The Characterization of Flavobacteriaceae using Multilocus Sequencing Typing

Presented to the faculty of Lycoming College in partial fulfillment of the requirements for Departmental Honors in Biology

By
KC Failor
Lycoming College
18 February 2011

Approved by:
[Signatures]
Abstract:

Publication of novel species requires a demonstration of genomic uniqueness in addition to the standard 16S rRNA sequencing. DNA-DNA hybridization is often done; however, analysis of protein coding genes may be preferred because it provides higher resolution between members of the same genus, and is archivable in an online database. Conserved amino acid sequences were identified from a multiple sequence alignment of essential proteins among the Flavobacteriaceae and were used to design degenerate oligonucleotide primers. The genes were PCR amplified from several Chryseobacterium, including potentially novel species. The range of similarity among different species within the genus was 82-92 percent for the gyrB, groEL and rpoB genes. This strategy provided clearer resolution between members of the genus as compared to the 16S rRNA sequencing. It also allowed for the submission and storage of the data in the online database GenBank.

Background:

Bacteria inhabit nearly every ecosystem on this planet, from the frigid poles, to the hydrothermal vents on the ocean floor. However, only a small fraction of the existing bacteria are capable of surviving in culture. While many are harmless, there are some that cause disease, while others provide beneficial factors for humans. Many of the first bacteria to be identified and studied extensively were those bacteria that exhibited pathogenicity; subsequently; species related to the pathogenic strains have since been studied as well.

DNA-DNA hybridization (DDH) was previously utilized to determine whether a new isolate was a novel bacterial species, but due to the complexity of performing this test, simpler methods have been developed. One such method was identified by Carl Woese as a possible replacement for DDH: the sequencing of the 16S rRNA gene. Because all microbes contain the small ribosomal subunit (16S), it is possible to use the nucleotide differences within the sequence to determine the species of the organism. A less than 97% identity match for 16S rRNA and less than 70% DDH are used to determine whether organisms are sufficiently divergent to be considered separate species. However, Stackebrandt and Ebers argue that less than 99% 16S
rRNA identity is equally sufficient to determine whether a species is novel. They demonstrated that even the 99% 16S rRNA still corresponded to less than 70% DDH (Stackebrandt and Ebers 2006). 16S rRNA sequencing, however, does not always have sufficient resolution to distinguish among similar species or strains (Erwin et al., 2008).

Multilocus Sequence Typing (MLST) was developed as a means to distinguish different isolates of the same bacterial species (Maiden et al. 1998). The process utilizes fragments of several genes necessary for cell function, known as housekeeping genes (Margos et al. 2008). Because the fragments are usually less than 1kb, they can be accurately sequenced from both ends. The process was initially developed for Neisseria meningitidis, but has been expanded to work in many genera including Yersinia (Chen et al. 2010), Streptococcus (Pullinger et al. 2006), Candida (Lin et al. 2007) and Escherichia (Tartof et al. 2005). Because this system only uses several genes, it is both faster and less expensive than both full genome DDH, and full genome sequencing. MLST can more accurately distinguish closely related species or different strains of the same species because protein-coding housekeeping genes show greater variation than the 16S rRNA gene. Because the small ribosomal subunit requires such specificity in its folding pattern, there cannot be as much variation in its primary structure. The housekeeping genes/proteins however, can tolerate more modifications without compromising their function. This increased variation, even between strains of the same species, allows for more specific identification.

MLST uses genes necessary for the survival of the bacteria, called housekeeping genes. These genes are highly conserved between species, allowing for the comparisons between closely related species, as well as between strains of the same species. This study focused on four main genes: glnA, groEL, gyrB, and rpoB. The 16S rRNA gene was also studied, but was not used for generating the MLST data. Since a MLST study has not been conducted on Chryseobacterium species, a collection of usable housekeeping genes taken from a previous MLST study done by Damian Mariano ’10 were utilized in this project.

The genetic code, comprised of combinations of four nucleotides: adenine, guanine, cytosine and thymine, is a degenerate code. The genetic code has redundancy but not ambiguity;
each codon, or collection of three nucleotides, codes for a single amino acid; however, an amino acid can be coded for by several different codons. Generally, the variation in the codons that code the same amino acid occurs in the third position, except in the cases of arginine, leucine and serine, where the variation occurs in both the first and third positions. The degeneracy results because there are more possible codons than amino acids for which to code. However, this property allows for some tolerance of point mutations, resulting in silent mutations that do not affect the structure or function of the final functional protein.

Glutamine Synthetase (glnA) is the protein responsible for catalyzing the condensation of glutamate and ammonia into glutamine using ATP. ATP phosphorylates glutamate to form ADP and an acyl-phosphate intermediate, γ-glutamyl phosphate, which reacts with ammonia, forming glutamine and inorganic phosphate. ADP and an inorganic phosphate do not dissociate until ammonia binds and glutamine is released. The full glutamine synthetase protein is composed of eight to twelve subunits that form two parallel concentric rings with 12 active sites between each monomer (Eisenberg et al. 2000).

Chaperonin protein GroEL (groEL) is responsible for aiding in the proper folding of proteins. GroEL functions alongside GroES, a second chaperonin protein, to fold proteins using ATP. Unfolded substrate proteins bind to a hydrophobic binding patch on the inner rim of GroEL. A conformational change due to this binding as well as the binding of ATP allows GroES to associate with the complex. Binding of GroES causes individual subunits of the chaperonin to rotate forcing the substrate protein into a hydrophilic chamber. The hydrophilic environment of the chamber causes hydrophobic residues of the substrate to be buried inside the protein, which in turn, induces substrate folding. Hydrolysis of ATP and binding of new substrate proteins to the opposite cavity sends an allosteric signal causing GroES and the encapsulated protein to be released into the cytosol. (Horwich et al. 2007)

DNA gyrase subunit B (gyrB), is the β subunit of a type II topoisomerase responsible for the unwinding of DNA for replication. The protein cleaves a section of the DNA and causes the negative supercoiling of the strand, causing a relaxation, or flattening out, of the DNA. The
protein utilizes ATP both to cut the DNA as well as to unwind the DNA (Gore et al. 2006). In bacteria the protein is also capable of removing any knots present in the circular chromosome. The protein is composed of two monomers with an ATPase domain on one side of the molecule and a DNA gate on the opposite side.

Recombination protein A (recA), also called recombinase A, is a globular DNA-dependent ATPase and an ATP-dependent DNA binding protein. Binding to DNA is a fundamental aspect of the role of RecA protein in DNA metabolism. RecA activity is a multistep process that results in the establishment of a nucleoprotein filament. The first step consists of the nucleation of a RecA monomer onto the DNA; this binding is the slow step in the activity of the recA protein. Nucleation is considerably faster on single-stranded DNA than on double-stranded DNA due to interference with the other DNA strand (Lusetti and Cox 2002).

DNA directed RNA polymerase subunit B (rpoB) is the β subunit of an enzyme which transcribes a DNA template sequence into a strand of RNA. The enzyme binds upstream of the target sequence, switching from a closed complex to an open complex. This shape change involves the separation of the DNA strands allowing for ribonucleotides to be paired with the template DNA. The β subunit has the polymerase activity including chain initiation and elongation. The polymerase enzyme is composed of five subunits (α2ββ′σ), (Opalka et al. 2010).

The Polymerase chain reaction (PCR) is a technique used to generate over one billion copies of DNA from a single template strand. PCR relies on thermal cycling in order to denature the DNA, anneal the primer sequences to the template, and then replicate the strands. The primers are small DNA fragments that have been previously designed to bind to a specific gene or region in the DNA. PCR require the use of a heat stable DNA polymerase, taken from *Thermus aquaticus*, as well as sufficient amounts of deoxynucleotide triphosphates (dNTPs) and both the forward and reverse primers.

Agencourt, a subsidiary of Beckman Coulter Biosciences, utilizes Sanger sequencing to determine the exact sequence of the DNA samples. Sanger sequencing utilizes chain termination to sequence the DNA. Fluorescently labeled dideoxynucleotides (ddNTPs) are added to a DNA
synthesis reaction and terminate each newly synthesized strand. Since each ddNTP is labeled with a different fluorescent dye, when the fragment is passed through a detector, color is recorded, producing a trace chromatogram, which can then be used to determine the DNA sequence.

With sequencing providing more and more data, computers have become an essential part of identifying novel species. Several bioinformatics tools were utilized during this project, including BLAST, CAP3, MEGA4 and the GenBank database. These tools are all available for free online to be utilized in a web browser, or for download as an executable program. BLAST, or Basic Local Alignment Search Tool is a web based program available through the National Center for Biotechnology Information (NCBI). The program compares a query sequence to all available sequences within NCBI’s database, and returns results with the highest degrees of similarity. The tool can also be used for the identification of an unknown gene in this same manner. BLAST2, a variant of BLAST can compare two known sequences to one another in order to show the similarities between the two sequences (Altschul et al. 1990). Another web based program, the Contig Assembly Program CAP3, allowed for the creation of contiguous consensus sequences from multiple sequencing reads. CAP3 also allows for the visualization on the overlap of the submitted sequences in order to determine the overall coverage that the PCR primers generate (Huang and Madan, 1999). The primary tool for this project, however, was a downloadable Molecular Evolutionary Genetics Analysis program called MEGA4 (Tamura et al. 2007). This program was primarily used for multiple sequence alignments, and construction of distance matrices and phylogenetic trees. Phylogenetic trees are graphical representations of the evolutionary hypotheses and allow for a visual model of the relatedness of species.

Methods:

Genes for MLST were selected from the housekeeping genes that are present in all species that are being analyzed, and ideally in most bacterial species. In a previous MLST study done by Damian Mariano '10 using members of the Yersinia genus, gyrB, recA, rpoB, and Y-HSP60 were all utilized. All of these genes were selected, including Y-HSP60, which is the same
as groEL. Unfortunately, the primer sets for the recA gene have not produced any results and will not be included in this report.

The prokaryotic genome database available through the National Center for Biotechnology Information (NCBI) was utilized to retrieve housekeeping gene sequences from Chryseobacterium gleum, the only Chryseobacterium species whose genome had been fully sequenced at the time this study began. After locating each C. gleum gene, a BLAST search was performed. Thirteen sequences from different genera on the best match list were selected for multiple alignment. The aligned sequences were then added to the Biology Workbench and analyzed using a box shade tool which identifies conserved sequence regions among all thirteen species (Subramaniam 1998). Highly conserved areas were manually screened to determine whether they would be viable primer sites. Each site needed to be at least six amino acids long. After a site was deemed usable, the degeneracy of each primer sequence was determined. Because primers with low degeneracy are preferred, arginine, lysine and serine were avoided since there are six codons that code for each of those amino acids. The primers were then located in the full gene, using the C. gleum sequence generated by the Open Reading Frame (ORF) finder. Forward primers corresponded to the coding strand and generally were towards the beginning (5' end) of the gene while reverse primers, complementary to the coding strand, were closer to the 3' end.

This study utilized the DNA of four potentially novel species: Chryseobacterium piperi, C. angstadti, C. diehli and Kaistella zaccaria, as well as several other members of the Chryseobacterium genus (a table of all species is available in the appendix). Cultures were made for each organism from their respective frozen permanents obtained from either the Lycoming College Culture Collection, or other national and international culture collections such as the American Type Culture Collection and the Belgian Culture Collection. The cultures were grown in Tryptic Soy Broth overnight at 30oC before a Qiagen Blood and Tissue kit was used to isolate the DNA following the manufacturer’s protocols for gram positive pretreatment and DNA purification.
After receiving the primers from Fisher Scientific, they were tested for their functionality using the genomic DNA of *Chryseobacterium piperi*. The thermocycler was allowed to run 35 cycles using a program with denaturation temperature of 94°C, annealing temperature of 55°C and extension temperature of 72°C. Further tests were done to determine the proper annealing temperatures for each specific primer pair. Once these temperatures were determined, PCR amplification was performed for each of the four genes for all available strains. For all of the PCR reactions ExTaq polymerase was used to replicate the DNA. ExTaq is a heat stable self checking DNA polymerase utilized for high quality PCR reactions. After completion of the thermocycling program, 5μl of each PCR product was run on a 1% agarose gel for 15-20 minutes along with 10 μl of λ BstEII marker.

DNA concentrations were determined using the molecular weight marker on the agarose gel photograph. A 96-well plate for sequencing was set up such that approximately 20ng/μl of DNA was present in each well. A second 96-well plate was also set up containing the specific 3μM primer for sequencing corresponding to the well in the first plate. Both plates were then sent to Beckman Coulter Genomics to be sequenced.

After receiving the data, each sequence was checked to determine if the sequence was usable. Any sequences for the same gene of a single species were then assembled using the CAP3 sequence assembly program (available at http://pbil.univ-lyon1.fr/cap3.php) in order to generate a contiguous consensus sequence. After this was performed for each single gene, the consensus sequences were then aligned to each other using the ClustalW algorithm in MEGA4. Each sequence was then manually checked to remove any primer sequence and to double check against the original sequencer files in order to correct for any gaps in the sequence.

Once each sequence was verified, the alignment data was used to create a phylogenetic tree using the neighbor joining algorithm, also in MEGA4. Bootstrap values for each branch of the phylogenetic tree and a difference table showing the percent difference between each sequence that was compared were generated.

After the alignment a GenBank submission form was also created for each batch of
sequences. GenBank utilizes the Sequin submission program which requires the input of the necessary information including the DNA and amino acid sequences.

Results:

Figure 1, shown below, depicts a portion of the GroEL protein multiple sequence alignment utilized in generating primer sequences. The regions in yellow show highly conserved regions in the majority of the sequences, whereas the regions colored in green are the regions where all the sequences are identical. The conserved regions in green were analyzed to determine if they would be usable for primer sequences. For example, the highly conserved region from residues 84-92 was not used for primer design due to the high degeneracy of the sequences. An average usable degeneracy value for a region of equal length would be 256; however this region’s degeneracy value is 32768.

Table 1, below, depicts the primers chosen from the multiple alignment. Each primer ranges in size from 17-23bp to specifically binding to the DNA in the bacterial genome. The sequences were selected due to the low degeneracy and the location within each gene. The number of primer sequences per gene varies due to the overall size of the gene. The main primers
utilized in this study are highlighted in bold below.

Table 1: Primer Sequences: Shown below are the primer sequences designed for use in this project. The total gene lengths as well as both the nucleotide and amino acid sequences are shown. Primers used in this study are highlighted

<table>
<thead>
<tr>
<th>Gene with lengths</th>
<th>Primer Nucleotide Sequences</th>
<th>Primer Amino Acid Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>GlnA 732aa 2196bp</td>
<td>390aaF 5'GGNAARCAAYAAYAYTGG 3'</td>
<td>390aaF G K H N N W</td>
</tr>
<tr>
<td></td>
<td>515aaR 5'TNCCNGTRAANCCRAA 3'</td>
<td>515aaR F A F T G N</td>
</tr>
<tr>
<td></td>
<td>390aaR 5'CCATRRTTRRTGYTTNCC 3'</td>
<td>390aaR</td>
</tr>
<tr>
<td></td>
<td>335aaF 5'CAAYAANGARGTNGCNCC 3'</td>
<td>335aaF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H N E V A P</td>
</tr>
<tr>
<td>GroEL 542aa 1626bp</td>
<td>55aaF 5'GTNGCNAARGARATHGA 3'</td>
<td>55aaF V A K E I E</td>
</tr>
<tr>
<td></td>
<td>190aaR 5'TCRAATygCATNCCYTTC 3'</td>
<td>190aaR</td>
</tr>
<tr>
<td></td>
<td>120aaF 5'GTNATHGAYAARGCNGT 3'</td>
<td>120aaF</td>
</tr>
<tr>
<td></td>
<td>120aaR 5'ACNGCYTTRTCDATNAC 3'</td>
<td>120aaR</td>
</tr>
<tr>
<td>GyrB 645aa 1935bp</td>
<td>350aaF 5'CARTTYGARGGNCARCANAA 3'</td>
<td>350aaF Q F E G Q T K</td>
</tr>
<tr>
<td></td>
<td>600aaR 5'TCATTGYNCCARTYTT 3'</td>
<td>600aaR</td>
</tr>
<tr>
<td></td>
<td>290aaF 5'TAYGTAAYAAYATHAAYAC 3'</td>
<td>290aaF</td>
</tr>
<tr>
<td></td>
<td>350aaR 5'TTRGTYTNGGTYTCRAAYTG 3'</td>
<td>350aaR</td>
</tr>
<tr>
<td></td>
<td>90aaF 5'GARATHTAYGGNCCNGA 3'</td>
<td>90aaF</td>
</tr>
<tr>
<td></td>
<td>225aaR 5'AACATNACNCCRTTAYT 3'</td>
<td>225aaR</td>
</tr>
<tr>
<td>RecA 334aa 1002bp</td>
<td>90aaF 5'GARATHTAYGGNCCNGA 3'</td>
<td>90aaF</td>
</tr>
<tr>
<td></td>
<td>225aaR 5'AACATNACNCCRTTAYT 3'</td>
<td>225aaR</td>
</tr>
<tr>
<td>RpoB 1274aa 3822bp</td>
<td>645aaF 5'TTYYTNGARCAYAYGGA 3'</td>
<td>645aaF</td>
</tr>
<tr>
<td></td>
<td>1245aaR 5'CCANACYTCATYTNCC 3'</td>
<td>1245aaR</td>
</tr>
<tr>
<td></td>
<td>260aaF 5'GARGAYTTYGTNGAYGA 3'</td>
<td>260aaF</td>
</tr>
<tr>
<td></td>
<td>530aaR 5'CDATRTTNNGYNCCYTCCNG 3'</td>
<td>530aaR</td>
</tr>
<tr>
<td></td>
<td>530aaF 5'CCNGARGGNCCNAAYTHG 3'</td>
<td>530aaF</td>
</tr>
<tr>
<td></td>
<td>645aaR 5'TCRTCRTGYTNCARRAA 3'</td>
<td>645aaR</td>
</tr>
</tbody>
</table>

Figure 2, below, shows two examples of PCR products run on a 1% agarose gel. Both gels showed DNA usable for sequencing, seen as yellow bands in the photographs. The yellow bands on the left side of each picture show a DNA marker; the length of each are known, allowing for determining both the gene fragment length as well as the DNA concentration. The first image (picture A) shows the specific binding of the primers resulting in a single band,
whereas picture B shows nonspecific binding of the primers. In the event of nonspecific binding, the samples were extracted from the gel and purified to guarantee the proper sequence.

Figure 2. PCR Gels: Picture A shows specific binding of the primers to the bacterial DNA. Samples include gyrB, rpoB and 16S rRNA primers. Picture B shows nonspecific binding of the primers to the bacterial DNA. Samples include groEL and glnA primers. Both samples were run on 1% agarose gels. The corresponding lane orders are shown below the gel photograph.

Figure 3, below, depicts the differences between a high and low quality sequencing read. The good sequence clearly shows the colored peaks corresponding to each individual nucleotide. As seen in the low quality read, however, it is impossible to determine the sequence information from the chromatogram.
Figure 3. Sequence chromatogram: Diagram A depicts the high quality sequencing read where each peak clearly corresponds to a single nucleotide. Diagram B depicts the low quality sequencing read where the DNA sequence is impossible to determine. Both sequencing reads were performed by Agencourt, and are viewed using Chromas Lite.

Figures 4 and 5, below, show the current phylogenetic tree generated from the available gyrB sequences for members of the *Flavobacteriaceae* family. These trees however, have removed two of the unpublished novel strains, JJC and BLS98, as well as *Chryseobacterium haifense* due to the possibility that these strains may be moved to a different genus. The gene fragment sequences which were aligned in MEGA4, were then used to generate these trees using the Neighbor Joining algorithm. Each branch also shows a bootstrap value, or confidence level. High bootstrap values correspond to a high confidence that the branch is in the proper place in association with all other branches. The length of each branch also shows the relative divergence of each species from its closest neighboring branch.
Figure 4: Neighbor-joining phylogenetic tree based on gyrB gene sequences showing the phylogenetic relationship between strain CTMT and other species of the Flavobacteriales. Bootstrap values (expressed as percentages of 1000 replications) >70% are given at branch nodes. Dots indicate branch nodes also present in either the maximum likelihood or maximum parsimony trees. Sequences generated in this study are shown in bold. Bar, 5 substitutions per 100 nucleotides. Branches in bold were a portion of this study (Strahan et al. 2011).
Figure 5: Neighbor-joining phylogenetic tree based on gyrB gene sequences showing the phylogenetic relationship between strain KMT and other species of the Flavobacteriales. Bootstrap values (expressed as percentages of 1000 replications) >70% are given at branch nodes. Dots indicate branch nodes also present in either the maximum likelihood or maximum parsimony trees. Sequences generated in this study are shown in bold. Bar, 5 substitutions per 100 nucleotides. Branches in bold were a portion of this study (Kirk et al. 2011).

Table 2, below depicts the difference table generated from the multiple sequence alignment. Only the first 20 sequence comparisons are shown. This table allows for a determination of the variation between the different species within the alignment. This difference table also allows for the comparison of every tested strain to every other, allowing for the visualization of the minimum differences between members of the same genus or species as well as showing the difference in the possibly novel species. As seen in the table below, the minimum difference between C. angstadti, and another Chryseobacterium species is 0.113 or 11.3%. In species of Chryseobacterium that have already been published, there are species whose range
around 8.1% to 10.8%.

Table 2: Difference Table for gyrB alignment. The first 10 sequences of the gyrB alignment done in MEGA4. Higher values correspond to a larger degree of variation among the compared sequences.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Difference Table for gyrB alignment. The first 10 sequences of the gyrB alignment done in MEGA4. Higher values correspond to a larger degree of variation among the compared sequences.
Figure 6, shown below, is an example of the GenBank page for one of the submitted gyrB sequences, *Chryseobacterium piperi*. Each entry gives information about the authors and the gene sequence, including where the sample was collected and both the nucleotide and amino acid sequences for the fragment. Because GenBank is free to the public, anyone with internet access worldwide is able to download and utilize the sequence information for free.

Table 3, below, shows the gene sequences that have been submitted to GenBank; as well as sequences that have not yet been submitted, sequences which only have one read, and/or are
currently being finalized. The sequences that have yet to be added to GenBank will be submitted in a batch file that contains the sequences for the same gene for all of the studied species.

Table 3: Gene fragment sequences in various states of submission. Sequences that have been submitted to GenBank are noted with their corresponding accession number. Partial sequences are noted to whether the forward or reverse sequence has produced usable data.

<table>
<thead>
<tr>
<th>Submitted Sequences (with accession number)</th>
<th>Ready for Submission to Genbank</th>
<th>Single sequencing read available</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. luteum</em> gyrB HQ011909</td>
<td><em>C. luteum</em> rpoB</td>
<td><em>C. soli</em> groEL (f)</td>
</tr>
<tr>
<td><em>C. piperi</em> gyrB HQ011910</td>
<td><em>C. piperi</em> rpoB</td>
<td><em>K. zaccaria</em> groEL (f)</td>
</tr>
<tr>
<td><em>C. soldanellicola</em> gyrB HQ011912</td>
<td><em>C. soldanellicola</em> rpoB</td>
<td><em>C. joosteii</em> groEL (f)</td>
</tr>
<tr>
<td><em>C. haifense</em> gyrB HQ011908</td>
<td><em>C. haifense</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>S. anhuiense</em> gyrB HQ011915</td>
<td><em>S. anhuiense</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. shigense</em> gyrB HQ011911</td>
<td><em>C. shigense</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. soli</em> gyrB HQ011913</td>
<td><em>C. soli</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. angstadti</em> gyrB HQ011906</td>
<td><em>C. angstadti</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. diehli</em> gyrB HQ011907</td>
<td><em>C. diehli</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>K. zaccaria</em> gyrB HQ011914</td>
<td><em>K. zaccaria</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. aquaticum</em> gyrB HQ645933</td>
<td><em>C. aquaticum</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. oranimense</em> gyrB HQ645932</td>
<td><em>C. oranimense</em> rpoB</td>
<td></td>
</tr>
<tr>
<td><em>C. joosteii</em> gyrB HQ645934</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. vrystaatense</em> gyrB HQ645931</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion:

The phylogenetic relationships of the novel species can be determined by examining both the sequence information as well as the phenotypic data from other tests. This study has contributed to showing that *Chryseobacterium piperi* and *C. angstadti* are sufficiently divergent from their closest neighbors on the phylogenetic trees to be considered separate species. However, the data from this study alone are not sufficient to identify a species as novel and must be combined with 16S rRNA data as well as several phenotypic tests, including metabolic tests,
differential and selective media tests, and fatty acid methyl ester analysis.

When looking at the phylogenetic tree for gyrB, it is possible to determine the evolutionary relationships of the individual genes for the novel species used in this study. *C. piperi*, as seen in Figure 4 groups most closely with *C. soli*, likewise *C. angstadti* groups most closely with *C. luteum*. The recently accepted article on *C. piperi* also groups *C. piperi* with *C. soli* in the 16S rRNA phylogenetic tree (Strahan *et al.* 2011).

The gyrB sequences generated for the CTMT study and the corresponding reference *Chryseobacterium* strains were aligned to GenBank database sequences of other *Flavobacteriaceae* gyrB genes using the Clustal W implementation in MEGA 4. The alignment was manually edited and used to construct a neighbor joining phylogenetic tree. All of the *Chryseobacterium* gyrB sequences clustered together and the pairwise similarities between strain CTMT gyrB sequence and those of other Chryseobacteria (85.8-87.8%) was less than the pairwise similarity between *C. gleum* and *C. indologenes* (88.6%) or between *C. luteum* and *C. shigense* (89.5%) (Strahan *et al.* 2011). Likewise the gyrB sequences generated in the KMT study and the corresponding four reference *Chryseobacterium* strains were used to construct a neighbor joining phylogenetic tree (Figure 5). All of the *Chryseobacterium* gyrB sequences clustered together and the pairwise similarities between strain KMT gyrB sequence and those of other *Chryseobacterium* species (81.6 - 87.7%) were less than the pairwise similarity between *C. shigense* and *C. vrystaatense* (91.5%) or between *C. shigense* and *C. luteum* (89.9%) (Kirk *et al.* 2011).

This study, along with several phenotypic tests, has noted that *Kaistella zaccaria* and *C. haifense* have divergent enough 16S rRNA and gyrB sequences to be considered outside of the genus *Chryseobacterium*.

To date this project has successfully completed the GenBank submissions for one of the genes (GyrB) for many of the currently studied species. This project will be continued to finish the sequencing and analysis of the remaining three genes (GlnA, GroEL, and RpoB), culminating in the eventual submission of the relevant sequence data to GenBank. As seen in table 3, the
groEL gene and rpoB gene fragments are nearly ready for submission. The sequences need to be aligned and then manually edited before being submitted to GenBank.

References:


Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russel, J.E., Urwin, R., Zang, Q.,


### Appendix Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer Design - Boxshade</td>
<td>1</td>
</tr>
<tr>
<td>GlnA</td>
<td>4</td>
</tr>
<tr>
<td>GroEL</td>
<td>7</td>
</tr>
<tr>
<td>GyrB</td>
<td>10</td>
</tr>
<tr>
<td>RecA</td>
<td>12</td>
</tr>
<tr>
<td>Primer Design - ORF Finder Sequences</td>
<td>17</td>
</tr>
<tr>
<td>GlnA</td>
<td>19</td>
</tr>
<tr>
<td>GroEL</td>
<td>20</td>
</tr>
<tr>
<td>GyrB</td>
<td>21</td>
</tr>
<tr>
<td>RecA</td>
<td>22</td>
</tr>
<tr>
<td>Primer Sequence Degeneracy</td>
<td>23</td>
</tr>
<tr>
<td>Organism Source (Culture Collection)</td>
<td>24</td>
</tr>
<tr>
<td>Good Sequences</td>
<td>25</td>
</tr>
<tr>
<td>GenBank Accession Reports</td>
<td></td>
</tr>
<tr>
<td><em>Chryseobacterium angstadti</em></td>
<td>27</td>
</tr>
<tr>
<td><em>Chryseobacterium diehli</em></td>
<td>28</td>
</tr>
<tr>
<td><em>Chryseobacterium haifense</em></td>
<td>29</td>
</tr>
<tr>
<td><em>Chryseobacterium luteum</em></td>
<td>30</td>
</tr>
<tr>
<td><em>Chryseobacterium piperi</em></td>
<td>31</td>
</tr>
<tr>
<td><em>Chryseobacterium shigense</em></td>
<td>32</td>
</tr>
<tr>
<td><em>Chryseobacterium soldanellicola</em></td>
<td>33</td>
</tr>
<tr>
<td><em>Chryseobacterium soli</em></td>
<td>34</td>
</tr>
<tr>
<td><em>Kaistella zaccaria</em></td>
<td>35</td>
</tr>
<tr>
<td><em>Sphingobacterium anhuiense</em></td>
<td>36</td>
</tr>
</tbody>
</table>
Chryseobacterium gleum glnA and primers

2196 atg tcaaccttaagattcaaagcgttagaaaccctaccatttaag

M S T L R F K A L T L P F K

2151 gatttcagaagagataactctgttgaagttcctgtaaaattatca

D F R R D N S V E V P V K L S

2106 gaattattctgccagaatgttttctcagaagaaaca

atg agagaa

E L F C Q N V F S E E T M R E

2061 tatttaacgaaagaagcattccagtctatt

G Y L T K E A F Q S I M D A V K

2016 aaagggactaaaatccagagacacattgcagatcaggtagctgta

K G T K I Q R H I A D Q V A V

1971 gca atg aaagactgggca atg agcaaaggagtaactcactatact

A M K D W A M S K G V T H Y T

1926 cactggtttcagcctttgacagggacaaccgcagaaaagcacgat

H W F Q P L T G T T A E K H D

1881 tctttcttcacccctatcgaaggaggaagagctattgaaagattc

S F F T P I E G G R A I E R F

1836 agcggaaatttattgatccagcaggagccggatgcctcttctttc

S G N L L I Q Q E P D A S S F

1791 ccgaacggag gaatcagaaacacgttcgaggcaagaggttatacc

P N G G I R N T F E A R G Y T

1746 gcatgggatcctacttctccggctttcatc atg ggaactacattg

A W D P T S P A F I M G T T L

1701 tgtatcccatctatcttcatctcttatacaggagaaactttagat

C I P S I F I S Y T G E T L D

1656 tataaagccccccttttaagagctttgaacgctgtagacgaggct

Y K A P L L R A L N A V D E A

1611 gcaaccaacgtatacgatttcgacaaaaatgtaacaaaagta

A T N V M Q Y F D K N V T K V

1566 actcctactttaggttgggagcaggaatatttccttgtagattca

T P T L G W E Q E Y F L V D S

1521 gcattataccaatcccgtccggaccttgtattaacaggtaagaca

A L Y Q S R P D L V L T G K T

1476 ttattaggacattctcctgcaaa aggacagcagctggatgatcac

L L G H S P A K G Q Q L D D H

1431 tatttcggttcaattcctacaagagta atg aacttc atg aaggag

Y F G S I P T R V M N F M K E

1386 ctggaaatcgaatgt atg aaattaggtattcctgtaactacaaga

L E I E C M K L G I P V T T R CAYAANGARGTNGCNCC

GlnA -335 aa

1341 cacaacgaggtagctccgaaccaatttgagctggctcca atg ttt

H N E V A P N Q F E L A P M F

1296 gaggaagtaaacgttgccgtagaccacaactctttgttg atg gat

E E V N V A V D H N S L L M D

1251 atc atg gcaagaattgcctcacagacaccacttccacattttattc

I M R I A H E F R K
Chryseobacterium gleum groEL and primers

1626 atggcaaaagaaataaaattcgatattgaatcaagagacgcgcta

1581 aagagaggggtagatgcattggctaatgcagtaaaagtaacttta

1536 ggtcctaaagggagaaacgtagtgatcgaaaaatctttcggtgct

1491 ccacacgtaactaaggacggtgtttctgtagcaaaagaaatcgaa

1446 cttgaagataaggtagaaaacatgggagctcagatggtaaaagaa

1401 gtagcttccaaaaccaatgatattgcaggagacggtactactacc

1356 gctactgtattggcacaggctatcgtaagagaaggtcttaagaac

1311 gtagctgccggtgctaacccgatggatcttaaaagagggatcgac

1266 aaagcggtaactgctgttgttgaaaaccttaaatctcaatctcag

1221 gaggttggagattcttcagaaaaaatcaagcaggttgcttctatc

1176 gtagctcagatcaaagctcagatggaaactacaacttctgattac

1131 gctgtagaagaaggtatcgttgcaggtggtggtgttgctttagta

1086 aagagctactccaggggtctgatgtggtagaatcaaaagaaatcgaa

1041 gacagaggataccagtcaccttatttcgtgacaaaccctgagaaa

996 atgttagctgaacttgaaagcccatatatccttttagtagagaag

951 aaaatctcttccatgaaagaattgttaccggttcttgagccaatt

906 gcacaaggtggtaagtctctattaattatctctgaagaagttgaa

861 ggtgaagctttagctactttggtagtaaacaaactaagaggttcc

816 cttaaaattgctgctgtaaaagctcctggattcggagacagaaga

771 aacagaactgtagctggagaaggttctgtaatcactgtagaagaggct

726 ttataagagctactccaggggtctgatgtggtagaatcaaaagaaatcgaa

681 atgcttggaactgctgagaaagtaacgatcgacaaagacaacaca

636 acggttgtaaacggaggtggtgacgaagctaaaatcaaaggaaga

591 gtagctcagatcaaagctcagatggaaactacaacttctgattac

546 gacagagaaaaacttcaggaaagactggctaagttagctggaggt

501 gttgctgtactttatgtaggagctgcttctgaagtggaaatgaaa

456 gtagctcagatcaaagctcagatggaaactacaacttctgattac

411 gctgtagaagaaggtatcgttgcaggtggtggtgttgctttagta

366 agagcaatttcttcacttgacaatctttcaggatctaacgctgat

321 gaaactacagggatcaaaatcgttaagagagctatcgaggagcca

276 ttataagagctactccaggggtctgatgtggtagaatcaaaagaaatcgaa

231 gttgctaaagtagcagaaggaaacggagacttcggatacaatgcg

186 aaaactgacgagtacgttcacatgcttgaagcaggaatcattgac

141 ccaactaaagtaacaagagttgcccttgaaaacgcagcttctgta

96 ttataagagctactccaggggtctgatgtggtagaatcaaaagaaatcgaa

51 scgc ggctgaaccagctatgccaatgggtggtggaatgccaggaatg

6 atgtaa 1

M *
Chryseobacterium gleum gyrB and primers

1 atgagtcaaaaacaatatacagctagtagtattcaggcattggaa

M  S  Q  K  Q  Y  T  A  S  S  I  Q  A  L  E

46 ggaatggaacacgttcgtatgcgtccttcaatgtacattggtgat

G  M  E  H  V  R  M  R  P  S  M  Y  I  G  D

91 gtaggagtcagaggacttcaccacttggtttatgaagtagtagat

V  G  V  R  G  L  H  H  L  V  Y  E  V  V  D

136 aactctattgacgaggcattggcagggtactgcgatacgatcttc

N  S  I  D  E  A  L  A  G  Y  C  D  T  I  F

181 gttagcatcaaggaaggaaacggaattgaagtaagtgataacggt

V  S  I  K  E  G  N  G  I  E  V  S  D  N  G

226 agaggtatcccggttgatttccacgaaaaggagcagaaatcagcc

R  G  I  P  V  D  F  H  E  K  E  Q  K  S  A

271 cttgaagttgtaatgactaaaatcggggccggaggtaagtttgat

L  E  V  V  M  T  K  I  G  A  G  G  K  F  D

316 aaagactcttataaggtttcaggaggtcttcacggggttggggtt

K  D  S  Y  K  V  S  G  G  L  H  G  V  G  V

361 tcgtgtgtgaatgcactttccaatgagatgatcactactgtttac

Q  T  G  V  E  E  I  G  H  S  E  R  R  G  T

406 agagacggaaacgtttatcagcagatttattccagaggaaaagcg

R  D  G  N  V  Y  Q  Q  I  Y  S  R  G  K  A

451 cagaccggtgttgaagagattggtcacagcgaaagaagaggaacc

Q  T  G  V  E  E  I  G  H  S  E  R  R  G  T

496 aagcagtttttccagccggatgatactatatttacagaattagtt

K  Q  F  F  Q  P  D  D  T  I  F  T  E  L  V

541 tacaattacgatacattagcaagccgtttaagagagctttcttac

Y  N  Y  D  T  L  A  S  R  L  R  E  L  S  Y

586 cttaataaaggaattaccattacgcttaccgatgaaagaaaaa

L  N  K  G  I  T  I  T  L  D  E  R  E  K

631 ttggaagacggatctttccgttcagaagttttccattctgaaggt

L  E  D  G  S  F  R  S  E  V  F  H  S  E  G

676 ggtttgaaggagtttgttgcttatattgacggtaaccgtgaatct

G  L  K  E  F  V  A  Y  I  D  G  N  R  E  S

721 atcatggagcatgtgatcttcatggaaggagaaagagatgatatc

I  M  E  H  V  I  F  M  E  G  E  R  D  D  I

766 ccggtagaagtagcaatgcgttacaatacttctttcaatgagaat

P  V  E  V  A  M  R  Y  N  T  S  F  N  E  N

811 cttcactcttacgttaacaatatcaacacccatgaagggggaact

GyrB-290   TAYGTNAAYAAYATHAAYAC

856 cacctggcaggtttcagacgtgctttgacgagaacccttaagaaa

H  L  A  G  F  R  R  A  L  T  R  T  L  K  K

901 tatgctgatgatttaggaattccgcagaaggaaaaagtagaagtt

Y  A  D  D  L  G  I  P  Q  K  E  K  V  E  V

946 acaggagatgacttccgtgaaggtttaacagctgtagtttccgta

T  G  D  D  F  R  E  G  L  T  A  V  V  S  V

1036 ggaaactctgaagtttcaggggctgtagataaaattgtaggggaa

G  N  S  E  V  S  G  A  V  D  K  I  V  G  E

1081 atgcttaccaatttccttgaagaaaacccaaatgaagcgaagcag

M  L  T  N  F  L  E  E  N  P  N  E  A  K  Q

1126 atcgttcagaaagtggttttggcagcaaaagcaagacaggccgct

I  V  Q  K  V  V  L  A  A  K  A  R  Q  A  A

1171 aagaaagcgcgtgaaatggttcagagaaaatctccgatgggaggt

K  K  A  R  E  M  V  Q  R  K  S  P  M  G  G

1216 tccggtcttcctggaaaactatctgactgttcatcaaaagatccg

S  G  L  P  G  K  L  S  D  C  S  S  K  D  P

1261 gcagaatctgaaatcttccttgtagagggagattccgcagggggt

A  E  S  E  I  F  L  V  E  G  D  S  A  G  G

1306 acagccaagcagggaagagacagacacttccaggctattcttccg

T  A  K  Q  G  R  D  R  H  F  Q  A  I  L  P

1351 ttaagaggtaaaatcctgaacgtagaaaaatctatgcttcacaaa

L  R  G  K  I  L  N  V  E  K  S  M  L  H  K

1396 gtatatgataacgaagagatcagaaatatctatacagctcttgga

V  Y  D  N  E  E  I  R  N  I  Y  T  A  L  G

1441 gtttccgtaggaactgaggaagacagcaaagcattgaacttatca

V  S  V  G  T  E  E  D  S  K  A  L  N  L  S

1486 aaactgagataccacaagatcgttatcatgaccgatgccgatatt

K  L  R  Y  H  K  I  V  I  M  T  D  A  D  I

1531 gacggatctcacatttccacactgattcttactttcttcttcaga

D  G  S  H  I  S  T  L  I  L  T  F  F  F  R

1576 tatatgaaagaacttattgagaatggttatatctatatcgctcaa

Y  M  K  E  L  I  E  N  G  Y  I  Y  I  A  Q

1621 ccacctttatacctgttaaagaaaggaaacaaaaaagtatacgct

P  P  L  Y  L  L  K  K  G  N  K  K  V  Y  A

1666 tacaacgaaaaagagcgtgaagagtttactctggaaatgtctccg

Y  N  E  K  E  R  E  E  F  T  L  E  M  S  P

1711 gacggaaaaggggtagaggtacagcgttacaaaggtcttggagaa

D  G  K  G  V  E  V  Q  R  Y  K  G  L  G  E

1756 atgaaccctgagcagctttgggaaacaacccttaaccctgaacac

M  N  P  E  Q  L  W  E  T  T  L  N  P  E  H

1801 agaattctgaagcaggtaaccattgataatgcagtggaagctgac

R  I  L  K  Q  V  T  I  D  N  A  V  E  A  D

1846 agtattttctccatgttgatgggtgacgaggttccgccaagaaga

S  I  F  S  M  L  M  G  D  E  V  P  P  R  R

1891 gaatttattgagaaaaatgcaaaatatgccaaaattgatgcataa

E  F  I  E  K  N  A  K  Y  A  K  I  D  A  *
<table>
<thead>
<tr>
<th>Protein</th>
<th>Total Degeneracy</th>
<th>Degenerate Code</th>
<th>Degenerate Code</th>
<th>Degenerate Code</th>
<th>Degenerate Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>GlnA</td>
<td>784</td>
<td>4 2 1 2 4</td>
<td>1 2 3 2 4</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
</tr>
<tr>
<td>GroEL</td>
<td>768</td>
<td>4 2 1 2 4</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
<td>2 4 3 2 1</td>
</tr>
<tr>
<td>GyrB</td>
<td>768</td>
<td>4 2 1 2 4</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
<td>2 4 3 2 1</td>
</tr>
<tr>
<td>RecA</td>
<td>784</td>
<td>4 2 1 2 4</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
<td>2 4 3 2 1</td>
</tr>
<tr>
<td>RpoB</td>
<td>120</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
<td>2 4 3 2 1</td>
<td>5 6 2 4 1</td>
</tr>
<tr>
<td>Organism</td>
<td>Organism Source</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. angstadti</em> KM</td>
<td>LCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. aquaticum</em> KCTC12488</td>
<td>KCTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. diehli</em> BLS98</td>
<td>LCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. greenlandensis</em> UMB10</td>
<td>Jean Brenchley</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. greenlandensis</em> UMB14</td>
<td>Jean Brenchley</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. greenlandensis</em> UMB34</td>
<td>Jean Brenchley</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. haifense</em> DSM19056</td>
<td>DSMZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. joostei</em> LMG18212</td>
<td>BCCMLMG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. jejuense</em> DSM19299</td>
<td>Jean Brenchley</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. luteum</em> DSM18605</td>
<td>DSMZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. oranimense</em> LMG24030</td>
<td>BCCMLMG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. piperi</em> CTM</td>
<td>LCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. shigense</em> DSM17126</td>
<td>DSMZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. soldanellicola</em> DSM17072</td>
<td>DSMZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. soli</em> DSM19298</td>
<td>DSMZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. vyrstaatense</em> LMG22846</td>
<td>BCCMLMG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>K. koreensis</em> CCUG49689</td>
<td>CCUG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>K. zaccaria</em> JJC</td>
<td>LCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>K. sp.</em> 3519-10</td>
<td>Brent Christner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. anhuiense</em></td>
<td>LCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BCCMLMG= Belgian Coordinated Collection of Microorganisms  
CCUG= Culture Collection, University of Göteborg (Sweden)  
DSMZ= Deutsche Sammlung von Mikroorganismen und Zellkulturen  
(German Collection for Microorganisms and Cell Cultures)  
KCTC= Korean Collection for Type Cultures  
LCCC= Lycoming College Culture Collection  
Jean Brenchley=Professor of Microbiology, Penn State University  
Brent Christner= Professor Biology, Louisiana State University
<table>
<thead>
<tr>
<th>Organism</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>GenBank Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. angstadt</td>
<td>i gyrb 350f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td>HQ011906</td>
</tr>
<tr>
<td>C. aquaticum</td>
<td>rpoB 530f, rpoB 645r, rpoB 1245r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. diehli</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. haifense</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. joosteii</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. luteum</td>
<td>rpoB 260f, rpoB 645r, rpoB 1245r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. oranimense</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. piperi</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. shigense</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. soldanellicola</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. soli</td>
<td>rpoB 260f, rpoB 530f</td>
<td>rpoB 530r, rpoB 645r, rpoB 1245r</td>
<td></td>
</tr>
<tr>
<td>C. vryssaatense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K. zaccaria</td>
<td>rpoB 306f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. anhuiense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. aquaticum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. diehli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. haifense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. joosteii</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. luteum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. oranimense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. piperi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. shigense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. soldanellicola</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. soli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. vryssaatense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K. zaccaria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. anhuiense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td>Forward Primer</td>
<td>Reverse Primer</td>
<td>GenBank Accession</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>C. angstadt</td>
<td>i groEL 55f</td>
<td>i groEL 190r</td>
<td></td>
</tr>
<tr>
<td>C. aquaticum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. diehli</td>
<td>groEL 55f</td>
<td>groEL 190r</td>
<td></td>
</tr>
<tr>
<td>C. haifense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. joosteii</td>
<td>groEL 55f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. luteum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. oranimense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. piperi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. shigense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. soldanellicola</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. soli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. vrystaatense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K. zaccaria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. anhuiense</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

27
Chryseobacterium sp. KM DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HM011906.1

FASTA  Graphics

Go to:

LOCUS  HM011906  727 bp  DNA  linear  BCT 28-SEP-2010
DEFINITION  Chryseobacterium sp. KM DNA gyrase subunit B (gyrB) gene, partial cds.
ACCESSION  HM011906
VERSION  HM011906.1  GI: 307592545
SOURCE  Chryseobacterium sp. KM
ORGANISM  Chryseobacterium sp. KM
            Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Chryseobacterium.
REFERENCE  1 (bases 1 to 727)
AUTHORS  Failor, K.C., Battles, A.M. and Newman, J.D.
TITLE  Multilocus sequence typing to delineate different species among the Flavobacteriaceae
JOURNAL  Unpublished
REFERENCE  2 (bases 1 to 727)
AUTHORS  Failor, K.C., Battles, A.M. and Newman, J.D.
TITLE  Direct Submission
JOURNAL  Submitted (03-AUG-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA
FEATURES
  Location/Qualifiers
  source  1..727
         /mol_type="genomic DNA"
         /strain="KM"
         /db_xref="taxon:559151"
         /PCR_primers="fwd_name: 350aaF, fwd_seq: carytgaggnccarcanaa, rev_name: 600aar, rev_seq: ttcatytoncnmacnnytnrrta"
    gene  <1..727
          /gene="gyrB"
    cds  <1..727
         /gene="gyrB"
         /codon_start=1
         /transl_table=11
         /product="DNA gyrase subunit B"
         /protein_id="AB963961.1"
         /db_xref="GI:307592546"
    translation="MIIISSGAVYKVIGMLTNLFLEENSPEAKTVQVKSLAAMARQAAA
                   KRAEMVQKESFNGWGLPGKLSDCSKDPASEELFLVEGDSAGTGARKQRDRHFQAI
                   LPLRKILNEVEMLHMYDNIEEIKNYIALGVSQTEQDSKALNMAELKTYHRIVINT
                   DADIDGHSIISSLFFRRMKKELIENGYYIAQAQPFLYLEKGNKGVAYNKRKREKF
                   TLNGVFGVWGSGRYKFCGKLK"
    origin
    1 aaatatatt taacgggcggc tgtagtataaa tagtgagg cagatgccct cggctttctag
26  61 gacagcac ccacgagcc tacaattatt gcacaag cctgaagca ctcaagtcctc gcggggcc
121 181 gacactctcccc tcaaatgc ttcgagtctc cttggttctc ctcggggtgtctt ctttaggttg
241 214 cgggtctctc tatctacgaca tctttggcatt cctgtctctc cgggtctctc ctttgctctc
301 301 cttatatatt gcgagttg ggtttgttat tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
**Chryseobacterium sp. BLS98 DNA gyrase subunit B (gyrB) gene, partial cds**

**GenBank: HQ011907.1**

**FASTA  Graphics**

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LOCUS</strong></td>
<td>HQ011907</td>
</tr>
<tr>
<td><strong>DEFINITION</strong></td>
<td>Chryseobacterium sp. BLS98 DNA gyrase subunit B (gyrB) gene, partial cds.</td>
</tr>
<tr>
<td><strong>ACCESSION</strong></td>
<td>HQ011907.1</td>
</tr>
<tr>
<td><strong>VERSION</strong></td>
<td>0.1 G2.307S92547</td>
</tr>
<tr>
<td><strong>KEYWORDS</strong></td>
<td>Chryseobacterium sp. BLS98</td>
</tr>
<tr>
<td><strong>SOURCE</strong></td>
<td>Chryseobacterium sp. BLS98 Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Chryseobacterium.</td>
</tr>
<tr>
<td><strong>REFERENCE</strong></td>
<td>1 (bases 1 to 526)</td>
</tr>
<tr>
<td><strong>AUTHORS</strong></td>
<td>Failor,K.C., Batters,A.M. and Newman,J.D.</td>
</tr>
<tr>
<td><strong>TITLE</strong></td>
<td>Multilocus sequence typing to delineate different species among the Flavobacteriaceae</td>
</tr>
<tr>
<td><strong>JOURNAL</strong></td>
<td>Unpublished</td>
</tr>
<tr>
<td><strong>REFERENCE</strong></td>
<td>2 (bases 1 to 526)</td>
</tr>
<tr>
<td><strong>AUTHORS</strong></td>
<td>Failor,K.C., Batters,A.M. and Newman,J.D.</td>
</tr>
<tr>
<td><strong>TITLE</strong></td>
<td>Direct Submission</td>
</tr>
<tr>
<td><strong>JOURNAL</strong></td>
<td>Submitted (03-JUL-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA</td>
</tr>
<tr>
<td><strong>FEATURES</strong></td>
<td>Location/qualifiers</td>
</tr>
<tr>
<td><strong>source</strong></td>
<td>1..526</td>
</tr>
<tr>
<td><strong>/organism</strong></td>
<td>Chryseobacterium sp. BLS98</td>
</tr>
<tr>
<td><strong>/mol_type</strong></td>
<td>genomic DNA</td>
</tr>
<tr>
<td><strong>/strain</strong></td>
<td>BLS98</td>
</tr>
<tr>
<td><strong>/db_xref</strong></td>
<td>taxon:605566</td>
</tr>
<tr>
<td><strong>/gene</strong></td>
<td>gyrB</td>
</tr>
<tr>
<td><strong>/codon_start</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>/transl_table</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>/product</strong></td>
<td>DNA gyrase subunit B</td>
</tr>
<tr>
<td><strong>/protein_id</strong></td>
<td>ABN61962.1</td>
</tr>
<tr>
<td><strong>/translation</strong></td>
<td>PNEAK1IVQVTVLAAARQPAAEAEKRMVRKRSREIGGSGLFGLS</td>
</tr>
<tr>
<td><strong>/translation</strong></td>
<td>DCSKREDLAEFLVEQVEQGPSTAQRNHDQPAILPGSFLKSNVESSHMLKRVYCEEE</td>
</tr>
<tr>
<td><strong>/translation</strong></td>
<td>KRHIYTLGVSVIEDESRSAIAMRKRKRMINTDADIG89HISTLILT7FPRMKE</td>
</tr>
<tr>
<td><strong>/translation</strong></td>
<td>LENGIYFLAPPFL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>ORIGIN</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 acctaaaga aggagaaaaat atcgatctgta aggtagtattt ggcagaaaaa gsaagacagg</td>
</tr>
<tr>
<td>61 caagcaagaa aggtcgagaa atcgatctgta gaaatgagct gatgggrggt tctggttttac121</td>
</tr>
<tr>
<td>111 attcggct策划 tatgttcctgac acctggtgtc atctctagt ggagtcct</td>
</tr>
<tr>
<td>161 tgcctggt tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>211 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>261 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>311 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>361 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>411 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>461 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
<tr>
<td>511 acctgggtc tggaggtgc acctggagct cagagacct catttcttctgag</td>
</tr>
</tbody>
</table>

//
Chryseobacterium haiense strain DSM 19056 DNA gyrase subunit B (gyrB) gene, partial cds

GenBank MG119081
PASTA Image

LOCUS  MG119081  7.1 Kbp DNA linear NCT 25-SEP-2010

DESCRIPTION  Chryseobacterium haiense strain DSM 19056 DNA gyrase subunit B (gyrB) gene, partial cds.

ACCESSION  MG119081
VERSION  MG119081.1  GI:397592949

KEYWORDS  .

SOURCE  Chryseobacterium haiense

ORGANISM  Chryseobacterium haiense

     bacteria; Eubacteriota; Flavobacteria; Flavobacteriales; Flavobacteriidae; Chryseobacteriaceae.

REFERENCE  1 (base 1 to 719)

AUTHORS  Failler, P.C., Battie, A.M. and Pavan, J.D.

TITLE  Multilocus sequence typing to delineate different species among the Flavobacteriaceae

JOURNAL  Unpublished

REFERENCE  2 (base 1 to 719)

AUTHORS  Failler, P.C., Battie, A.M. and Pavan, J.D.

TITLE  Direct Submission

JOURNAL  Submitted (03-APR-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA

FEATURES  Location/Qualifiers

    source   1..1170
      /organism="Chryseobacterium haiense"
      /mol_type="genomic DNA"
      /strain="DSM 19056"
      /db_xref="taxon:101026"
      /PCR_primers="tmd_name: 802AR, tmd_seq: cctcgccgagccacacece, rev_name: 650AR, rev_seq: ttgcttctggccgacctyccct"/
    /source_type="type strain of Chryseobacterium haiense"

    gene   1..719
      /gene="gyrB"
    /gene="gyrB"
      /codon_start=1
      /translation="15"
      /product="DNA gyrase subunit B"
      /protein_id="A2CL3G981.1"
      /db_xref="GI:207592860"

    CDS   1..719
      /gene="gyrB"
      /codon_start=1
      /transl_table=11
      /product="DNA gyrase subunit B"
      /protein_id="A2CL3G981.1"
      /db_xref="GI:207592860"

ORIGIN  1 acacaccgtcttggg cagccttgacctctgctcctgctctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctgcta

30
Chryseobacterium luteum strain DSM 18605 DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ115909.1
Fasta Graphics

Quick

LOCUS      HQ115909                     715 bp    DNA linear DEC 26-SEP-2010
DEFINITION Chryseobacterium luteum strain DSM 18605 DNA gyrase subunit B (gyrB) gene, partial cds.
ACCESSION HQ115909
VERSION   HQ115909.1             GI: 807592851
KEYWORDS  .
SOURCE Chryseobacterium luteum
          Chryseobacterium luteum
          Bacteroides; Bacteroidales; Flavobacteriaceae; Flavobacteriales;
          Flavobacteriacea; Chryseobacterium.
REFERENCE 1 (dates 1 to 715)
AUTHORS Fullor, R.C., Rappaz, A.M. and Newman, J.D.
TITLE Multilocus sequence typing to delineate different species among the Flavobacteriaceae.
JOURNAL Unpublished
REFERENCE 2 (dates 1 to 715)
AUTHORS Fullor, R.C., Rappaz, A.M. and Newman, J.D.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA
FEATURES Location/Qualifiers
source 1..715
| /organism="Chryseobacterium luteum"
| /mol_type="genomic DNA"
| /strain="DSM 18605"
| /db_xref="GenPept:411511"
| /fam淋名="CWD_name: 550aa; CWD_seq;
| CTRtygseqrencnm; CWD_seq: 800nkb; CWD_seq:
| ttreynomsncncnrytmt"
| /note="Type strain of Chryseobacterium luteum"

gene  1..715
| /gene="gyrB"
| /cdd="gycB"
| /codon_start=1
| /transl_table=11
| /product="DNA gyrase subunit B"
| /protein_id="AD609944.1"
| /db_xref="InterPro:INTERPRO104000001; Enzyme:00091"
| /translation="MTLQSEVSGAVDIEKDLNIEFIEHNPDNEKITYKMYVLLASHK
| FQHAEAKNDHQGQERMDN0002LFHLDSSD320FMAESEL1VLGDDGNN0224R0N0H"
| Q9QUR4RQV7YKREK18YTV1AQP04;1Q000295Y0A2691560"
| DAEFPLESFCGRLGTCEAV协会"

ORIGIN 1 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
61 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
121 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
181 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
241 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
301 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
361 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
421 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
481 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
541 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
601 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
661 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
721 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
781 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
841 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
901 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
961 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq
1021 accsaasagq wquqgggq arcwqsgaq awqqgggq agawgggq awqqgggq

//
Chryseobacterium sp. CTM DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ011910.1

FASTA

LOCUS       HQ011910    711 bp DNA linear BCT 28-SEP-2010
DEFINITION  Chryseobacterium sp. CTM DNA gyrase subunit B (gyrB) gene, partial cds.
ACCESSION  HQ011910
VERSION     HQ011910.1 GI:307592553
SOURCE      Chryseobacterium sp. CTM
ORGANISM    Chryseobacterium sp. CTM
            Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Chryseobacterium.
REFERENCE  1 (bases 1 to 711)
AUTHORS     Failor,K.C., Satties,A.M. and Newman,J.D.
TITLE       Multilocus sequence typing to delineate different species among the Flavobacteriaceae
JOURNAL     Unpublished
REFERENCE  2 (bases 1 to 711)
AUTHORS     Failor,K.C., Satties,A.M. and Newman,J.D.
TITLE       Direct Submission
JOURNAL     Submitted (02-AUG-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA
FEATURES    Location/Qualifiers
            source 1..711
            /organism="Chryseobacterium sp. CTM"
            /mol_type="genomic DNA"
            /strain="CTM"
            /isolation_source="Loyalsock Creek"
            /db_xref="taxon:558152"
            /PCR_primers="fwd_name: 350aaF, fwd_seq: cattyygarggnncarncanaa, rev_name: 600aaR, rev_seq: ttcctynonncnnrncytrrrta"

CDS

/locus=gyrB
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="ADN59651.1"
/db_xref="GI:307592554"
/translation="TYLGNSEV3GADV3KIVGSEM3LNFLEL3NHEAKLVQKVVLAA8RA
RQAARERAM7QRKSPMG5SGLPKLSDC3SPEHEL3PUEGDSBB00G1FQ0QDR8H
QPQ1LPL3K1L73K13NK3VL3V3DKEIKY3NITALGV3GVG3E3DSKL3N4AKS3LKV3H
VIMA3DAD3IGMSLTHL3L13FFYR3M3ELE3N0Y3I3AQ3P13Y1L3K3GNK3AM3AHNE3E
RE3FTL3M3P3D3G3K3G3V3E3V3R"
Chryseobacterium shigense strain DSM 17126 DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ011911.1

```
LOCUS    HQ011911  711 bp  DNA    linear  BCT 28-SEP-2010
DEFINITION Chryseobacterium shigense strain DSM 17126 DNA gyrase subunit B (gyrB) gene, partial cds.
ACCESSION HQ011911
VERSION   HQ011911.1  01-07-2006
KEYWORDS  Chryseobacterium shigense
ORGANISM Chryseobacterium shigense
          Bacteria;  Deltaproteobacteria;  Flavobacteria;  Flavobacteriales;  Flavobacteriaceae;  Chryseobacterium.
REFERENCE 1  (bases 1 to 711)
          Faller, K.C., Batters, A.M. and Newman, J.D.
          TITLE  Multilocus sequence typing to delineate different species among the Flavobacteriaceae
          JOURNAL  Unpublished
REFERENCE 2  (bases 1 to 711)
          Faller, K.C., Batters, A.M. and Newman, J.D.
          TITLE  Direct Submission
          JOURNAL  Submitted (02-AUG-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA
FEATURES Location/Qualifiers
  gene      1..711
  /locusname="Chryseobacterium shigense"
  /mol_type="genomic DNA"
  /strain="DSM 17126"
  /db_xref="taxon:257344"
  /PCR_primer="fwd_name: 25522F, fwd_seq: GCTCTTGAGCGTACAC, rev_name: 650GAR, rev_seq: TATATTACCCTGC"
  /note="type strain of Chryseobacterium shigense"
  /gene=gyrB
  /gene=gyrB
  /cdd=1..711
  /product="DNA gyrase subunit B"
  /protein_id="A0U65286.1"
  /db_xref="GI:100729250"
  /translation="TTTAEAEVGLVLVEMCVLNELEPNEEPVQIKEYQLTVLLAMRA
             AQARERKQERIGREJPN93LYPNPDDESEDDAEI2ISLEVD30AASTQQRQHNL
             FYQVPSLACDDDDLEVNEEMLuVYDEEHJIQUIVCTVCYTEDGRLMENL3HVR
             EVQNYLAQSLNRFRTETERTQI1PS4YFFILKFLRNVQQNED
             REDTLDGDFQHVTQK"

ORIGIN
  1 accaaagtgg agagccggtt gccagcttgc aggagaccttg gcggagaagtgc
  61 accaattgc tgaagagaag gaaacgattga aagtggaggc agaagttgg
  121 gggaagagct ggcggagaat aggagctatg atagctggag aagatcggc
  181 atggagcg ctgacctgtg ggtgtgc tggagagaagt gcagctgctg
  241 tggagcg ctgacctgtg ggtgtgc tggagagaagt gcagctgctg
  301 agaagttgg  aagatcggc gcgaagttgg  aagatcggc gcgaagttgg
  361 ggtgtgc tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
  421 ggtgtgc tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
  481 tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
  541 ggtgtgc tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
  601 ggtgtgc tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
  661 ggtgtgc tggagagaagt gcagctgctg  ggtgtgc tggagagaagt gcagctgctg
```
Chryseobacterium solanelllicola strain DSM 17072 DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ011912.1

**FASTA**

```plaintext
LOCUS    HQ011912.1    771 bp    DNA    linear    DCT 29-SEP-2006

DEFINITION    Chryseobacterium solanelllicola strain DSM 17072 DNA gyrase subunit B (gyrB) gene, partial cds.

ACCESSION    HQ011912.1

VERSION    1.0 07-FEB-2007

SOURCE    Chryseobacterium solanelllicola

ORGANISM    Chryseobacterium solanelllicola

REFERENCE    1 (bases 1 to 771)

AUTHORS    Kistler, K.C., Battles, A.M. and Newman, J.D.

TITLE    Multilocus sequence typing to delineate different species among the Flavobacteriaceae

JOURNAL    Unpublished

REFERENCE    2 (bases 1 to 771)

AUTHORS    Kistler, K.C., Battles, A.M. and Newman, J.D.

TITLE    Direct Submission

JOURNAL    Submitted (09-AUG-2006) Biology Department, Lycoming College, 700 College Drive, Williamsport, PA 17701, USA

FEATURES

**Location/Qualifiers**

<table>
<thead>
<tr>
<th>Location</th>
<th>Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1..771</td>
<td></td>
</tr>
</tbody>
</table>

**Description**

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gyrB</td>
</tr>
</tbody>
</table>

**CDs**

<table>
<thead>
<tr>
<th>Start</th>
<th>Stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>771</td>
</tr>
</tbody>
</table>

**Protein**

<table>
<thead>
<tr>
<th>Protein ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP_696097.1</td>
</tr>
</tbody>
</table>

**GenBank**

<table>
<thead>
<tr>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:207592555</td>
</tr>
</tbody>
</table>

**Organ**

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>gyrB</td>
<td>1..771</td>
</tr>
</tbody>
</table>

```
Chryseobacterium soli strain DSM 19298 DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ11913.1

**LOCUS** HQ11913

**DEFINITION** Chryseobacterium soli strain DSM 19298 DNA gyrase subunit B (gyrB) gene, partial cds.

**ACCESSION** HQ11913

**VERSION** 1

**SOURCE** Chryseobacterium soli

**ORGANISM** Chryseobacterium soli

**REFERENCE** 1

**AUTHORS** Failor, K.C., Batters, A.M. and Newman, J.D.

**TITLE** Multilocus sequence typing to delineate different species among the Flavobacteriaceae

**JOURNAL** Unpublished

**FEATURES**

- **source**: 1..672
  - **mol_type**: genomic DNA
  - **strain**: DSM 19298
  - **db_xref**: taxon:349361
  - **PCR_product**: fwd: 1..672, rev: 673-1

- **gene**: gyrB
  - **product**: DNA gyrase subunit B

- **CDS**:
  - **gene**: gyrB
  - **translation_table**: 11
  - **protein_id**: LN65268.1

**ORIGIN**

```
1  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
31  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
61  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
91  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
121  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
151  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
181  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
211  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
241  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
271  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
301  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
331  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
361  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
391  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
421  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
451  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
481  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
511  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
541  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
571  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
601  tctgtctgg gggagcagc tcatcatt acaatcaag aggagagatt
631  atagcata atgaaggtg ggtatttgg cggcggcct cggcggcct cggcggcct cggcggcct
```

//
Flavobacteriaceae bacterium JJC DNA gyrase subunit B (gyrB) gene, partial cds

GenBank H0011914.1

FASTA  Graphics

LOCUS       H0011914.1  S15 bp  DNA   linear  20-SEP-2010
DEFINITION  Flavobacteriaceae bacterium JJC DNA gyrase subunit B (gyrB) gene, partial cds.
ACCESSION  H0011914
VERSION     H0011914.1  GI:807592561
SOURCE      Flavobacteriaceae bacterium JJC
ORGANISM    Flavobacteriaceae bacterium JJC
            Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae.
REFERENCE   1 (bases 1 to 515)
AUTHORS     Fauri, E.C., Bacteria, A.M. and Newman, J.D.
TITLE       Multilocus sequence typing to delineate different species among the Flavobacteriaceae
JOURNAL     Unpublished
REFERENCE   2 (bases 516 to 516)
AUTHORS     Fauri, E.C., Bacteria, A.M. and Newman, J.D.
TITLE       Direct Submission
JOURNAL     Submitted (01-AUG-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA
FEATURES   Location/Qualifiers:
            source 1..515
            /organism="Flavobacteriaceae bacterium JJC"
            /strain="JJC"
            /db_xref="taxon:512012"
            /PCR_primer="fwd_name: 330aaf, fwd_seq: ccttygargncacnana, rev_name: 600aaf, rev_seq: tctctynconmcrnocytyrtaa"
            gene <1..515
            /gene="gyrB"
            /codon_start=1
            /transl_table=11
            /product="DNA gyrase subunit B"
            /protein_id="ATP6976.1"
            /db_xref="GI:807592562"
            /translation="ARQAARAEACELYHRKSMCOC2LGC8LSCS8KEPLASFLVHE
GDBAGTQAKQRHRQFQLPLRMLINVEKSMHKCVYNEEIKNYTIALGVSUGT
DEVALNITKLHHRKWVMTDADIODGHS1LIIITLIFPPFRYREELIENGYIIAQPFLYL
LKRSNKNVYTA""

ORIGIN
1 gcaaggacq cagctaaaaa agcaggaata cctgttctca gaaattctcc atcggtcagg
11 agggattc caggyctat cctcgacgt gccttcaag acggcgaat tccagcatt
21 ttcctttac cagggattc cagggattc agcaggaata cggcagcttc cggcctttcc
31 cggcagttc aagttgatc caagatatt ttactgtgct tggcggttct tggcgtgttct
41 cgagctgtt gcaagcctg aatcattc gatattcgt gattttggtg tttctttcct
51 ttcgtttag ggtcctggt taatcattc gatattcgt gtttctttct ccaatattttc
tctctgcg taa

//
Sphingobacterium anhuiense DNA gyrase subunit B (gyrB) gene, partial cds

GenBank: HQ011915.1

**LOCUS** HQ011915

**DEFINITION** Sphingobacterium anhuiense DNA gyrase subunit B (gyrB) gene, partial cds.

**ACCESSION** HQ011915

**VERSION** HQ011915.1 GI:207502562

**KEYWORDS** .

**SOURCE** Sphingobacterium anhuiense

**ORGANISM** Sphingobacterium anhuiense

**REFERENCE** 1 (bases 1 to 708)

**AUTHORS** Failor,K.C., Battles,A.M. and Newman,J.D.

**TITLE** Multilocus sequence typing to delineate different species among the Flavobacteriaceae

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 708)

**AUTHORS** Failor,K.C., Battles,A.M. and Newman,J.D.

**TITLE** Direct Submission

**JOURNAL** Submitted (03-AUG-2010) Biology Department, Lycoming College, 700 College Place, Williamsport, PA 17701, USA

**FEATURES**

**source**

```
1..708
/organism="Sphingobacterium anhuiense"
/sol_type="genomic DNA"
/strain="CRC"
/db_xref="taxon:49370"
/PCR_primer="fwd name: 350aaF, fwd_seq: cacttygarggncarcanaa, rev_name: 600aaR, rev_seq: twctstcctctcnsncctcctta"
```

**gene**

```
<1..708
'/gene="gyrB"
```

**cds**

```
<1..708
'/gene="gyrB"
'/product="DNA gyrase subunit B"
'/protein_id="ABM66970.1"
/db_xref="GI:207502564"
/transl_table=11
/translation="TGLNHEVMDAVVAGVEILLNYLENEFREAKIYQKQITVIAQAARAAKAKRQDLQKTVDMMGG6LPGKLDLCDQNDPPKCIYTFEGOSAGTARKSERDRKVGAMPL6RILVPERMAKYYENIKIMTCLGAVLGFDPFLARLEKLAYHRIMDAV60SHKTTTILTYFPYNNKLELREKGYTYATHPFLVLKVQDKHRQAWTEDQ
RINHKBQMBGSGKREDSVH"
```

**ORIGIN**

```
1 aacacttgg gccactctgg gctgctggag tgcagttgg agaccttttt
118 aacgtatatt tgaagccgaa tcttccgca gcaaaagaac aagccaaac
d 61 aatctattt tgaagcagaac tcttccgca gcaaaagaac aagccaaac
d 121 gcactcagc ctctctcgcc gctccattaa ggctcctgaaatc
d 181 atcygctcc cagctcttt gcattaggg ccagatcgg ggcctctcgaaatc
d 241 tgtgagatct acatcttgc ggcacctgg gctgctggag tgcagttgg agaccttttt
d 301 gcagctcatt tgcatttgc ggcctctctg gcaaaagaac aagccaaac
d 361 gcactcagc ctctctcgcc gctccattaa ggctcctgaaatc
d 421 tgtgagatct acatcttgc ggcacctgg gctgctggag tgcagttgg agaccttttt
d 481 atcygctcc cagctcttt gcattaggg ccagatcgg ggcctctcgaaatc
d 541 gcagctcatt tgcatttgc ggcctctctg gcaaaagaac aagccaaac
d 601 gcagctcatt tgcatttgc ggcctctctg gcaaaagaac aagccaaac
d 661 atcygctcc cagctcttt gcattaggg ccagatcgg ggcctctcgaaatc
```

//