FINAL REPORT

Pretesting of molecular identification tests for *Bactericera cockerelli* (Šulc, 1909)

Contents

1. Introduction

The European Reference Laboratory for Insects and Mites has to select, adapt or develop reliable identification tests for the phytosanitary insect and mite species that are relevant for the European Union (included in the Commission Delegated Regulation (EU) 2019/1702 and in the EURL for Insects and Mites working programmes). One of the tasks of the EURL is to validate available diagnostic protocols before recommending their use to the National Reference Laboratories of the European Union. Pretesting of available tests is necessary to select the most reliable ones for the validation study.

Fig. 1 *Bactericera cockerelli* adult (Pest and Diseases Image Library, Bugwood.org)

The potato psyllid *Bactericera cockerelli* (Fig. 1) is of interest to the European Union as a vector to the bacterium '*Candidatus* Liberibacter solanacearum' (haplotypes A, B, F). *B. cockerelli* originates from North America, with its four biotypes correlating to certain regions of the western part of the United States. These biotypes are associated with polymorphisms of the mitochondrial cytochrome c oxidase subunit I (*COI*) gene, which is typically used for barcoding. The internal transcribed spacer region 2 (*ITS2*) contains enough interspecific variation to distinguish *B. cockerelli* from closely related species and seems to contain less intraspecific variation than the *COI* gene locus.

In addition to North America, *B. cockerelli* has been found in parts of Oceania. Currently it is absent in Europe, but due to its polyphageous nature and climatic requirements, establishment in parts of Europe could be possible (current distribution see Fig. 2).

B. cockerelli is a European Union regulated species, listed among the EU quarantine pests (Annex II of the Commission Implementing Regulation (EU) 2019/2072) and among the EU priority pests (Commission Delegated Regulation (EU) 2019/1702).

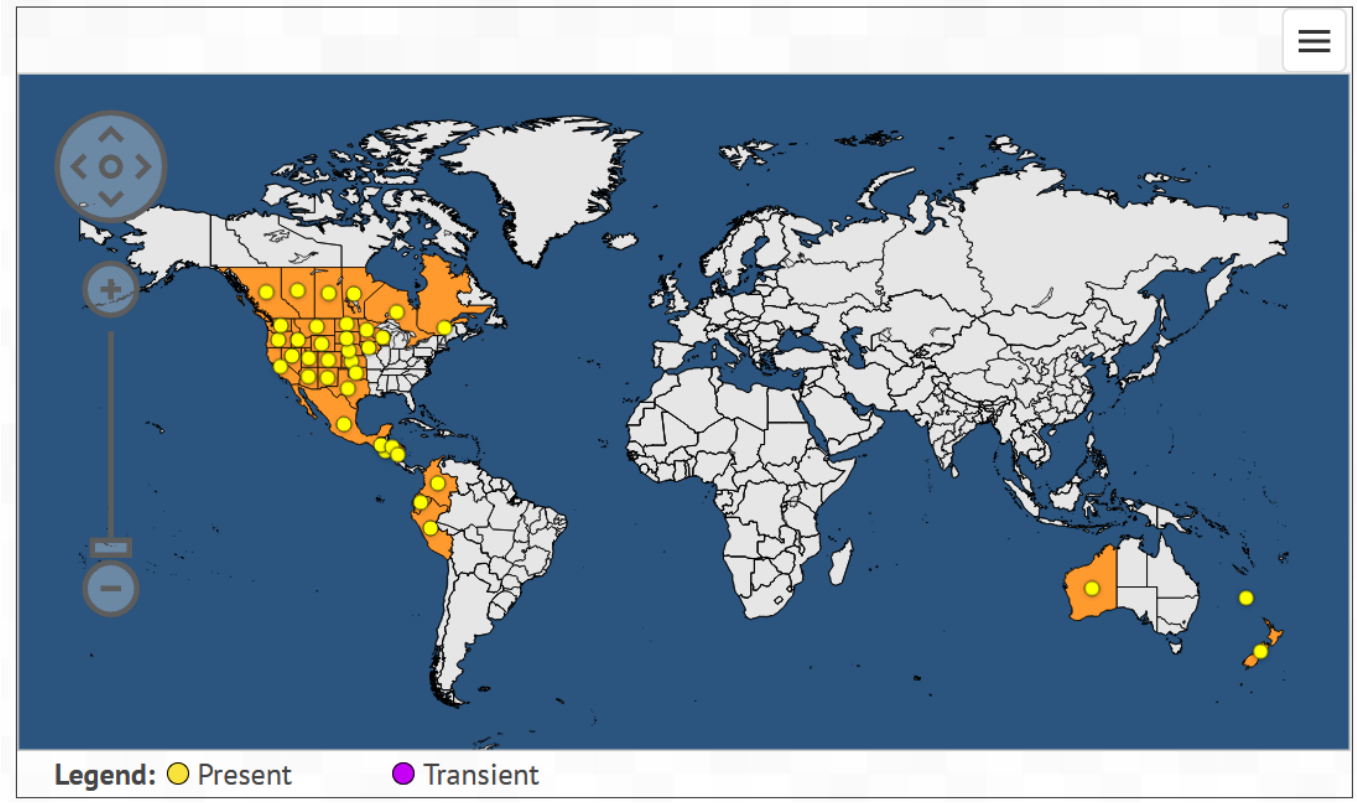


Fig. 2: Distribution of *B. cockerelli* according to the EPPO global database (query date January 2023)

2. Scope of pretesting

The scope of this preliminary study was to review available molecular tests which are appropriate for the identification of *B. cockerelli*. Additionally, a database inventory for sequence records should led light on the application possibilities of barcoding as identification method.

3. Literature review

For this pest species no published specific diagnostic protocols in an international standard are available yet. However, *B. cockerelli* is mentioned in diagnostic protocols of '*Candidatus* Liberibacter solanacearum' as its vector.

However, identification of insect quarantine pests is covered in the EPPO PM 7/129 (2) DNA barcoding as an identification tool for a number of regulated pests (EPPO, 2021) (Appendix 1), which includes tests for the DNA barcoding of arthropods in general. Additionally, Sumner-Kalkun *et al.* (2020) describe a *B. cockerelli*-specific real-time PCR.

4. Evaluation of available tests

Following tests were selected for *in silico* pretesting (details of which can be found in Appendix 2):

- Real-time PCR test based on Sumner-Kalkun *et al*. (2020) targeting the *ITS2* region
- Barcoding according to EPPO PM 7/129(2)

5. Database inventory for sequence records

Frequently used DNA barcoding markers for the discrimation of insect species are loci such as the cytochrome C oxidase subunit 1 (*COI*) or the internal transcribed spacer region 2 (*ITS2*). DNA barcoding relies on comparison of those sequences to a database of reference sequences (Armstrong and Ball 2005). Applying barcoding for insect identification requires enough sequence records from the species within the genus for a reliable comparision. Not only the number, but also the genetic and geographic diversity of the records and the quality are potential issues that should be taken into account.

Three different databases (NCBI GenBank, Bold and EPPO Q-Bank) were consulted for the inventory. As search parameters the genus and species name and the gene locus (*COI* and synonyms) were used. If no records for the *COI* locus were available, the search was extended to other gene loci. In addition the reliability of the records were checked.

5.1 Results

At the time of query, only NCBI GenBank and Bold have sequence records available for *B. cockerelli* on the *COI* locus (55 and 53, respectively), while none have been deposted in EPPO Q-Bank. Additionally, GenBank also offered 52 records on the *ITS2* locus and 123 sequence records for other genes are present (query date October 2022). It has to be noted, that the presence of 13 whole mitochondrial genome as well as whole genome shotgun sequencing records indicates possible additional barcoding success with single gene loci sequences.

Table 1: Number of sequence records per gene for each database (query date October 2022)

Gene	GenBank	Q-Bank	Bold
COI	55	0	53
COII	0	0	0
ITS2	52	0	0
18S	1	0	0
28S	2	0	0
60S	1	0	0
complete genome (mitochondrium)	13	0	0
whole genome shotgun sequencing	15	0	0
other genes	123	0	0

4.2 Detailed information

The Bactericera genus encompasses a high species diversity (123 species, assessed October 2022, Wikipedia https://en.wikipedia.org/wiki/Bactericera). Bold database holds 20 species, nine of these with barcodes. For *B. cockerelli* 53 public records are available in Bold, 45 of which are mined from GenBank. These specimens were collected mainly in the western part of the United States, as well as single collection sites in Canada, Australia and New Zealand. 33 *B. cockerelli* records had sequences of at least 500 bp.

4.3 Tree-based identification

To evaluate the species divergences within the genus, Neighbor Joining (NJ) trees of distance were constructed using NCBI GenBank (max. seq. diff. of 0.75) on both the *COI* and *ITS2* locus.
Fig 3 revealed *B. cockerelli* clustering with one entry for *B. maculipennis* and some *Psyllopsis discrepans,* whereas other sequences of the latter do not fall into this cluster.
On the *ITS2* locus, however, no such irregularities were observed (Fig. 4).

Fig 3: Neighbor Joining tree of distance for the *COI* locus, NCBI Genbank

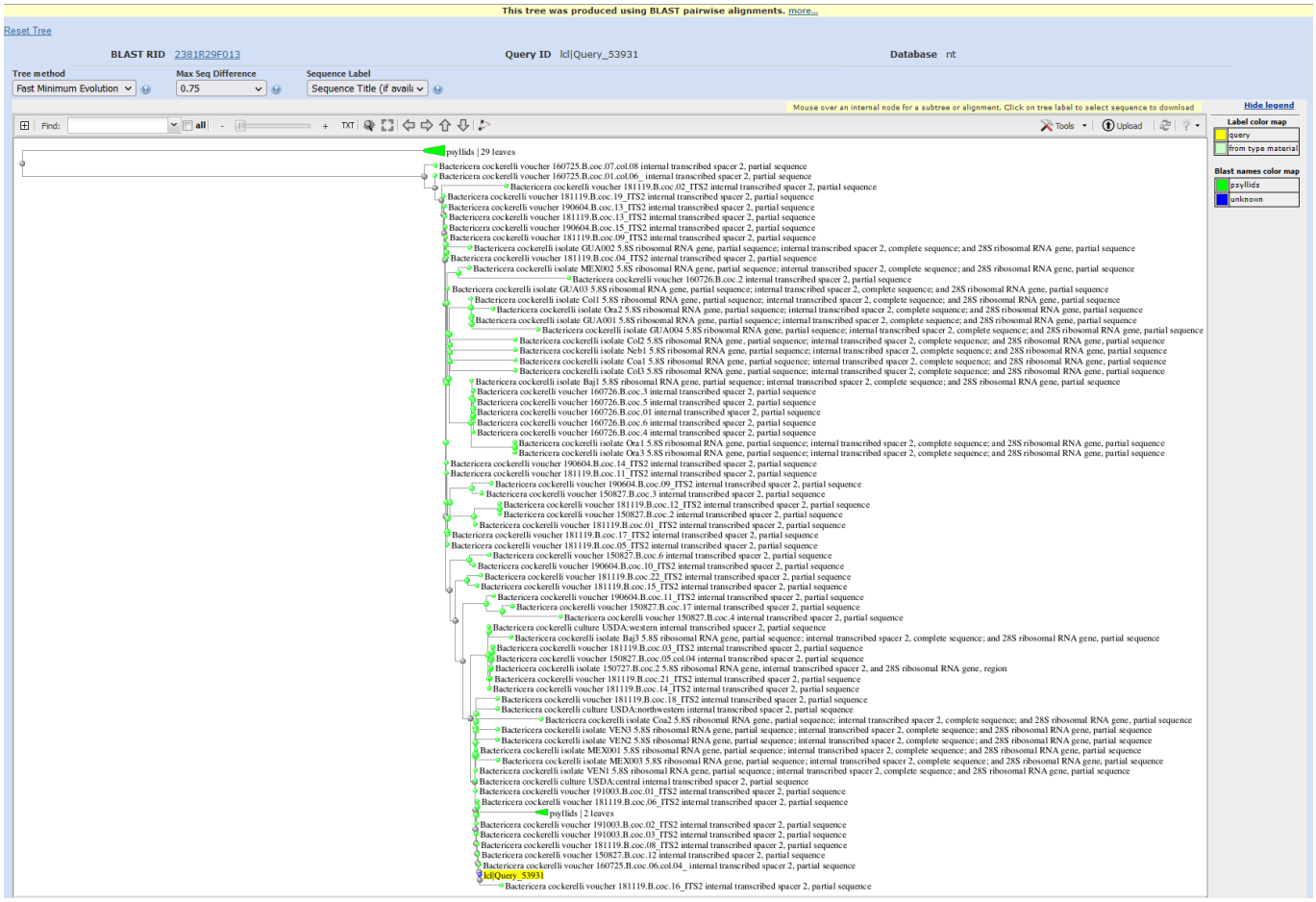


Fig 4: Neighbor Joining tree of distance for the *ITS2* locus, NCBI Genbank

6. Results

For the *COI* gene locus a sufficient number of sequences has been deposited in the Bold and GeneBank database. However, *in silico* data of this pretesting suggests that issues might occur during barcoding on the *COI* locus due to one *B. maculipennis* sequence as well as four *Psyllopsis discrepans* sequences clustering with *B. cockerelli* on this locus. Primer blast of the LepF/R primer pair according to Hajibabaei *et al*. (2006) with restricting the search to *B. cockerelli* suggests suitability of identification with this test.
Investigating tree-based identification on the *ITS2* locus revealed sufficient clustering of *B. cockerelli* that is not interspersed with any other psyllid species. *In silico* data for the real-time PCR according to Sumner-Kalkun *et al*. (2020) targeting the *ITS2* locus indicates it as suitable for identification of *Bactericera cockerelli*. For more details of the *in silico* data see App. 2.

7. Discussion

Molecular identification of *B. cockerelli* investigated in this pretesting (*COI* barcoding, *ITS2* sequencing, TaqMan real-time PCR) is highly promising based on database inventory of sequence records and *in silico* data.
However, even though there are indiciations that the *COI* locus is suitable for identification of *B. cockerelli*, it is recommended to proceed with caution and to employ tree-based identification.
In silico trees on the *ITS2* locus suggest the suitability of *ITS2* sequencing for identification of *B. cockerelli* as well, as the closely related species in this genus seem to cluster separately from each other, indicating sufficient interspecies variability.
Both *in silico* data and the validation data given in the publication by Sumner-Kalkun *et al.* (2020) demonstrates this TaqMan-based real-time PCR test as specific for *B. cockerelli*.

Given the acquisition of target and non-target psyllid samples in the future, a validation study would compare barcoding of the *COI* locus and *ITS2* sequencing as well as the TaqMan real-time PCR according to Sumner-Kalkun *et al*. (2020).

Participating staff:

Experts/ Supervisors: Richard Gottsberger, Helga Reisenzein
Technical staff/ Operators: Chiara Pohn

Date: December 13th, 2022

Helga Reisenzein
EURL Deputy Director

Richard Gottsberger
Scientific Project Leader – Molecular Unit

Appendix 1 - References

EPPO (2021). EPPO standards PM 7/129 (2) DNA barcoding as an identification tool for a number of regulated pests. Bulletin OEPP/EPPO Bulletin, 51 (1): 100–143.

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. (2006). DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences, 103(4), 968-971.

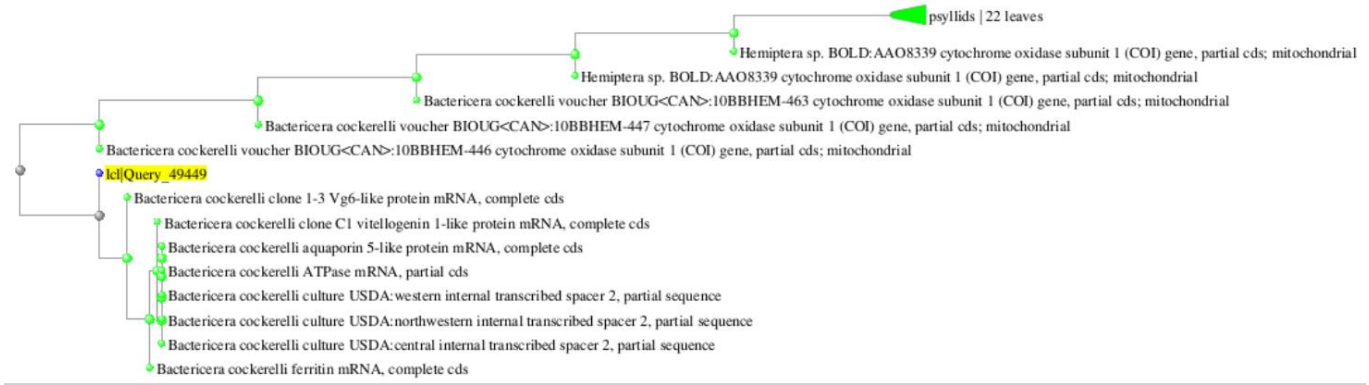
Sumner-Kalkun, J. C., Sjölund, M. J., Arnsdorf, Y. M., Carnegie, M., Highet, F., Ouvrard, D., ... & Kenyon, D. M. (2020). A diagnostic real-time PCR assay for the rapid identification of the tomato-potato psyllid, Bactericera cockerelli (Šulc, 1909) and development of a psyllid barcoding database. Plos one, 15(3), e0230741.

Armstrong, K. F., & Ball, S. L. (2005). DNA barcodes for biosecurity: invasive species identification. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1462), 1813-1823.

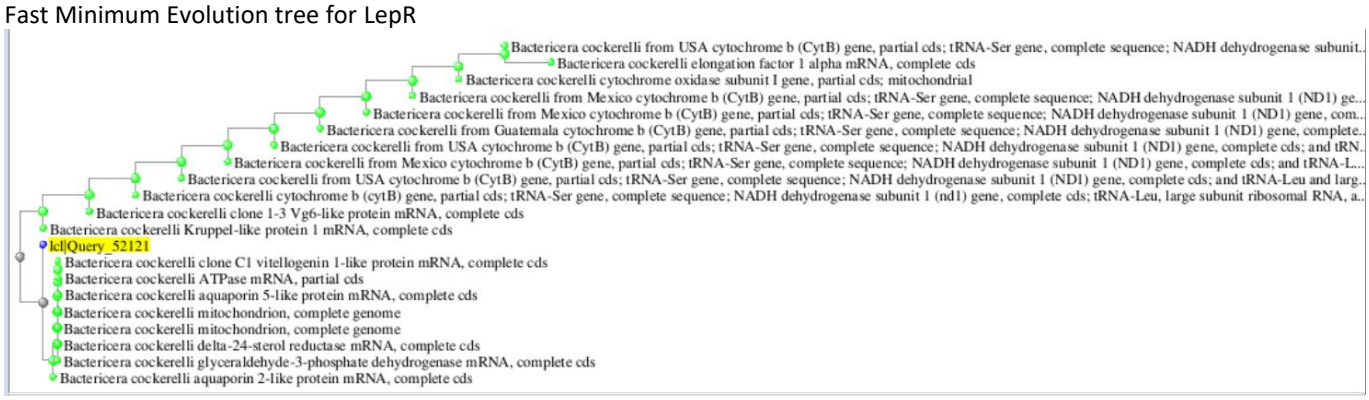
Appendix 2 – *In silico* data

Barcoding according to EPPO PM7/129(2): Search was restricted to *Bactericera cockerelli*

Fast Minimum Evolution tree for LepF



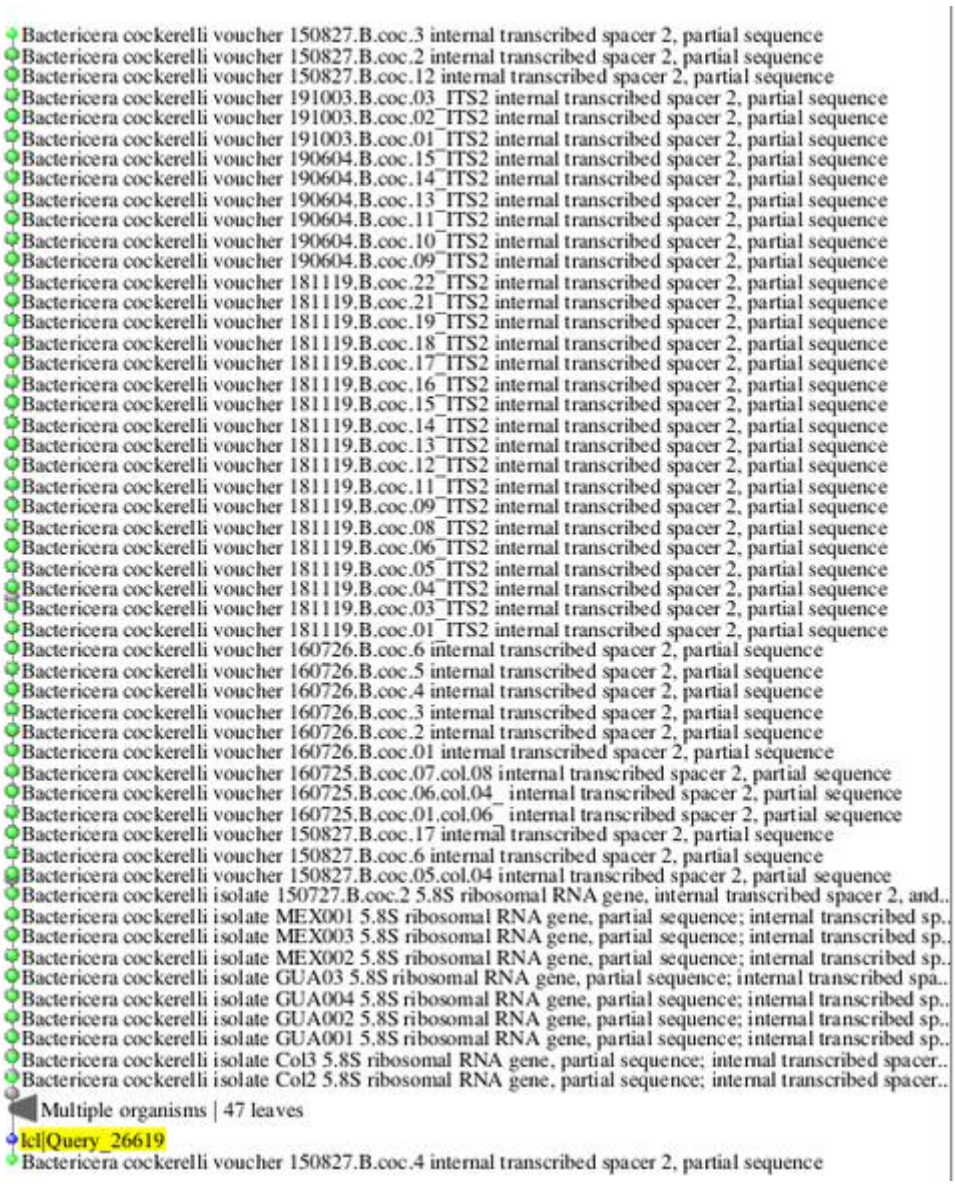
Fast Minimum Evolution tree for LepR



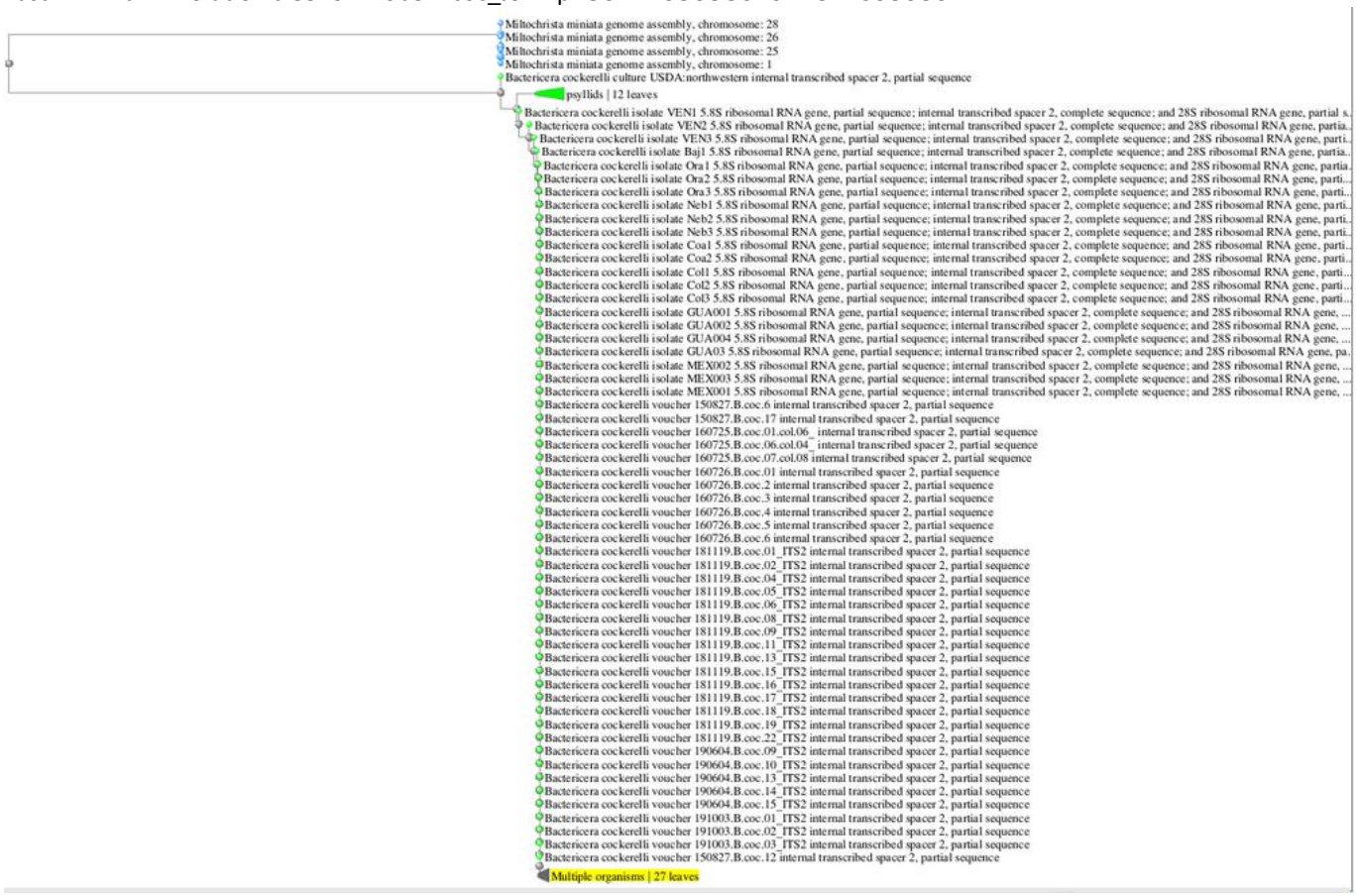
TaqMan real-time PCR according to Sumner-Kalkun *et al*. (2020), expected amplicon: 187bp

Fast Minimum Evolution tree for Bcoc_JSK2-f: GAGGTCTCCTCATCGTGCGT

Fast Minimum Evolution tree for Bcoc_JSK2-r: GGACGAGCATTGCTGCTGC



Fast Minimum Evolution tree for Probe: Bcoc_JSK2-p: GCAAACGCGGCACAAGTACCGCGC



In silico PCR Bcoc_JSK2_f/r

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GAGGTCTCCTCATCGTGCGT	20	61.66	60.00	5.00	0.00
Reverse primer	GGACGAGCATTGCTGCTGC	19	62.39	63.16	7.00	5.00

Products on target templates

>MW641991.1 Bactericera cockerelli culture USDA:western internal transcribed spacer 2, partial sequence

```
product length = 187
Forward primer  1    GAGGTCTCCTCATCGTGCGT  20
Template        638  ....................  619

Reverse primer  1    GGACGAGCATTGCTGCTGC  19
Template        452  ...................  470
```

>MW641990.1 Bactericera cockerelli culture USDA:northwestern internal transcribed spacer 2, partial sequence

```
product length = 187
Forward primer  1    GAGGTCTCCTCATCGTGCGT  20
Template        638  ....................  619

Reverse primer  1    GGACGAGCATTGCTGCTGC  19
Template        452  ...................  470
```

>MW641989.1 Bactericera cockerelli culture USDA:central internal transcribed spacer 2, partial sequence

```
product length = 187
Forward primer  1    GAGGTCTCCTCATCGTGCGT  20
Template        638  ....................  619

Reverse primer  1    GGACGAGCATTGCTGCTGC  19
Template        452  ...................  470
```

>MT027598.1 Bactericera cockerelli voucher 150827.B.coc.3 internal transcribed spacer 2, partial sequence

```
product length = 187
Forward primer  1   GAGGTCTCCTCATCGTGCGT  20
Template        26  ....................  45

Reverse primer  1    GGACGAGCATTGCTGCTGC  19
Template        212  ...................  194
```