User Guide to Assignment of rpoB or sigB allelic types for Bacillales or Listeria isolates

FILE NAME: rpoB_sigB_ATassignment_User.doc

Authored by: Steven Warchocki

Last Modified on: 11/17/2014 by Steven Warchocki

Approved by: Martin Wiedmann

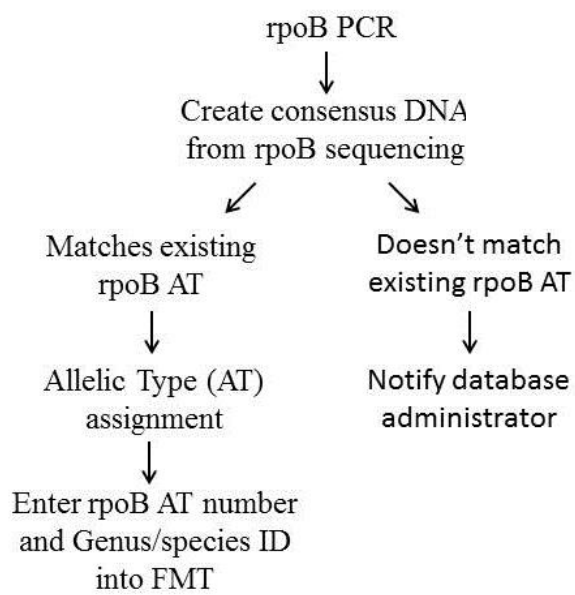EFFECTIVE DATE: Date of Approval

APPROVED BY:

_____   _____

Dr. Martin Wiedmann   (date)

AUTHORED BY:

_____   _____

Steven Warchocki   (date)

```
                    rpoB PCR
                       ↓
             Create consensus DNA
              from rpoB sequencing
                ↙              ↘
     Matches existing      Doesn't match
         rpoB AT          existing rpoB AT
            ↓                   ↓
     Allelic Type (AT)     Notify database
        assignment          administrator
            ↓
   Enter rpoB AT number
   and Genus/species ID
         into FMT
```

Figure 1: Flow Chart of steps needed to provide isolate with an rpoB sequence type number (AT number) and Genus/species identification. sigB can take place of rpoB in chart.

TABLE OF CONTENTS

SECTION 1   INTRODUCTION

1.1   Purpose

To determine rpoB or sigB sequence allelic types (AT) of bacteria isolated in the Milk Quality Improvement Program (MQIP) and Food Safety Laboratory (FSL). rpoB/sigB sequences obtained through DNA sequencing of PCR products will be compared to an existing local database of unique rpoB or sigB ATs through Food Microbe Tracker (FMT).

1.2   Scope

This SOP applies to the MQIP and the FSL.

1.3   Definitions

AT: allelic type; defined as one specific DNA sequence of (or within) a gene, in this case a 632 nucleotide region of the rpoB gene in Bacilleaceae. With *Listeria*, a 660 nucleotide region of the sigB gene.

bp: base pair

BLAST: Basic Local Alignment Sequence Tool

Consensus: a single sequence derived from a set of overlapping DNA segments originating from one genetic source

FMT: Food Microbe Tracker; WWW-based tool for information exchange on bacterial subtypes and strains, containing a large amount of bacterial gene information

PCR: Polymerase Chain Reaction, used to amplify a specific region within a DNA sequence.

Percentage sequence identity: proportion of identical nucleotides between two sequences multiplied by 100.

Phylogeny – The evolutionary history of taxonomic groups.

rpoB: RNA polymerase beta subunit

sigB: sigma factor that coordinates the stress response of *Listeria* species

SECTION 2   MATERIALS

- **Computer**

- **Partial rpoB or sigB gene sequences:** From PCR products of isolates; sequences are obtained after PCR products are sent to the BRC facility (Biotechnology Research Center). Raw data files are in .ab1 format and edited consensus sequences are saved as .fas files

- **Internet Access:** For accessing Food Microbe Tracker.

- **Local rpoB or sigB database via Food Microbe Tracker** This can be accessed by logging into FMT, searching by DNA sequence and selecting rpoB allelic typing or sigB allelic typing.

SECTION 3   PROCEDURES

3.1 rpoB/sigB Allelic Type Assignment with Food Microbe Tracker

If you're working with rpoB sequences, ignore all sigB references. If you're working with sigB sequences, ignore all rpoB references.

1. Obtain edited (final) rpoB or sigB sequence data (see *CreatingConsensusDNA.doc*). These sequences should be minimally 632bp in length for rpoB and 660bp in length for sigB, but are often longer.

2. Open website for Food Microbe Tracker: http://www.foodmicrobetracker.com
   a. Log-in, or request account for Log-in.

3. In Food Microbe Tracker (FMT), on the left-hand side of the main page, under "Search By", click on "DNA Sequence".

   b. Once on this page use the pull-down menus to adjust your search parameters:
      i. "Number of Results", *default=10*, you may wish to increase/decrease this.
      ii. "Genus", *default=Unspecified*, leave this.
      iii. "Species", *default=Unspecified*, leave this.
      iv. "Sequence Type", *default=Unspecified*, this must be changed to "rpoB allelic typing" or "sigB allelic typing" in the pull-down menu.

4. Open the rpoB or sigB consensus (final) sequence file (.fas) you wish to find an allelic type for. This can be done in either Notepad or Sequencher.

5. Copy and paste your rpoB or sigB sequence into the space labeled "Enter DNA sequence".

6. Click Submit.
   a. Once your results page ("Search Results from DNA Sequence Search") has loaded choose the first Alignment file by clicking on "See Report" in red.

7. When new page/tab has appeared click to view it and review some key details:
   a. "Identities", this should read 632/632 (rpoB) or 660/660 (sigB) for a 100% allelic type match. Unique AT sequences will have less than a 100% AT match.
   b. "Sbjct", which is the rpoB or sigB allelic type sequence your entered "Query" is being compared to, should start at 1 and end at 632 (660 for sigB). (See Troubleshooting section for sequences showing all nucleotides 1 through 635bp, or 635 bp total).

c. If there is no 100% match (out of 632bp for rpoB and 660 bp for sigB) for "Identities" you may have a new rpoB/sigB allelic type. Check in Sequencher if the differences between your sequence and the sequence in the database are legitimate, i.e., they are not artifacts introduced during editing. If the differences are real, send an email to the database manager, currently Dave Kent (dk657@cornell.edu) for rpoB and Steven Warchocki (skw59@cornell.edu) for sigB. Additionally, add raw and consensus rpoB sequence (final) data to FOOD-MQIP→rpoB Database→Possible new ATs. Add raw and consensus sigB sequence (final) data to BoorWiedmannLab→LAB STUFF→sigB allelic types→Potential New sigB Allelic Types. If more than one new rpoB/sigB AT are found, a list or file of consensus sequences may be compiled to facilitate easy data transfer.

8. If you have confirmed a 100% match between an <u>rpoB/sigB allelic typing sequence</u> currently listed in FMT and your query sequence, return to the search results page and click on the "Bacterial ID" in red. When directed to the isolate information page scroll down to "Additional Characteristics" and look for the "rpoB allelic type" or "sigB allelic type" number in bright red. This number is the allelic type of the sequence you have searched.

9. The rpoB/sigB allelic type number may be entered into FMT for your isolate by going to the information page for your isolate, scrolling down to "Additional Characteristics" and clicking "Edit" and entering the correct AT number in the box for "rpoB allelic type" or "sigB allelic type". For 100% matches, Genus/species information from the type strain should also be added to your query strain's FMT information.

10. At this time rpoB/sigB sequence data (raw and consensus) must also be added to Food Microbe tracker (see step 10 below). NOTE: *Only sequences with both "rpoB allelic type" and "rpoB allelic typing" sequences are representative rpoB type strains. This means <u>you will never enter "rpoB allelic typing sequence" data; this is only done by database managers</u>. The same is true with sigB.* In order to add sequence data for an isolate, open its page, scroll down to "DNA sequences" and click "Add". On the "Add DNA Sequence page", under "Type" select "rpoB" or "sigB". Where "Sequence" is listed, use the "Browse" button to upload your consensus (final) sequence (.fas). Under "Raw data files" add the corresponding two raw files (AB1, for both Forward and Reverse). Click "Submit".

11. As a check, search FMT for all entries with the same rpoB AT or sigB AT number you have found for your rpoB/sigB sequence (You can do this using the advanced search option, under "additional characteristics" and using quotation marks around the AT number). If all Genus/species identifications match, then leave your new entry alone, if there are discrepancies, email the database manager with the rpoB AT number.

SECTION 4   TROUBLESHOOTING

4.1 If "Identities" in your results read anything but out of 632 (e.g. 630/630 or 626/631), the BLAST algorithm has somehow trimmed your sequence. First check the second best match, if that one isn't out of 632 either, then it is best to pull out rpoB sequences from each isolate's FMT page and align them (ClustalW or Mesquite can do this) and see if the complete length matches for 632bp. The same should be done for sigB, which should be out of 660 bp.

    4.1.1 Matches showing 635/635 usually indicate a match with a *Staphylococcus sp*. Close attention should be paid to these isolates. Record percentage identities to the best rpoB AT match. Report to database manager, or Martin Wiedmann. These sequences may be added to the database so that these Genera can be identified, however if you find many of these in your project there may be a breakdown in laboratory methods that may need to be investigated.