The ecological success of Burkholderia terrae BS001 and related strains in the mycosphere

Nazir, Rashid

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2012

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Insights into the genome of *Burkholderia terrae* BS001: potential fungal-interactive bacterial gene systems

Rashid Nazir, Martin A. Hansen*, Jan A. Warmink, Lars Hansen*, Soren Sorensen*, Jan Dirk van Elsas

Department of Microbial Ecology, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

*Molecular Microbial Ecology Group, University of Copenhagen, Sølvgade 83H DK-1307K, Denmark.

Partially published as:

Abstract

Burkholderia terrae BS001 is a soil bacterium, which was originally isolated from the mycosphere of the ectomycorrhizal fungus Laccaria proxima. It exhibits a range of fungal-interactive traits which are consistent with its propensity to actively interact at fungal interfaces. In this chapter, we describe the draft genome of B. terrae BS001. The genome consists of five replicons, of 4.1, 3.5, 2.3, 0.9 and 0.5 Mb, for a total of about 11.5 Mb, and > 80% of the sequence is predicted to be coding. This is one of the largest genomes of unicellular bacteria that have been found so far, indicating a remarkable functional versatility. The genome’s average G+C content was found to be 61.52%. A closer analysis of the draft genome of B. terrae BS001 revealed a range of genetic systems that are potentially involved in energy generation (about 27%). Moreover, a plethora of genes involved in interactions of the bacterium with other organisms, such as soil fungi, was found. Thus, the B. terrae BS001 genome was found to possess ample genetic systems that encode chemotaxis, motility and signaling. It also contained colonization-specific genetic systems, such as type two, three, four and six secretion systems (several) next to type four pilus and biofilm formation gene systems. Furthermore, the draft genome of B. terrae BS001 revealed the presence of genes that are presumably involved in the capture and metabolism of fungal-released carbonaceous compounds, like glycerol and trehalose. The presence of antibiotic biosynthesis as well as resistance genes indicated the capacity of this bacterium to build fences as well as to stand potentially hostile conditions established by growing fungi. The collective bacterial genetic load is consistent with the ecological success that B. terrae BS001 attains in the mycosphere, where a colonizable surface emerges, moves and may provide nutrients in an erratic manner.
Introduction

Burkholderia terrae strain BS001 is a Gram-negative rod-shaped bacterium which belongs to the beta-proteobacteria. It was originally isolated from acidic sandy soil underneath the foot of Laccaria proxima fruiting bodies (Warmink and van Elsas 2008). Similar Burkholderia terrae types were positively selected in soil by the growing fungus Lyophyllum sp. strain Karsten (Warmink & van Elsas, 2009; Nazir et al., 2012). Remarkably, B. terrae BS001 is able to migrate along with growing fungal hyphae and assist other bacteria, e.g. Dyella japonica strain BS003, to migrate as well (Warmink et al., 2011). Interestingly, it relates to the group of so-called beneficial (plant-associated) Burkholderia species (Saurez-Moreno et al., 2011), which is also related to the endomycotic B. rhizoxinica (Partida et al., 2005). These fungal-interactive Burkholderia species exhibit characteristics like swimming and twitching motility, chemotaxis (Furuno et al., 2010) and biofilm formation (Warmink and van Elsas, 2009). They also reveal the presence of diverse protein or nucleoprotein secretion systems (Nazir et al., 2010), which may allow them to attach to, colonize and infest fungal surface structures and tissue. The ecological success of B. terrae BS001 in the mycosphere may be related to the presence on its genome of a rich plethora of genetic systems which enable it to be optimally fit at the fungus.

Given the expected benefit to our understanding of its ecology, we decided to undertake whole genome sequencing of B. terrae BS001. We here describe the genome sequence and interpret it in terms of functioning of the bacterium in soil and mycosphere. The draft sequence of the genome, encompassing 336 contigs, has been deposited at DDBJ/EMBL/GenBank under accession number AKAU00000000. The information contained in the first version, AKAU01000000, is described here.
Materials and methods

Bacterial strain and genome sequencing

The bacterial strain *Burkholderia terrae* BS001, isolated from the *Laccaria proxima* mycosphere (Warmink and van Elsas 2008), is present in the bacterial collection of the Microbial Ecology Department of Groningen University, The Netherlands. The strain *B. terrae* BS001 was cultured from the -80 °C stock overnight (at 23 °C) and fully-grown cultures were used for high molecular weight (HMW) whole genome extraction by the ‘GNOME DNA kit’ (MP Bio Illkirch, France) according to the supplier’s instructions.

The genome sequence of *B. terrae* BS001 was determined via paired-end sequencing using the Illumina GAIIx platform (BASECLEAR Leiden, NL) with average insert of ~260 bases (min 175, max 375), attaining 7,014,080 reads. Sequencing was also done using the Roche 454 pyrosequencing platform (Molecular Microbial Ecology Group, Copenhagen) which yielded 541,595 reads (of 363 bp each). The genome was de novo assembled on the basis of the 454 generated data, supported by the Illumina data, using a range of different assemblers like the Newbler (version 2.5.3) and Velvet (version 0.7.59) assembler.

De novo assembly of genome sequence data

The Illumina genome shotgun sequencing effort yielded 7,014,080 (paired-end) reads, which were evaluated for quality. The mean quality scores were all over Q30, indicating good quality overall. Next, all sequences were cleaned of remaining adapter parts, using the standard Illumina programme used for this purpose. All reads were then trimmed in a two-step process. First, we end-trimmed all, by removing residues from the ends with quality scores equal to or below 15 (Q15), and discarded all reads shorter than 30 bases. Next, we ran a 5-base sliding window over the sequences to locate valleys with mean quality scores below 15, discarding such reads. Using the remaining 6,714,528 “healthy” reads, which encompassed about 335.7 Mb of sequence (average coverage of 29-fold) de novo assembly was performed using either of three assemblers, IDBA, Velvet and Ray. We found the
Raw Velvet - pair assembly to be the best, as it yielded the lowest number of contigs and, on average, the largest contigs. The partially closed genome data were sent to the annotation pipeline ‘RAST’ for automatic annotation.

Next to the Illumina run, we performed shotgun sequencing with the Roche 454 pyrosequencing platform. On this platform, we achieved 541,595 (single-end) reads of 363 bp on average (maximum 851 bp). Thus, another 196.73 Mb of sequence information (coverage roughly 17) was obtained. Using this dataset, we performed assembly using Newbler (version 2.5.3). The best assembly achieved around 336 contigs (Table 1).

In a second stage, we attempted to combine both the Illumina and Roche 454 generated datasets using the CLC work bench. The best combined assembly resulted from shredding the contigs from one assembly (the Illumina one) into pseudo reads and feeding them to the Velvet software. The contigs from the Newbler assembly were then subjected to closure using IMAGE2 software for gap closure with 5 rounds of iterations. The results of this approach did not improve our analysis, as roughly a similar number of contigs was the best result that could be achieved. Thus, the original assembly in which only the 454 data were used was the best assembly that could be obtained. It served as the basis for all further analyses.

**Annotation**

Annotation was performed automatically using RAST (Aziz et al., 2008) with the set default parameters of the software. Manual annotation was also performed for some selected gene regions to improve the quality of automated annotation for this genome project.

**In silico genomics techniques**

Selected gene regions were compared to database entries using BLAST-N. In particular, we compared the 16S rRNA genes found in our genome to those in databases. Moreover, comparisons of putative ORFs were done at the protein level using BLAST-P. We used described genetic regions (operons) to fit our predicted
**Table 1** De-novo assembly for 454 data performed using Newbler-assembler

<table>
<thead>
<tr>
<th>N50</th>
<th>Gene coverage</th>
<th>Min. length (bp)</th>
<th>Max. length (bp)</th>
<th>Mean</th>
<th>Number of contigs</th>
<th>Total length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>73653</td>
<td>425563</td>
<td>500</td>
<td>498088</td>
<td>34502</td>
<td>336</td>
<td>11316694</td>
</tr>
<tr>
<td>67582</td>
<td>425478</td>
<td>500</td>
<td>498085</td>
<td>35689</td>
<td>417</td>
<td>11313633</td>
</tr>
<tr>
<td>42577</td>
<td>184220</td>
<td>500</td>
<td>212092</td>
<td>18463</td>
<td>609</td>
<td>11244284</td>
</tr>
</tbody>
</table>

Operons. A range of bacterial effector proteins were found and compared with the available database (http://www.effectors.org/) and ‘T3SE db’ (http://effectors.bic.nus.edu.sg/T3SEdb/predict.php). Non-ribosomal peptide synthases (NRPS) were screened by using the software http://www.nrpssp.com/index.php.

**Results**

**Confirmation of strain identity**

We found three copies of the 16S rRNA gene in the *B. terrae* strain BS001 genome, which were 99.5% homologous to each other. Based on the canonical sequence, the total 16S rRNA gene sequence shared high similarity (>99%) with that of the type strain of *B. terrae*, i.e. strain KMY02 (Fig. 1). Other closest strains, with respectively 99.3, 98.5 and 96.7% homology, were *B. hospita* LMG20598, *B. caribensis* MWAP64, and *B. phymatum* STM815. This analysis placed *B. terrae* BS001 in a cluster composed of *B. terrae*, *B. hospita* and *B. caribensis* (Fig. 1). As by the phylogenetic tree the strain BS001 is more close to *B. hospita* type strain LMG20598 than to *B. terrae* type strain KMY02, there is a need in future to reconsider the taxonomy of this strain BS001. Reclassification should await further work with a suit of more and similar strains.
Chapter 6

Fig. 1 16S rRNA gene based phylogenetic analysis of B. terrae BS001 and relevant type strains. Neighbour joining tree was built by Kimura 2-parameter model. Bootstrapping was performed with 1000 replications.

General features of the B. terrae BS001 genome

On the basis of pulsed-field gel electrophoresis (PFGE), the genome of B. terrae BS001 was found to consist of five replicons, of respectively 4.1, 3.5, 2.3, 0.9 and 0.5 Mb, giving a total of (estimated) of roughly 11-12 Mb. We termed the three large replicons chromosomes, whereas the two smaller ones were termed (mega) plasmids. We did not obtain evidence for the presence of any plasmids of smaller size. On the basis of both the Roche 454 and the Illumina runs, the size of the sequenced genome was estimated to be, respectively, 11.5 and 11.3 Mb. Taken together with the PFGE data, and given the relative uncertainty of the smaller estimates, we estimate the total genome size of B. terrae BS001 to be 11.5 Mb.

The resulting best assembly yielded 336 contigs (Table 1). The draft genome had an average G+C content of 61.52%. RAST-based automatic annotation
predicted the existence of 10,975 protein-encoding sequences (CDS), with 67.27% assigned to recognizable functional genes. Moreover, this analysis roughly predicted that about 9.33 Mb of genome sequence was coding, amounting to about 81.2% of the total genome sequence. A whole suit of predicted gene-classes is summarized in figure 2. Furthermore, 91 tRNA genes were predicted to be present in the draft genome of \textit{B. terrae} BS001.

In comparison to the genomes of related \textit{Burkholderia} species, the genome of \textit{B. terrae} BS001 had the largest size. Interestingly, about the same fraction, i.e. 32%, of the genome of \textit{B. terrae}, \textit{B. phymatum} and \textit{B. rhizoxinica} was found to contain hypothetical genes (Fig. 3). Surprisingly there was a significant increase in the number of CDSs in the \textit{B. terrae} BS001 genome for stress responses, amino acid and carbohydrate metabolism as compared to the genomes of \textit{B. phymatum} \textit{B. rhizoxinica} and \textit{B. xenovorans} (Fig. 4).

Core genome, central metabolism and stress responses

The draft genome of \textit{B. terrae} BS001 was found to have a modest core involved in building the cell structure and its basic functioning.

\textit{Cell division and cell cycle} – A total of 43 functional genes were annotated by RAST as presumably involved in cell division and the cell cycle. Several of these predicted genes, \textit{i.e. MreB, MreC, MreD, Maf, RodA, FtsZ, FtsA, FtsQ, FtsW, FtsK, FtsL, MraZ, FtsI, FtsB, FtsN, ZipA, ZapA, MinC, MinD, MinE, DivIVA, ParA and ParB} are involved in producing the bacterial cytoskeleton. Furthermore, there were about 164 CDSs presumably involved in DNA metabolism, including replication (40), repair (101) and recombination (4). Remarkably, there were six genes for non-homologous end-joining in the DNA repair category. Along with DNA processing and uptake (4), there was a subset of 15 genes potentially involved in structure, modification and restriction of DNA at different phases of the cell cycle. There was a full suite of DNA-interactive CDSs related to the canonical \textit{recA, recX, mutS, mutL} and \textit{lexA} genes, next to DNA polymerases I through IV.
Furthermore, 156 CDSs found in the draft genome were possibly involved in the processing of nucleosides and nucleotides. In addition, a set of 154 CDSs was predicted to be involved in RNA metabolism, including essential genes like \( rpoS \) (1), \( rpoN \) (4), \( rpoH \) (1), \( rpoE \) (4) and \( rpoD \) (6). A subset of 118 and 36 genes of this category are likely dedicated to RNA processing, modification and transcription respectively.

For energy acquisition, the \( B.\ terrae\ BS001\) genome was found to contain 365 CDSs related to aerobic and anaerobic respiration. This includes electron transport processes (238) and proton-motive force (12). There are 131 CDS related to only electron accepting reactions. In this sub-category, there are 48 CDS serving as putative anaerobic respirator reductases. Anaerobic respiration is likely to occur on the basis of nitrate as the terminal electron acceptor.

With respect to responses to stress, there are 401 CDSs in the draft genome of \( B.\ terrae\ BS001\) that are potentially devoted to cope with any kind of stress that a cell may encounter. The main subsets of this category include CDSs that provide resistance to osmotic (50), oxidative (191), cold (13), heat (19) and periplasmic (6) stresses. Some other six genes potentially allow sensing of carbon starvation, giving signals to the cell to induce differentiation to a stress-adapted form.

**Fermentation**

The \( B.\ terrae\ BS001\) draft genome contained 178 CDSs that are potentially involved in fermentation. This category has subsets of CDSs devoted to butanol biosynthesis (43), acetyl-CoA fermentation to butyrate (89), lactate fermentation (28) and acetoin and butanediol metabolism (18).

**Presence of genetic systems that are presumably involved in nutritive strategies and interactions with (fungal) hosts**

Several predicted genes and/or operons were found that are potentially important in the fungal-interactive behaviour of \( B.\ terrae\ BS001\). These encompass genetic systems that are presumably involved in (regulation of) motility and chemotaxis.
Table 2 Salient predicted functional genes in draft genome of *B. terrae* BS001; particularly the ones potentially involved in bacterial-fungal interactions in soil

<table>
<thead>
<tr>
<th>Name of operon</th>
<th>Number of CDSs</th>
<th>% of whole genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defense and resistance</td>
<td>192</td>
<td>1.75</td>
</tr>
<tr>
<td>Potassium metabolism</td>
<td>58</td>
<td>0.53</td>
</tr>
<tr>
<td>Phages</td>
<td>3</td>
<td>0.03</td>
</tr>
<tr>
<td>Membrane transport (sec. systems)</td>
<td>208</td>
<td>1.89</td>
</tr>
<tr>
<td>Iron acquisition and metabolism</td>
<td>35</td>
<td>0.32</td>
</tr>
<tr>
<td>Cell division and cell cycle</td>
<td>43</td>
<td>0.39</td>
</tr>
<tr>
<td>Motility and chemotaxis</td>
<td>159</td>
<td>1.45</td>
</tr>
<tr>
<td>Regulation and cell signaling</td>
<td>159</td>
<td>1.45</td>
</tr>
<tr>
<td>Nitrogen metabolism</td>
<td>112</td>
<td>1.02</td>
</tr>
<tr>
<td>Stress response (mainly oxidative)</td>
<td>401</td>
<td>3.65</td>
</tr>
<tr>
<td>Carbohydrates metabolism</td>
<td>1207</td>
<td>10.99</td>
</tr>
<tr>
<td>NRPS*</td>
<td>34</td>
<td>0.3</td>
</tr>
</tbody>
</table>

* NRPS: non Ribosomal peptide synthase

(159), cell signalling (155), iron acquisition and metabolism (35), transport over the membrane and secretion (208). In particular, the draft genome revealed the presence of several tens of putative ABC transporters (86), as well as type II, III, IV and VI secretion systems. Among the almost 2,000 predicted genes involved in metabolism, the genome revealed the presence of 367 CDSs for the metabolism of aromatic compounds, next to 1,207 ones for carbohydrate metabolism. This remarkable investment of *B. terrae* BS001 in systems that allow it to feed on different C compounds (over about 16% of its total genome) indicates it was selected to deal with fluctuating amounts and types of carbon sources that become available in its niche, potentially being released from a colonized host. Moreover, using the ‘Effective’ (http://www.effectors.org/) and ‘T3SE db’ (http://effectors.bic.nus.edu.sg/T3SEdb/predict.php) softwares for rapid prediction and identification, we detected a total of 15 effector proteins that are potentially secreted by the type III secretion system. Finally, several genes were found for the biosynthesis of bacteriocins (14), resistance to antibiotics and toxic compounds (178).
Host-interactive gene systems in the *B. terrae* BS001 genome

**Sensing systems and cell viability**

**Regulation and cell signaling** - A total of 155 predicted genes were found that presumably encode functions involved in cell signaling and regulation. The majority of these genes were recognized as being regulatory, including two-component sensing/regulatory systems. Subsets of this category include CDSs involved in cAMP signaling (61), HPr catabolite repression system (2), DNA-binding regulatory proteins (18), orphan regulatory proteins (27), global two-component regulator PrrB,A of *Proteobacteria* (2), Zinc regulated enzymes (36) and the stringent response alarmone (p)ppGpp metabolism (6), while a subset of these is designated to quorum sensing.

![Graph showing the distribution of different classes of gene systems](image)

Fig. 2 A visual display of different classes of gene systems present in the draft genome of *B. terrae* BS001, annotated by RAST.

**Chemotaxis and motility** - Fungal mycelium can release a number of carbonaceous compounds that may become available for bacteria that live in the vicinity. The *B. terrae* BS001 genome was shown to contain 55 different predicted genes that are potentially involved in bacterial chemotaxis. These genes normally work in operons that determine the functionality of the flagellum. A close analysis revealed that these
clusters of chemotaxis-related genes, particularly \textit{fliC}, \textit{fliD}, \textit{flhC}, \textit{flhD}, \textit{motA}, \textit{motB} and \textit{cheY} were similar to those of \textit{B. xenovorans} LB400.

The genome of \textit{B. terrae} BS001 was further found to harbour 104 different predicted functional genes involved in motility. Sixty six of these 104 CDSs are structural, enabling flagellar assembly while the remaining 38 are involved in functionality. These flagellar motility genes shared a high nucleotide sequence similarity (>80%) with similar systems found in \textit{B. cepacia} R181 and \textit{B. xenovorans} LB400.

\textbf{Genetic systems involved in interaction with (fungal) hosts}

\textit{Type three secretion system (T3SS):} The T3SS is a molecular tool that a bacterium uses to interact with host cells/tissue or a surface (He et al., 2004). The genome of the fungal-interactive \textit{B. terrae} BS001 revealed the presence of a full operon representing a typical T3SS (Fig. 5). In detail, almost all of the genes described in the canonical T3SS of \textit{Ralstonia solanacearum} GMI1000 are present in this operon. One structural gene, \textit{i.e.} \textit{hrcC} (\textit{R. solanacearum} nomenclature), of the T3SS machinery was divided over two contigs with a possible gap in its sequence of 100-150 bp. Next to a match with the \textit{Ralstonia solanacearum} T3SS, the one of \textit{B. rhizoxinica} revealed a close match as well (data not shown here). Other related bacterial strains with sequence homology for T3SS genes are \textit{B. pseudomallei} 668, \textit{Xanthomonas campestris} ATCC33913 and \textit{Acidovorax avenae} AAC00-1.

Furthermore, using the ‘Effective’ (http://www.effectors.org/) and ‘T3SE db’ (http://effectors.bic.nus.edu.sg/T3SEdb/predict.php) softwares, we found 15 candidate effector proteins that are potentially secreted by the T3SS. Two of these effector proteins, \textit{i.e.} CDS278 (161 amino acids) and CDS7944 (460 amino acids) belong to the protein superfamily DUF1537 and share 57% and 79% identity with similar proteins from \textit{B. phytofirmans} PsJN and \textit{B. graminis} C4D1M, respectively. Another predicted candidate effector protein was CDS8404 (114 aa), which belongs to the HopJ superfamily and shares maximum identity of 49% with a similar protein of \textit{Methylphaga aminisulfidivorans} MP.
**Type four secretion system (T4SS):** The T4SS is a bacterial secretion system that is involved in the export of either protein or nucleoprotein complexes. The *B. terrae* BS001 genome was found to contain one copy of the T4SS. We surmised this system might potentially be involved in the association of this organism with hosts such as fungi. All genes of the canonical T4SS (denoted *vir* genes), *i.e.* *virB1* to *virB11*, were found to be located in one operon while a *virD4* homologue is located in the flanking region of *virB11*. *VirD4* is a coupling protein for DNA transfer and its presence in *B. terrae* BS001-genome indicates the presence of the capacity for gene transfer and mobility in this organism. However, the *virB4* gene of the T4SS cluster was only partially present, *i.e.* it was divided over two different contigs of the *B. terrae* BS001 draft genome. The T4SS gene cluster found, shared the highest amino acid homology with similar systems found in *B. phymatum* strain STM815 (≥ 70%) and - to some extent – with one in *Vibrio fischeri* ES114 (≥ 40%).

![Fig. 3 Comparison of the B. terrae BS001 draft genome with genomes from related Burkholderia strains; Red: functional part of the genomic coding regions; Blue: hypothetical part of the genomic coding regions.](image-url)
Type six secretion system (T6SS): We found 2 or 3 clusters of CDSs that constituted canonical T6SSs (11-17 genes in each) in the draft genome of *B. terrae* BS001. The canonical T6SS consists of about 21 proteins (*e.g.* in *B. mallei*; Schell et al., 2007) that span the inner and outer membranes of Gram-negative cells (Records, A.R., 2011; Pukatzi et al., 2009). All 45 annotated genes in the operons of *B. terrae* BS001 putatively encode structural as well as functional proteins of this secretion system. The functional involvement of these gene clusters in interactions with fungal hosts should be further explored in the near future. A prominent gene involved in T6SS functioning is vgrG; it works as a sensor/indicator and is the very first protein molecule to hit the host surface. The draft genome of *B. terrae* BS001 was found to contain eight different partial or full gene copies of vgrG which are either situated in the T6SS clusters or were found scattered across the whole genome outside of the T6SS operons. However, all T6SS clusters contained at least one copy of the vgrG gene.

Type IV pilus: The type IV pilus is normally used to export DNA from a bacterial cell to the outside milieu or into a eukaryotic cell, that may facilitate the horizontal gene transfer (HGT) even across kingdom (Lawrence et al., 2011). The genome of *B. terrae* strain BS001 was shown to contain the so-called Flp pilus, which is a type IVb pilus. Unfortunately, the type IV pilus related gene cluster was spread over two different contigs with sequence gaps in between. However, the available information revealed the presence of nine CDSs which potentially are involved in the assembly of this Flp pilus in *B. terrae* BS001.

The DNA secreted by the type IV pilus has been implicated in enhancement of the matrix in emerging biofilms. We speculate that the biofilm that is visibly formed by *B. terrae* strain BS001 on the hyphae of *L. sp.* strain Karsten may also contain DNA excreted by the type IV pilus.

Biofilm formation: After adhering to a colonizable surface, bacteria often develop biofilms which are constituted of matrices of cells encompassed in a polysaccharide layer produced by the bacterial cells (Brandl et al., 2011; Wang et al., 2004). In these polysaccharide matrices, bacterial cells live in thin layers. The draft genome of *B.
**terrae** BS001 was found to contain the predicted genes *pgaA, pgaB, pgaC* and *pgaD*, which are presumably involved in the active secretion of polysaccharides important for biofilm formation. These genes are situated in a cluster, flanked by the PmbA family gene TidE and a glucosamine-fructose-6-phosphate aminotransferase gene. The gene cluster was closely related to a similar cluster in *R. solanacearum* GMI1000 (50%) (on the basis of amino acid homology) and also to ones of *B. multivorans* CGD2M (49%) and *B. parapertussis* 12822 (49%).

**Glycerol uptake and utilization:** We found four predicted gene clusters in the draft genome of *B. terrae* BS001 that are potentially involved in the uptake and utilization of glycerol by this bacterium. Glycerol uptake can be passive, but this is a slow process. It can also involve ABC transporters along with so-called ‘*ugp*’ genes. Two distinct gene clusters consisting of the *ugpA, ugpC* and *ugpE* genes, next to other

![Fig. 4](attachment:image_url)
ABC transport related genes, were found. On the other hand, the other two clusters contained only ugpA and other glycerol-3-phosphate related ABC transporter genes. Surprisingly, ugpB genes were never found for any of the systems. Such genes normally make part of the glycerol uptake system in bacteria. A suite of predicted genes (14) for glycerol catabolism was further found.

**Sugar alcohols and organic acids:** Growing fungi may release sugar alcohols and/or organic acids in the surrounding milieu which serve as carbon sources for the bacteria living nearby. The genome of *B. terrae* BS001 was found to contain 10 and 23 predicted genes, respectively, for mannitol and inositol uptake and utilization along with the glycerol-related (14) genetic systems as described above. Trehalose biosynthesis and utilization was encoded by respectively 17 and 5 predicted genes (scattered over the genome). Furthermore, the genome was found to contain 88 CDSs involved in organic acid metabolism.

**Chitin degradation:** one predicted gene involved in chitin depolymerization, chiA, was found in the *B. terrae* BS001 draft genome. Along with this, eight different genes - spread over five different contigs of *B. terrae* BS001 genome – were apparently designated for the utilization of the chitin monomer N-acetylglucosamine.

**Antibiotic related functions allowing survival in fungal-infested habitats**

*Ribosomally synthesized antibacterial peptides and non-ribosomal peptide synthetases (NRPSs)*

We questioned whether *B. terrae* BS001 can express antimicrobial characteristics allowing it to build fences which might enhance the ecological success of itself and its fungal partner. Interestingly, we found 14 different predicted genes for ribosomally synthesized antibacterial peptides, particularly for Colicin V and related bacteriocins. Several of these genes e.g. Colicin V production protein *DedE* and related ones like *DedA* and *DedD* were found to share sequence homology with similar ones in *B. xenovorans* strain LB400. Furthermore, by using
http://www.nrpssp.com/index.php, we found 34 CDSs, with 37 conserved domains in them, which bind to specific substrates *i.e.* phenylalanine (27), dihydroxybenzoate (2), glutamic acid (1), tryptophan (1), alanine (1), dihydroxy phenylglycine (1), aspartic acid (1), serine (1), valine (1) and glycine (1). These multi-modular enzymes may potentially biosynthesize many important peptide compounds in this bacterium, which would, subsequently, make it ecologically successful in the soil environment. Some of the tentative NRPS's of *B. terrae* BS001 have high amino acid similarity (up to 64%) with those of *B. xenovorans* species.

**Fig. 5** T3SS operon in *B. terrae* BS001 genome compared to related T3S systems. Nomenclature of predicted genes: (1) hrcJ, (2) hrcN/yscN/escN, (3) yscU/hrcU, (4) yscT/hrcT, (5) hrcV/escV, (6) yscR/spaR/hrcR, (7) yscS, (8) hpaB, (9) yscL, (10) sctQ, (11) hrcC, (12) araC/hrpB, (13) sctD, (21) hrpB7, (22) hrpB1 and (23) hpaP. Digits refer to the numbers in the figure. Similar number across the organisms indicates the same predicted gene.

**Resistances**

A high number of predicted genes on the *B. terrae* BS001 draft genome, *i.e.* 178, are presumed to be involved in resistance to antibiotics and other toxic compounds. This category could be split into three sub-categories: (A) metal resistance including copper homeostasis (23), cobalt-zinc-cadmium resistance (57), mercuric reductases (3), copper homeostasis: copper tolerance (6), arsenic resistance (6) and resistance to chromium compounds (7); (B) resistance to fluoroquinolones (4) and beta-lactamase (8) and (C) multidrug resistance, tripartite systems found in Gram-
negative bacteria (30) and multidrug resistance efflux pumps (34). Resistance to beta-lactam antibiotics like penicillin and other fungal-produced antibiotics were found, as the *B. terrae* BS001 genome contains eight different beta-lactamase genes, *i.e.* BL, BLc and BLI scattered over the genome. Along with this, it was found to contain four predicted genes, *i.e.* parC, parE, gyrA and gyrB, that confer resistance against fluoroquinolones. There are also around 30 predicted genes that are tentatively involved in multidrug resistance (MDR) efflux systems. In this efflux system, an inner membrane component, a membrane fusion protein and an outer membrane protein are thought to work together to eliminate undesirable chemical compounds from the cell into the outside environment. This provides resistance to the cell against putative chemical pressures from its host. The draft genome revealed the presence of eight intact clusters of these tripartite gene systems, while another two are divided among different contigs of the genome.

**Discussion**

To be successful in soil and the mycosphere, a bacterium ideally contains the functionality that allows it to survive in these highly challenging environments. In particular in the mycosphere, the selective pressures provided by the host fungus, both positive and negative ones, need to be dealt with. In our previous work with *B. terrae* strain BS001 in interactions with the soil saprotrophic fungus *Lyophyllum sp.* strain Karsten, we found that *B. terrae* BS001 exhibits a range of different ecological characteristics that allows it to live with this fungus. Thus, *B. terrae* BS001 cells may move with/towards growing fungal hyphae but not with older ones (Warmink and van Elsas, 2009). This indicates the organism is capable of performing chemotaxis towards/with fungal hyphae. Indeed, it was shown to migrate along the mycelial network and also help other bacteria to do so (Warmink et al., 2011). In its interaction with the host, the organism also forms biofilms around the hyphal surfaces (Warmink and van Elsas, 2009). Furthermore, strain BS001 was found to induce exudation and inhibit mushroom formation in *L. sp.* strain Karsten (chapter 4), thus obtaining glycerol and possibly other nutrients from the fungal host.
B. terrae BS001 also helps to protect fungal mycelium against antifungal agents present in the environment (chapter 5).

Given the foregoing remarkable behavioural features, we expected a plethora of interactive and nutrient acquisition functions to be encoded by the B. terrae BS001 genome. In the draft genome, we could detect 67.32% of function-encoding sequence, which is very close to the values found in the genomes of B. phymatum and B. rhizoxinica, i.e. 67.72% and 65.23%, respectively (Fig. 3). The remaining 32.68%, which was found to encode hypothetical genes, should be further investigated to come closer to the actual potential functionality of B. terrae BS001. Moreover, other Burkholderia types, like B. xenovorans LB400, have been found to devote up to 77.43% of their genome to functional genes.

In the light of the aforementioned versatility of functions carried by B. terrae strain BS001, it came as no surprise that this bacterium carries such a huge (about 11.5 Mb) genome. The current study affirmed previous estimates of the genome size of B. terrae BS001. Next to the expected functions that make part of the core genome, a plethora of functions was found that presumably allow strain BS001 to survive in the complexity of environments offered by the bulk soil (absence of fungal tissue) and the mycosphere (presence of emerging fungal tissue). In bulk soil, local conditions may fluctuate on a daily or hourly basis, requiring a bacterium to be constantly “on the watch” for danger or resources. For a successful association with emerging fungi, a sequence of events may be depicted, i.e. signaling and recognition of growing fungal hyphae, sensing of nutritional sources in the vicinity, migration towards the source of nutrition (fungus) and then colonization of the surface provided by the fungus. The draft genome of B. terrae BS001 appears to exhibit a plethora of molecular systems that may be involved in these phenomena. Furthermore, bacterial colonization of the hyphal network may ensue, given the biofilm formation and activity of gene systems like the T3SS, T2SS, T4SS, T6SS and T4 pili. Interestingly, all these highly evolved molecular tools are present in B. terrae BS001. We found some T3S effector proteins as well in the draft genome of B. terrae BS001 and these secreted proteins may be important in host modification. Furthermore, the draft genome of B. terrae BS001 has a full suite of predicted genes.
for the utilization of fungal-released carbonaceous compounds like glycerol, trehalose and also the fungal chitin made available by dying fungal hyphae. *B. terrae BS001* further devotes a major portion of its genome to carbohydrate metabolism (numerous different compounds), which indicates its physiological versatility and ecological potential in a range of environments. Furthermore, the presence of antibiotic biosynthesis as well as resistance genes reveals the capability of this bacterium to tolerate the harsh conditions provided by the growing fungus as well as to build fences. The bacterial genes involved in antibiosis are thought to explain the ecological success that *B. terrae BS001* attains in the mycosphere where other bacteria without such systems are unable to survive or at least to be selected in this particular environment.

Our current analysis of the draft genome of *B. terrae BS001* has already provided good insight into the molecular strategies and mechanisms that fungal-interactive bacteria may use to be ecologically successful in soil and mycosphere. However, more work needs to be done in this regard. A detailed analysis of the current genome compared to other (smaller) *Burkholderia* genomes is likely to teach us more about how *B. terrae* strain BS001 evolved into the behavioral type it currently exhibits. Furthermore, deep analysis of gene clusters that are poorly homologous with *Burkholderia* spp. but show high homology with other bacteria might indicate horizontal transfer of the genomic material, pinpointing strategies/mechanisms that caused these phenomena to happen over evolutionary time in soil and/or the mycosphere.

**Acknowledgements**

We gratefully acknowledge the help of Pilar Puentes and Anna Kielak in different phases of this work. We also thank the HEC, Pakistan which financially assisted RN. RN was further supported by the Soil Biotechnology Foundation.
Chapter 6

References


Nazir,R., Zhang,M., de Boer,W., van Elsas,J.D., 2012. The capacity to comigrate with *Lyophyllum* sp. strain Karsten through different soils is spread among several phylogenetetic groups within the genus *Burkholderia*. *Soil Biol Biochem* 50: 221-233.


