HIGH SCHOOL EDITION
Journal of Student Research
Volume 14 Issue 1 (2025)

Function Of Pseudomonas Fluorescens L5.1-96 Proteins MPPE And NADK In Wheat Root Colonization

Julia Zhang[1] and Stacy Alvares[#]

[1]Newport High School, USA
[#]Advisor

ABSTRACT

Pseudomonas fluorescens strain L5.1-96 is exceptional at colonizing wheat roots and resisting desiccation, both characteristics valuable to combatting Take-All disease, a devastating wheat root disease caused by the fungal pathogen *Gaeumannomyces graminis var. triciti* that drastically decreases wheat crop yield in the state of Washington. As a result, the USDA-ARS has generated a genomic library from the genome of strain L5.1-96 to identify genes which may confer these characteristics. In this study, DNA from the clone PF-25A 8 was cultured, purified using the Qiagen Miniprep kit, and sequenced using SL1 and SR2 primers through Eurofins Genomics. Bioinformatics tools BLAST, ORF Finder, MAFFT, and SWISS Model were used to analyze unambiguous base calls. Two notable proteins that may contribute to enhanced survival and colonization were identified: metallophosphoesterase (MPPE) and NAD+ kinase (NADK). The former functions in DNA repair and may improve PF L5.1-96's colonization by increasing longevity in the wheat rhizosphere, while the latter may enhance survival by conferring a degree of antibiotic resistance and providing protection against bactericides used on farms.

Introduction

Take-All is a disease caused by the fungus *Gaeumannomyces graminis var. triciti* that drastically reduces wheat yield by infecting wheat roots and choking off their water supply, killing the plants. Wheat is especially susceptible due to being farmed as a monoculture, meaning that there is no genetic diversity, making it difficult for wheat fungal resistance characteristics to arise. In terms of overall agricultural production, the spread of Take-All has forced farmers to relocate to areas where the disease is less prevalent, resulting in the termination of most wheat farming on the Western portion of mountains in Washington (Bangera & Thomashow, 1999). Despite this, the disease is still rampant in many wheat fields and remains a large problem. Take-All greatly reduces crop yields, which threatens both the continued cultivation of important wheat crops that are exported from Washington state such as soft white wheat and hard red wheat, and the livelihoods of smaller farmers that grow these crops (Bangera & Thomashow, 1999).

Potential methods that could be used to combat the effects of Take-All disease include using pesticides, chemicals, rotating the crops, burning infected crops, generating disease-resistant cultures of wheat, tilling the land, or other methods of biological control. While all of these methods have been tested to some degree, methods using pesticides, chemicals, burning, and tilling are environmentally harmful and show little success, while rotating crops and generating resistant cultures are extremely expensive and not viable in most cases. Therefore, the method that is under current investigation for its viability in combatting the disease in a cost effective and eco-friendly manner is biological control of the fungus, via a bacterium named *Pseudomonas fluorescens* (PF). In trials testing soils that exhibit Take-All Decline (TAD), it was found that soils containing *P. fluorescens* were most effective in preventing the wheat grown from being infected (Bangera & Thomashow, 1999). This was found to be due to the presence of a compound named DAPG that stops the growth of *G. graminis* and is produced by some strains of *P. fluorescens*. DAPG is a polyketide compound that is produced by *Pseudomonas* species containing the *phl* operon, and kills

pathogenic organisms by cleaving the plasma membrane, increasing the level of ROS, and damaging their DNA (Suresh et al. 2022).

Of particular interest is the *P. fluorescens* strain named L5.1-96, as it makes up 90% of PF strains that produce DAPG in the soil tested (named Lind TAD soil). Upon running some tests on how well the strain did in dry conditions and in populating the rhizosphere of wheat plants, it was shown that the strain is both significantly better at recovering from drought and at remaining in the rhizosphere than other DAPG-producing *Pseudomonas* strains (Bangera & Thomashow, 1999). In previously characterized *Pseudomonas* strains, it has been shown that the ability to resist heat and desiccation is largely due to the action of protein chaperones, proteases, and thermosensors that work in conjunction with each other to replace denatured proteins and maintain structural stability of proteins during construction (Craig et al. 2021). Additionally, the ability of *P. fluorescens* strains to adapt to colonizing a variety of plants has been attributed to the upregulation of certain genes used in colonization - such as the *sif* genes - upon encountering a new environment (Varivarn et al. 2013). The ability to withstand drought and to populate the rhizosphere are both valuable traits in combatting the *G. graminis* fungus, as it allows the bacteria to survive through the seasons or unpredictable weather and to remain associated with the wheat rhizosphere over long periods of time to provide consistent protection. Therefore, it is vital to sequence the genome of L5.1-96 to determine what genes confer these characteristics to it in the hopes of finding out what makes the strain so suitable to act as biological control for Take-All. In particular, the investigation will focus on what genes are responsible for its enhanced colonization abilities & enhanced drought resistance, as well as whether said gene(s) are located in the main genome of *Pseudomonas fluorescens* or if they are specific to L5.1-96. By characterizing clones from PF L5.1-96's genomic library, we aim to determine if any of the clones contain genes relevant to colonization or drought resistance so that the role of these genes in biological control of Take-All can be further investigated.

Methods

Using a clean pipette tip, a small portion of PF-25A 8 colonies were scraped from the USDA-ARS-funded genomic library for PF L5.1-96. The bacteria were inoculated onto the edge of a fresh agar plate and streaked outwards in several different directions, decreasing the bacterial load so that single, genetically homogenous colonies would grow. The isolation of a single colony is important as it ensures that the plasmid copies isolated from them in later steps will be identical to each other. The plate was then incubated upside down overnight at 37°C.

A sterile falcon tube was filled with 6mL of LB broth containing Kanamycin diluted to a 1x concentration and inoculated with a single colony scraped from the agar plate. As the plasmid we are trying to purify contains a Kanamycin resistance gene on it, the addition of Kanamycin to the liquid bacterial culture will kill any bacteria that do not contain the plasmid (Supplemental Figure 1). The falcon tube was sealed immediately and incubated in a shaker incubator overnight at 37°C.

The liquid bacterial culture was spun down to isolate the pelleted cells, which were then lysed using alkaline lysis and washed using PB and PE in order to purify the pSMART HC Kan plasmid DNA containing the PF L5.1-96 genomic insert (Supplemental Figure 1). The solution contained in the lysis buffer both breaks down cellular components and separates strands of genomic DNA (which is unable to re-anneal after the addition of neutralization buffer), so that only pure plasmid DNA remains.

Samples of the purified DNA were digested using Apa1 & Dra1 from NEB in preparation for gel electrophoresis to confirm proper purification based on consistent band size. A negative control sample was also prepared. All 3 samples were incubated overnight at 37°C.

After incubation, the Apa1 restriction digest, Dra1 restriction digest, negative control, and 3 samples of plasmid concentration gradient (5 microliters, 2.5 microliters, 1 microliter) were combined with loading dye and run through a 1% agarose gel containing GelGreen (Sigma-Aldrich) and compared to an Axygen 1kB ladder to assess insert size. The finished gel was imaged using a BioRad ChemiDoc in order to determine whether purification was successful, the size of the genomic insert, and what concentration of DNA to send for Sanger sequencing to Eurofins

Genomics. The DNA samples were sent with both SL1 and SR2 primer, which sequence from opposite ends of the plasmid insert, in order to return the maximum amount of sequence information.

After sequencing, 967 unambiguous base calls were returned from the SL1 sample and 1050 unambiguous base calls were returned from the SR2 sample (Supplemental Table 1). For each returned sample, the sequence was inputted into BLASTn and interesting hits (1 strain of each unique species in the top 100 hits) were selected and pasted into MAFFT along with the query sequence to determine phylogenetic relationships. Additionally, the query sequence was inputted into BLASTx to determine probable protein products and ORF Finder to identify the specific amino acid sequence of the protein. Legitimate ORFs that returned consistent, previously characterized proteins from an organism in the genus *Pseudomonas* when inputted into BLASTp were recorded and inputted into SWISS Model to generate a protein model and better analyze protein function.

Results

The data resulting from the experiment will characterize clone PF-25A 8 to identify proteins that may confer superior colonization and survival abilities to it. Running the DNA samples and restriction digest through the 1% agarose gel confirmed proper DNA purification due to the appearance of distinct, unsmeared bands (Figure 1). The total length of the plasmid is around 4000bp when compared to the Axygen 1kB ladder (Figure 1).

The SL1 primer sequence was determined to be most closely related to similar sequences from 4 strains of P. brassicacearum (Figure 2). ORF Finder results showed that the SL1 primer sequence contains 9 ORFs (Supplemental Figure 2), but after inputting each ORF into BLASTp to check for the legitimacy of each ORF, only ORF6 returned consistent hits that matched the hits of the original query sequence inputted into BLASTx (Supplemental Table 2, Table 1). As a result, only ORF6 of the SL1 primer sequence was selected for further analysis. The sequence's most likely protein product, derived from ORF6 (Supplemental Figure 2, Supplemental Table 1) and confirmed with BLASTx results (Supplemental Table 2) was determined to be metallophosphoesterase (MPPE) with a 100% query cover and e-value of 0. This result was consistent throughout the rest of the top 100 hits, except for the 25th hit that displayed the protein Calcineurin-like phosphoesterase (Table 1). A SWISS Model of the protein with a 95.72% sequence identity and 100% coverage reveals the structure of the protein to be primarily alpha helices with 2 small portions of beta sheets on the inside (Figure 3).

The SL2 primer sequence was determined to be most closely related to a sequence from P. brassicacearum as well, but the other 3 most closely related strains are from the P. fluorescens and P. ogarae species (Figure 4). The query sequence for PF-25A 8 SR2 returned 9 ORFs when inputted into ORF Finder (Supplemental Figure 3), though only ORF6 was able to return consistent hits that confirmed BLASTx results of the original primer sequence (Supplemental Table 3, Table 2). Although ORF9 also returned consistent hits when inputted into BLASTp, the only hit was DUF1853, which has not been characterized before (Supplemental Table 4). Further analysis of ORF9 would therefore not be possible within the limits of this research, so only ORF6 of the SR2 primer was selected for continued analysis. BLASTp results of ORF6 confirm that the most likely protein encoded by the query sequence was NAD+ kinase (NADK) (Table 2).

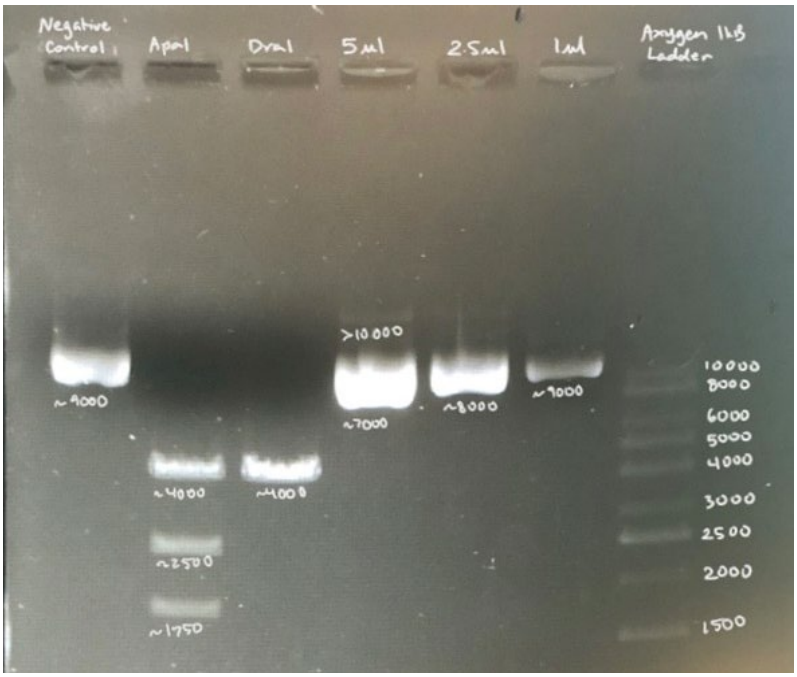


Figure 1. Image of 1% agarose gel run with 6 samples from PF-25A 8. From left to right, the wells are negative control, Apa1 digest, Dra1 digest, 5 microliter concentration, 2.5 microliter concentration, 1 microliter concentration, and Axygen 1kB DNA Ladder.

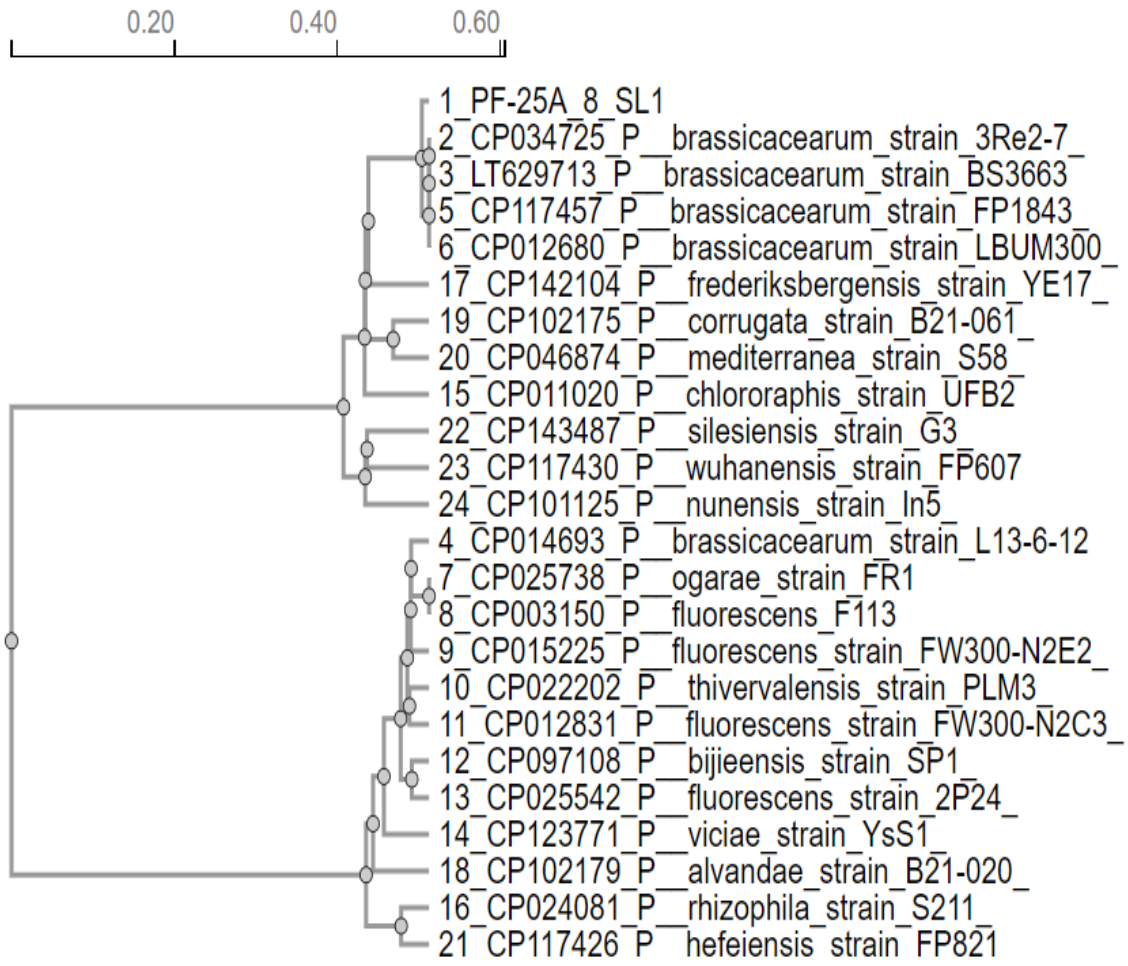


Figure 2. Guide tree generated using MAFFT from 23 hits of interest (obtained from inputting the PF-25A 8 SL1 primer query sequence into BLASTn) in relation to the query sequence (topmost entry). All unique species hits, the top 5 hits, and 5 *P. fluorescens* hits were selected.

Table 1. Top 5 hits resulting from inputting ORF6 of the PF-25A 8 SL1 primer sequence into BLASTp and an additional unique 25th hit.

Description	Organism Name	E value	Percent Identity	Hit Number
Metallophosphoesterase	*Pseudomonas*	0.0	100.00%	1
Conserved hypothetical protein	*P. brassicacearum*	0.0	100.00%	2
Metallophosphoesterase	*P. brassicacearum subsp. Brassicacearum*	0.0	99.61%	3
Metallophosphoesterase	*P. brassicacearum*	0.0	99.61%	4
Metallophosphoesterase	*Unclassified Pseudomonas*	0.0	99.22%	5
Calcineurin-like phosphoesterase	*Pseudomonas sp.*	0.0	98.05%	23

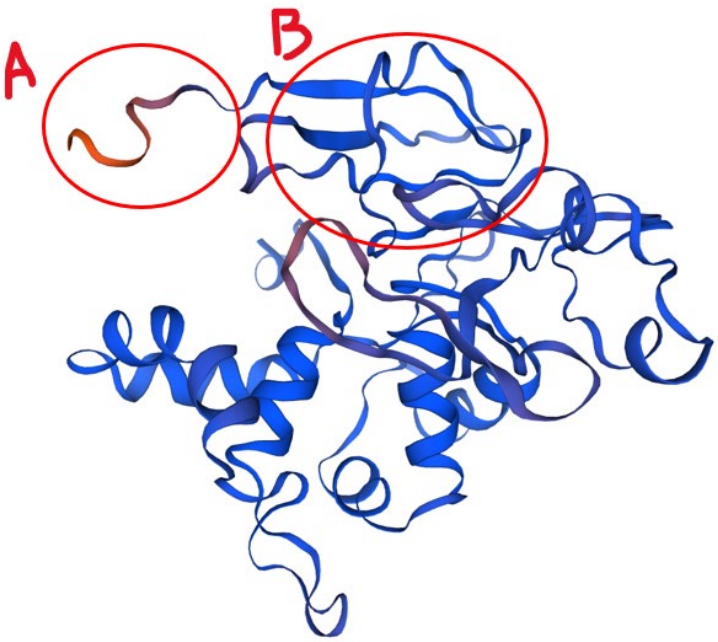


Figure 3. SWISS Model generated from the amino acid sequence of PF-25A 8 SL1 primer ORF6.

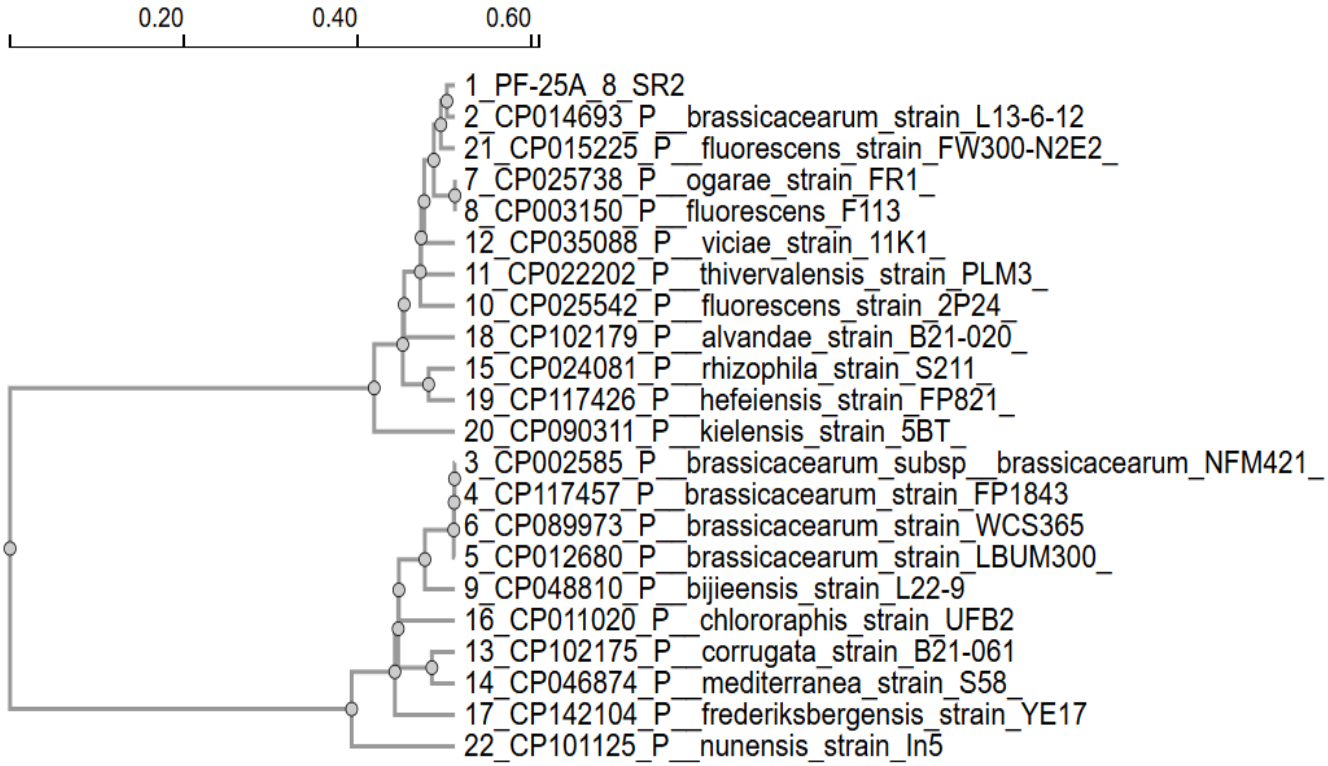


Figure 4. Guide tree generated using MAFFT from 21 hits of interest (obtained from inputting the PF-25A 8 SR2 primer query sequence into BLASTn) in relation to the query sequence (topmost entry). All unique species hits, the top 5 hits, and 3 *P. fluorescens* hits were selected.

Table 2. Top 5 hits resulting from inputting ORF6 of the PF-25A 8 SR2 primer sequence into BLASTp.

Description	Organism Name	E value	Percent Identity	Hit Number
Hypothetical protein	*P. brassicacearum subsp. brassicacearum*	8e-156	100.00%	1
NAD(+) kinase	*Pseudomonas*	9e-153	100.00%	2
NAD(+) kinase	*Pseudomonas*	9e-153	100.00%	3
NAD(+) kinase	*P. mediterranea*	9e-153	100.00%	4
NAD(+) kinase	*Pseudomonas sp.*	9e-153	100.00%	5

Discussion

The insert is estimated to be around 2212bp in length, as the restriction digest bands of the gel electrophoresis estimate the total plasmid size to be around 4000bp (Figure 1), while the base pSMART HC Kan plasmid without inserts is 1788bp in length (Supplemental Figure 1).

Previous research conducted on the protein encoded for by ORF6 of the SL1 primer sequence, metallophosphoesterase (MPPE) (Table 1), shows that in *Pseudomonas,* it is a manganese-dependent DNA nuclease that acts primarily on single-stranded areas of DNA and is involved in DNA repair (Ejaz et al., 2019). Despite MPPE's identity as a relatively conserved protein across both prokaryotic and eukaryotic species, which have a region of highly conserved double-beta sheets in a "sandwich" shape (Panel B of Figure 3) and a di-metal active site at the C-terminal end (Panel A of Figure 3), small alterations to the active site structure result in large changes to the function of the protein and greatly change its catalytic ability (Matange et al. 2015). Therefore, looking at the *specific function* of MPPEs of organisms not in the *Pseudomonas* genus will likely provide little insight into the MPPE of PF L5.1-96 due to their great variation.

However, analyzing the *mechanism* by which MPPEs benefit their host in more distantly related organisms may still be of use in understanding general principles by which MPPEs function. Research on bacteria in the genus *Bdellovibrio* showed that the gene encoding MPPE mutated frequently and tended to accumulate the most nonsynonymous mutations as opposed to synonymous mutations, and that the population containing these mutations were better at resisting starvation than the ancestral population (Mulvey et al. 2023). This research provides evidence that mutations in MPPE are involved in positive selection and that unique mutations may contribute to positive selection. Interestingly, although the SWISS Model generated from the PF-25A 8 SL1 primer was generally high confidence, the small region surrounding the C-terminus di-metal active site (Panel A of Figure 3) only returned a confidence of around 30-60%. This area of low confidence may be attributed to a mutation specific to this L5.1-96 strain and could potentially confer an evolutionary benefit related to the MPPE's known function of DNA repair. Based on this, it may be speculated that a mutation in the MPPE of L5.1-96 confers a characteristic such as enhanced DNA repair that allows it to better survive and colonize wheat roots.

As for the second protein, BLASTp results showed that ORF6 of the SR2 primer sequence most likely encodes an NAD+ kinase (NADK) (Table 2). Although NADKs are nearly ubiquitous due to their essential function in energy production and utilization, some research has shown that *Pseudomonas aeruginosa* is resistant to certain competitive inhibitor bactericides that target the NAD+ kinase as a means of killing the bacteria due to the unique crystalline structure of their NADKs (Rahimova et al., 2023). Even though *P. aeruginosa* is more distantly related to clone PF-25A 8, as it did not show up in the top 100 hits of either primer's BLASTn searches (Figure 2, Figure 4), the

conserved nature of the protein means that the function and characteristics of the NAD+ kinase should be fairly similar between bacteria (especially those of the same species). This could point towards the role of NADK in providing a degree of antibiotic resistance to L5.1-96, which may be valuable in a farm environment in which they are often used. Additionally, other research studies show that NADK has an important role in managing oxidative stress, such as in aluminum-contaminated environments commonly caused by lower soil pH that increases aluminum solubility (Lemine et al. 2008). This may also be a valuable characteristic that allows the bacteria to survive different growing seasons and any accidental drops in soil pH, such as those caused by soil erosion, a common occurrence on farmland.

Limitations

Altogether, although there is data that supports the function of both MPPE and NADK in enhancing L5.1-96's colonization and survival in wheat roots and the farm environment, it is difficult to analyze whether these characteristics are specific to this strain or if they are instead a characteristic of most other *Pseudomonas* strains. Some data from Figure 3 that reveals a differently shaped active site from other known models of the same protein provides some evidence that the MPPE's structure may actually be unique and specific to L5.1-96, though it is also possible that it is due to an error in sequencing or modeling. Future work in identifying the specific metal that the C-terminal end of the MPPE binds to may provide further insights into whether or not it is actually unique. The case for the NADK being unique to L5.1-96 is much weaker because of how ubiquitous NADK is, though it is certain that it plays a necessary role in the survival of L5.1-96 regardless of whether or not it is unique.

Finally, because the total insert size is 2212bp, but only a total of 2017bp were sequenced successfully using both the SL1 primer and SR2 primer (Supplemental Table 1) due to the limitations of Sanger sequencing, there are around 195bp in between the two sequences that were not able to be recovered. However, because this length of base pairs is relatively short in relation to the successfully sequenced DNA and that each of the selected ORFs were able to consistently return the same hit when their query sequences were inputted into BLASTp (Table 1, Table 2), it is unlikely that this short missing stretch of base pairs affected bioinformatic analysis significantly. Nevertheless, future work involving a more complete sequencing and characterization of the insert that includes these missing base pairs may be a beneficial next step in uncovering potentially overlooked proteins whose sequences could have been interrupted by the missing segment between the two primer sequences.

Acknowledgments

I would like to thank Bellevue College, the USDA-ARS, and all anonymous peer reviewers for their contributions towards this research paper.

References

Achouak, W., Conrod, S., Cohen, V., & Heulin, T. (2004). Phenotypic variation of Pseudomonas brassicacearum as a plant root-colonization strategy. *Molecular plant-microbe interactions : MPMI*, *17*(8), 872–879. https://doi.org/10.1094/MPMI.2004.17.8.872

Bangera, M. G., & Thomashow, L. S. (1999). Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2,4-diacetylphloroglucinol from Pseudomonas fluorescens Q2-87. *Journal of bacteriology*, *181*(10), 3155–3163. https://doi.org/10.1128/JB.181.10.3155-3163.1999

Craig, K., Johnson, B. R., & Grunden, A. (2021). Leveraging *Pseudomonas* Stress Response Mechanisms for Industrial Applications. *Frontiers in microbiology*, *12*, 660134. https://doi.org/10.3389/fmicb.2021.660134

Ejaz, A., Goldgur, Y., & Shuman, S. (2019). Activity and structure of *Pseudomonas putida* MPE, a manganese-dependent single-strand DNA endonuclease encoded in a nucleic acid repair gene cluster. *The Journal of biological chemistry*, *294*(19), 7931–7941. https://doi.org/10.1074/jbc.RA119.008049

Matange, N., Podobnik, M., & Visweswariah, S. S. (2015). Metallophosphoesterases: structural fidelity with functional promiscuity. *The Biochemical journal*, *467*(2), 201–216. https://doi.org/10.1042/BJ20150028

Mulvey, K., Brosnan, K., Galvin, M., Mohr, S., Muldowney, L., Oser, M., & Williams, L. E. (2023). Parallel Evolution in Predatory *Bdellovibrio* sp. NC01 during Long-Term Coculture with a Single Prey Strain. *Applied and environmental microbiology*, *89*(1), e0177622. https://doi.org/10.1128/aem.01776-22

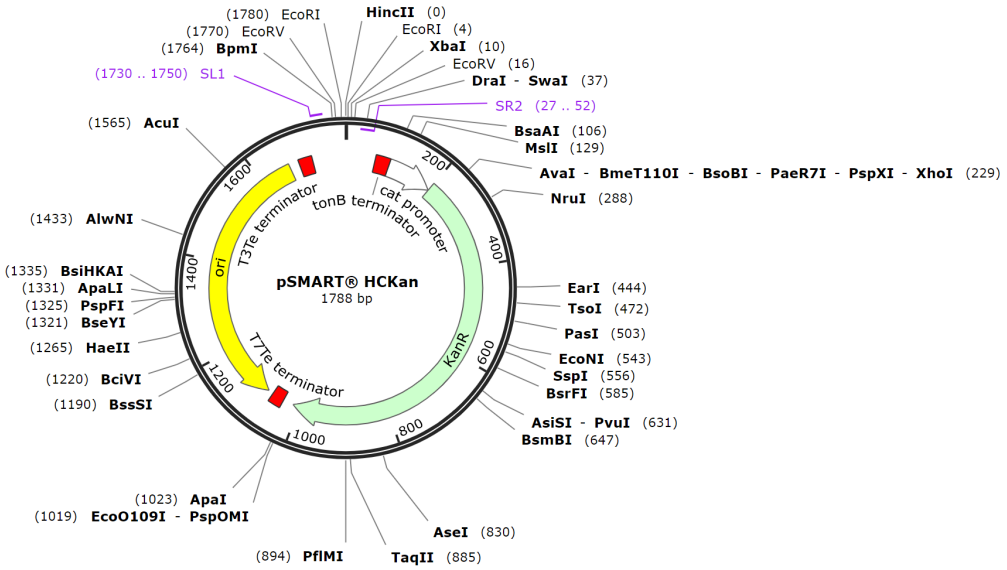
Lemire, J., Kumar, P., Mailloux, R., Cossar, K., & Appanna, V. D. (2008). Metabolic adaptation and oxaloacetate homeostasis in P. fluorescens exposed to aluminum toxicity. *Journal of basic microbiology*, *48*(4), 252–259. https://doi.org/10.1002/jobm.200800007

Rahimova, R., Nogaret, P., Huteau, V., Gelin, M., Clément, D. A., Labesse, G., Pochet, S., Blanc-Potard, A. B., & Lionne, C. (2023). Structure-based design, synthesis and biological evaluation of a NAD+ analogue targeting Pseudomonas aeruginosa NAD kinase. *The FEBS journal*, *290*(2), 482–501. https://doi.org/10.1111/febs.16604

Suresh, P., Rekha, M., Gomathinayagam, S., Ramamoorthy, V., Sharma, M. P., Sakthivel, P., Sekar, K., Valan Arasu, M., & Shanmugaiah, V. (2022). Characterization and Assessment of 2, 4-Diacetylphloroglucinol (DAPG)-Producing *Pseudomonas fluorescens* VSMKU3054 for the Management of Tomato Bacterial Wilt Caused by *Ralstonia solanacearum*. *Microorganisms*, *10*(8), 1508. https://doi.org/10.3390/microorganisms10081508

Varivarn, K., Champa, L. A., Silby, M. W., & Robleto, E. A. (2013). Colonization strategies of Pseudomonas fluorescens Pf0-1: activation of soil-specific genes important for diverse and specific environments. *BMC microbiology*, *13*, 92. https://doi.org/10.1186/1471-2180-13-92

Supplemental Materials



Supplemental Figure 1. Map of the pSMART HC Kan plasmid in which the PF L5.1-96 genomic insert was inserted.

Supplemental Table 1. Table of all nucleotide and amino acid sequences used in bioinformatic analysis.

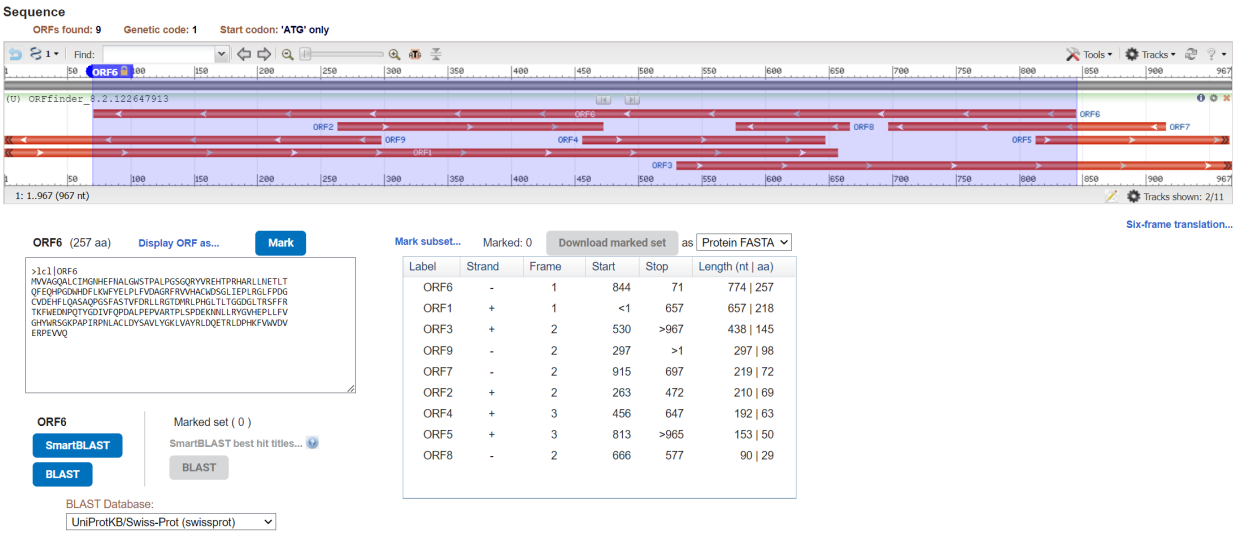
Description	Sequence
PF-25A 8 SL1 primer nucleotide sequence (967bp)	ATGCGTTGCAGCAACTTGACGAAGCCGCTCAGGTCCACTGCCAAGGGCAGGCG- CAAAAC- CGCGACGACGCTCATTGCACCACCTCCGGCCGCTCGACATCCACCCAGACGAACTT GTGCGGATCCAGGCGGGTTTCCTGGTCGAGGCGATAGGCGACCAGCTTGCCG- TAGAGCAC- CGCGCTGTAATCCAGGCAGGCCAGGTTCGGGCGAATGGGCGCGGGCTTGCCGCTG CGCCAATAGTGGCCGACGAACAGCAACGGTTCATGGACGCCATAGCGCAG- CAGGTT- GTTTTTTTCATCCGGCGACAGCGGTGTGCGGGCCACCGGTTCTGGCAAGGCGTCGG GCTGGAACACGATGTCGCCGTAGGTCTGCGGGTTGTCTTCCCAGAACTT- GGTGCGAAAAAAC- GAACGGGTCAGGCCATCGCCACCGGTCAGCGTCAGGCCATGGGGCAGGCGCATGT CGGTGCCACGTAGCAAACGATCGAACACCGTGCTGGCAAAACTGCCCGGTT- GCGCCGAGGCCTGGAGGAAATGCTCGTCGACACAACCGTCCGGGAACAGGCCACG CAGCGGTTCGATCAGGCCCGAATCCCAACAGGCATGCACCACCCGGAAC- CGGCCGGCGTCGACAAACAGCGGCAGCTCGTAGAACCACTTGAGAAAGTCATGCC AGTCGCCGGGATGCTGCTCGAACTGGGTCAGGGTCTCGTTGAGCAGGCGAG- CATGAC- GCGGCGTGTGTTCGCGCACGTAGCGCTGGCCGCTGCCGGGCAAAGCCGGTGTGCTC CAGCCCAGGGCGTTGAACTCGTGGTT- GCCCATGATGCACAGCGCCTGGCCGGCCACGAC- CATGTCGTGGACGATATGCAGCGCCTCGCGAATCCGCGGGCCACGGTCGATGATGT CCCCAGGAACACAGCCATGCGCGACGGATGCCGCCAGACCCCGCCCTGCTTGGG- GAACCGAGCCGGCCCA

PF-25A 8 SR2 primer nucleotide sequence (1050bp)	GATGGATGCCGTCACGATCGCGCAGCAGCAGGTCCAGTTCGCCCAGCG- TATGGCCTTCGTGGCGGATCGGCAGGTTCGCGGCAATCAGCTCGATGCCAGGCGCG TGCTGCACGGCAAACTGCCACAGCCTTTCG- TAATACAGGCCCAGGCGTCGCGTCCGTCCCAGGGACAGGCAGTGCAGCAGCGGGT AGCTGTCGCGGTCCAGTTGTCGCAGCCAGTGCTCCAGCTGATCAGGCGCCTG- TACCCAG- TCGCTGCCGGCCAGGGGATGGCGCTGGGGCCAGGGCGTGTCGATGAGCATCGGCG GGGCGATGATCGCCCACGCGAGGTCGCGCACTTCGG- GATGGCGAAGGCGGCGGGG- CAAGTCCAGCAATTCGGGGAACAGGATCATTTTGCGAGCATAGCTGCTGCGAAAG GCTTAAAGGATTTTGTCTATGGAGCGCTTTCGCCCA- TAATCGTGGTTTTCGCCCGTCG- CAGACCCTCGCAGGAGCCTCATGGAGCAATTTCGCAATATCGGCATCATCGGTCGC CTGGGCAGTTCGCAGGTGTTGGACACCGTTCGCCGTCTCAAACGTTTCCTCCTG- GAGCGACACCTGCATGTGATCCTCGAAGATACCATCGCCGAAGTCTTGCCGGGCCA TGGCCTGCAGACGTCGTCGCGCAA- GATGCTGGGCGAGGTCTGCGACATGGTGATCGTGGTCGGCGGCGACGGCAGCCTG CTGGGGGCCGCCCGGGCCCTGGCGCGGCACAATATCCCGGTGCTGGGGATCAAC- CGGGG- CAGCCTGGGGTTCCTCACCGATATCCGCCCTGACGAGCTGGAAACCAAGGTCGCCG AGGTGCTGGACGGCCACTACCTGGTGGAAAACCGCTTCCTGCTGCAA- GCCGAGGTCCGTCGTCACGCCGAGGCCATCGGCCAGGGCGATGCCCTCAATGACGT GGTGCTGCACCCTGGCAATCCACGCGCATGATCGAATTCGAGCTGTACATCGAC- GGCCAG- TTCGTCTGCAGCCAGAAGGCCNACGGCCTGATCGTCGCGACCCGACGGTTCACGGC
PF-25A 8 SL1 ORF6 amino acid sequence (257aa)	MVVAGQALCIMGNHEFNALGWSTPALPGSGQRYVREHTPRHARLLNETLT QFEQHPGDWHDFLKWFYELPLFVDAGRFRVVHACWDSGLIEPLRGLFPDG CVDEHFLQASAQPGSFASTVFDRLLRGTDMRLPHGLTLTGGDGLTRSFFR TKFWEDNPQTYGDIVFQPDALPEPVARTPLSPDEKNNLLRYGVHEPLLFV GHYWRSGKPAPIRPNLACLDYSAVLYGKLVAYRLDQETRLDPHKFVWVDV ERPEVVQ
PF-25A 8 SR2 ORF6 amino acid sequence (157aa)	MEQFRNIGIIGRLGSSQVLDTVRRLKRFLLERHLHVILEDTIAEVLPGHG LQTSSRKMLGEVCDMVIVVGGDGSLLGAARALARHNIPVLGINRGSLGFL TDIRPDELETKVAEVLDGHYLVENRFLLQAEVRRHAEAIGQGDALNDVVL HPGNPRA
PF-25A 8 SR2 ORF9 amino acid sequence (128aa)	MILFPELLDLPRRLRHPEVRDLAWAIIAPPMLIDTPWPQRHPLAGSDWVQ APDQLEHWLRQLDRDSYPLLHCLSLGRTRRLGLYYERLWQFAVQHAPGIE LIAANLPIRHEGHTLGELDLLLRDRDGIH

Supplemental Table 2. Top 10 hits resulting from inputting the PF-25A 8 SL1 primer sequence into BLASTx and an additional unique 25th hit.

Description	Organism Name	E value	Percent Identity	Hit Number
Metallophosphoesterase	*Pseudomonas*	0.0	100.00%	1

Conserved hypothetical protein	*Pseudomonas brassicacearum subsp. brassicacearum NFM421*	0.0	100.00%	2
Metallophosphoesterase	*Pseudomonas brassicacearum*	0.0	99.64%	3
Metallophosphoesterase	*Pseudomonas brassicacearum*	0.0	99.64%	4
Metallophosphoesterase	*Pseudomonas brassicacearum*	0.0	99.64%	5
Metallophosphoesterase	*Pseudomonas*	0.0	99.28%	6
Serine/threonine protein phosphatase	*Pseudomonas fluorescens*	0.0	99.28%	7
Serine/threonine protein phosphatase	*Pseudomonas ogarae*	0.0	99.28%	8
Metallophosphoesterase	*Pseudomonas*	0.0	99.28%	9
Metallophosphoesterase	*Pseudomonas*	0.0	98.91%	10
Calcineurin-like phosphoesterase	*Pseudomonas sp. NFACC17-2*	0.0	98.19%	25

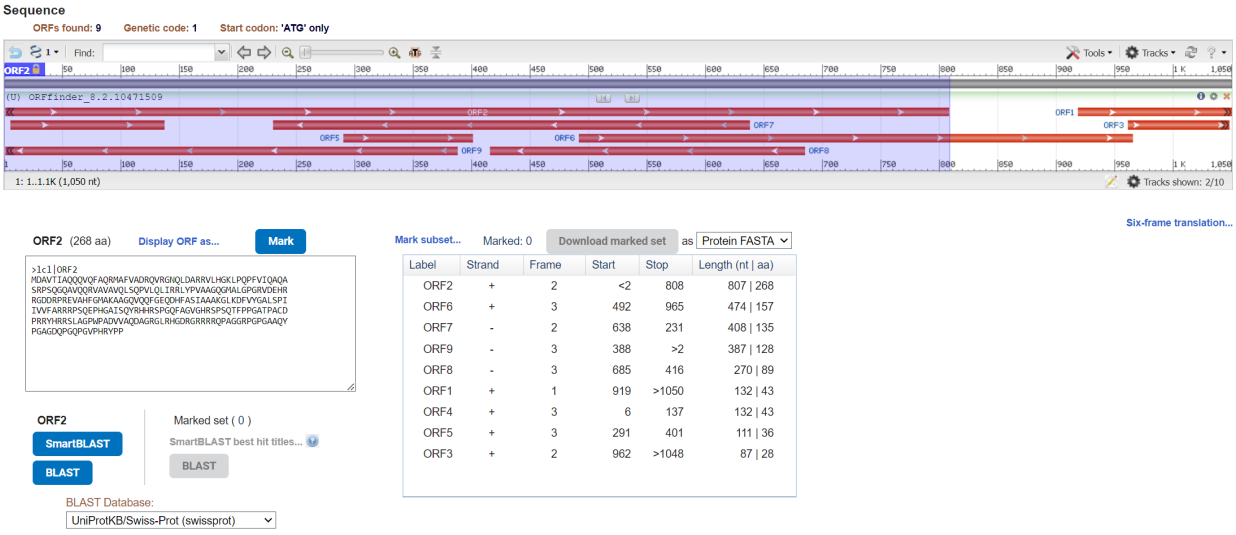


Supplemental Figure 2. ORF Finder results screen after inputting the PF-25A 8 SL1 primer sequence.

Supplemental Table 3. Top 10 hits resulting from inputting the PF-25A 8 SR2 primer sequence into BLASTx.

Description	Organism Name	E value	Percent Identity	Hit Number
Hypothetical protein GCM10020185_69440	*Pseudomonas brassicacearum subsp. brassicacearum*	5e-87	99.35%	1
Inorganic polyphosphate/ATP-protein kinase	*Pseudomonas savastanoi pv. glycinea*	2e-86	88.37%	2

Inorganic polyphosphate/ATP-protein kinase	*Pseudomonas savastanoi pv. glycinea*	6e-86	88.37%	3
NAD(+) kinase	*Pseudomonas*	6e-86	99.35%	4
NAD(+) kinase	*Pseudomonas sp. Pf153*	6e-86	99.35%	5
NAD(+) kinase	*Pseudomonas mediterranea*	7e-86	99.35%	6
NAD(+) kinase	*Pseudomonas*	7e-86	99.35%	7
NAD(+) kinase	*Pseudomonas*	9e-86	98.70%	8
NAD(+) kinase	*Pseudomonas citri*	3e-85	98.70%	9
NAD(+) kinase	*Pseudomonas*	4e-85	98.05%	10



Supplemental Figure 3. ORF Finder results screen after inputting the PF-25A 8 SR2 primer sequence.

Supplemental Table 4. Top 5 hits resulting from inputting ORF9 of the PF-25A 8 SR2 primer sequence into BLASTp.

Description	Organism Name	E value	Percent Identity	Hit Number
DUF1853 family protein	*Pseudomonas*	8e-134	100.00%	1
DUF1853 family protein	*Pseudomonas brassicacearum*	8e-134	100.00%	2
DUF1853 family protein	*Pseudomonas*	8e-134	100.00%	3
DUF1853 family protein	*Pseudomonas*	9e-134	100.00%	4
DUF1853 family protein	*Pseudomonas brassicacearum*	9e-134	100.00%	5