Identification of potential CepR regulated genes using a cep box motif-based search of the *Burkholderia cenocepacia* genome

Catherine E Chambers, Erika I Lutter, Michelle B Visser, Peggy PY Law and Pamela A Sokol*

Address: Department of Microbiology and Infectious Diseases, University of Calgary Health Sciences Center, Calgary, Alberta, Canada

Email: Catherine E Chambers - psokol@ucalgary.ca; Erika I Lutter - elutter@ucalgary.ca; Michelle B Visser - m.visser@uq.edu.au; Peggy PY Law - peggy_py@hotmail.com; Pamela A Sokol* - psokol@ucalgary.ca

* Corresponding author

Abstract

**Background:** The *Burkholderia cenocepacia* CepIR quorum sensing system has been shown to positively and negatively regulate genes involved in siderophore production, protease expression, motility, biofilm formation and virulence. In this study, two approaches were used to identify genes regulated by the CepIR quorum sensing system. Transposon mutagenesis was used to create lacZ promoter fusions in a cepI mutant that were screened for differential expression in the presence of N-acylhomoserine lactones. A bioinformatics approach was used to screen the *B. cenocepacia* J2315 genome for CepR binding site motifs.

**Results:** Four positively regulated and two negatively regulated genes were identified by transposon mutagenesis including genes potentially involved in iron transport and virulence. The promoter regions of selected CepR regulated genes and site directed mutagenesis of the cepI promoter were used to predict a consensus cep box sequence for CepR binding. The first-generation consensus sequence for the cep box was used to identify putative cep boxes in the genome sequence. Eight potential CepR regulated genes were chosen and the expression of their promoters analyzed. Six of the eight were shown to be regulated by CepR. A second generation motif was created from the promoters of these six genes in combination with the promoters of cepI, zmpA, and two of the CepR regulated genes identified by transposon mutagenesis. A search of the *B. cenocepacia* J2315 genome with the new motif identified 55 cep boxes in 65 promoter regions that may be regulated by CepR.

**Conclusion:** Using transposon mutagenesis and bioinformatics expression of twelve new genes have been determined to be regulated by the CepIR quorum sensing system. A cep box consensus sequence has been developed based on the predicted cep boxes of ten CepR regulated genes. This consensus cep box has led to the identification of over 50 new genes potentially regulated by the CepIR quorum sensing system.
Background

Burkholderia cenocepacia, belongs to a group of nine related species with common phenotypes, but distinct genotypes collectively named the "Burkholderia cepacia complex" (Bcc) [1,2]. The Bcc are opportunistic pathogens in immunocompromised and cystic fibrosis (CF) patients but have also been isolated from plant rhizospheres as well as urban and suburban soils [1-3].

The ability of bacteria to adapt to diverse environments is dependent on the coordinate regulation of factors required to survive and proliferate in each niche. The CepR quorum sensing system is one regulatory network that contributes to the response of B. cenocepacia to environmental signals (reviewed in [4,5]). Quorum sensing allows bacterial populations to coordinate gene expression in response to population density. CepR belongs to a group of more than 50 quorum sensing systems that are homologous to the LuxR system of Vibrio fischeri [6,7]. LuxI homologs are N-acyl homoserine lactone (AHL) synthases that generate AHL signal molecules that are released into the environment. LuxR homologs are transcriptional regulators that complex with AHL and typically bind to a lux-box overlapping the -35 sequence of a promoter to regulate transcription. The lux-box consensus sequence recognized by LuxR homologs typically consists of an inverted repeat with significant consensus among quorum sensing systems [6,8-10].

The CepIR system was originally identified in B. cenocepacia (formerly B. cepacia) K56-2 [11] and has subsequently been shown to be widely distributed throughout the Bcc [12,13]. CepI directs the synthesis of N-octanoyl homoserine lactone (OHL) and N-hexanoyl homoserine lactone (HHL) and cepR encodes for the transcriptional regulator [11,14]. CepR has been shown to negatively regulate its own expression, but positively regulate genes positively regulated by the CepIR quorum sensing system and in at least some systems a functional CepR is required for ornibactin biosynthesis [14], the zmpA and zmpB extracellular metalloproteases [15,16], the aidA gene involved in virulence in Caenorhabditis elegans [17-20], swarming motility and in at least some systems a functional CepIR quorum sensing system is necessary for biofilm formation [21-23]. The CepIR system has been shown to contribute to virulence in both plant and animal models. In B. cepacia ATCC 25416 mutations in cepI and cepR attenuated maceration in the onion-rot model [24]. The contribution of CepIR to the severity of B. cepacia infections has been demonstrated in two different animal models, a chronic respiratory infection model in rats and an acute intranasal infection model in Cfrt(-/-) mice [16]. CepIR have also been shown to be important for virulence in C. elegans [25].

Proteomics and promoter based approaches have been used to identify genes regulated by the CepIR quorum sensing system. Proteome analysis was used to compare the protein profiles of B. cepacia strain H111 and an H111 cepI mutant [19]. Differences in expression were observed for 55 out of 585 proteins and partial N-terminal amino acid sequences were determined for peptide fragments of 11 proteins including AidA, FimA, and SodB. A promoter trap approach was used to identify positively regulated OHL-CepR dependent promoters in B. cepacia ATCC 25416 [17]. A library of ATCC 25416 fragments cloned upstream of a promoterless lacZ gene in a vector that also contained cepR was screened in E. coli in the presence and absence of OHL. Twenty-eight clones with genes upregulated in the presence of OHL were identified. The genes belonged to several functional classes; however, the only overlap in genes identified between the two studies was aidA [17,19]. Mutagenesis with a transposon containing a promoterless lacZ reporter was used to identify seven genes positively regulated by the cepIR quorum sensing system in B. cenocepacia strain K56-2, including cepI and aidA [20].

Identification of genes directly and indirectly regulated by CepR is a key step to understanding this regulatory system and the regulatory hierarchies that mediate the adaptation B. cenocepacia to the diverse environments it encounters. The above approaches search for genes regulated under defined in vitro conditions and therefore may not identify genes induced only in specific environmental niches including the plant or animal host. Only the study by Aguilar et al. [17] attempted to identify genes that are regulated by the direct interaction of CepR at the promoter.

LuxR homologs have been shown to bind to specific sequences referred to as lux boxes or the boxes for the gene designation of the respective luxR homolog such as tra boxes in the case of recognition sequences for Agrobacterium tumefaciens TraR [26-28]. These sequences have dyad symmetries and generally overlap the -35 RNA polymerase binding site. Lewenza et al. demonstrated that CepR was required for the expression of cepI in B. cepacia ATCC [11,14] and identified a putative lux-box like sequence with imperfect repeats that overlapped the -35 region of the putative cepI promoter [11]. Weingt et al. [20] demonstrated that CepR directly bound to a DNA fragment that contained the cepI promoter using electrophoretic mobility shift assays. They also mapped the transcriptional start site of cepI and using DNase I footprinting experiments localized the CepR binding site to a region that closely corresponded to the cep box predicted by Lewenza et al.[11]. In the present study, we used a functional genomics approach to identify genes in the B. cenocepacia J2315 sequence with a cep box-like sequence in their promoters. We confirmed by site-directed mutagen-
esis the cep box sequence located upstream of the cepI gene that is necessary for cepI transcription. Using selected B. cenocepacia CepR regulated genes we predicted a consensus cep box motif sequence and used that motif to search the B. cenocepacia J2315 genome to identify promoters potentially regulated by CepR.

Results

Functional analysis of the CepR binding site

Lewenza et al. identified a potential cep box sequence upstream of cepI [11]. Weingart et al. demonstrated using DNAasel footprinting of the cepI promoter that CepR protected a region of DNA that corresponded to the predicted cep box [20]. To confirm that the cep box is required for cepI transcription, eleven mutations, each with a 4 bp substitution, were introduced into the region -59 bp to -18 bp from the transcriptional start site of cepI (Fig. 1A). BamHI-Xhol fragments containing the mutations were subcloned into pMS402 directly upstream of the promotorless luxCDABE operon [29]. The luxCDABE fusions (pCPI302 to pCPI313) were introduced into K56-2 and expression determined by measuring luminescence (Fig. 1B). The K56-2 cepI::luxCDABE fusions with mutations within the 24 bp inverted repeat (pCPI304-310) had luminescence levels below 20% of the wild type K56-2 (pCPI301), whereas promoter fusions containing mutations flanking the inverted repeat (pCPI303, and pCPI311-314) expressed at levels either similar to or higher than wild-type.

Identification of CepR regulated genes by transposon mutagenesis

Nine Tn5-OT182 transposon insertion mutants in K56-12 were identified with differences in β-galactosidase activity on TSA-DC agar with AHL extract and TSA-DC agar without AHL extract. Expression of β-galactosidase activity was increased in the presence of OHL in six mutants and, expression was decreased in three mutants. To locate the Tn5-OT182 insertions in these mutants, the flanking genomic DNA was cloned, sequenced and the sequence obtained was used to search the B. cenocepacia J2315 genome with BlastN to identify the gene containing the insertion (Table 1). A total of 7 distinct genes in 5 regions of the genome were identified. K56-12-P12, K56-12-2PB2 and K56-12-P9 had three distinct insertions within a few hundred base pairs of each other. The P12 transposon inserted into a hemin specific ATPase similar to the phuV gene of Pseudomonas aeruginosa involved in heme iron acquisition [30]. The phuV homolog was predicted to be in an operon with phuk and phusSTUVE homologs of P. aeruginosa. The phuR gene has been shown to be positively regulated by quorum sensing in P. aeruginosa [31]. The insertion in K56-12-2PB2 transposon was also located in phuV; however, in this case the lacZ fusion was in the opposite orientation to that of the gene. K56-12-P9 had an insertion in a hypothetical protein which appears to be in an operon and located directly downstream of a pbp1 homolog. K56-2-P1 and P2 were sibling insertions within a predicted acyltransferase that may be involved in lipid metabolism (COG1835). Directly upstream of the acyltransferase is a class D β-lactamase, likely an oxacillin hydrolase. The insertion in K56-12-P3 was located in a gene, subsequently designated scpB, which belongs to the serine-carboxyl proteinase family [32]. K56-12-P5 and K56-12-P10 contained insertions located in aidA, which was also identified in the transposon mutagenesis screen used by Weingart et al. [20]. K56-12-NB12 contained an insertion in an ORF that has a conserved domain (COG4774) shared with several outer membrane receptors involved in uptake of catechol siderophores, although the other genes flanking this insertion do not appear to be involved in iron acquisition. The insertion in K56-2-2PB2 did not appear to be in a gene. This insertion may result in creation of an artificial promoter-lacZ fusion or influence expression of a regulatory RNA.

To confirm the observations in the plate assay, expression of the unique AHL responsive lacZ fusions was examined over a 24 hr time course in the presence and absence of OHL extract. The growth rates for each mutant were similar to the parent strain K56-12 (Fig. 2A), indicating that the insertions did not result in growth defects that might influence lacZ expression. Expression of the Tn5-OT182 fusions in K56-12-P1 (Fig. 2C) and K56-12-P10 (Fig. 2D) were similar to that observed for a cepI-lacZ fusion (Fig. 2B). There was little expression in the absence of OHL and expression increased in the presence of OHL. The expression of the K56-12-P12 fusion was also increased by the presence of OHL in the culture medium but expression started slightly earlier in growth and decreased after 10 hr (Fig. 2E). Three of the insertions appear to be negatively regulated by cepR since their expression was higher in the absence of OHL and decreased markedly when OHL was added to the culture medium (Fig. 2FGH). Positive regulation of β-galactosidase activity was observed for the K56-12-P3 insertion in the presence of AHL on the plate assay; however, this fusion expressed very poorly in liquid medium (data not shown). When K56-12-P3 grown on agar plates was analyzed for β-galactosidase activity, expression was significantly higher in cultures from plates supplemented with AHL (data not shown).

The predicted promoter regions for the three positively regulated genes containing the Tn5-OT182 insertions, phuV, aidA and the acyltransferase, were cloned into pMS402 and expression of the resulting promoter-luxCDABE fusions was determined in K56-2, K56-R2 (cepR) and K56-d12 (cepI) with and without OHL in the medium. The aidA promoter fusion, pAiD301, had an expression pattern similar to the cepI promoter with significant activity...
Functional analysis of the CepR binding site. A. Site directed mutagenesis was used to determine the effects of mutations on the luminescence activity of a cepI::luxCDABE fusion. The sequence upstream of the CepI ORF is shown. The ATG start codon is indicated by bold lettering and the predicted -10 hexamer is underlined. A series of 4 bp substitutions used to mutate the promoter region are indicated as bx303-313 and the cep box consensus sequence is enclosed in the rectangle. B. Expression of the cepI::luxCDABE fusions in B. cenocepacia K56-2. Luminescence (CPM) was measured at 22 hours and is represented as CPM/O.D. The numbers on the x axis indicate K56-2 (pCPI303-313) respectively. WT is K56-2 (pCPI301) and the vector control is K56-2 (pMS402).
Table 1: OHL responsive genes identified by Tn5-OT182 mutagenesis of K56-12

| Transposon Mutant | Orf\(^a\) | Predicted start codon\(^b\) | Location of insert\(^b\) | Gene/domain homology\(^c\) | OHL effect on expression\(^d\) |
|------------------|----------|--------------------------|-------------------------|---------------------------|-----------------------------|
| K56-12-P1, K56-12-P2 | BCAM0392 | 2:445357 | 2:444971 | COG1835: Predicted acyltransferases | + |
| K56-12-P3 | BCAM0957 | 2:1062298 | 2:1060868 | scpB: serine-carboxyl protease precursor | + |
| K56-12-P5, K56-12-P10 | BCA05293 | 3:328037 | 3:328810 | aidA | + |
| K56-12-P9 | BCAM2631 | 2:2901279 | 2:2900753 | COG2860: predicted membrane protein | - |
| K56-12-P12 | BCAM2630 | 2:2977974 | 2:2980345 | phuV: hemin specific ATP-binding protein | + |
| K56-12-2PB2 | no gene | 2:2980336 | 2:2980336 | - |
| K56-12-NB12 | BCAM1187 | 2:1298085 | 2:1297891 | COG4774: Outer membrane receptor | - |

\(\text{a}\) Open reading frame designation from the unpublished annotation of the \textit{B. cenocepacia} J2315 genome.

\(\text{b}\) Locations reported as chromosome:nucleotide.

\(\text{c}\) Gene and domain homologies determined using BLASTP.

\(\text{d}\) Effect of OHL on expression of the \emph{lacZ} fusion created by insertion of the transposon. +, positive regulation or greater expression in the presence of OHL; -, negative regulation or lower expression in the presence of OHL.

in K56-dI2 only when OHL was added to the medium (Fig. 3A and 3B). This expression pattern was similar to the chromosomal Tn5-OT182 \emph{lacZ} fusion. Expression of the acyltransferase was increased in K56-dI2 in the presence of OHL; however, expression of this fusion in K56-R2 was intermediate between that in K56-dI2 and the parent strain (Fig. 3C). The \emph{phuV} homolog was predicted to be in an operon with the promoter upstream of a \emph{phuR} homolog and therefore the \emph{phuR} promoter was cloned into pMS402. Expression of the \emph{phuR} promoter was similar to K56-2 until early stationary phase where expression was significantly lower in K56-R2 and K56-dI2 in the absence of OHL (Fig. 3D). Expression of \emph{phuR::luxCDABE} was slightly enhanced in the presence of OHL in stationary phase. The pattern of expression of the \emph{phuR::luxCDABE} was similar to that of the \emph{phuV::lacZ} chromosomal fusion (compare Fig. 2E and Fig. 3D). Expression of the \emph{scpB} promoter was very weak in both the presence and absence of OHL suggesting different growth conditions are required for \emph{scpB} expression (data not shown).

To determine if the putative \textit{cep} box sequences identified were potentially involved in CepR regulation of downstream genes, eight of the promoter regions identified that were located within 40–250 bp upstream of a predicted ORF were cloned into pMS402 and expression of the resulting \emph{luxCDABE} fusions was compared in K56-2, K56-dI2 and K56-R2. The three matching motifs with the lowest E-values and five arbitrarily selected motif matches were selected for analysis. When the motifs were located between two putative divergent promoters, one promoter region was chosen for further analysis. The predicted promoters containing putative \textit{cep} box motifs were located upstream of the following orfs: BCAL0340, a TPR repeat protein (MST005); BCAL0715, a LysR-type transcriptional regulator (MST011); BCAL1354, a conserved hypothetical protein (MST028); BCAL2739, \emph{fusA} (MST052); BCAL3191, \emph{caiA} (MST059); BCAM0009, a transcriptional regulator (MST068); BCAM077, hydroxylase (MST072); and BCAM1943, a transcriptional regulator (MST112). The \emph{luxCDABE} fusions containing the MST005, MST011, MST028, MST059 and MST072 sequences had expression patterns similar to \emph{cep} in that expression was higher in K56-2 than in K56-dI2 or K56-R2 and expression was increased in K56-dI2 in the presence of OHL (Fig. 4A,4B,4C,4E and 4G), although expression varied for some fusions depending on the stage of growth. For example, expression of the MST028 fusion peaked at 6 hours and decreased over the remainder of the assay (Fig. 4C). Expression of MST068 was only decreased in K56-R2 in stationary phase although expression was lower in K56-dI2 than in K56-2 and expression in K56-dI2 increased
Figure 2
Effect of OHL on \(\beta\)-galactosidase activity in K56-I2 Tn5-OT182 mutants. A: Growth curves for strains shown in panels B-H. (●) CLW101; (▲) K56-I2-P1; (▼) K56-I2-P3; (□) K56-I2-P9; (○) K56-I2-P10; (■) K56-I2-P12; (◇) K56-I2-NB12; and (*) K56-I2-2PB2. Panels B-H: \(\beta\)-galactosidase activity with (■) and without (□) OHL. Fifty \(\mu\)l of OHL obtained from extracts from a 50 ml culture purified by FPLC and resuspended in 1 ml were added to 10 ml broth. This volume of OHL was shown to restore cepI expression to maximum levels. B: CLW101, C: K56-I2-P1, D: K56-I2-P10, E: K56-I2-P12, F: K56-I2-NB12, G: K56-I2-P9 and H: K56-I2-2PB2.
when the medium was supplemented with OHL (Fig. 4F). MST112, did not appear to be affected by the cepR mutation although expression was lower in K56-dI2 without OHL (Fig. 4H). MST052 did not demonstrate any regulation by CepR in the conditions examined (Fig. 4D).

Construction of the second generation cep box motif and search of the B. cenocepacia genome for potential cep boxes

To improve the specificity of the cep box motif the six promoters with cep box motifs identified by the MAST program with expression patterns similar to that expected for cepIR regulated genes (MST005, MST011, MST028, MST059, MST068 and MST072) were used with the promoters for cepI, phuR, aidA and zmpA to generate a second generation cep box consensus motif using MEME (Table 2). The promoters for scpB, the acyltransferase, MST052 and MST112 did not share the same expression pattern, and therefore were not included. The resulting second generation cep box had the same sequence as the original motif; however the specific score for each position had changed (Fig. 5). The most conserved residues in the second generation motif were in positions 6 (A), 8 (A), 10 (T), 16 (G) and 18 (T).

The new PSSM file was used to search the B. cenocepacia J2315 genome, resulting in 72 sequences matching the motif. Fifty-five of these matches (76%) were potentially

Figure 3
Expression of promoter::luxCDABE fusions for OHL responsive genes identified by K56-dI2 Tn5-OT182 mutagenesis. The promoter fusions in pMS402 were introduced into strains K56-2 (●), K56-dI2 with no OHL (▼), K56-dI2 with 25 nM OHL (▲) and K56-R2 (■). Strains were grown in triplicate in 96 well microtitre plates for 24 hours. Luminescence and optical density were measured at various timepoints and the activity of the promoter was calculated as CPM/O.D. A. pCPI301 (cepI), B. pAID301 (aidA), C. pAYL301(acyltransferase), and D. pHMV301(phuR).
within a promoter region although it must be noted that the transcriptional start sites of these genes have not been experimentally determined. The genes or operons predicted to be downstream of these matching sequences are listed in Table 3. Both MST designsations are included in Table 3 for the six first generation MSTs used in the second generation motif search. Several of the cep boxes identified in the second search had more significant E-values than at least one of the input sequences (data not shown). A cep box was identified upstream of cepR (MST2058), using the second motif. This was the only gene previously determined to be regulated by CepR identified. MST112, which was identified with the first motif, but did not appear to be CepR regulated (Fig. 4H), was not identified with the second motif. Potential cep box sequences were identified on all three chromosomes and the plasmid, suggesting that CepR regulated genes are distributed throughout the genome. Genes downstream of promoters containing cep boxes were classified into seven categories: cell surface or membrane protein genes, genes encoding hypothetical proteins, phage genes, regulatory genes, genes involved in secretion or transport, and genes encoding proteins of unknown function (Table 3). In ten cases the putative cep boxes were located between predicted divergent promoters. In these situations orfs located both downstream and upstream of the cep box are included in Table 3 since it would be possible that CepR regulates genes in both directions. An alignment of the putative cep boxes for each of the MST sequences listed in Table 3 is shown in Fig. 6. The most conserved residues are in position six (A), eight (A), ten (T), sixteen (G) and eighteen (T) which correlates with the motif used in the MEME input file. Further studies are needed to determine if the genes downstream of these predicted promoters and cep box motifs are regulated by CepR.

**Discussion**

In this study we used a computational genome screen and experimental approaches to identify cepR regulated genes in *B. cenocepacia*. Transposon mutagenesis was used to
Figure 4
Expression of promoter::luxCDABE fusions identified in the first cep box motif screen. The promoter fusions in pMS402 were introduced into strains K56-2 (○), K56-dI2 with no OHL (▼), K56-dI2 with 25 nM OHL (▲) and K56-R2 (■). A. MST005, B. MST011, C. MST028, D. MST052, E. MST059, F. MST068, G. MST072, H. MST112. Strains were grown in triplicate in 96 well microtitre plates for 24 hours. Luminescence and optical density were measured at various timepoints and the activity of the promoter was calculated as CPM/O.D.
identify OHL responsive genes in an approach similar to that described by Weingart et al[20]. Since we had previously determined that genes involved in production of the siderophore ornibactin were cepIR regulated [14], we performed our screen in low iron medium in an attempt to identify other iron regulated genes that were responsive to OHL. We also had previously determined that cepR could both positively and negatively regulate gene expression, and therefore, the transposon library was screened for insertion mutants in which β-galactosidase activity was either turned on or off in the presence of exogenous AHLs. Four unique positively regulated and three negatively regulated lacZ fusions were identified. We identified two genes potentially involved in iron transport, a putative outer membrane receptor (BCAM1187) and phuV, a hemin ATP binding protein (BCAM2630). Interestingly, expression of the outer membrane receptor gene was negatively influenced by OHL, whereas phuV expression was positively influenced.

In a screen of approximately 25,000 transposon mutants we only identified six loci with AHL responsive genes. The screening assay was dependent on the visual identification of colonies that were either blue or white in the presence or absence of AHL on medium with X-gal. Although we were able to detect as little as two-fold differences in expression with this assay, we would not detect gene fusions expressed in both the presence and the absence of AHL since we did not attempt to identify mutants with varying shades of blue. For example, although CepR positively regulates zmpA, the CepIR system is not required for its expression since zmpA is expressed at low levels in the absence of AHL and in cepI or cepR mutants [16]. The lacZ fusions in the positively regulated genes identified with transposon insertions were only expressed at significant levels in the presence of OHL. The three negatively regulated fusions had very low expression in the presence of OHL (Fig. 2). It was surprising that we did not identify cepI since CepR tightly regulates cepI expression [14] and cepI

Figure 5
Graphical representation of the cep box consensus sequence. Nucleotide sequence logos are derived from the sequences in Table 2. The relative sizes of the letters indicate their likelihood of occurring at a particular position. The upper logo is based on the six sequences used in the first generation consensus search and the lower logo is based on the ten sequences used in the second generation consensus search.
Table 3: *B. cenocepacia* J2315 genes identified using the second generation cep box motif

| Motif name  | Position | bp | gene | Gene/domain and predicted function |
|-------------|----------|----|------|-------------------------------------|
| **Cell Surface or Membrane** |

MST2008 (-) 1:806161 45 BCAL0738 (-) COG0793: Periplasmic protease; cell envelope biogenesis

MST2009 (+) 1:901874 295 BCAL0831 (+) phpA: phasin-like protein

MST2031 (-) 1:2662911 104 BCAL2406 (-) COG0859, rfaL, LPS heptosyltransferase (rfaL, rfaG; LPS biosynthesis genes)

MST2048 (-) 2:211218 106 BCAM0183 (+) COG3468, autotransporter type V secretion, shdA homolog: adhesin

MST2050 (+) 2:1129604 172 BCAM1015 (-) COG3203: Outer membrane protein

MST2068 (-) 3:174253 153 BCAS0156 (+) COG1680: β-lactamase class C

| **Hypothetical Protein** |

MST2014 (-) 1:1228119 131 BCAL1124 (+) Conserved hypothetical protein

MST2020 (+) MST028f 1:148174 140 BCAL1354 (-) COG1404: conserved hypothetical protein (vgrG related protein)

MST2030 (-) 1:2567308 41 BCAL2313 (+) hypothetical protein

MST2052 (+) 2:1249946 118 BCAM1149 (+) hypothetical protein

MST2056 (-) 2:1667312 57 BCAM1502 (+) hypothetical protein (Chemoreceptor mcpA)

MST2063 (-) 2:2720454 -19 BCAM2417 (+) hypothetical protein

MST2067 (+) 2:3070180 254 BCAM2713 (-) hypothetical protein

MST2071 (+) 3:836110 63 BCAS0753 (+) hypothetical protein

| **Metabolism** |

MST2002 (-) 1:273243 21 BCAL0232 (+) Elongation factor Tu

MST2005 (+) 1:390962 47 BCAL0358 (-) COG0308: Aminopeptidase N

MST2010 (-) 1:778996 101 BCAL0716 (+) COG1250: fadB, 3-hydroxyacyl-CoA dehydrogenase; lipid metabolism

MST2022 (-) 1:1602043 50 BCAL1448 (+) hypothetical protein

MST2027 (+) 1:1626201 104 BCAL1468 (-) COG0644: fvaC, electron transfer flavoprotein-ubiquinone oxidoreductase

MST2029 (+) MST052 1:3351536 -15 BCAL3058 (+) COG0480, fusA: Translation elongation factor

MST2034 (+) MST059f 1:348874 117 BCAL3191 (+) COG1960: acyl CoA dehydrogenase

MST2043 (+) 1:3745369 60 BCAL3191 (+) COG0757: orqQ: 3-dehydroquininate dehydratase II

MST2045 (+) 2:11426 137 BCAM0010 (+) ikbB homolog, AKB ligase

MST2046 (+) MST072f 2:84847 55 BCAM0077 (-) COG0654: ubh or mphA, hydroxylase

MST2055 (+) 2:1564008 139 BCAM1405 (-) sobB: Levansucrase (sacC: Levanase precursor)

MST2059 (+) 2:2088113 71 BCAM870 (+) cepA: homoserine lactone synthase

MST2061 (+) 2:2134837 112 BCAM1922 (+) repB: replication protein

MST2064 (±) 2:2839793 44 BCAM2502 (-) COG0757: orqQ: 3-dehydroquininate dehydratase II (aroE: Shikimate 5-dehydrogenase)

MST2065 (-) 2:2938113 48 BCAM2588 (-) menG: putative S-adenosylmethionine:2 demethylmenaquinone methyltransferase
Table 3: B. cenocepacia J2315 genes identified using the second generation cep box motif (Continued)

| Phage genes | | |
|-------------|-------------|-------------|
| MST2024 (+) | 1:1735446 | 71 | BCAL1564 (-) | Hypothetical proteins Mup46, Mup47 and Mup48 [phage tail protein] |
| MST2060 (+) | 2:2096677 | 28 | BCAM1879 (+) | Phage antirepressor |

| Regulatory gene | | |
|-----------------|-------------|-------------|
| MST2006 (-) | 1:616909 | 88 | BCAL0562 (-) | COG2747, flgM: Negative regulator of flagellar synthesis (flgN; Flagellar biosynthesis/type III secretory pathway) |
| MST2007 (+) MST011 | 1:778996 | 59 | BCAL0715 (-) | Phage antirepressor |
| MST2013 (+) | 1:1085981 | 40 | BCAL0999 (+) | COG3073: RseA; Negative regulator of sigma E activity (RseB or MucB, negative regulator for alginate biosynthesis) |
| MST2019 (-) | 1:437591 | 385 | BCAL1318 (+) | COG3073, nasR: Nitrate-and nitrite-responsive positive regulator |
| MST2026 (+) | 1:2016418 | 259 | BCAL1826 (+) | COG3707, gltF: regulator of gltBDF operon, glutamate synthase enzymes |
| MST2036 (-) | 1:3153030 | 18 | BCAL2871 (-) | COG3073, rseA: Negative regulator of sigma E activity (mucB/rseB, mucD) |
| MST2039 (+) MST059 | 1:3488874 | 102 | BCAL3190 (-) | COG1414: Transcriptional regulator, IclR family |
| MST2040 (-) | 1:3502381 | 36 | BCAL3205 (-) | COG1396: hipB homolog, Putative transcription regulator |
| MST2045 (+) MST060 | 2:11142 | 22 | BCAM0009 (-) | COG1396: hipB homolog, Predicted transcription regulator |
| MST2046 (+) MST072 | 2:84847 | 58 | BCAM0076 (-) | COG1846: Transcriptional regulator, MarR family |
| MST2055 (+) | 2:1564008 | 64 | BCAM1406 (+) | COG: phuR, HTH-type transcriptional regulator |
| MST2057 (+) | 2:1959876 | 36 | BCAM1750 (+) | COG1846: Transcriptional regulator, MarR family |
| MST2058 (+) | 2:2087487 | 31 | BCAM1868 (-) | cepR: Transcriptional regulator, LuxR family |
| MST2071 (+) | 3:836110 | 40 | BCAS0752 (-) | COG0583: LysR type Transcriptional regulator |

| Secretion or secreted product | | |
|-----------------------------|-------------|-------------|
| MST2003 (+) | 1:351306 | 25 | BCAL0321 (+) | COG3671: Predicted membrane protein (tatA, tatB, tatC secretion pathway) |
| MST2004 (+) MST005 | 1:366026 | 206 | BCAL0340 (+) | COG0457: TPR repeat, (evpA and evpB, evpC, evpE, evpF, and evpG virulence and possible secretion) |
| MST2070(-) | 3:478440 | 108 | BCAS0509 (+) | zmpA: extracellular zinc metalloprotease |

| Transport | | |
|------------|-------------|-------------|
| MST2001 (-) | 1:61816 | 272 | BCAL0051 (+) | COG0834: ABC-type amino acid transport/signal transduction systems |
| MST2066 (+) MST005 | 2:2974227 | 115 | BCAM2626 (+) | phuR: Haem/Haemoglobin uptake outer membrane receptor precursor (phuS, phuT, phuU phuV) |
| MST2072 (+) | P:55610 | 113 | BCAS053 (-) | COG1638, dctP homolog: TRAP-type C4-dicarboxylate transport system, |

| Unknown | | |
|----------|-------------|-------------|
| MST2004 (+) MST005 | 1:366026 | 181 | BCAL0339(-) | COG3521: Uncharacterized protein conserved in bacteria |
| MST2025 (+) | 1:1979817 | 274 | BCAL1791 (-) | COG2606: Uncharacterized conserved protein |
| MST2047 (-) | 2:169540 | 86 | BCAM1048 (+) | Putative vgr-related protein (pldA: Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthases) |
| MST2051 (-) | 2:1150338 | -37 | BCAM1044(-) | no homolog (COG1536: Flagellar motor switch protein) |
| MST2053 (-) | 2:1467792 | 28 | BCAM1328-329 (+) | Unknown proteins |
| MST2061 (+) | 2:2134837 | 59 | BCAM1921-919 (-) | no homologs |
| MST2069 (+) | 3:329197 | 160 | BCAS0293(-) | aidA, intracellular protein of unknown function involved in nematode virulence; (second aidA) |

*a* MSTs were identified by searching the B. cenocepacia J2315 genome with the position specific scoring matrix (PSSM) from the second generation motif. Only genes with a motif match within a potential promoter and within 300 bp of the predicted start codon are reported.

*b* The location center of the predicted motif is reported as chromosome:nucleotide. (+) or (-) refers to the DNA strand encoding the motif sequence. The motif names in bold were used to generate the PSSM file.

*c* Number of base pairs between the centre of the motif and the predicted translational start site.

*d* Open reading frame number from the unpublished annotation of the B. cenocepacia J2315 genome. (+) and (-) refer to the DNA strand.

*e* Gene and domain homologies were obtained using the standard protein-protein BlastP program as described in the methods. Genes in parentheses are downstream of the first orf following the motif and may be in the same operon.

*f* Confirmed to be CepR regulated by either lux or lacZ transcriptional fusions.
appeared to be a hotspot for transposon insertions in the study by Weingart et al. [20]. The aidA gene which is tightly regulated by cepIR was identified in both transposon screens, as well as the proteomics and promoter trap approaches [17,19].

Lewenza et al [11] identified a putative CepR binding site in the cepI promoter. During the course of this current study it was reported that CepR directly interacted with a cep box that overlapped this region and directly bound to a cep box within the aidA promoter [20]. We demonstrated using site directed mutagenesis of the cep box region that a 24 bp sequence that contained the cep box was required for cepI expression. All cepI::luxCDABE promoter fusions with mutations in the 24 bp cep box had levels of expression less than or equal to 20% of K56-2 (pCPI301). Similar mutations constructed flanking the cep box had either no effect or in one case increased transcription.

The use of bioinformatics to identify CepR regulated genes has several advantages that are complementary to the experimental methods used to search for CepR regulated genes. Procedures such as transposon mutagenesis, promoter libraries, microarray analysis or proteomics are dependent on the transcription and expression levels of the genes and on the conditions used in the study. Furthermore, the genes and proteins identified by these approaches may be regulated directly or indirectly by CepR. The use of a motif in a genome-wide search for CepR regulated genes may identify niche specific genes that may only be expressed in certain conditions. Identification of a cep box motif may also be used to predict whether CepR genes are directly regulated by interaction with CepR at the promoter or indirectly by CepR interaction with a promoter for an intermediate regulatory gene. In fact, 14 of the 55 putative cep boxes identified were in the predicted promoter regions for regulatory genes. We are currently characterizing some of these regulatory genes to confirm that they are cepR regulated and to determine their regulatory properties.

When searching the genome using the first generation cep box motif we identified some sequences that were not identified with the refined motif used in the second screen of the genome (data not shown). It is possible that these genes are regulated by CepR but have less conserved cep box sequences. Of the eight promoter-lux fusions constructed from sequences identified in the first generation search, six were determined to have cepR regulated expression. There was no difference between the expression of the pMST112 in K56-R2 and K56-2; however, luminescence was increased in K56-dI2 in medium with OHL. The MST112 motif was not detected in the second cep box motif, suggesting that this BCAM1943 may not be cepR.

Figure 6
Alignment of the putative cep box sequences. The MST sequences listed are described in Table 3. Bases conserved in at least 70% of the sequences are shown in red and indicated by an upper case letter in the consensus sequence at the bottom of the alignment, and those conserved in at least 50% of the sequences are shown in blue and indicated by a lower case letter in the consensus sequence. Other bases are indicated in black.

|    | 1 | 10 | 18 |
|----|---|----|----|
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
| MST2001 | CACCATGTATACGCTT |
| MST2007 | CACCATGTATACGCTT |
| MST2009 | CACCATGTATACGCTT |
| MST2002 | CACCATGTATACGCTT |
| MST2006 | CACCATGTATACGCTT |
| MST2005 | CACCATGTATACGCTT |
Interestingly, the MEME program identified a motif farther upstream of the *aidA* promoter region identified by Aguilar et al. [17] and Weingart et al. [20]. It is possible that there is more than one CepR binding site upstream of *aidA*. The additional site might contribute to its tight regulation by CepR and dependency on OHL for expression, features that may have resulted in *aidA* being detected in all of the approaches to date to identify CepR regulated genes.

We identified a *cep* box in the *cepR* promoter region that contains all of the most conserved bases. We have previously shown that *cepR* negatively regulates itself [14]. This is the first confirmed negatively regulated gene identified in the motif search.

It is difficult to compare the extent of overlap between the genes identified using the bioinformatics approach to those identified by Aguilar et al. [17] and Weingart et al.

---

### Table 4: Bacterial strains and plasmids used in this study.

| Strain or plasmid | Description and relevant genotype | Source |
|-------------------|-----------------------------------|--------|
| DH5α              | Δ80delZΔM15 (lacZΔM15 (lacZYA-argF) recA1 endA gyrA96 thi-1 hsdR17 supE44 relA1 deoR U169 | Invitrogen |
| SM10              | Mobilizing strain, RP4 trans genes integrated in the chromosome, Km’ | [50] |
| DH10B             | F’ proA Δ(mrr-hsdRMS-mcrBC) Δ80delZΔM15 ΔlacY74 endA recA1 deoR Δ(ara, leu)7697 araD139 galU galK supG rpsL Δ- | Invitrogen |
| JM110 F’          | endA recA1 gyrA96 thi hsdR17 (r c+ m c+) relA1 supE44 Δ(lac-proAB) [F’ traD36 proA8 lacIqZΔM15] | Promega |
| K56-2             | Cystic fibrosis respiratory isolate | [51] |
| K56-2 R2          | cepR::Ts-OT182 derivative of K56-2, Tc’ | [11] |
| K56-12            | cepI::sp derivative of K56-2, Tp’ | [11] |
| K56-d12           | cepI deletion mutant of K56-2 | [35] |
| CLW101            | cepI::Tn5-OT182 derivative of K56-2, Tc’, Tp’ | [20] |
| K56-12 P1         | BCAM03092:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P3         | BCAM0957:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P5         | BCAS0293:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P9         | BCAM2631:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P12        | BCAM2630:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P2PB2      | BCAM1187:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| K56-12 P2PB12     | BCAM1187:: Tn5-OT182 derivative of K56-12, Tc’ | This study |
| Plasmids          | | |
| pCR<sup>®</sup>2.1 TOPO | PCR cloning vector, pUC ori, Ploc, lacZΔM15, Km<sup>®</sup> Ap<sup>®</sup> | Invitrogen |
| pOT182            | pSU102(GM):: Tn5-OT182, Cm<sup>®</sup>, Tc’, Gm<sup>®</sup>, Ap<sup>®</sup> | [40] |
| pALTER<sup>®</sup>-<sup>®</sup>-<sup>®</sup>-<sup>®</sup> | mutagenesis plasmid, Tc’ | Promega |
| pSL5225           | pUCP26 with 1.5 kb Shh-KpnI fragment containing the cepI gene, Tc’ | [11] |
| pCPI101           | pCR<sup>®</sup>2.1 TOPO with a 266 bp fragment containing the cepI promoter, Ap’, Km’ | This study |
| pCPI201           | pALTER<sup>®</sup>-<sup>®</sup>-<sup>®</sup>-<sup>®</sup>-<sup>®</sup> with the BamHI-XhoI fragment from pCPI101, Tc’ | This study |
| pMS402            | Broad host range vector with promoterless luxCDABE operon, Tpr, Kmr | This study |
| pCPI301           | pMS402 with the BamHI-XhoI fragment containing the wild type cepI promoter region from pCPI101, Tpr’, Km’ | This study |
| pCPI303-313       | pMS402 containing the BamHI-XhoI fragments containing the cepI promoter region with the cep box mutations designated 303-313, Tpr’, Km’ | This study |
| pRK2013           | ColE1 Tra (RK2)*, Km’ | [52] |
| pPHU301           | pMS402 containing the phuR promoter region | This study |
| pAYL301           | pMS402 containing the acyltransferase promoter region | This study |
| pSCP301           | pMS402 containing the scpB promoter region | This study |
| pAD1301           | pMS402 containing the aidA promoter region | This study |
| pMST005           | pMS402 containing the MST005 promoter region | This study |
| pMST011           | pMS402 containing the MST011 promoter region | This study |
| pMST028           | pMS402 containing the MST028 promoter region | This study |
| pMST052           | pMS402 containing the MST052 promoter region | This study |
| pMST059           | pMS402 containing the MST059 promoter region | This study |
| pMST068           | pMS402 containing the MST068 promoter region | This study |
| pMST112           | pMS402 containing the MST112 promoter region | This study |
We have identified several new CepR regulated genes that may be further studied to increase our understanding of the CepR regulon.

**Methods**

**Reagents, bacterial strains and culture conditions**

Unless otherwise stated all molecular biology reagents were purchased from Invitrogen Life Technologies (Burlington, Ontario) and all chemicals purchased from Sigma Chemical Co. (St. Louis, Mo.). The strains and plasmids used in this study are listed in Table 4. For genetic manipulations, *B. cenocepacia* and *Escherichia coli* strains were grown at 37°C in Luria-Bertani (LB) broth (Invitrogen) or on 1.5% LB agar plates. Concentration of antibiotics in selective medium for *E. coli* were 100 μg/ml ampicillin, 1.5 mg/ml trimethoprim and 15 μg/ml tetracycline, and for *B. cenocepacia* were 200 μg/ml tetracycline and 100 μg/ml trimethoprim. For transcription assays, *B. cenocepacia* strains were grown in tryptic soy broth (TSB, Difco, Detroit, Mich.) or TSBD-C [37].

**AHL extraction and OHL purification**

AHLs were extracted from culture supernatants of K56-2 as previously described [14]. The extract from 50 ml culture fluid was resuspended in 1 ml distilled water and 20 μl aliquots of this stock solution were spread onto agar plates to screen for mutants in which lacZ expression was altered in the presence of AHL. This quantity of AHL extract was found to restore wild-type protease activity to *B. cenocepacia* K56-12 as indicated by the zones of clearing observed on skim milk plates. OHL was purified from culture supernatants of *B. cenocepacia* K56-12 (pSLS225), a strain that carries the cepI gene in trans as previously described [38].

**Molecular biology and sequence analysis**

DNA manipulations were performed generally as described by Sambrook et al. [39]. T4 DNA ligase was purchased from Promega Corporation (Madison, WI) and New England Biolabs Inc. (Beverly, MA). Custom oligonucleotides were synthesized by Invitrogen Life Technologies. DNA sequencing was performed at the University of Calgary Core DNA Services (Calgary, Canada) using an ABI1371A DNA sequencer or at Macrogen Inc. (Seoul, Korea) on an ABI3730 XL automatic DNA sequencer.

**Transposon mutagenesis**

Mutagenesis of *B. cenocepacia* K56-12 (Tp') with Tn5OT182 was performed as described by Lewenza et al. [11]. Tn5-OT182 is a self-cloning transposon with a promoterless lacZ gene that is transcribed from the promoter of a host gene when it is fused in the direction of transcription [40]. Transposon insertion mutants were picked using a robot (Norgren Systems, Palo Alto, CA) into Becton Dickinson microtest flat bottom polystyrene 96 well micro-
### Table 5: Oligonucleotide primers

| Primer     | Sequence                                      | Restriction Site or size of product (bp) |
|------------|-----------------------------------------------|----------------------------------------|
| **PCR Oligonucleotides**                      |                                              |                                        |
| cepIfor    | CAGGCGCGGATAGCCTTG                           |                                        |
| cepIrev    |CACAGATCCAGGACATCCA                           |                                        |
| EXcepR3    | CCGGAAGGGAAAGGCC                             | BamHI                                  |
| EXcepR2    | CCGGAAGGGAAAGGCC                             | BamHI                                  |
| **Oligonucleotides used to clone promoters**  | position of 5’ base | Size (bp) |
| aidA       | CAGATTCAATGTCGCG                             | 3:329288                                | 272 |
|            | GCACATCGGTAAACGCG                            |                                        |
| scpB       | CTGCAACGAAGCGACCG                            | 2:1062555                               | 294 |
|            | GACGGAAGGGAAAGGCC                            |                                        |
| cepI       | GCCTGACGGGCAACAGCAGCGCTATCATG               | 2:2087932                               | 267 |
|            | GAACGAAAGTCTGATGACTG                        |                                        |
| PBP        | CTGCGGAAACGAGGCC                             | 2:2983704                               | 313 |
|            | CGATGGTTGGCGCTGGAG                         |                                        |
| phuR       | CTTGCGATGATCGCGG                             | 2:2973940                               | 404 |
| acyltransferase | CAGATACGTGACGCGCC |                                        |
|            | ACAGGTGATGCTCC                              | 2:2974344                               |    |
| zmpA       | TCTGACGGGCTGTCGACTG                         | 3:478051                                | 638 |
|            | GAATCCAGACGGGACAGGAC                      |                                        |
| MST005     | GCACGGCCCGCTGCCG                            | 1:366108                                | 325 |
|            | CGCAAGCCACACTACCC                            |                                        |
| MST011     | CTTTCTGCTGATGGTCCGGCG                      | 1:779005                                | 321 |
| MST028     | CTGCTGCTGCCCCGCC                            | 1:148174                                | 451 |
|            | GTCTGCTGCTGCCG                              |                                        |
| MST052     | CCGTCAATTTCGTCGCG                            | 1:3009328                               | 341 |
|            | CAGATCCATGTGCGCG                            |                                        |
| MST059     | GCCTTCGGACCAGGCG                            | 1:3488873                               | 315 |
|            | GCTGCTGCTGCCG                              |                                        |
| MST068     | CTGCGAGCTCAGCTGCTGCG                       | 2:11203                                 | 325 |
|            | GTGCGAGCTGCCGCC                            |                                        |
| MST072     | GCATCCAGCCAGGCGC                            | 2:84846                                 | 398 |
|            | CCGACGGGACACCGCC                            |                                        |
| MST112     | GCAGGTGCCTACGTCCGGCG                       | 2:2156170                               | 441 |
|            | ACCAGGCCTACGTCCGGCG                       |                                        |

### Mutagenic Oligonucleotides

| CepBx103   | GCGTCTTTACGCGTGCACCTGTAAGTACGTTACC          | Ssf |                  |
| CepBx104   | GTCTTTACGCGGTCATATGTAAGGATTACCGAG           | Ndel|                  |
| CepBx105   | CGGCCGTCACCCCTGTGACATAGGTTACGATTACGAG      | PstI|                  |
| CepBx106   | GCCGTCACCCCTGTGACCTGATTACGATTACGAG         | PstI|                  |
| CepBx107F  | ACACGCGTGACCCCTGTGAACTAGTACGTTACGATTACG    | Spel|                  |
| CepBx107R  | GAGGAGCGCTGAACTGAGTTACATTACAGGTCG          |      |                  |
| CepBx108F  | CGCGTACCTGTGTAAGGTTACGTTACGATTACG          | Xbal|                  |
| CepBx108R  | GCAGCGGACGCGCTGTAACGTTACGATTACG            | Xbal|                  |
| CepBx109F  | ACACGCGTGACCCCTGTGAACTAGTACGTTACGATTACG    | Spel|                  |
| CepBx109R  | GAGGAGCGCTGAACTGAGTTACATTACAGGTCG          |      |                  |
| CepBx109R  | GCGGCCGTCACCCCTGTGACATAGGTTACGATTACGAG    | Ssf |                  |
| GCGGCCGTCACCCCTGTGACATAGGTTACGATTACGAG       | Ssf |                  |
| CepBx109R  | GCGGCCGTCACCCCTGTGACATAGGTTACGATTACGAG    | Ssf |                  |
| CepBx109R  | GCGGCCGTCACCCCTGTGACATAGGTTACGATTACGAG    | Ssf |                  |
| CepBx111F  | AGAGGTACCCCTGTGACCTGTTACGTTACGATTACG       | EcoRV|                 |
| CepBx111R  | ATAGACCTGACCCCTGTGACCTGTTACGATTACG         | EcoRV|                 |
| CepBx112F  | GTGACCTGTGACCCCTGTGACCTGTTACGATTACG        | BspVI|                |
| CepBx112R  | GTGACCTGTGACCCCTGTGACCTGTTACGATTACG        | BspVI|                |
| CepBx113F  | CCAGGTACCCCTGTGACCTGTTACGATTACG            | Ssf |                  |
| CepBx113R  | GCGTGCATACACGCGGCGTGCACCTGTTACGATTACG      | Ssf |                  |

*a Locations reported as chromosome:nucleotide*
titer plates containing 200 μl medium per well and grown overnight at 37°C with shaking at 200 rpm. Cultures were stamped onto TSBD-C (200 μg/ml tetracycline, 100 μg/ml trimethoprim and 40 μg/ml X-gal) agar with and without the addition of AHL extract and grown for 48 hours at 37°C. β-galactosidase expression was visually monitored at 24 and 48 hours for differences in blue color. Approximately 25,000 tetracycline and trimethoprim resistant transposon insertion mutants from five independent mutagenesis experiments were screened. Positively regulated insertion mutants appeared blue in the presence of AHL and X-gal and white in the absence of AHL. The reverse is true in the case of negatively regulated genes. Nine mutants exhibiting reproducible differences in AHL dependent β-galactosidase expression were chosen for further characterization. The DNA flanking the Tn5-OT182 insertions was self-cloned from XhoI or EcoRI digests of genomic DNA and sequenced using oligonucleotides OT182-LT and OT182-RT [41].

**Construction of cepI promoter mutations**

The Altered Sites® II in vitro Mutagenesis System (Promega) and the Quick Change® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) were used to create mutations spanning the proposed cepI box in the cepI promoter (Fig. 2). The template used with the Altered Sites® II in vitro Mutagenesis System was created by ligating a BamHI-XhoI fragment containing the cepI promoter region from pCP101 into pALTER-Ex1 (pCP1201). The Altered Sites® II System was used with mutagenic oligonucleotides CepBx103-106 (Table 5). These oligonucleotides were 5'-phosphorylated using T4 DNA Kinase (Promega) and annealed to single stranded DNA prepared according to the manufacturers instructions from cultures of JM109 F' (pCP1201). The remaining mutagenic oligonucleotides were used with plasmid pCP101 and the Stratagene Quick Change® Site-Directed Mutagenesis Kit. Mutagenic oligonucleotides were designed with 4 base pair substitutions that resulted in the introduction of a new restriction enzyme site (Table 5). Mutations were confirmed by restriction enzyme analysis and sequencing. To construct the cepI::luxDCABE fusions, the mutated promoter regions were excised from pCP101 and pCP201 by digestion with BamHI-XhoI and ligated into the BamHI-XhoI site of pMS402.

**In vitro transcription assays**

Putative promoters identified in this study were PCR amplified using the primers listed in Table 5 from K56-2 genomic DNA and cloned into the vector PC2.1*-TOPO. The promoters were excised from the PCR2.1*-TOPO clones using BamHI-XhoI and ligated into pMS402 to create plasmids pCP301, pPHU301, pAYL301, pSCP301 and pAID301, respectively. The eight promoters identified in the first genome search for cep box motifs were cloned using the primers listed in Table 5 for each MST promoter as described above and named pMST005, pMST011, pMST028, pMST052, pMST059, pMST068, pMST072, and pMST112, respectively.

Five ml overnight cultures of K56-2, K56-dl2 and K56-R2 hosting the luxDCABE fusions were grown in TSB supplemented with 100 μg/ml trimethoprim to maintain pMS402. Overnight cultures were diluted with TSB to an A600 of 0.05 and aliquots of 150 μl were placed in wells of 96 well clear bottom plates (Costar, Corning Incorporated, Corning, NY). The plates were covered and incubated at 37°C with shaking and the luminescence and absorbance was measured in a Victor™ multilabel counter at various intervals for 24 hours. Each strain was assayed at least three times in triplicate.

**Bioinformatics**

Nucleotide sequence obtained from DNA flanking the transposon insertions was used with BLASTN to determine the location of the insertion in the unpublished genome sequence of B. cenocepacia J2315 [42], a strain of the same lineage as K56-2. Homologues of open reading frames were predicted using BLASTP [43]. Potential promoter elements were identified using BPMO [44]. The cep box consensus sequence was predicted by analyzing the promoter regions of selected positively regulated genes with the motif discovery tool MEME [45]. The MEME program [46] represents motifs as position-dependent letter-probability matrices which describe the probability of each possible letter at each position in the pattern. The output from the MEME program provides a position-specific scoring matrix (PSSM) for the predicted motif. The PSSM for the predicted cep box consensus sequence was used to search the B. cenocepacia J2315 genome with the motif alignment search tool MAST [45,47]. The cep box motifs identified by MAST were also aligned using Multalin [48,49].

**Authors’ contributions**

CC designed the cep box mutagenesis, performed the analysis of the cep box consensus sequence and screening of the genome, contributed to the promoter fusion expression experiments, and helped draft the manuscript. EL performed the transposon mutagenesis, expression experiments on the mutants, cep box alignments and helped draft the manuscript. MV constructed cep box mutants, analyzed genome sequence data, contributed to promoter fusion assays, and helped draft the manuscript. PL cloned MST promoters and performed lux fusion assays, PS participated in the experimental design and data analysis, coordinated the study and drafted the manuscript. All authors read and approved the final manuscript.
Acknowledgements
This study was supported by a grant from the Canadian Cystic Fibrosis Foundation to PS. The authors thank J. Parkhill and M. Holden at the Welcome Trust Institute for access to the annotation data of B. cenocepacia J2315 genome sequence prior to publication.

References
1. Coenye T, LiPuma JJ: Molecular epidemiology of Burkholderia species. Front Biosci 2003, 8:e55-67.
2. Mahenthiralingam E, Urban TA, Goldberg JB: The multifarious, nontypical Burkholderia cepacia complex. Nat Rev Microbiol 2005, 3(2):144-156.
3. Coenye T, Vandamme P: Diversity and significance of Burkholderia species occupying diverse ecological niches. Environ Microbiol 2003, 5(9):719-729.
4. Eberl L: Quorum sensing in the genus Burkholderia. Int J Med Microbiol 2006, 296(2-3):103-110.
5. Venturi V, Frisicna A, Bertani I, Devescovi G, Aguilar C: Quorum sensing in Burkholderia cepacia complex. Res Microbiol 2006, 157(4-5):244.
6. Fuqua C, Greenberg EP: Listening in on bacteria: acyl-homoserine lactone signalling. Nat Rev Mol Cell Biol 2003, 3(9):685-695.
7. Waters CM, Bassler BL: Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol 2003, 21:319-346.
8. Eberhard A, Eberl L: Quorum sensing in Vibrio Fischeri: elements of the LuxI promoter. Mol Microbiol 1999, 31(4):1197-1204.
9. Fuqua C, Winans SC: Conserved cis-acting promoter elements are required for density-dependent transcription of Agrobacterium tumefaciens conjugal transfer genes. J Bacteriol 1996, 178(2):435-440.
10. Gray KM, Passador L, Iglewski BH, Greenberg EP: Interchangeability and specificity of components from the quorum-sensing regulatory systems of Vibri,Fischeri and Pseudomonas aeruginosa. J Bacteriol 1994, 176(10):3076-3080.
11. Lewenza S, Conway B, Greenberg EP, Sokol PA: Quorum sensing in Burkholderia cepacia: identification of the LuxR homologs CepR1. J Bacteriol 1999, 181(3):748-756.
12. Gotschlich A, Huber B, Geisenberger O, Toyg A, Steidle A, Riedel K, Hill P, Tummler B, Vandamme P, Middleton B, Camara M, Williams P, Hardman A, Eberl L: Synthesis of multiple N-acylhomoserine lactones is widespread among the members of the Burkholderia cepacia complex. Syst Appl Microbiol 2001, 24(1):1-11.
13. Lutter E, Lewenza S, Dennis J Jr, Visser MB, Sokol PA: Distribution of quorum-sensing genes in the Burkholderia cepacia complex. Infect Immun 2001, 69(7):4661-4666.
14. Lewenza S, Sokol PA: Regulation of ornibactin synthesis and N-acyl-L-homoserine lactone production by CepR in Burkholderia cepacia. J Bacteriol 2001, 183(22):2112-2118.
15. Kool C, Subsin B, Chen R, Pohorelic B, Sokol PA: Burkholderia cepacia ZmpA is a broad specificity zinc metalloproteinase involved in virulence. Infect Immun 2006, 74:4083-4093.
16. Sokol PA, Sajan U, Visser MB, Ginges S, Forster J, Kool C: The CepIR quorum-sensing system controls the virulence of Burkholderia cepacia, a respiratory pathogen. Microbiology 2003, 149:3649-3658.
17. Aguilar C, Frisicna A, Devescovi G, Kojic M, Venturi V: Identification of quorum-sensing-regulated genes of Burkholderia cepacia. J Bacteriol 2003, 185(21):6456-6462.
18. Huber B, Feldmann F, Kotho M, Vandamme P, Wupperer J, Riedel K, Eberl L: Identification of a novel virulence factor in Burkholderia cepacia H111 required for efficient slow killing of Caenorhabditis elegans. Infect Immun 2004, 72(12):7220-7230.
19. Peggan K, Arsenievich F, Huber B, Feldmann F, Gorg A, Lottspeich F, Eberl L: Analysis of the quorum-sensing regulon of the opportunistic pathogen Burkholderia cepacia H111 by proteomics. Electrophoresis 2003, 24(4):740-750.
20. Weingart CL, White CE, Liu S, Chai Y, Cho H, Tsai CS, Wei Y, Delay NR, Gronquist MB, Beherard A, Winans SC: Direct binding of the quorum-sensing regulator CepR of Burkholderia cepacia to two target promoters in vitro. Mol Microbiol 2005, 57(2):452-467.
21. Huber B, Riedel K, Hentzer M, Heydorn A, Gotachlich A, Givskov M, Molin S, Eberl L: The cep quorum-sensing system of Burkholderia cepacia H111 controls biofilm formation and swelling motility. Microbiology 2001, 147(Pt 9):2517-2528.
22. Lewenza S, Visser MB, Sokol PA: Interspecies communication between Burkholderia cepacia and Pseudomonas aeruginosa. Can J Microbiol 2002, 48(7):707-716.
23. Tomlin KL, Malott RJ, Ramage G, Storey DG, Sokol PA, Ceri H: Quorum-sensing mutations affect attachment and stability of Burkholderia cepacia biofilms. Appl Environ Microbiol 2005, 71(9):5208-5218.
24. Aguilar C, Bertani I, Venturi V: Quorum-sensing system and stationary-phase sigma factor (rpoS) of the onion pathogen Burkholderia cepacia genomovar I strain, ATCC 25416. Appl Environ Microbiol 2003, 69(3):1739-1747.
25. Kotho M, Aml M, Huber B, Stoecker K, Ebrecht D, Steinmetz I, Eberl L: Killing of Caenorhabditis elegans by Burkholderia cepacia is controlled by the cep quorum-sensing system. Cell Microbiol 2003, 5(5):343-351.
26. Wooster MS, Urbanowski ML, Greenberg EP: Promoter specificity in Pseudomonas aeruginosa quorum sensing revealed by DNA binding of purified LasR. Proc Natl Acad Sci U S A 2004, 101(45):15833-15839.
27. Urbanowski ML, Lostroh CP, Greenberg EP: Reversible acylhomoserine lactone binding by purified Vibrio Fischeri LuxR protein. J Bacteriol 2004, 186(3):631-637.
28. Zhu J, Visser SCA: Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. Proc Natl Acad Sci U S A 1999, 96(13):7482-7487.
29. Duan K, Dammel C, Steil J, Rabin H, Surette MG: Modulation of Pseudomonas aeruginosa gene expression by host microflora through interspecies communication. Mol Microbiol 2003, 50(5):1477-1491.
30. Ochsnier UA, Johnson Z, Yasil ML: Genetics and regulation of two distinct haem-uptake systems, phu and has, in Pseudomonas aeruginosa. Microbiology 2000, 146( Pt 1):185-198.
31. Arevalo-Ferro C, Hentzer M, Reil A, Kjelleberg S, Givskov M, Riedel K, Eberl L: Identification of quorum-sensing regulated proteins in the opportunistic pathogen Pseudomonas aeruginosa by proteomics. Environ Microbiol 2003, 5(12):1350-1369.
32. Oda K, Takahashi T, Tokuda Y, Shibano Y, Takahashi S: Cloning, nucleotide sequence, and expression of an isovaleryl peptatin-insensitive carboxyl proteinase gene from Pseudomonas sp. 101. J Biol Chem 1994, 269(42):26518-26524.
33. Subsin B, Chambers CE, Visser MB, Sokol PA: Identification of genes regulated by the cepIR quorum sensing system in Burkholderia cepacia cenocepacia by high-throughput screening of a random promoter library. J Bacterial 2007, in press.
34. Baldwin A, Sokol PA, Parkhill J, Mahenthiralingam E: The Burkholderia cepacia epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in Burkholderia cepacia. Infect Immun 2004, 72(3):1537-1547.
35. Malott RJ, Baldwin A, Mahenthiralingam E, Sokol PA: Characterization of the cciIR quorum-sensing system in Burkholderia cepacia. Infect Immun 2005, 73(8):4982-4992.
36. Kool C, Corbett CR, Sokol PA: Functional analysis of the Burkholderia cepacia ZmpA metalloprotease. J Bacteriol 2005, 187(13):4421-4429.
37. Ohman DE, Sadowski JC, Iglewski BH: Toxin A-deficient mutants of Pseudomonas aeruginosa PA103: isolation and characterization. Infect Immun 1980, 28(3):899-908.
38. Chambers CE, Visser MB, Schwagel U, Sokol PA: Identification of N-acylhomoserine lactones in mucopurulent respiratory secretions from cystic fibrosis patients. FEMS Microbiol Lett 2005, 242(4):297-304.
39. Fries JF, Frisch EF, Maniatis T: Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor, N.Y.; Cold Spring Harbor Press; 1989.
40. Merriman TR, Lamont IL: Construction and use of a self-cloning promoter probe vector for gram-negative bacteria. Gene 1993, 126(1):17-23.
41. DeShazer D, Brett PJ, Carloyon R, Woods DE: Mutagenesis of Burkholderia pseudomallei with Tn5-OT182: isolation of motility mutants and molecular characterization of the flagellar structural gene. J Bacteriol 1997, 179(7):2116-2125.
42. Welcome Trust Sanger Institute B. cenocepacia sequencing project [http://www.sanger.ac.uk/b_cenocepacia].
43. BLASTP (http://www.ncbi.nlm.nih.gov/BLAST). 
44. Softberry (www.softberry.com). 
45. The MEME/MAST system motif discovery and search (http://meme.sdsc.edu/meme). 
46. Bailey TL, Elkan C: Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 1994, 2:28-36.
47. Bailey TL, Gribskov M: Combining evidence using p-values: application to sequence homology searches. Bioinformatics 1998, 14(1):48-54.
48. Corpet F: Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 1988, 16(22):10881-10890.
49. MultAlin http://prodes.toulouse.inra.fr/multalin/multalin.html.
50. Simon R, Priefer U, Puhler A: A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Bio/Technology 1983, 1:784-791.
51. Mahenthiralingam E, Coenye T, Chung JW, Speert DP, Govan JRW, Taylor P, Vandamme P: Diagnostically and experimentally useful panel of strains from the Burkholderia cepacia complex. J Clin Microbiol 2000, 38(2):910-913.
52. Figuski DH, Helenki DR: Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci USA 1979, 76:1648-1652.