burkholderia pseudomallei*, but also discriminated *Burkholderia cepacia* genomovars I and III, *Burkholderia multivorans*, *Burkholderia pyrrocinia*, *Burkholderia stabilis* and *Burkholderia vietnamiensis*. Present data showed a sufficient degree of congruence with previous DNA–DNA reassociation techniques. As part of a polyphasic taxonomic scheme, this straightforward approach is proposed to improve species definition, especially for application in the rapid screening necessary for large numbers of clinical or environmental isolates.

The genus *Burkholderia*, formerly *Pseudomonas* homology group II (Yabuuchi *et al.*, 1992), comprises 29 validly published species at the time of writing. It includes soil and rhizosphere bacteria, as well as plant and human pathogens (Gilligan, 1995). Among them are several species of particular interest in clinical microbiology. A review on *Burkholderia* taxonomy is available (Coenye *et al.*, 2001c).

Burkholderia pseudomallei, the causative agent of melioidosis in humans, is a natural saprophytic micro-organism present in soils and stagnant waters in tropical regions (Dance, 2002). The ability of this species to invade and survive within free-living amoebae might affect its environmental survival and subsequent human exposure (Inglis *et al.*, 2000). Melioidosis is transmitted by contact with contaminated soil or water through skin wounds, ingestion or inhalation (Benenson, 1995). *Burkholderia mallei* is responsible for glanders, a highly communicable disease of horses that is transmissible to humans (Benenson, 1995). This disease no longer occurs in the Western hemisphere, except for sporadic occupational cases (Srinivasan *et al.*, 2001). Both *B. pseudomallei* and *B. mallei* are potential biological warfare and terrorism agents that necessitate efforts for preparedness, including rapid and accurate diagnostics (Rotz *et al.*, 2002).

A major outcome of recent taxonomic studies of this genus has been the division of the *Burkholderia cepacia* complex into at least nine discrete genomic species, genomovars I–IX (Coenye *et al.*, 2001a, b; Vandamme *et al.*, 1997, 2000, 2002) that inhabit major environmental reservoirs but are frequently involved in nosocomial infections in patients with cystic fibrosis (CF) and other vulnerable individuals (Heath *et al.*, 2002; Speert *et al.*, 2002). *B. cepacia* complex infections contribute significantly to morbidity and mortality in CF patients. By 18 years of age, 80 % of patients harbour *Pseudomonas aeruginosa* and 3·5 % harbour *B. cepacia* (Rajan & Saiman, 2002). A study of 905 isolates from the *B. cepacia* complex recovered from 447 CF patients in Canada confirmed the prevalence of *B. cepacia* genomovar III (80 %), *Burkholderia multivorans* (formerly

Published online ahead of print on 19 September 2002 as DOI 10.1099/ijs.0.02483-0.

Abbreviations: CF, cystic fibrosis; MF, microfluorimetry; NS1, nuclease S1; *R*, hybridization signal ratio; SP, spectrophotometry.

The EMBL/GenBank accession number for the 16S rRNA gene sequence of CEB 01056 is AJ491304.

genomovar II) (10 %) and other genomovars of the *B. cepacia* complex or other *Burkholderia* species (9 %) (Speert *et al.*, 2002). On the other hand, *B. cepacia* genomovar III is a common plant-associated bacterium (Balandreau *et al.*, 2001). Moreover, human-pathogenic strains may not necessarily be distinct from environmental strains (LiPuma *et al.*, 2002). There is some evidence that the health risks associated with infection are dependent on the nature of the genomovar, so the ability to differentiate genomovars is important for clinical microbiologists (Brisse *et al.*, 2000; Henry *et al.*, 2001; Mahenthiralingam *et al.*, 2000). The accurate diagnosis of *Burkholderia* isolates is a challenge for the prevention and control of microbial infection caused by opportunistic pathogens. Moreover, the agricultural use of *B. cepacia* as a biopesticide for protecting crops against fungal diseases and as a bioremediation agent for decontamination of remanent pesticides is the subject of controversy (Holmes *et al.*, 1998; Jones *et al.*, 2001). The risks of the biotechnological uses of the *B. cepacia* complex and the ecology of the bacteria have been reviewed by Parke & Gurian-Sherman (2001).

Commercial bacterial identification systems are not always able to determine the genomovar status, nor accurately confirm the identification of *B. cepacia* isolates while differentiating them from closely related species. We have also observed ambiguous identification of some isolates of *B. mallei* and *B. pseudomallei*, thus rendering PCR- or antibody-based procedures valuable to confirm species identification (Bauernfeind *et al.*, 1998; Steinmetz *et al.*, 1999). A combination of phenotypic and molecular tests are recommended for differentiation among the genomovars of the *B. cepacia* complex (Henry *et al.*, 2001). Nucleotide sequence variation within 16S rDNA is not sufficient to enable all genomovars to be discriminated by RFLP (Segonds *et al.*, 1999). Nucleotide sequence analysis of the orthologous single-copy gene *recA* provides a means of identifying the current genomovars and newly delineated species within the *B. cepacia* complex (Mahenthiralingam *et al.*, 2000), but single-locus-based identification schemes are risky because of the prevalence of lateral transfer (Mougel *et al.*, 2002; Ochman *et al.*, 2000).

The determination of the DNA–DNA relatedness of the whole genome remains an irreplaceable step in the delineation of bacterial species (Stackebrandt *et al.*, 2002; Wayne *et al.*, 1987). Nevertheless, there is a strong tendency to replace it because DNA–DNA reassociation experiments are not easily carried out, even though numerous methods with simplified and more reproducible steps have been described (Johnson, 1991). Moreover, the polyphasic taxonomic approach fulfils taxonomists' requirements by studying a wide range of genotypic and phenotypic information, but makes the identification procedure more complex (Rosselló-Mora & Amann, 2001; Vandamme *et al.*, 1996). Developments in the field of high-density DNA arrays show great promise for bacterial systematics (Stackebrandt *et al.*, 2002), but there is no routinely available application at present. As it is not easy for any laboratory to invest the efforts required to ensure a polyphasic approach for bacterial identification, and considering that species definition should still be based on determination of DNA–DNA relatedness, we felt the necessity to return to this approach but in a more simple manner. Here, we report a prospective work concerning a simplified way to implement species delineation, based on whole genomic DNA–DNA hybridization using DNA macro-arrays. This paper describes the method and its evaluation with previously characterized strains of the genus *Burkholderia*.

DNA macro-arrays were constructed from DNA extracted, as described by Yabuuchi *et al.* (1992), from reference strains of the *Burkholderia* species listed in Table 1. *Burkholderia* sp. CEB 01056, isolated in Centre d'Etudes du Bouchet (CEB) as a laboratory contaminant, was included in the study in an attempt to clarify its identification. The almost-complete 16S rDNA sequence (GenBank no. AJ491304) suggests that this isolate is closely related to the members of the *B. cepacia* complex, with similarity values ranging from 98·6 to 99·7 %. *Staphylococcus epidermidis* (an environmental isolate, CEB 01074) was used as an outgroup. A DNA macro-array consisted of a 18×18 mm square of positively charged nylon transfer membrane Hybond-N+ (Amersham Biosciences) onto which 15 different alkali-denaturated genomic DNAs (250 ng DNA in 10 µl 200 mM NaOH, 30 µM bromophenol blue per dot) and a blank (all reagents but no DNA) were spotted using the Hydra-96 Microdispenser coupled to its vacuum manifold (Robbins Scientific). Dots were 2 mm in diameter. Macro-arrays were produced in 10 batches, each consisting of 24 membranes (240 macro-arrays in total). Genomic DNA (100 ng) from all the strains listed (Table 1) was labelled, in a final volume of 20 µl, using 370 kBq [α-^{32}P]dCTP (110 TBq mmol^{-1}; Amersham Biosciences) and the Random Primed DNA Labeling kit (Roche Diagnostics), according to the manufacturer's instructions. Unincorporated nucleotides were removed by Sephadex G-50 filtration with the Multiscreen separation system MAHV-N45 (Amersham Biosciences) according to the manufacturer's instructions. Based on the observed mean of the incorporation rate (>50 %) and on the rate of DNA recovery after filtration (>65 %), the procedure routinely gives >65 ng labelled DNA with a specific activity of 2 MBq µg^{-1}. Individual hybridization chambers consisted of 1·8 ml Nunc CryoTube vials (Nalge Nunc International) into which a DNA macro-array was introduced and allowed to pre-hybridize for 2 h at 65 °C under agitation (8 rotations min^{-1}) with 500 µl Rapid-hyb buffer (Amersham Biosciences) before introduction of the purified probe (20 µl) and overnight hybridization under the same conditions. Free labelled DNA was removed by successive washing steps at 65 °C with 2 × SSC, 0·1 % SDS then 0·1 × SSC, 0·1 % SDS washing buffer (Amresco). DNA macro-arrays were scanned at 100 µm resolution using the Molecular Imager system GS-525 (Bio-Rad). For each DNA macro-array hybridized with a given labelled DNA, DNA–DNA relatedness was expressed as the hybridization signal

Table 1. DNA homology as determined by DNA macro-array analysis among 13 *Burkholderia* strains

Strains: 1, *B. cepacia* (I) ATCC 25416^T; 2, *B. cepacia* (III) LMG 12614; 3, *B. multivorans* (II) LMG 13010^T; 4, *B. pyrrocinia* ATCC 15958^T; 5, *B. stabilis* (IV) LMG 14294^T; 6, *B. vietnamiensis* (V) LMG 10929^T; 7, *Burkholderia* sp. CEB 01056; 8, *B. caribensis* LMG 18531^T; 9, *B. gladioli* ATCC 10248^T; 10, *B. glathei* ATCC 29195^T; 11, *B. plantarii* LMG 9035^T; 12, *B. mallei* ATCC 23344^T; 13, *B. pseudomallei* ATCC 23343^T; 14, *S. epidermidis* CEB 01074; 15, *S. epidermidis* CEB 01074 (duplicate); 16, No DNA. Strains were obtained from ATCC (American Type Culture Collection), Manassas, VA, USA; LMG (Culture Collection of Laboratorium voor Microbiologie), State University of Ghent, Ghent, Belgium; CEB (Culture Collection of Centre d'Etudes du Bouchet), Vert le Petit, France.

DNA	R (%)* with labelled DNA (100 ng) from:													
on arrays	1 ($n=4$)	2 ($n=4$)	3 ($n=4$)	4 ($n=4$)	5 ($n=4$)	6 ($n=4$)	7 ($n=4$)	8 ($n=4$)	9 ($n=5$)	10 ($n=4$)	11 ($n=4$)	12 ($n=15$)	13 ($n=3$)	14 ($n=8$)
1	**54·00±3·90**	12·29±2·03	8·06±0·71	14·63±4·16	11·12±1·97	6·13±2·87	10·59±1·92	2·33±1·34	2·72±1·15	2·00±0·24	8·19±1·91	6·90±2·29	5·20±1·64	0·18±0·07
2	7·53±1·86	**41·23±7·43**	5·77±1·27	7·98±2·17	6·00±1·87	4·93±2·05	5·46±1·91	2·08±0·78	1·58±0·46	1·29±0·33	5·16±1·63	4·22±1·06	3·14±0·98	0·57±0·47
3	4·88±0·47	6·69±1·15	**44·99±7·26**	5·82±1·24	4·12±1·29	3·77±2·43	5·92±0·89	2·02±1·14	2·28±1·25	2·09±0·57	7·30±1·84	6·03±1·43	4·44±1·88	0·74±0·23
4	6·07±1·89	6·69±1·48	7·87±3·70	**40·12±8·90**	6·24±0·52	4·74±0·88	5·45±0·78	1·23±0·99	1·65±1·09	1·07±0·67	4·03±2·55	3·47±1·10	2·74±1·25	0·17±0·05
5	8·22±0·73	9·51±2·33	8·42±1·40	10·15±2·26	**57·66±3·35**	3·69±2·48	8·21±2·01	1·66±1·35	2·53±1·50	1·01±0·44	5·97±2·01	6·03±2·28	4·30±2·32	0·31±0·13
6	6·40±1·96	8·38±3·41	7·61±4·54	6·04±0·71	4·47±0·86	**56·69±6·58**	9·01±3·70	2·25±1·86	2·66±1·98	1·62±1·11	9·59±7·27	9·89±4·77	5·30±2·97	0·47±0·29
7	6·04±1·57	6·79±1·64	5·97±1·10	7·94±2·89	4·79±1·64	5·98±2·82	**46·29±7·16**	1·90±0·70	2·33±0·84	1·83±0·86	4·97±1·05	3·65±0·90	3·67±2·14	0·13±0·04
8	0·98±0·22	1·41±0·58	1·72±1·01	1·06±0·39	0·72±0·03	1·16±0·29	1·58±0·22	**79·72±11·17**	1·33±0·55	1·73±0·54	2·84±0·86	1·87±0·82	1·61±0·74	0·21±0·06
9	1·56±0·53	1·76±0·53	2·52±0·32	1·68±0·88	1·11±0·24	1·85±0·46	2·32±0·44	1·83±0·86	**75·77±12·10**	1·66±0·54	12·70±2·45	3·05±1·26	2·65±1·76	0·08±0·02
10	1·17±0·72	1·31±0·78	1·49±0·58	1·23±0·80	0·84±0·41	1·21±0·82	1·60±0·88	2·40±1·01	2·33±1·64	**81·74±4·10**	3·37±0·81	3·75±1·68	2·55±1·25	0·40±0·44
11	0·42±0·08	0·47±0·08	0·87±0·21	0·46±0·08	0·53±0·09	0·83±0·39	0·46±0·09	0·49±0·42	1·87±0·32	0·49±0·13	**27·89±15·16**	0·97±0·47	0·87±0·29	0·22±0·11
12	0·91±0·46	1·13±0·26	1·63±0·50	0·86±0·17	0·75±0·19	1·33±0·38	1·13±0·26	0·71±0·43	0·94±0·32	1·92±0·75	3·05±0·69	**30·43±8·29**	25·58±7·19	0·59±0·28
13	1·02±0·13	1·39±0·32	1·77±0·44	1·15±0·38	0·80±0·09	2·34±1·13	1·22±0·44	0·87±0·54	0·99±0·58	0·85±0·26	3·56±0·89	18·60±6·75	**36·70±10·26**	0·62±0·25
14	0·31±0·13	0·43±0·17	0·54±0·13	0·32±0·19	0·15±0·03	2·27±3·57	0·18±0·01	0·25±0·13	0·33±0·26	0·38±0·10	0·41±0·19	0·53±0·47	0·62±0·19	**44·45±3·03**
15	0·26±0·11	0·27±0·08	0·54±0·40	0·25±0·13	0·51±0·08	0·48±0·33	0·42±0·33	0·18±0·08	0·33±0·34	0·14±0·09	0·44±0·32	0·39±0·27	0·36±0·17	**50·73±2·76**
16	0·23±0·10	0·24±0·17	0·23±0·21	0·32±0·10	0·19±0·07	2·60±4·61	0·15±0·03	0·09±0·09	0·36±0·28	0·17±0·07	0·53±0·44	0·22±0·21	0·28±0·18	0·11±0·04
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100

*R values at position x [$R=100(S_x/\Sigma S_{1-16})$] are given as the mean±SD from n separate experiments. R_{max} values are indicated in bold. The total signal detected on the 16 positions of the array is 100 %.

ratio (R) between the maximum-intensity signal of each individual spot (S_x) and the total signal from the 16 spots (ΣS_1–S_{16}) and expressed as a percentage. All data in the text are R values. Standardized data were calculated using the formula $R_{std} = 100(S_x - S_{min})/(S_{max} - S_{min})$, where S_{min} and S_{max} are respectively the minimum and maximum signal observed for an individual spot among all the spots present on the DNA macro-array. To avoid any confusion between classical DNA–DNA hybridization format (which is a true pairwise comparison where only two different DNAs are involved) and multiple DNA–DNA hybridization format on macro-arrays, we deliberately made the choice to display R, not R_{std}. The usefulness of displaying one of these two parameters is discussed below. DNA macro-array hybridizations were repeated 4–10 times for a given labelled DNA. Values of Spearman's correlation coefficient (r) were calculated to find relationships between hybridization patterns of macro-arrays. Microsoft Excel software was used for statistical calculations.

Reproducibility

A preliminary experiment was designed to assess the observed dispersion of R values that could result from both the homogeneity of the DNA macro-arrays produced and the reproducibility through the entire procedure, from DNA hybridization to data integration. Ten aliquots corresponding to one-tenth of the labelled DNA of *B. mallei* ATCC 23344^T (1 μg) were hybridized as described above, independently but simultaneously, to 10 different DNA macro-arrays originating from each of the 10 batches. This yielded 10 macro-array images, each containing 16 spots defined by a circular area of 304 pixels corresponding to the diameter of dots. The signal was extracted using the following parameters: volume, volume percentage and maximum pixel value. The maximum intensity level of 65 524 counts per pixel was never reached. Each of the three types of parameter yielded the same results without significant variation of R (data not shown). Maximum pixel value was used for subsequent calculations of R and R_{std}. For each position that refers to a given DNA spotted onto the macro-arrays, the arithmetic mean of R and the standard deviation were: *B. mallei* ATCC 23344^T, 27·4 ± 4·0 %; *B. pseudomallei* ATCC 23343^T, 15·9 ± 3·8 %; *B. multivorans* LMG 13010^T, 6·8 ± 0·9 %; *S. epidermidis*, 0·3 ± 0·1 %; no DNA, 0·2 ± 0·1 % (n = 10) (whole data not shown). Homogeneity of DNA macro-arrays hybridized to the DNA of *B. mallei* ATCC 23344^T and experimental conditions yielded reproducible hybridization patterns (r = 0·9541).

Resolution

The strains chosen for this study represent a portion of the variety of the genus *Burkholderia* (Table 1). To investigate the DNA–DNA relatedness of these strains, labelled DNA from each reference strain was hybridized individually to a DNA macro-array that allowed multiple nucleic acid hybridization. R values were calculated to express the proportion of the signal measured from a spot, compared to the total signal measured for the 16 spots (Table 1). The unique R_{max} value measures the homologous DNA–DNA reassociation, and the 15 remaining R values measure heterologous DNA–DNA reassociations. Consequently, R_{max} should decrease as the DNA–DNA relatedness among strains increases. As expected, R_{max} ranged from 81·74 % (*Burkholderia glathei* ATCC 29195^T) to 27·89 % (*Burkholderia plantarii* LMG 9035^T) for homologous pairs of DNAs. The higher R_{max} values observed for labelled DNA of *Burkholderia caribensis* LMG 18531^T, *Burkholderia gladioli* ATCC 10248^T and *B. glathei* ATCC 29195^T (respectively 79·72, 75·77 and 81·74 %), coupled to the weakness of R values for heterologous DNA pairing, may denote that these strains are distinct. Conversely, labelled DNA of other species yielded low R_{max} values (27·89–57·66 %), but non-negligible R values, as the probable result of cross-hybridization of the probe to multiple positions on the array. The *B. mallei* ATCC 23344^T probe yielded an R_{max} value of 30·43 ± 8·29 % (at the position of *B. mallei*) and markedly lower, an R value of 18·60 ± 6·75 % at the position of *B. pseudomallei* ATCC 23343^T. Similarly, the probe for *B. pseudomallei* ATCC 23343^T yielded an R_{max} value of 36·70 ± 10·26 % (at the position of *B. pseudomallei*) and markedly lower, an R value of 25·58 ± 7·19 % at the position of *B. mallei* ATCC 23344^T. It is known that *B. mallei* and *B. pseudomallei* are closely related, to the extent that published DNA–DNA relatedness values, obtained by the classical fluorimetric method in microdilution wells, are higher than 90 % (Yabuuchi *et al.*, 1992). However, the distinction between these two species is justified, based on pathogenicity. Originally included in the study design, *Burkholderia thailandensis* ATCC 700388^T was not available when the work started. It is reasonable to assume that labelled DNA from *B. thailandensis* ATCC 700388^T would have hybridized strongly to itself, to a lower extent to *B. mallei* and *B. pseudomallei*, and to a non-negligible extent to the members of the *B. cepacia* complex.

The symmetry of R values between *B. mallei* and *B. pseudomallei* was further investigated by determining R values with labelled DNA from two additional strains of *B. mallei* and 12 strains of *B. pseudomallei* (Table 2). The hybridization patterns observed for the three different strains of *B. mallei* are homogeneous (r = 0·9578) and consistent with that of *B. mallei* ATCC 23344^T (n = 10). The R values observed for positions of *B. mallei* ATCC 23344^T and *B. pseudomallei* ATCC 23343^T are not significantly different, but both differ significantly from R values observed at other positions (i.e. other *Burkholderia* species). The hybridization patterns observed for the 12 different strains of *B. pseudomallei* are more homogeneous (r = 0·9781). Moreover, significant differences of R values are observed for positions of *B. mallei* ATCC 23344^T and *B. pseudomallei* ATCC 23343^T on the array. For labelled DNA from both *B. mallei* and *B. pseudomallei*, the hybridization profiles at the remaining positions were similar. The R values are high for *B. cepacia* genomovar I

Table 2. Dispersion of *R* values observed with labelled DNA from three different strains of *B. mallei* and 12 different strains of *B. pseudomallei*

In addition to the type strains of *B. mallei* and *B. pseudomallei*, the following strains were tested: *B. mallei* CIP A192 and CIP A193 (obtained from the Collection de l'Institut Pasteur, Paris, France); *B. pseudomallei* CIP 58-52-238, CIP 58-56-91, CIP 59-62-28, CIP 60-68-3 (obtained from the CIP), CRSSA 14/97 and 29/97 (formerly ATCC 11668 and ATCC 15682, respectively), 8/93, 41/97, 42/97, 43/97 and 47/97 (obtained from Centre de Recherche du Service de Santé des Armées, La Tronche, France). Each strain was tested once.

DNA on arrays*	*R* (%) with labelled DNA from:	
	B. mallei ($n=3$)	*B. pseudomallei* ($n=12$)
1	7·3±0·9	7·4±1·2
2	7·3±1·9	9·2±1·6
3	8·4±1·1	8·4±1·2
4	3·3±0·5	3·4±0·4
5	5·5±0·9	5·4±0·8
6	8·2±0·3	6·5±0·5
7	4·3±1·4	6·3±0·8
8	3·5±1·7	4·2±1·0
9	3·3±1·1	4·4±0·7
10	4·0±1·8	4·1±0·7
11	1·5±0·6	1·5±0·4
12	22·3±5·9	15·1±2·7
13	20·3±5·4	23·4±4·7
14	0·3±0·2	0·3±0·2
15	0·2±0·1	0·2±0·1
16	0·3±0·1	0·1±0·1
Total	100	100

*See legend to Table 1 for strains.

ATCC 25416^T, *B. cepacia* genomovar III LMG 12614, *B. multivorans* LMG 13010^T, *Burkholderia pyrrocinia* ATCC 15958^T, *Burkholderia stabilis* LMG 14294^T, *Burkholderia vietnamiensis* LMG10929^T and *Burkholderia* sp. isolate CEB 01056, and low at other positions. One could have expected that these cross-hybridizations should have yielded reciprocal values (by a significant increase of *R* at positions of *B. mallei* and *B. pseudomallei*) when using labelled DNA from the species *B. cepacia* genomovar I ATCC 25416^T, *B. cepacia* genomovar III LMG 12614, *B. multivorans* LMG 13010^T, *B. pyrrocinia* ATCC 15958^T, *B. stabilis* LMG 14294^T, *B. vietnamiensis* LMG 10929^T and *Burkholderia* sp. isolate CEB 01056. Curiously, this was not the case, thus generating asymmetrical sections of the matrix (Table 1). This was also observed for *B. plantarii* LMG 9035^T. The genome of *B. cepacia* ATCC 25416^T (Rodley *et al.*, 1995) and that of *Burkholderia fungorum* LB400 (Joint Genome Institute, http://www.jgi.doe.gov), both estimated at 8·1 Mb, are markedly larger than those of *B. pseudomallei* clinical isolate K96943, estimated at 6·5 Mb (Songsivilai & Dharakul, 2000) and *B. mallei* ATCC 23344^T, estimated at 6·0 Mb (The Institute for Genomic Research, http://www.tigr.org). The homologous hybridizing proportion of the larger genomes will thus be correspondingly smaller than that of the smaller genomes, and may result in non-reciprocity. Such non-symmetry of matrices has already been observed for classical DNA–DNA homology studies (Lisdiyanti *et al.*, 2002; Takeuchi & Hatano, 1998).

The *B. cepacia* complex strains tested (genomovars I, II, III, IV and V), *B. pyrrocinia* ATCC 15958^T and the isolate CEB 01056 of *Burkholderia* sp. constitute a distinct DNA homology subgroup (Table 1): $40·12\% < R_{max} < 57·66\%$, $3·69\% < R < 12·29\%$ within this subgroup (mean 7·06±2·31 %), $R < 2·52\%$ with other species (mean 0·99±0·65 %). All these type strains are fully differentiated. This does not prejudge about the differentiation of *B. cepacia* members other than those tested herein.

DNA homology values relevant to the strains used herein have been obtained previously (Achouak *et al.*, 1999; Coenye *et al.*, 1999; Vandamme *et al.*, 1997; Viallard *et al.*, 1998; Yabuuchi *et al.*, 1992) based on three different techniques: nuclease S1 (NS1), microfluorimetry (MF) and spectrophotometry (SP) (Fig. 1). These data and current *R* values were plotted and a linear regression curve was then calculated (Fig. 1). The r^2 value of 0·740 demonstrates a weak positive correlation between DNA homology measures obtained by two radically different approaches. The slope of the tendency curve is therefore greatly affected by the value of *R* at 100 % DNA–DNA relatedness, because R_{max} values are directly lowered by the extent of cross-hybridization at other positions on the array. The use of the standardized ratio R_{std} eliminates this bias, thus strengthening the correlation ($r^2 = 0·904$). The correlation between the NS1, MF, SP and macro-array methods by comparing DNA similarity values for several DNA pairs remains to be assessed. Presently available data are rare but encouraging, and highlight the asymmetry mentioned above (Table 3). The qualitative (e.g. diversity of species) and quantitative (e.g. number of species) nature of the genetic material spotted on the array remains to be addressed, to determine how it affects *R* values.

B. mallei and *B. pseudomallei* are discriminated by PCR on the basis of a single nucleotide difference in the 23S rDNA sequence (Bauernfeind *et al.*, 1998). The ongoing identification of virulence determinants by subtractive hybridization will provide new tools in specific detection of these two species (DeShazer *et al.*, 2001). Proposition of *B. thailandensis* (Brett *et al.*, 1998), formerly a *Burkholderia pseudomallei*-like species that is able to assimilate L-arabinose contrary to *B. pseudomallei*, has been reinforced by 16S rDNA sequence analysis of numerous Ara$^+$ or Ara$^-$ clinical and environmental isolates (Dharakul *et al.*, 1999). Differentiation of *Burkholderia* species, most particularly those belonging to the *B. cepacia* complex, has been intensively addressed in the last few years, and a combination of phenotypic and molecular tests such as *recA* PCR and 16S rDNA RFLP are recommended for

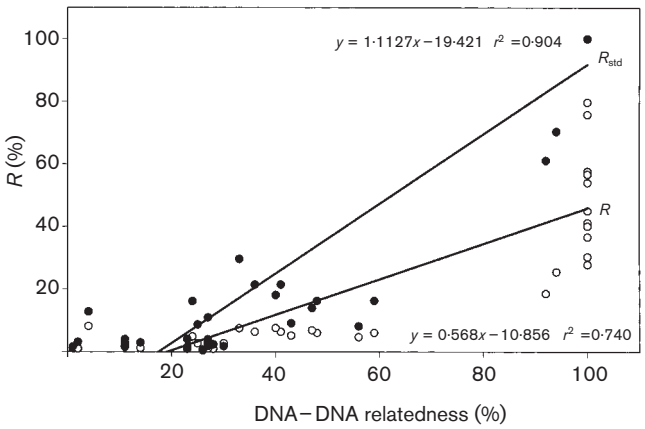

Fig. 1. Relationship between previously reported DNA–DNA relatedness values (references in the text) and R (open circles) or R_{std} (filled circles) values, obtained by the DNA macro-array method.

differentiation among the genomovars; the reliability of these approaches has been reviewed recently (Speert *et al.*, 2002). All these methods target different cellular constituents and are thus complementary; however, none of them consider the whole genome. Delineation of four *Pseudomonas* species by DNA–DNA hybridization, which uses hundreds of genome fragments of each species spotted onto DNA micro-arrays, offsets laborious cross-hybridization but needs tedious cloning and handling of genome fragments (Cho & Tiedje, 2001). The overall homology between two genomes remains an important feature in the identification process for atypical isolates or species never encountered before. The sequencing of complete genomes has opened the way for micro-arraying nearly all ORFs of a bacterium, and for future evaluation of whole-genome homology in a gene-by-gene manner (Dziejman *et al.*, 2002; Schoolnik, 2002).

DNA–DNA hybridization using genomic DNA macro-arrays looks like an interesting tool at the species level, and should be thoroughly addressed at the subspecies level by investigating several clinical and environmental isolates. Low intraspecies DNA–DNA hybridization values (up to 54 %) have been reported in several members of the *B. cepacia* complex, especially for genomovars I and III (Vandamme *et al.*, 1997). Moreover, comparison of data in Tables 1 and 2 suggests that R values (at the *B. mallei* and *B. pseudomallei* positions) tend to overlap more as the number of strains examined increases. This suggests a low clonality of the population structure that could be a result of 'rampant interspecific recombination' (Ochman *et al.*, 2000).

DNA–DNA hybridization using genomic DNA macro-arrays is simple and constitutes a straightforward approach for species identification. Automation of arraying enables the production of large quantities of macro-arrays, and the most demanding efforts consist of preparing the reference DNA stocks from type strains and updating macro-arrays for their species content as novel species are described. Once available, these macro-arrays could be used to investigate the nature of numerous unknown labelled DNAs extracted from various sources. There are at least two potential applications in whole-genome-based bacterial identification, as an initial step preceding the use of more appropriate or targeted tests. The first is the species identification of a large number of strains isolated from clinical infection or recovered from natural environments. Identification and monitoring of bacteria by reverse sample genome probing (RSGP) has already been described, for example in soil (Léveillé *et al.*, 2001; Shen *et al.*, 1998). The second application concerns the monitoring of bacteria in the environment by RSGP and a possible way to analyse spatio-temporal variations in bacterial communities within a biotope. DNA–DNA hybridization using genomic DNA macro-array analysis may be modified to use non-radioactive labelling and thus become a high-throughput alternative for existing DNA–DNA hybridization methods.

Acknowledgements

We would like to thank M. A. Labarre and I. Rébillat for maintaining cultures, C. Simoes for DNA extraction, and J. Chatoux and C. Pannetier who made sequencing facilities available. This work was funded by Délégation Générale pour l'Armement through grants PEA983601 and PEA983602 related to bio-defence.

Table 3. Comparison of DNA similarity values obtained by various methods

Methods used were nuclease S1 (NS1) (Viallard *et al.*, 1998), microfluorimetry (MF) (Yabuuchi *et al.*, 1992), spectrophotometry (SP) (Vandamme *et al.*, 1997) and macro-arrays (MA), with labelled DNA of *B. cepacia* (I) ATCC 25416^T. NA, Not available.

Strain	DNA similarity (%) with labelled DNA from *B. cepacia* ATCC 25416^T:				
	NS1	MF	SP	MA	MA_{rev}*
B. gladioli ATCC 10248^T	27	23	11	1·56 ± 0·53	2·72 ± 1·15
B. pyrrocinia ATCC 15958^T	48	NA	59	6·07 ± 1·89	14·63 ± 4·16
B. vietnamiensis (V) LMG 10929^T	41	NA	36	6·40 ± 1·96	6·13 ± 2·87

*MA_{rev} indicates the reverse labelling situation.

References

Achouak, W., Christen, R., Barakat, M., Martel, M.-H. & Heulin, T. (1999). *Burkholderia caribensis* sp. nov., an exopolysaccharide-producing bacterium isolated from vertisol microaggregates in Martinique. *Int J Syst Bacteriol* **49**, 787–794.

Balandreau, J., Viallard, V., Cournoyer, B., Coenye, T., Laevens, S. & Vandamme, P. (2001). *Burkholderia cepacia* genomovar III is a common plant-associated bacterium. *Appl Environ Microbiol* **67**, 982–985.

Bauernfeind, A., Roller, C., Meyer, D., Jungwirth, R. & Schneider, I. (1998). Molecular procedure for rapid detection of *Burkholderia mallei* and *Burkholderia pseudomallei*. *J Clin Microbiol* **36**, 2737–2741.

Benenson, A. S. (1995). *Control of Communicable Diseases Manual*, 6th edn, pp. 335–338. Washington, DC: American Public Health Association.

Brett, P. J., DeShazer, D. & Woods, D. E. (1998). *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *Int J Syst Bacteriol* **48**, 317–320.

Brisse, S., Verduin, C. M., Milatovic, D. & 7 other authors (2000). Distinguishing species of the *Burkholderia cepacia* complex and *Burkholderia gladioli* by automated ribotyping. *J Clin Microbiol* **38**, 1876–1884.

Cho, J.-C. & Tiedje, J. M. (2001). Bacterial species determination from DNA-DNA hybridization by using genome fragments and DNA microarrays. *Appl Environ Microbiol* **67**, 3677–3682.

Coenye, T., Holmes, B., Kersters, K., Govan, J. R. W. & Vandamme, P. (1999). *Burkholderia cocovenenans* (van Damme *et al.* 1960) Gillis *et al.* 1995 and *Burkholderia vandii* Urakami *et al.* 1994 are junior synonyms of *Burkholderia gladioli* (Severini 1913) Yabuuchi *et al.* 1993 and *Burkholderia plantarii* (Azegami *et al.* 1987) Urakami *et al.* 1994, respectively. *Int J Syst Bacteriol* **49**, 37–42.

Coenye, T., LiPuma, J. J., Henry, D., Hoste, B., Vandemeulebroecke, K., Gillis, M., Speert, D. P. & Vandamme, P. (2001a). *Burkholderia cepacia* genomovar VI, a new member of the *Burkholderia cepacia* complex isolated from cystic fibrosis patients. *Int J Syst Evol Microbiol* **51**, 271–279.

Coenye, T., Mahenthiralingam, E., Henry, D., LiPuma, J. J., Laevens, S., Gillis, M., Speert, D. P. & Vandamme, P. (2001b). *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. *Int J Syst Evol Microbiol* **51**, 1481–1490.

Coenye, T., Vandamme, P., Govan, J. R. W. & LiPuma, J. J. (2001c). Taxonomy and identification of the *Burkholderia cepacia* complex. *J Clin Microbiol* **39**, 3427–3436.

Dance, D. A. (2002). Melioidosis. *Curr Opin Infect Dis* **15**, 127–132.

DeShazer, D., Waag, D. M., Fritz, D. L. & Woods, D. E. (2001). Identification of a *Burkholderia mallei* polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. *Microb Pathog* **30**, 253–269.

Dharakul, T., Tassaneetrithep, B., Trakulsomboon, S. & Songsivilai, S. (1999). Phylogenetic analysis of Ara^+ and Ara^- *Burkholderia pseudomallei* isolates and development of a multiplex PCR procedure for rapid discrimination between the two biotypes. *J Clin Microbiol* **37**, 1906–1912.

Dziejman, M., Balon, E., Boyd, D., Fraser, C. M., Heidelberg, J. F. & Mekalanos, J. J. (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.

Gilligan, P. H. (1995). *Pseudomonas* and *Burkholderia*. In *Manual of Clinical Microbiology*, 6th edn, pp. 509–519. Edited by P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover & R. H. Yolken. Washington, DC: American Society for Microbiology.

Heath, D. G., Hohneker, K., Carriker, C., Smith, K., Routh, J., LiPuma, J. J., Aris, R. M., Weber, D. & Gilligan, P. H. (2002). Six-year molecular analysis of *Burkholderia cepacia* complex isolates among cystic fibrosis patients at a referral center for lung transplantation. *J Clin Microbiol* **40**, 1188–1193.

Henry, D. A., Mahenthiralingam, E., Vandamme, P., Coenye, T. & Speert, D. P. (2001). Phenotypic methods for determining genomovar status of the *Burkholderia cepacia* complex. *J Clin Microbiol* **39**, 1073–1078.

Holmes, A., Govan, J. & Goldstein, R. (1998). Agricultural use of *Burkholderia* (*Pseudomonas*) *cepacia*: a threat to human health? *Emerg Infect Dis* **4**, 221–227.

Inglis, T. J. J., Rigby, P., Robertson, T. A., Dutton, N. S., Henderson, M. & Chang, B. J. (2000). Interaction between *Burkholderia pseudomallei* and *Acanthamoeba* species results in coiling phagocytosis, endamebic bacterial survival, and escape. *Infect Immun* **68**, 1681–1686.

Johnson, J. L. (1991). DNA reassociation experiments. In *Nucleic Acid Techniques in Bacterial Systematics*, pp. 21–44. Edited by E. Stackebrandt & M. Goodfellow. Chichester: Wiley.

Jones, A. M., Dodd, M. E. & Webb, A. K. (2001). *Burkholderia cepacia*: current clinical issues, environmental controversies and ethical dilemmas. *Eur Respir J* **17**, 295–301.

Léveillé, S. A., Leduc, L. G., Ferroni, G. D., Telang, A. J. & Voordouw, G. (2001). Monitoring of bacteria in acid mine environments by reverse sample genome probing. *Can J Microbiol* **47**, 431–442.

LiPuma, J. J., Spilker, T., Coenye, T. & Gonzalez, C. F. (2002). An epidemic *Burkholderia cepacia* complex strain identified in soil. *Lancet* **359**, 2002–2003.

Lisdiyanti, P., Kawasaki, H., Widyastuti, Y., Saono, S., Seki, T., Yamada, Y., Uchimura, T. & Komagata, K. (2002). *Kozakia baliensis* gen. nov., sp. nov., a novel acetic acid bacterium in the α-*Proteobacteria*. *Int J Syst Evol Microbiol* **52**, 813–818.

Mahenthiralingam, E., Bischof, J., Byrne, S. K., Radomski, C., Davies, J. E., Av-Gay, Y. & Vandamme, P. (2000). DNA-based diagnostic approaches for identification of *Burkholderia cepacia* complex, *Burkholderia vietnamiensis*, *Burkholderia multivorans*, *Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. *J Clin Microbiol* **38**, 3165–3173.

Mougel, C., Thioulouse, J., Perrière, G. & Nesme, X. (2002). A mathematical method for determining genome divergence and species delineation using AFLP. *Int J Syst Evol Microbiol* **52**, 573–586.

Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304.

Parke, J. L. & Gurian-Sherman, D. (2001). Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annu Rev Phytopathol* **39**, 225–258.

Rajan, S. & Saiman, L. (2002). Pulmonary infections in patients with cystic fibrosis. *Semin Respir Infect* **17**, 47–56.

Rodley, P. D., Romling, U. & Tummler, B. (1995). A physical genome map of the *Burkholderia cepacia* type strain. *Mol Microbiol* **17**, 57–67.

Rosselló-Mora, R. & Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiol Rev* **25**, 39–67.

Rotz, L. D., Khan, A. S., Lillibridge, S. R., Ostroff, S. M. & Hughes, J. M. (2002). Public health assessment of potential biological terrorism agents. *Emerg Infect Dis* **8**, 225–230.

Schoolnik, G. K. (2002). Functional and comparative genomics of pathogenic bacteria. *Curr Opin Microbiol* **5**, 20–26.

Segonds, C., Heulin, T., Marty, N. & Chabanon, G. (1999). Differentiation of *Burkholderia* species by PCR-restriction fragment

length polymorphism analysis of the 16S rRNA gene and application to cystic fibrosis isolates. *J Clin Microbiol* **37**, 2201–2208.

Shen, Y., Stehmeier, L. G. & Voordouw, G. (1998). Identification of hydrocarbon-degrading bacteria in soil by reverse sample genome probing. *Appl Environ Microbiol* **64**, 637–645.

Songsivilai, S. & Dharakul, T. (2000). Multiple replicons constitute the 6·5-megabase genome of *Burkholderia pseudomallei*. *Acta Trop* **74**, 169–179.

Speert, D. P., Henry, D., Vandamme, P., Corey, M. & Mahenthiralingam, E. (2002). Epidemiology of *Burkholderia cepacia* complex in patients with cystic fibrosis, Canada. *Emerg Infect Dis* **8**, 181–187.

Srinivasan, A., Kraus, C. N., DeShazer, D., Becker, P. M., Dick, J. D., Spacek, L., Bartlett, J. G., Byrne, W. R. & Thomas, D. L. (2001). Glanders in a military research microbiologist. *N Engl J Med* **345**, 256–258.

Stackebrandt, E., Frederiksen, W., Garrity, G. M. & 10 other authors (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* **52**, 1043–1047.

Steinmetz, I., Reganzerowski, A., Brenneke, B., Häussler, S., Simpson, A. & White, N. J. (1999). Rapid identification of *Burkholderia pseudomallei* by latex agglutination based on an exopolysaccharide-specific monoclonal antibody. *J Clin Microbiol* **37**, 225–228.

Takeuchi, M. & Hatano, K. (1998). Proposal of six new species in the genus *Microbacterium* and transfer of *Flavobacterium marinotypicum* ZoBell and Upham to the genus *Microbacterium as Microbacterium maritypicum* comb. nov. *Int J Syst Bacteriol* **48**, 973–982.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. & Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**, 407–438.

Vandamme, P., Holmes, B., Vancanneyt, M. & 8 other authors (1997). Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. *Int J Syst Bacteriol* **47**, 1188–1200.

Vandamme, P., Mahenthiralingam, E., Holmes, B., Coenye, T., Hoste, B., De Vos, P., Henry, D. & Speert, D. P. (2000). Identification and population structure of *Burkholderia stabilis* sp. nov. (formerly *Burkholderia cepacia* genomovar IV). *J Clin Microbiol* **38**, 1042–1047.

Vandamme, P., Henry, D., Coenye, T., Nzula, S., Vancanneyt, M., LiPuma, J. J., Speert, D. P., Govan, J. R. W. & Mahenthiralingam, E. (2002). *Burkholderia anthina* sp. nov. and *Burkholderia pyrrocinia*, two additional *Burkholderia cepacia* complex bacteria, may confound results of new molecular diagnostic tools. *FEMS Immunol Med Microbiol* **33**, 143–149.

Viallard, V., Poirier, I., Cournoyer, B., Haurat, J., Wiebkin, S., Ophel-Keller, K. & Balandreau, J. (1998). *Burkholderia graminis* sp. nov., a rhizospheric *Burkholderia* species, and reassessment of [*Pseudomonas*] *phenazinium*, [*Pseudomonas*] *pyrrocinia* and [*Pseudomonas*] *glathei* as *Burkholderia*. *Int J Syst Bacteriol* **48**, 549–563.

Wayne, L. G., Brenner, D. J., Colwell, R. R. & 9 other authors (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approach to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.

Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T. & Arakawa, M. (1992). Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiol Immunol* **36**, 1251–1275.