

**Elucidating the 16S rRNA 3' boundaries and defining optimal SD/aSD pairing in *Escherichia coli* and *Bacillus subtilis* using RNA-Seq data**

Yulong Wei<sup>1†</sup>, Jordan R. Silke<sup>1†</sup>, and Xuhua Xia<sup>12\*</sup>

1. Department of Biology, University of Ottawa, 30 Marie Curie, P.O. Box 450, Station A, Ottawa, Ontario, Canada, K1N 6N5. Tel: (613) 562-5800 ext. 6886, Fax: (613) 562-5486.
2. Ottawa Institute of Systems Biology, Ottawa, Ontario, Canada K1H 8M5.

† Equal contribution.

\* Corresponding author: E-mail: [xxia@uottawa.ca](mailto:xxia@uottawa.ca)

## Supplementary Information

```

          10          20          30          40          50          60          70          80
SeqGr119948_1  -----UUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCCC-
SeqGr1450749_4  -----UUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC--
SeqGr1496175_3  -----UUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCCC-
SeqGr3193684_1  --AAGGUGGGACAGAUGAUUGGGGUGAAGUCGAAACAAGGUAGCCGUUACGGAAAGGUGCGGCUGGAUCACCUCCUUU-----
SeqGr3837735_1  --UUGGUGGGACAGAUGAUUGGGGUGAAGUCGAAACAAGGUAGCCGUUACGGAAAGGUGCGGCUGGAUCACCUCCUUU-----
SeqGr3961119_1  ---CUUUUGAUUCAUGAAUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr452173_1   ---UUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr4611819_1  CCACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCU-----
SeqGr46365_22   ---CUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr509358_11  --ACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCU-----
SeqGr69143_4    --CACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCC---
SeqGr134077_56  ---CUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr144116_27  --ACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCU-----
SeqGr1815518_1  -CACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCC-----
SeqGr1970881_3  CCACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCU-----
SeqGr2342583_1  --ACAUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCU-----
SeqGr2989770_1  -----UGUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCCCC
SeqGr3159133_1  ---UUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr3279744_1  ---CUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGUGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr3543924_1  --AUGGUGGGACAGAUGAUUGGGGUGAAGUCGAAACAAGGUAGCCGUUACGGAAAGGUGCGGCUGGAGCACCUCCUUU-----
SeqGr366483_1   ---CUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr448293_2   -----UUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCCC-
SeqGr682676_8   -----UUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC--
SeqGr78824_7    -CACUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCC-----
SeqGr841891_1   --AAGGUGGGACAGAUGAUUGGUGAAGUCGAAACAAGGUAGCCGUUACGGAAAGGUGCGGCUGGAUCACCUCCUUU-----
SeqGr851203_1   ---CUUAGUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr97624_1    ---UUUUUUUUUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr1024682_2  ---CUUUUGCAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
SeqGr1054000_2  ---UUUUUGAUUCAUGACUGGGGUGAAGUCGUAAACAAGGUAAACCGUAGGGGAAACCGCGGUUGGAUCACCUCCUCCUCC---
rrnA16S        CGAAGGUGGGACAGAUGAUUGGGGUGAAGUCGUAAACAAGGUAGCCGUUACGGAAAGGUGCGGCUGGAUCACCUCCUUUCUA-

```

Figure S1. Sequence alignment between the annotated 16S rRNA (rrnA16S, NC\_000964) and 75 nt RNA-Seq reads (SRR3466199). RNA-Seq segment IDs are in FASTA+ format (e.g. SeqGr134077\_56 represents the 134077<sup>th</sup> unique sequence, and it has 56 identical copies).