Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 27:1-11

Characterising the insecticidal capability of *Pseudomonas fluorescens* from the *Fragaria x ananassa* leaf endosphere

Ali Anwari, Adam C. Wilcockson, Christopher Yap, Fione Yip

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY *Drosophila Suzukii* is a major pest to the berry industry in the Pacific Northwest, resulting in a need for cost effective solutions that are safe for the surrounding ecosystem. Natural biocontrol strategies can offer a more efficient and pest-specific control measure than traditional pesticides. Plant symbiotic *Pseudomonas fluorescens* are known to express the insecticidal *fitD* toxin which makes them an ideal candidate for a biocontrol solution. Based on this, we hypothesise that *P. fluorescens* strains found in the fruiting bodies, leaves, and stems of strawberry and blackberry plants will express the insecticidal *fitD* gene. We isolated a bacterial strain from the strawberry leaf endosphere, named SL1 and confirmed it is a member of the genus *Pseudomonas* by 16S rRNA sequencing. Whole genome sequencing of *Pseudomonas* sp. SL1 and subsequent assembly and polishing returned a genome of 6.8 Mbps. We were unable to locate the *fitD* gene in this assembled genome using Prokka or BlastKOALA but were able to find the response regulator gene *gacA*, responsible for encoding virulence factors in *Pseudomonas* species. This process allowed for the development of a detailed approach towards isolating and characterising possible insecticidal endophytic species of Pseudomonas from above-ground plant tissues.

INTRODUCTION

The fruit fly *Drosophila suzukii*, commonly known as spotted wing drosophila (SWD), is an invasive agricultural pest which infests the fruiting bodies of plants, such as blueberries, blackberries, strawberries, raspberries, and cherries (1,2). It uses pre-ripe and ripe berries as an egg-laying reservoir, rendering them inedible and unmarketable (1,2). In BC, garden strawberries (*Fragaria x ananassa*) are a critical source of income for both farmers and the province, and are currently affected by SWD (1). Washington, Oregon and California together had an estimated $300 million loss in strawberry crops in 2008 due to SWD, suggesting a need for the protection of plants from SWD in BC (3). The Himalayan blackberry (*Rubus armeniacus*), a non-crop plant present in British Columbia is also susceptible to SWD infestation (4). As an abundant invasive species, Himalayan blackberries tend to grow in the field margins adjacent to commercial crops and are a concern due to their potential to serve as a reservoir for SWD. They can act as a refuge for SWD, allowing them to populate and infest crops in adjacent fields (4,5).

Strategies used against SWD include traps, spray insecticides, and sanitary farm practices which are costly, time-consuming and can harm plant-beneficial bacterial species (1,2). Biocontrol strategies have been shown as an efficient, less environmentally damaging solution against pests, including the use of *Pseudomonas* species against root rot pathogens in many plants and fungal infection of strawberry plants (6,7). *Pseudomonas fluorescens* is a gram-negative, symbiotic bacterium commonly associated with plants, including strawberry plants (8). It has been shown to demonstrate insecticidal capabilities against insects including the SWD-related *Drosophila melanogaster* (1,2,9). Strains of *P. fluorescens* have been shown to demonstrate pathogenicity to insects using different mechanisms of action (9). *P. fluorescens* KPM-018P, isolated from the leaves of tomato plants exhibited lethal effects against the larvae of ladybird beetles due to an increase in the enzymatic activity of chitinase and protease (10). Furthermore, *P. fluorescens* Pf-5 was found to cause dose-dependent lethality coupled with delayed metamorphosis and morphological defects in *D. melanogaster* that was dependent on the gene, *gacA*. The *gacA* gene, part of a two-component regulatory

Published Online: September 2022

Citation: Ali Anwari, Adam C. Wilcockson, Christopher Yap, Fione Yip. 2022. Characterising the insecticidal capability of *Pseudomonas fluorescens* from the *Fragaria x ananassa* leaf endosphere. UJEMI 27:1-11

Editor: Andy An and Gara Dexter, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

system, has been shown to control the expression of virulence factors in *P. fluorescens* Pf-5 and results in decreased mortality of larvae when deleted (11).

Certain strains of *P. fluorescens* possess a complex of 8 genes that confer strong insecticidal activity named the *fit* complex (12). The *fitD* toxin itself is found in the gene cluster flanked by the *fitA*, *fitB,* and *fitC* genes which have shown resemblance to putative ABC-type toxin transporters and *fitE* which shows similarities to a type I secretion outer membrane efflux protein. The other three genes named *fitF*, *fitG,* and *fitH* encode regulatory functions (12). *P. fluorescens* strains, such as the biocontrol strains CHA0 and Pf-5, have been found to contain the *fit* gene cluster which has been correlated with insecticidal activity (12,13). When injecting these two strains in the haemocoel of *Galleria mellonella*, both CHA0 and Pf-5 killed the insects within 40 hours (12). To validate the insecticidal capabilities of the *fit* gene cluster, Pechy-Tarr *et al.* introduced *P. fluorescens* Q2-87 and P3 strains to the *Galleria* larvae, both of which do not possess the *fit* gene cluster. When Q2-87 and P3 were injected into the larvae at similar concentrations to Pf-5 and CHA0, only 40% and 5% of the larvae were killed, respectively, while the *fit*-containing strains killed all the larvae within 24 hours (12). These results indicate that the presence of the *fit* gene cluster correlates with insecticidal activity and thus provides a marker to identify novel insecticidal strains.

Although insecticidal genes have been studied in certain *P. fluorescens* strains in detail, there are no studies, to our knowledge, that have isolated *P. fluorescens* directly from blackberry and strawberry plants, and annotated their genomes for insecticidal genes. Since blackberries and strawberries are in direct contact with SWD, the aim of this study was to investigate and identify *P. fluorescens* strains with the insecticidal *fit*D toxin growing endophytically within those berries and above-ground tissues to serve as a biocontrol agent against SWD and other insect pests. As *P. fluorescens* are symbiotic bacteria found in plants with insecticidal capabilities against *Drosophila melanogaster*, we hypothesize that *P. fluorescens* strains found in the fruiting bodies, leaves, and stems of strawberry and blackberry plants will express the insecticidal *fitD* gene. We performed 16S ribosomal RNA gene amplification and Sanger sequencing on a fluorescing isolate cultured from a strawberry leaf to determine its identity. We also performed whole genome sequencing using the MinION Nanopore sequencer (Nanopore, UK), which we annotated using both Prokka and BLASTKoala (14, 15). Our results show that the isolated strawberry microbe can be classified as a member of the *Pseudomonas* genus using BLAST, and that the *fitD* gene was not present in the genome.

METHODS AND MATERIALS

Sample collection and processing. Berries of different ripening stages were collected from the wild blackberry plant (*Rubus armeniacus*) at three different sites at the University of British Columbia (UBC) and were grouped by its source location (NW Marine Dr. (49.268962, -123.262166), Westbrook Building (49.264858, -123.249039), and Thunderbird parkade (49.261300, -123.243993)). Blackberries were collected in Ziploc bags and stored overnight at 4ºC. To sterilize the phylloplane and allow for the selection of endophytes, berries were rinsed with a 70% ethanol solution for 30s, a 0.5% bleach solution for 60s, and three times with sterile H_2O.

Berry samples were cut into small pieces, placed into Eppendorf tubes, crushed and mixed with sterile H_2O and sterile 100% glycerol to yield a final homogenized, 20% glycerol berry solution which were stored at -70ºC. Strawberry stems and leaves were collected from the UBC Farm and underwent the same processing as the blackberries.

Identification and purification of *Pseudomonas spp.* 8 total plant samples were chosen for the isolation of bacteria: 2 blackberry samples from each of the 3 UBC locations, 1 strawberry stem sample, and 1 strawberry leaf sample. Serial dilutions were performed by diluting 300uL of extract in 4.5mL sterile H_2O, followed by two subsequent 10x serial dilutions. 75µL of the final dilution was spread on King's B media agar plates and incubated for 2-6 days at 30ºC.

Pseudomonas protegens CHA0 and *Escherichia coli* DH5α were also grown on King's B agar plates and used as positive and negative controls for fluorescence, respectively. Plates were exposed to UV light using a BioRad Chemidoc imager (Biorad, USA) to identify

fluorescing colonies and obtain images. A handheld long-wavelength UV light was also used to verify the observations. Fluorescing colonies were restreaked onto King's B media and grown at 30ºC for 1-2 days until the colonies were apparent. To purify the colonies, this process was repeated 3 times by selecting from the most recently streaked plate, generating 4 subsequently streaked plates downstream of the original spread plate.

Genomic DNA Extraction. Single colonies from the 4th purification streak were suspended in 5 mL of LB broth (10 g tryptone, 5 g yeast extract and 10 g NaCl per 1 L of distilled H_2O) and incubated overnight at 30°C on an orbital shaker. 1 mL of overnight culture was used the following day for genomic DNA extraction according to the manufacturer's protocol using the PureLink™ Genomic DNA Mini Kit (Invitrogen, USA) and DNA was eluted using the provided elution buffer. Genomic DNA was later resuspended in water using the DNA Clean & Concentrator-5 kit (ZYMO RESEARCH, USA) and stored at -20°C to optimize sample conditions for MinION sequencing.

16S rRNA PCR. PCR was performed with Taq DNA polymerase and $(NH_4)_2SO_4$ Taq buffer following the provided manufacturer's instructions using ThermoFisher Taq DNA Polymerase, recombinant (Thermofisher, USA). The following 3 pairs of primers were used: 1) Universal 16S rRNA: 27F primer 5′-AGA GTT TGA TCM TGG CTC AG-3′ and 1492R primer 5'-TAC GGY TAC CTT GTT ACG ACT T-3′; 2) *Pseudomonas*-specific 16S rRNA: PA-GS-F primer 5′-GAC GGG TGA GTA ATG CCT A-3' and PA-GS-R primer 5′-CAC TGG TGT TCC TTC CTA TA-3′; 3) *P. fluorescens*-specific 16S rRNA: 16SPSEfluF primer 5′-TGC ATT CAA AAC TGA CTG-3′ and 16SPSER primer 5′-AAT CAC ACC GTG GTA ACC G-3′ (16,17). Universal 16S rRNA amplification products were produced with the following setup: 1) initial denaturation at 95°C for 30 sec; 2) 34 cycles consisting of denaturation at 95°C for 10 sec, annealing at 46°C for 30 sec, and extension at 72°C for 2 min; 3) final extension at 72°C for 7 min (18). *Pseudomonas*-specific 16S rRNA amplification products were produced with the following setup: 1) initial denaturation at 95°C for 5 min; and 2) 10 cycles consisting of denaturation at 94°C for 15 sec, annealing at 53°C for 30 sec, and extension at 72°C for 45 sec. This was repeated for another 25 cycles with the exception of a 5 sec increase in extension time per cycle for extension at 72°C; 3) final extension at 72°C for 10 min (18). *P. fluorescens*-specific 16S rRNA amplification products were produced with the following setup: 1) initial denaturation at 94°C for 3 min; 2) 34 cycles consisting of denaturation at 94°C for 30 sec, annealing at 47°C for 30 sec, and extension at 72°C for 1 min; 3) final extension at 72°C for 5 min (17). PCR products were run in 1.5% agarose gels and a running buffer made from 1x TAE.

Sanger Sequencing 16S rRNA. PCR products amplified by universal 16S rRNA primers were purified using the PureLink™ PCR Purification Kit (Invitrogen, USA) prior to sanger sequencing by GENEWIZ (Azenta, USA). Contigs were generated from the sequence reads obtained after Sanger sequencing using Genestudio Professional Edition.

Whole genome sequencing. Whole genome sequencing samples were prepared using the Oxford Nanopore rapid sequencing kit (SQK- RAD004, Nanopore, UK) and sequenced on a MinION™ Nanopore sequencer (Nanopore, UK) (19). The sequencing kit included: Fragmentation Mix (FRA), Rapid Adapter (RAP), Sequencing Buffer (SQB), Flush Buffer (FB), and Flush Tether (FT).

To prepare the genomic library, 0.3 µL of genomic DNA suspended in water, 3.45 µL of nuclease-free water and 1.25µL of FRA was combined in a 1.5 mL DNA LoBind tube. The mixture was then incubated at 30ºC for 1 minute and 80ºC for 1 minute using a Biorad thermocycler (Biorad, USA) and then placed on ice. 0.5 µL of RAP was added to the tube and was incubated at room temperature for 5 minutes before placing back on ice. SQB, FB and FLT tubes were vortexed and spun down, then placed on ice. 117µL of FB and 3 µL of FLT were combined in a separate 1.5 mL Eppendorf DNA LoBind tube and mixed by pipetting. 110 µL of the mixture was pipetted into the Flongle™ flow cell (Nanopore, UK), ensuring no air bubbles were added.

To prepare the sequencing mix, the vial of LB was vortexed and 11.0 µL was added along with 13.5µL of sequencing buffer to a new 1.5mL Eppendorf DNA LoBind tube. 5.5 µL of the genomic library was added as well. 25 µL of this sequencing mix was pipetted into the Flongle. The MinION was then set to a 16 hour sequencing run using MinKNOW (Nanopore, UK), and the sample was run for 4 days including the base calling step. The raw output from the MinION was then assembled using the Flye v2.9 protocol (20), and polished using Racon v1.4.(22) and Medaka v1.4.4 (21). Genome annotation was carried out using the softwares Prokka v1.12 and BLASTKoala v2.2 (14,15, 22-24). For a brief overview of this process, see Figure 1.

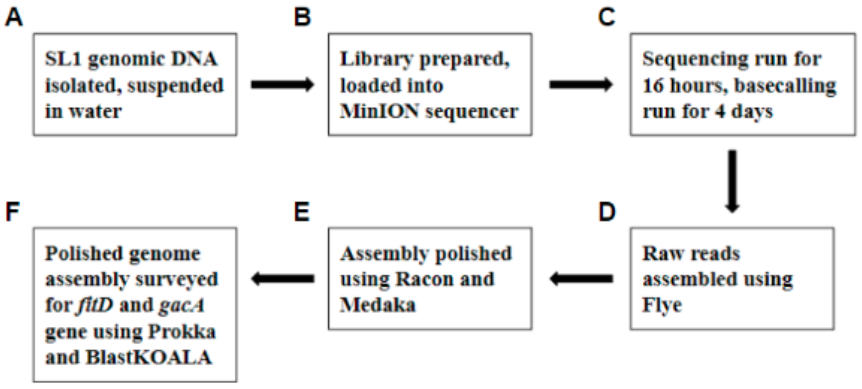

FIG. 1 Whole genome sequencing and bioinformatic process. To assess whether the SL1 isolate contained *fitD*, we had to assemble and annotate its whole genome. **(A)** First, we isolated genomic DNA using the Purelink™ and suspended it in water, which is required for proper functioning of the Nanopore ion current reader. **(B)** We then prepared the genomic library and **(C)** sequenced for 16 hours, including 4 days of basecalling using the MinION™ sequencer (Nanopore, UK). **(D)** The raw reads were then assembled using Flye and, **(E)** trimmed and polished using Racon and Medaka. **(F)** Using Prokka and BlastKOALA, the assembled genome was surveyed for the presence of the *fitD* and *gacA* genes. Although we assembled a high quality whole genome sequence using the MinION™ sequencer (Nanopore, UK), *fitD* was not found within the genome.

RESULTS

Fluorescent colonies were isolated from strawberry leaves but not strawberry stems or blackberries. In order to obtain environmental *P. fluorescens*, we cultured the endophytic material of strawberry leaves, stems, and blackberries. To identify cultured *P. fluorescens*, we utilized the bacteria's inherent ability to fluoresce under UV light when subjected to low iron conditions, as provided by King's B agar (25). The 3 types of plant samples examined were: 1) blackberries, 2) strawberry leaves, and 3) strawberry stems. Minimal growth was observed from the plated blackberry extracts. Of the blackberry extract plates that contained growth, the colonies were opaque and white or yellow in color, both of which did not fluoresce when exposed to UV light (Figure 2C). A fluorescent colony was found amongst a lawn of pale translucent hyphae (Figure 2D) on the strawberry leaf plate and was restreaked to obtain a pure colony. This fluorescing colony of interest was labeled Strawberry Leaf colony 1 (SL1) and was used for further downstream identification processes. No fluorescing colonies were found on the plates from strawberry stems (Figure 2A).

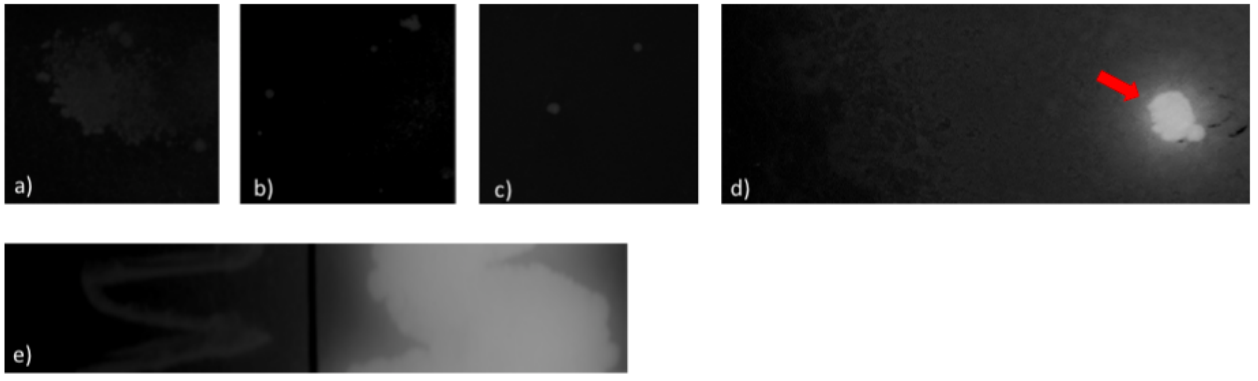

FIG. 2 Cultured strawberry stem, strawberry leaf, and blackberry extracts on King's B agar plates. The samples of negative fluorescent results are shown in the above images: **a)** cultured strawberry stem extracts at the lowest dilution; **b)** cultured strawberry leaf extract at the lowest dilution; **c)** cultured blackberry extract at the lowest dilution. **d)** Image shows cultured strawberry leaf extract at the second dilution, with visible fluorescing growth, indicated by the red arrow. This fluorescing growth sits atop a lawn of hyphanous growth, seen by slight autofluorescence in contrast with the black at the far left. **e)** image shows the negative control (*E. coli* DH5α) and positive control (*P. protegens* CHA0) for fluorescence (left and right respectively).

SL1 isolate identified as *P. fluorescens* using a PCR assay. To help determine the identity of SL1, *Pseudomonas*- and *P. fluorescens*-specific primers were utilized to amplify the 16S DNA region and gel electrophoresis was performed as an initial screening tool. *Pseudomonas*- and *P. fluorescens*-specific primers were identified from two previous papers which are capable of amplifying and recognizing regions of the 16S rRNA gene that were specific to bacteria part of the *Pseudomonas* genus bacteria or to *P. fluorescens* bacteria, respectively (16,17). *Pseudomonas protegens* CHA0, *Pseudomonas syringae* Cit-7, *Escherichia coli* DH5α and water were used as controls. When amplifying with the *P. fluorescens* primers, a bright amplification brand was obtained from the SL1 isolate sample, suggesting that the isolate is indeed a *P. fluorescens* species (Figure 3C). There was also a faint band observed in the *P. protegens* lane which can be a result of the high degree of similarity found in the 16S regions of *P. fluorescens* and *P. protegens* (Figure 3C). We did not observe any bands in the other negative controls, as expected. The *Pseudomonas*-specific primers amplified the SL1 isolate as well as the positive controls *P. syringae* Cit-7 and *P. protegens* CHA0, further confirming the identity of the SL1 isolate as part of the genus *Pseudomonas* (Figure 3B). However, when looking at the *E. coli* DH5α negative control, a faint band was observed, likely due to non-specific amplification as a result of low annealing temperatures and weak primer binding (Figure 3B).

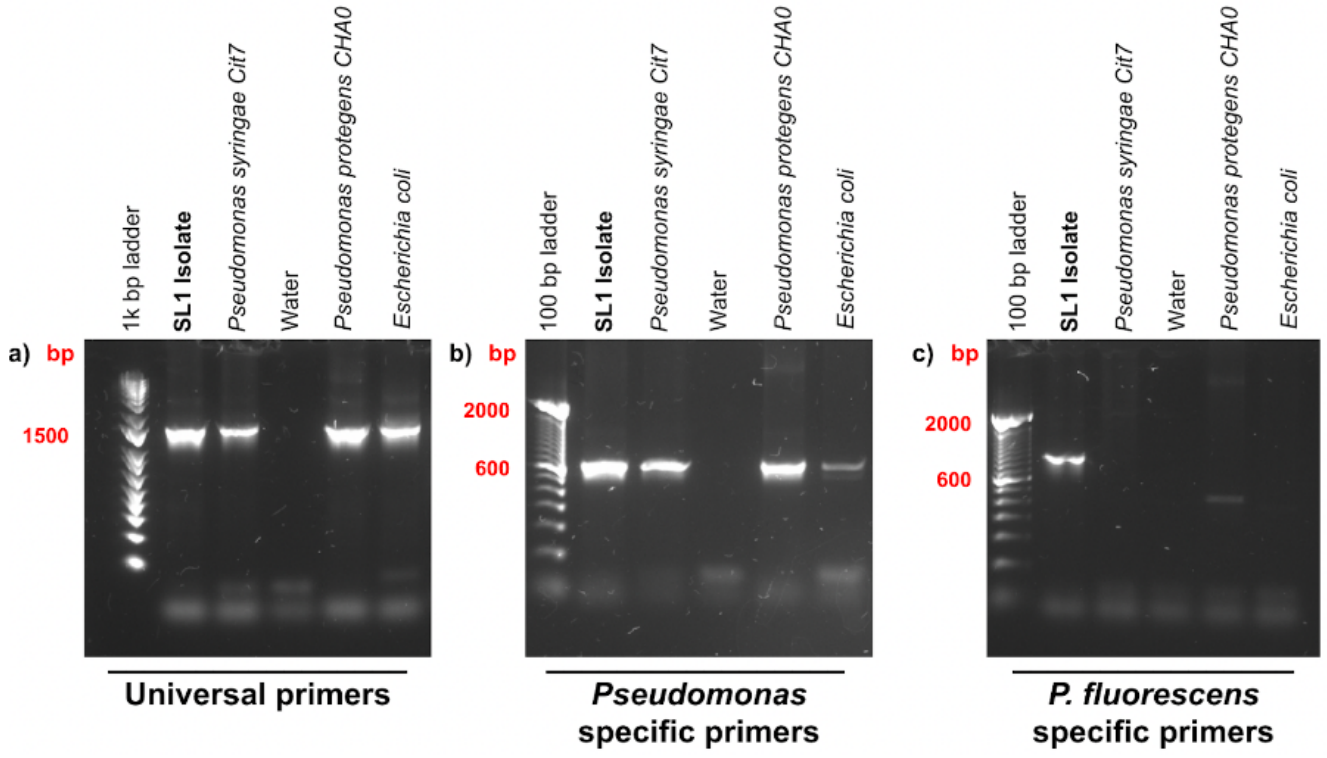

FIG. 3 16S rRNA amplification products using 3 primer pairs. Genomic DNA from the SL1 isolate, *P. syringae* Cit-7, *P. protegens* CHA0 *and E. coli* DH5α were used as template for each of the 3 primer sets. **a)** 1500 bp amplification products produced by PCR using universal primers that amplify the 16S rRNA region of bacteria. **b)** 600 bp amplification products produced by PCR using *Pseudomonas* specific 16S rRNA primers. **c)** 850 bp amplification products produced by PCR using *P. fluorescens* specific 16S rRNA primers.

SL1 is a member of the genus *Pseudomonas*. To confirm the identity of the SL1 isolate, the universal primers 27F and 1492R (**Figure 3A**) were utilized to amplify a larger region of the 16S rRNA gene, which were then Sanger sequenced. Amplification products were observed in the SL1 isolate, *P. syringae* Cit-7, *P. protegens* CHA0 and *E. coli* DH5α lanes, suggesting that the universal primers were able to amplify the 16S regions for different bacterial species

(Figure 3A). Both forward and reverse reads were obtained from Sanger sequencing and underwent a 16S rRNA search against the curated Ribosome Database Project. The quality control score of the forward and reverse reads provided by Genewiz were 29 and 40 respectively, as there were several low-lying background peaks when visualizing the chromatogram. Running a search against the Ribosome Database Project using the contigs generated from the forward and reverse reads indicated that the SL1 isolate is from the genus *Pseudomonas*, with a confidence threshold over 95%. As a result of the low quality control score of the forward read, we were not able to make a firm conclusion on the identity of the SL1 isolate at the species level but were confident at the *Pseudomonas* genus level.

***Fit* gene cluster not found in the SL1 whole genome sequence with genome annotation using Prokka and BLASTKoala.** To determine if our strawberry isolate contained the *fit* gene cluster, we performed whole genome sequencing from the isolated genomic DNA using the MinION Nanopore sequencer and annotated the genome using Prokka and BLASTKoala (Figure 1). We ran both the isolated genomic DNA of SL1 and a control *P. protegens* CHA0. The MinION sequencer utilizes a flow cell that allows genomic DNA from our strawberry leaf sample to move through a biological pore. As the DNA molecule moves through the pore, the MinION device measures changes in electrical conductivity to identify DNA bases (Figure 4). From our raw output, we were able to assemble the reads using Flye and polished the assembly using Racon and Medaka. The assembled genome contained a total of 12 contigs with a genome size of 6.8 Mbps and a total G-C content of approximately 60.4% (Supplemental Figure 1). BLASTkoala identified 3859 genes annotated within the genome. Genome annotation with Prokka and BLASTkoala did not identify the *fit* gene cluster within the SL1 genome; however, it was identified when annotating a secondary *P. protegens* CHA0 reference genome. When looking deeper into the SL1 genome, we identified a response regulator *gacA,* found in contig 10, which has been characterized as a regulator of insecticidal genes (11). This gene was annotated in both BLASTkoala and Prokka and closely resembles the *Pseudomonas fluorescens GacA* protein sequence found in the NCBI database. Altogether, these results suggest that the *fit* gene cluster is not encoded in the genome of the SL1 but does contain the *gacA* gene responsible for regulating insecticidal genes.

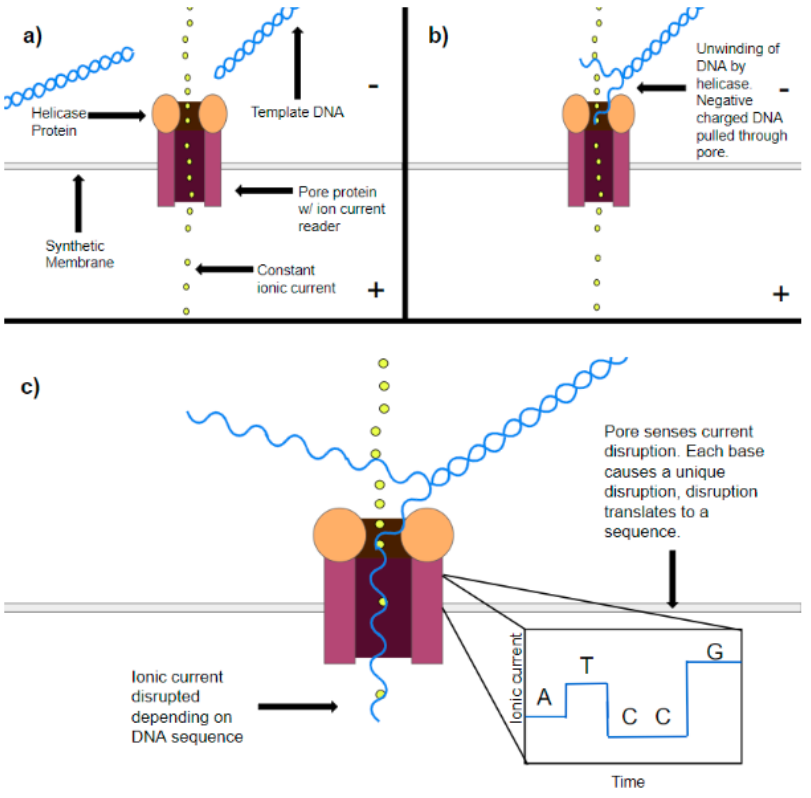

FIG. 4 Mechanism of a nanopore sequencer. This is a brief schematic of the theory and function of a nanopore sequencer. Each Flongle flow cell is covered in 156 nanopore structures, with one shown here. **a)** A nanopore consists of a helicase protein and pore protein with a current reader that allows passage of charged ions across an electromagnetic gradient, separated by a synthetic membrane. DNA is on the negatively charged side of the membrane. **b)** DNA is unwound by the helicase protein and begins to pass through the membrane, as the negatively charged backbone follows the electromagnetic gradient. **c)** The ionic current is disrupted as the strand passes through, to different degrees depending on the bases present. A current reader in the pore protein produces a disruption pattern graph as the strand passes through, which is then parsed into a sequence.

DISCUSSION

The principal goal of this investigation was to determine whether susceptible berry-producing plants harbored *P. fluorescens* containing the insecticidal gene *fitD* as a potential biocontrol agent for SWD. We isolated a fluorescing strain from a strawberry leaf (SL1) and performed 16S rRNA gene amplification using *Pseudomonas*-and *P. fluorescens*-specific primer sets for initial characterization. To further confirm the identity of the SL1 isolate, we sequenced the 16S region and ran a nucleotide BLAST search against the 16S ribosomal RNA sequences (bacteria and archaea) database. Lastly, we performed whole genome sequencing using the MinION Nanopore sequencer and annotated the SL1 genome using Prokka and BLASTKoala.

Growth was observed in all the cultured stem and leaf extracts but few of the berry extracts. We saw growth from cultured dilutions of the strawberry leaf and stem extracts, and limited growth on the berries, with some berries having no growth at all. Previous studies have shown enteric pathogen and oral bacterial growth to be inhibited by the phenolic compounds present in berries such as blackberries, strawberries, and blueberries (26-28). These are likely to inhibit the growth of bacteria in general, considering that the periodontal pathogens include a wide range of both gram-positive and gram-negative bacteria (27). The environment within the berry endosphere could prevent the growth of insecticidal *P. fluorescens* strains, thus explaining the lack of growth we saw when culturing berry extracts, and possibly allowing the SWD eggs and larvae to flourish in that environment.

***Pseudomonas*- and *P. fluorescens*-specific primers identified SL1 isolate as a *P. fluorescens* biotype**. We were able to amplify the 16S rRNA region of the SL1 isolate using both *Pseudomonas*- and *P. fluorescens*-specific primers as indicated by the amplification products in the SL1 lanes of their respective gels (**Figure 3B and 3C**). This suggests that the SL1 isolate is specific to the genus *Pseudomonas* and possibly to the *Pseudomonas fluorescens* family, which corresponds with our initial prediction. Previous studies have utilized these primers as a rapid and effective tool to differentiate *P. fluorescens* biotypes and to investigate the diversity and prevalence of *Pseudomonas* species in the gut of children with Crohn's disease (16,17). However, it cannot be used to fully differentiate and classify the SL1 isolate due to non-specific amplifications that may occur, which was observed when these primers were used with *Pseudomonas protegens CHA0* DNA (**Figure 3C**). Previous studies have shown high similarities in the 16S rRNA gene between certain *Pseudomonas* species, rendering the use of the 16S rRNA region inadequate for differentiating between species (29). It is possible that the binding sites of these primers happen to occur at sequences that are conserved between closely related species, resulting in non-specific amplification products. As a result, the *P. fluorescens*-specific primers are an effective screening tool to identify *P. fluorescens* but must be used with caution. The non-specific amplification observed when *E. coli* DNA was amplified with *Pseudomonas*-specific primers (**Figure 3B)** may have been due to the low annealing temperature selected. In the future, using a PCR protocol that reduces nonspecific binding, such as nested PCR, may improve the specificity of this assay (30).

16S rRNA data confirmed the isolate to be *Pseudomonas* at the genus level, but did not allow species to be determined. Based on the 16S rRNA sequencing data from the forward and reverse primers, we were unable to align the SL1 sequence to any previously sequenced *Pseudomonas* species with at least a percent identity of 99%, a threshold obtained from previous literature (29). However, the conventional genus percent identity of 97% was surpassed, allowing us to confirm we had a species within *Pseudomonas*. The *Pseudomonas* genus has a relatively conserved 16S rRNA region among species, with less than a 1% nucleotide sequence difference between many of the species (31). This means that the 16S rRNA sequence alone does not offer a high degree of certainty for species, even if the percent identity scores were to exceed 99%. 16S rRNA sequences could be used in conjunction with other conserved *Pseudomonas* genes for more specificity. Some studies have used primers to amplify the *rpoD* gene, which allowed differentiation between different strains within the *P. fluorescens* subgroup (32,33). Sequencing both the 16S rRNA region as well as the *rpoD* region may allow for the species of our isolate to be determined with certainty in the future. This could also be accomplished by surveying our assembled genome for the *rpoD* gene.

SL1 whole genome sequencing and annotation did not show presence of the *fitD* gene. After a survey of our assembled genome using Prokka and BLASTKoala, we were not

able to identify the *fitD* gene in our isolate; however, we were able to find it in the assembled genome of our control strain *P. protegens* CHA0. Since CHA0 is known to contain *fitD,* this suggested that our whole genome sequencing and assembly was successful, and that *fitD* was not present in SL1. Although *fitD* is found in many strains in the *P. fluorescens* subgroup such as Pf-5 and CHAO, it is not found in every strain, and has yet to be found in strains isolated from strawberry plants (7). Two studies analysing plant-associated *Pseudomonas* species found that *fitD* was only present in *P. protegens* and *P. chlororaphis* strains (34,35). Both of these species are part of the *P. fluorescens* subgroup, but neither were in the ten highest percent identity 16S rRNA sequences from our BLAST search. For the *P. fluorescens* species, *fitD* has yet to be identified in any strain. Thus, unless the exact species of our isolate can be determined, there is a possibility that it is a species of the *P. fluorescens* subgroup that is missing *fitD* as well as the other genes found in the *fit* locus.

However, we did locate a gene with similarity to *gacA*, which encodes a response regulator protein, known to control the expression of virulence factors in *P. fluorescens* Pf-5 (11). The *gacA* gene is required for *P. fluorescens* Pf-5 to infect *D. melanogaster,* and its deletion results in decreased *D. melanogaster* mortality. As a result, SL1 may still have insecticidal capabilities with further characterization of insecticidal properties needing to be performed.

Limitations Due to the time of year this study was conducted, environmental samples were limited. As this experiment began in October, the British Columbia strawberry and blueberry season was over and only wild blackberries were available. It would be beneficial to repeat this experiment with in-season strawberries and blueberries as endophytic *P. fluorescens* in berries would likely contact SWD including eggs, larvae, and adults at a higher rate than leaf or stem endophytes.

The processing of the environmental samples into freezable extracts also proved to be a challenging task**.** The woodiness of the leaves and stems from the strawberry plant were difficult to homogenize. To repeat the experiment, an improved protocol would be to homogenize a larger quantity of environmental extracts in a small sterile mortar and pestle. Additionally, the serial dilutions method can be improved using dilutions occurring with less sample and less volume while still plating the same amounts and can be performed in a deep well plate. These particular improvements would allow for faster processing of samples and a wider array of samples to be screened for *P. fluorescens*.

Although we were able to annotate our genome using Prokka and BLASTKoala, we were not able to apply additional annotation tools for further confirmation and characterization of the SL1 genome. One annotation tool that has been recently developed to identify novel insecticidal genes is the software ORFograph (36). However, the installation process is complex and requires substantial computer skills. Applying this tool to identify insecticidal genes can further increase certainty that the *fitD* gene is not present within the SL1 genome.

Conclusions This paper outlines a framework procedure to extract and cultivate endophytic *P. fluorescens* and sequencing to characterize the genome. From a small starting environmental sample, we found a *Pseudomonas* that we named SL1, and determined whether it had insecticidal properties for use as a biocontrol agent. Future studies need to be conducted to further characterize the SL1 strain. That includes further annotation for other genes, with the confirmation of the expression of those genes conducted experimentally. Annotation for genes relating to pathogenicity or symbiosis would be highly beneficial for the characterization of isolated strains, and their potential in the use for biocontrol of SWD and other pests.

Future Directions Although the genome has been annotated using the Prokka and BLASTKoala softwares, there are several other annotation tools that could have been utilized to annotate the genome further and confirm the presence or absence of insecticidal genes. These tools can be used to further characterize the hypothetical proteins that are present in the PROKKA output to get a more comprehensive understanding of the SL1 genome. One interesting pipeline is the NCBI Prokaryotic Genome Annotation Pipeline that utilizes an ORFfinder to identify potential reading frames and several databases including TIGRFam,

Pfam, PRK HMMs and NCBIfams for annotation purposes (37-39). However, the NCBI Prokaryotic Genome Annotation Pipeline has several requirements that must be met in order to execute, which was a limitation of our study. A secondary tool that can be utilized to further identify insecticidal genes is called ORFograph. It is a new method that has been developed to search for novel insecticidal protein genes by utilizing assembly graphs that are generated from assembly softwares such as SPAdes and metaSPAdes to identify novel genes rather than using individual contigs (37). Furthermore, it has been shown to identify insecticidal protein genes that have been previously hidden from pre-existing identification tools.

The potential insecticidal capabilities of the SL1 isolate can also be used in a future experiment to test directly on SWD. Although we did not identify the fit gene cluster in the genome of the isolate, there may be other insecticidal genes that can induce the killing of these insects such as the gene *gacA*. A feeding assay involving the SL1 isolate and measuring the period of induced mortality can further characterize this isolate as insecticidal. Additionally, performing a deletion experiment involving *gacA* can be beneficial to draw a firm conclusion of its insecticidal activity against the SWD.

Lastly, this experiment should be repeated with a wider range of environmental samples such as additional blackberries, strawberries, blueberries, along with their leaves and stems so that a library of *P. fluorescens* residing in the Vancouver, British Columbia area may be created. These strains would then be candidates for inserting insecticidal genes as a way of creating a biological pesticide.

ACKNOWLEDGEMENTS

We would like to thank the Haney lab for providing the sample sterilization protocol and various Pseudomonas species including *P. protegens* CHA0 and *P. syringae* Cit-7. We would also like to thank the MICB 401 teaching team including Daniela Yanez Ortuno, Dr. David Oliver, Dr. Cara Haney and Jade Muileboom for support and guidance throughout the semester and the Microbiology and Immunology Department for funding and lab space.

CONTRIBUTIONS

Lab experiments were done collectively between all the authors. Christopher Yap and Adam C. Wilcockson were involved in the sample sterilization, isolation and purification of the different berry stems, leaves and fruits. Ali Anwari and Fione Yip were involved in the DNA isolation, PCR amplification, DNA gel electrophoresis and Sanger sequencing. Whole genome sequencing using the MinION Nanopore sequencer was done collaboratively as a group and assembly and annotation was done by Daniela Yanez Ortuno. The writing of the manuscript was done evenly between all the authors.

REFERENCES

1. Ministry of Agriculture F and F. Spotted Wing Drosophila (Fruit Fly) Pest Alert - Province of British Columbia.
2. Management Guidelines: Spotted Wing Drosophila in Ontario.
3. **Bolda, MP, Goodhue RE, Zalom FG**. 2010. Spotted Wing Drosophila: Potential Economic Impact of a Newly Established Pest. ARE Update 13(3):5-8.
4. **Lee JC, Dreves AJ, Cave AM, Kawai S, Isaacs R, Miller JC, Van Timmeren S, Bruck DJ**. (2015) Infestation of wild and ornamental noncrop fruits by *Drosophila suzukii* (Diptera: Drosophilidae). Ann. Entomol. Soc. Am. 108:117-129.
5. **Klick J, Yang WQ, Walton VM, Dalton DT, Hagler JR, Dreves AJ, Lee JC, Bruck DJ**. 2015. Distribution and activity of *Drosophila suzukii* in cultivated raspberry and surrounding vegetation. J. Appl. Entomol. 140:37-46.
6. **Cartieaux F, Thibaud M-C, Zimmerli L, Lessard P, Sarrobert C, David P, Gerbaud A, Robaglia C, Somerville S, Nussaume L. 2003**. Transcriptome analysis of Arabidopsis colonized by a plant-growth promoting rhizobacterium reveals a general effect on disease resistance. The Plant Journal 36:177–188.
7. **Swadling IR, Jeffries P**. 1996. Isolation of Microbial Antagonists for Biocontrol of Grey Mould Disease of Strawberries. Biocontrol Science and Technology 6:125–136.
8. **Palleroni NJ**. 1984. Pseudomonadaceae, p 141-199. *In* Krieg NR, Holt JG(ed), Bergey's Manual of Systematic Bacteriology. The Williams and Wilkins Co, Baltimore, MD
9. **Teoh M-C, Furusawa G, Veera Singham G**. 2021. Multifaceted interactions between the pseudomonads and insects: mechanisms and prospects. Arch. Microbiol. 203:1891–1915.

10. **Otsu Y, Matsuda Y, Mori H, Ueki H, Nakajima T, Fujiwara K, Matsumoto M, Azuma N, Kakutani K, Nonomura T, Sakuratani Y, Shinogi T, Tosa Y, Mayama S, Toyoda H**. 2004. Stable phylloplane colonization by entomopathogenic bacterium *Pseudomonas fluorescens* KPM-018P and biological control of *Phytophagous* ladybird beetles Epilachna vigintioctopunctata (Coleoptera: Coccinellidae). Biocontrol. Sci. Technol. 14:427–439.
11. **Loper JE, Henkels MD, Rangel LI, Olcott MH, Walker FL, Bond KL, Kidarsa TA, Hesse CN, Sneh B, Stockwell VO, Taylor BJ**. 2016. Rhizoxin analogs, orfamide A and chitinase production contribute to the toxicity of *Pseudomonas protegens* strain Pf-5 to *Drosophila melanogaster*. Environmental Microbiology 18:3509–3521.
12. **Péchy-Tarr M, Bruck DJ, Maurhofer M, Fischer E, Vogne C, Henkels MD, Donahue KM, Grunder J, Loper JE, Keel C**. 2008. Molecular analysis of a novel gene cluster encoding an insect toxin in plant-associated strains of *Pseudomonas fluorescens.* Environmental Microbiology 10:2368–2386.
13. **Hansen AK, Vorburger C, Moran NA**. 2012. Genomic basis of endosymbiont-conferred protection against an insect parasitoid. Genome Res 22:106–114.
14. **Seemann T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
15. **Kanehisa M, Sato Y, Morishima K**. 2016. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. Journal of Molecular Biology 428:726–731.
16. **Wagner J, Short K, Catto-Smith AG, Cameron DJS, Bishop RF, Kirkwood CD**. 2008. Identification and Characterisation of Pseudomonas 16S Ribosomal DNA from Ileal Biopsies of Children with Crohn's Disease. PLoS One 3:e3578.
17. **Scarpellini M, Franzetti L, Galli A**. 2004. Development of PCR assay to identify Pseudomonas fluorescens and its biotype. FEMS Microbiology Letters 236:257–260.
18. JoVE Science Education Database. 2021. 16S rRNA Sequencing: A PCR-based Technique to Identify Bacterial Species. JoVE, Cambridge, MA.
19. **Quick J, Quinlan AR, Loman NJ**. 2014. A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. Gigascience. 3:22.
20. **Kolmogorov M, Yuan J, Lin Y, Pevzner PA**. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546.
21. **Vaser R, Sović I, Nagarajan N, Šikić M**. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746.
22. **Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ**. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. 11:119.
23. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J**. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624.
24. **Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FSL, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK-S, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock REW, Lory S, Olson MV**. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature. 406:959–964.
25. **Meyer JM, Abdallah MA**. 1978. The Fluorescent Pigment of *Pseudomonas fluorescens*: Biosynthesis, Purification and Physicochemical Properties. J. Gen. Microbiol. 107(2):319-328
26. **Khalifa HO, Kamimoto M, Shimamoto T, Shimamoto T**. 2015. Antimicrobial Effects of Blueberry, Raspberry, and Strawberry Aqueous Extracts and their Effects on Virulence Gene Expression in Vibrio cholerae. Phytother Res. 29(11):1791-7.
27. **González OA, Escamilla C, Danaher RJ, Dai J, Ebersole JL, Mumper RJ, Miller CS**. 2013. Antibacterial effects of blackberry extract target periodontopathogens. J Periodontal Res. 48(1):80-6.
28. **Shen X, Sun X, Xie Q, Liu H, Zhao Y, Pan Y, Hwang C, Wu VCH**. 2014. Antimicrobial effect of blueberry (vaccinium corymbosum L.) extracts against the growth of listeria monocytogenes and salmonella enteritidis. Food control. 35(1):159-165.
29. **Hofstetter V, Buyck B, Eyssartier G, Schnee S, Gindro K**. 2019. The unbearable lightness of sequenced-based identification. Fungal Diversity. 96:243–284
30. **Green MR, Sambrook J**. 2019. Nested Polymerase Chain Reaction (PCR). Cold Spring Harb Protoc. 2019(2).
31. **Mulet M, Lalucat J, García-Valdés E**. 2010. DNA sequence-based analysis of the Pseudomonas species. Environmental Microbiology. 12:1513–1530.
32. **Mulet M, Bennasar A, Lalucat J, García-Valdés E**. 2009. An rpoD-based PCR procedure for the identification of Pseudomonas species and for their detection in environmental samples. Molecular and Cellular Probes. 23:140–147.
33. **Lauritsen JG, Hansen ML, Bech PK, Jelsbak L, Gram L, Strube ML**. 2021. Identification and Differentiation of Pseudomonas Species in Field Samples Using an rpoD Amplicon Sequencing Methodology. mSystems. 6(4):e007042.1
34. **Rangel LI, Henkels MD, Shaffer BT, Walker FL, Ii EWD, Stockwell VO, Bruck D, Taylor BJ, Loper JE**. 2016. Characterization of Toxin Complex Gene Clusters and Insect Toxicity of Bacteria Representing Four Subgroups of *Pseudomonas fluorescens*. PLOS ONE. 11:e0161120.

35. **Loper JE, Hassan KA, Mavrodi DV, Ii EWD, Lim CK, Shaffer BT, Elbourne LDH, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, Mortel JE van de, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Iii LSP, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT**. 2012. Comparative Genomics of Plant-Associated Pseudomonas spp.: Insights into Diversity and Inheritance of Traits Involved in Multitrophic Interactions. PLOS Genetics. 8:e1002784.
36. **Dvorkina T, Bankevich A, Sorokin A, Yang F, Adu-Oppong B, Williams R, Turner K, Pevzner PA**. 2021. ORFograph: search for novel insecticidal protein genes in genomic and metagenomic assembly graphs. Microbiome. 9(1):149.
37. **Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F**. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res. 49(D1):D1020-D1028.
38. **Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD**. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res. 46(D1):D851-D860.
39. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J**. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44(14):6614-24.