Nitrate reductase dissimilarity in different strains of *Klebsiella pneumoniae* as a benchmark for virulence

Case Study III Report

November 5, 2019

Submitted: November 5, 2019

Report Written Individually By: Alex Yang

Group: [Member Names Edited Out for Privacy]

Tables and Figures Generated by: [Member Names Edited Out for Privacy]

Table of Contents

Alex Yang, *[Group Member Names Withdrawn for Privacy].* Nitrate reductase dissimilarity in different strains of *Klebsiella pneumoniae* as a benchmark for virulence. Case Study III, 11/5/2019

Abstract

Klebsiella pneumoniae IA565 and *Klebsiella pneumoniae* 43816 are two strains of the strains of *K. pneumoniae* bacteria which are associated with virulence in humans. However, environment-specific virulence has been shown to occur in previous studies in which the IA565 strain has is rapidly cleared in rodent pneumonia models[7]. The team has hypothesized that sequence differences in Nitrogen utilization loci between the two strains of the same species likely has led to differing or increased nitrate reductase activity in IA 565 in inflamed mucosal environments. To test this claim, three nitrate utilization operons (loci) were aligned using the NCBI GenBank software, with the "blastn" option specifically selected. The nitrate-reducing operons investigated were: *narGHJI, narZYWV,* and *napFDAGHBC*. In particular, it is believed that Nitrogen utilization loci in strains of *E. coli* can serve as a model as to how IA565 is able to survive within environments possessing inflammatory cells. Through the NCBI GenBank, *narZYWV, narGHJI,* and *napFDAGHBC* were all isolated through searching the whole genomes of the bacterial types: *E. coli* K12, *E. coli* Nissile, *K. pneumoniae IA565*, *K. pneumoniae* 43816. The alignment software tlbastn was also used to convert the compare homologies of protein residues based on the genetic sequence and protein homology matrices were constructed accordingly. The residue sequences were then aligned and inserted into ClustalX to construct protein dendrograms. UPGMA trees were constructed using the DendroUPGMA software[9]. Extremely high levels of both genetic and protein homology were shown both between and within species. No protein homologies were below 88% for *narZYWV* or *narGHJI*, and *napFDAGHBC* was not even found in its entirety in either of the *Klebsiella* pneumoniae strains. UPGMA genetic trees and protein dendrograms all showed low phylogenetic distances based upon nitrate-reductase dissimilarities. Given that the *napFDAGHBC* operon was not even found in *K. pneumoniae* of either strain and that *narZYWV* along with *narGHJI* were both highly homologous, it was determined that the hypothesis was not supported. Within the bounds of this investigation, no clear link to IA 565 virulence within the nitrate reductase loci were found.

(339 words)

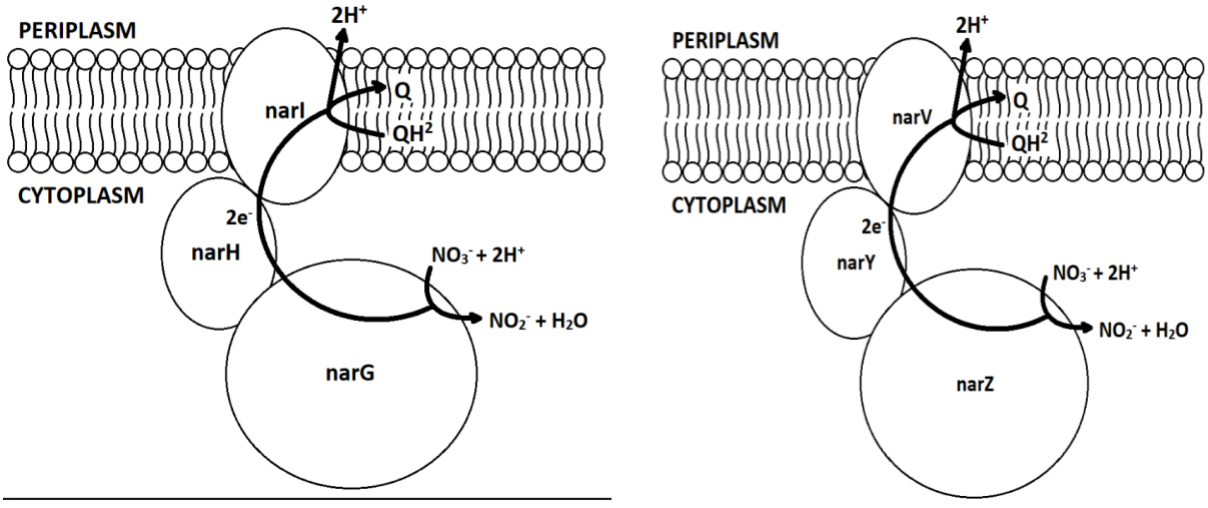
Introduction

Klebsiella pneumoniae is a rod-shaped, facultative anaerobic, gram-negative species of bacteria found as a part of the human microbiome. Two strains in particular: *Klebsiella pneumoniae* 43816 and *Klebsiella pneumoniae* IA565 are virulent in affected humans[7]. Both strains cause infection through colonization of mucosal surfaces such as the intestines and lungs[5]. IA565 and 43816, while virulent in diseased humans, showed differences in murine models. While 43816 expressed virulence in both humans and mice, IA565 was rapidly cleared in mice unless the mice had intestinal inflammation [7]. In a series of studies done by Winter et al, *E. coli* was found to take advantage of its nitrate utilization capabilities in high-inflammation environments with low Oxygen (O_2) present. Three operon loci encoding for nitrate reductases, or enzymes which promote Nitrate (NO_3^-) reduction as a final electron acceptor into Nitrite (NO_2^-) were investigated in the *E. coli*[11]. These three loci were: *narGHJI*, *narZYWV*, and *napFAGHBC*.

The *narGHJI* operon encodes the genetic sequence for the enzyme nitrate reductase A, which is a bacterial inner-membrane-associated enzyme responsible from catalyzing the reduction of nitrate to nitrite and can allow nitrate to serve as the final electron acceptor when the bacteria are in anaerobic environments[3]. The operon consists of genes *narG, narH, narJ*, and *narI* as is evident through the operon naming. *narG, narH,* and *narI* code for subunits α, β, and γ of the nitrate reductase A enzyme, respectively, with *narJ* coding for a chaperone responsible for proper assembly of the three subunits. Nitrate reductase A operates only during anaerobic growth in the presence of nitrate. It is responsible for 98% of nitrate-reducing activity when activated[3].

The *narZYWV* operon is homologous to the *narGHJI* operon[2] in that it too codes for a nitrate reductase (nitrate reductase Z). Nitrate reductase Z is similar in structure to nitrate reductase A but differs in the manner by which its reductive activities are regulated. While nitrate reductase Z also serves as an inner-membrane-associated catalyst for the reduction of NO_3^-, its synthesis through *narZYWV* is constitutively expressed at low levels under both aerobic and anaerobic conditions and its activity is only increased through the presence of higher concentrations of nitrate and upon detection of anaerobic conditions, but still only accounts for less than two percent of nitrate-reducing activity when activated[2]. *narZ, narY,* and *narV* respectively code for the associated subunits α, β, and γ of nitrate reductase Z, with *narW* coding for the protein assembly chaperone.

The two *nar*- operons can be compared as such:

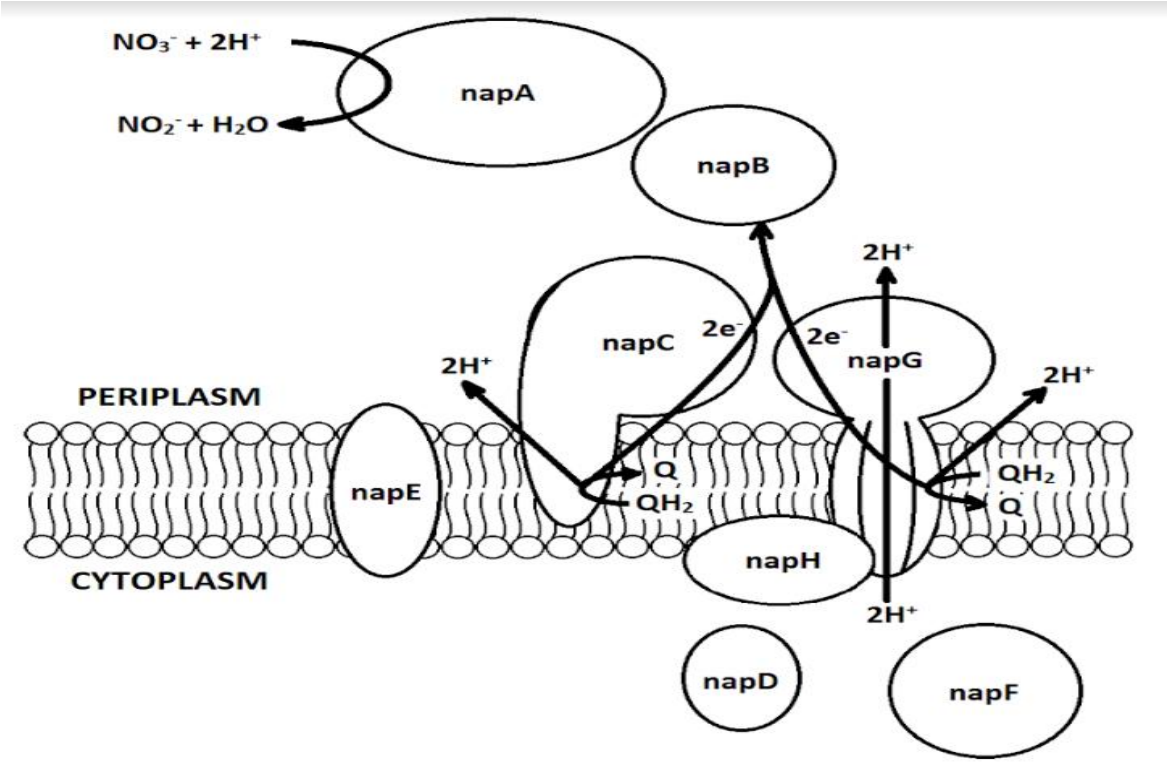


(Figures made ----------); Left: NarGHI mechanism; Right: NarZYV mechanism

As shown above, both *nar*- systems are very much homologous mechanistically except for the aforementioned manners in which they are regulated and their activity levels in anaerobic conditions.

A periplasmic nitrate-reductase-coding operon was examined as well, known as *napFDAGHBC*. The associated enzyme is known as NapABC or periplasmic nitrate reductase. NapABC, unlike nitrate reductases A and Z, is not completely bound to the inner membrane of the bacterial cell, but also has parts located within the periplasm. It too functions in anaerobic conditions but is also able to display comparatively high nitrate-reducing activity when environmental nitrate concentrations are low and the bioenergetic efficiencies of nitrate reductase A and nitrate reductase Z are limited[8] for *E. coli*.

The Nap- protein cluster can be shown as such:

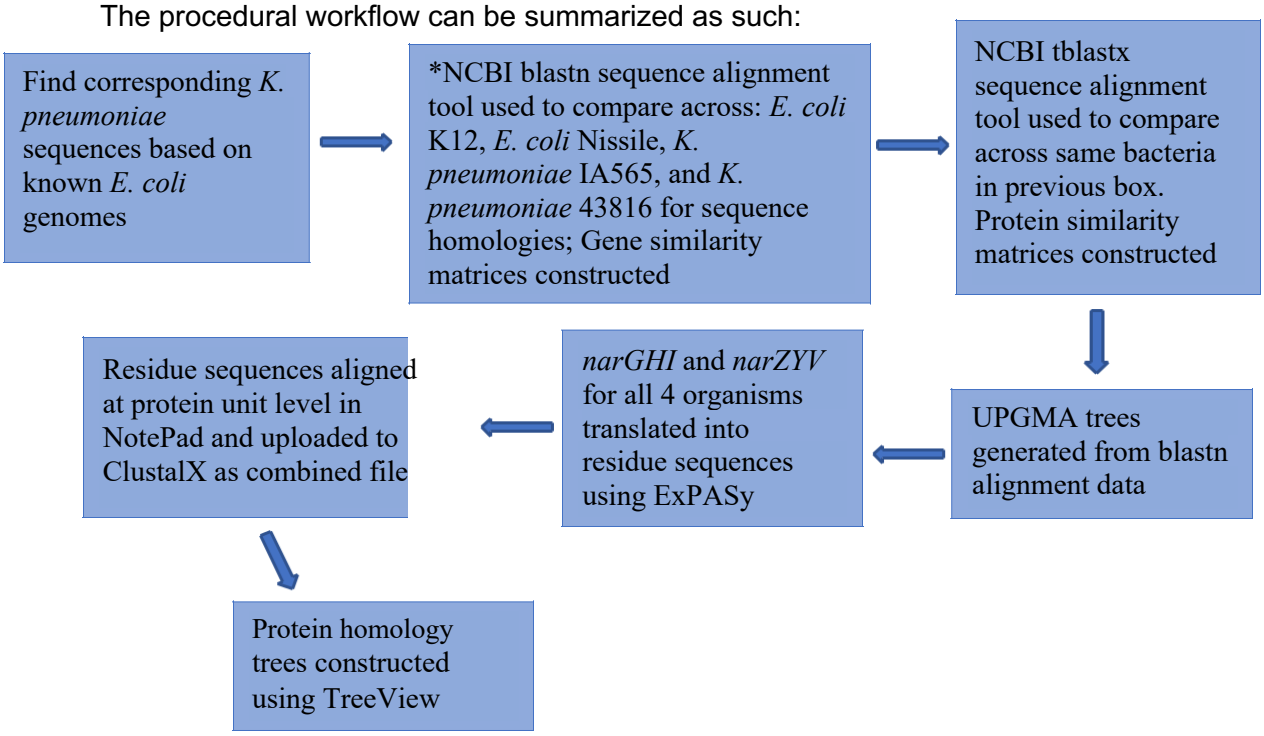


(Figure made by --------) Note the protein subunits within the periplasmic space.

Given the aforementioned and elucidated models of nitrate utilization in *E. coli,* it was of interest as to whether similar models could also apply to the proliferation of *Klebsiella pneumoniae* IA565 in inflamed murine intestines. With this *Escherichia coli* model in mind, it was hypothesized that significant differences within the genetic sequence of these nitrate utilization proteins in *Klebsiella pneumoniae* 43816 versus IA565 could explain the difference in virulence between the two strains.

Methods

The procedural workflow can be summarized as such:



*The bacterial genome GenBank identification numbers used for the blastn and tblastx sequence alignment analyses were as follows:

- NC_000913.3- Escherichia coli K12
- 316435- *Escherichia coli* Nissle
- CP009208.1- *Klebsiella pneumoniae* ATCC 43816 KPPR1
- NZ_JPIQ01000001.1- *Klebsiella pneumoniae* IA565

*Individual subunit genes (i.e. narG) were aligned in blastn and tblastx for each bacterial species as well as entire operons (narGHI)

*To construct the trees, dissimilarity matrices were created from the similarity matrices mentioned in the workflow diagram above. This was done by subtract the similarity value from 1. For example, two bacteria with a similarity value of 0.90 would have a dissimilarity value of 1-0.90=**0.10**.

The methods can ultimately be summed up into three overarching phases:

1. Finding genetic sequences for *K. pneumoniae* strains based upon *E. coli* genomes
2. Alignment of nucleotide sequences using blastn and tblastx to find genetic and protein homologies
3. Construction of dendrograms, or phylogenetic trees based upon sequence data, derived from genetic and protein homology data.

Results

Homologies in nucleotide sequence for two of the nitrate-utilization loci between *Klebsiella pneumoniae* 43816 and IA565 were very high, at 99.59% for *narGHJI*, 99.25% for *narZYWV* (Tables 1.1 and 2.1). The same was true for protein homologies between the same two loci and bacterial strains: with 99% homology for *narGHJI* and *narZYWV* (Tables 1.2 and 2.2). Even at the individual subunit level, identities were above 98% for blastn results (Tables 1.1 & 2.1) and for tblastx results (Table 1.2 & 2.2).

UPGMA dendrograms generated from dissimilarity matrices yielded no significant data in understanding genetic distance between the two strains but did show a taxonomic relationship grouped by the bacterial species and strains. The *Klebsiella pneumoniae* strains were grouped with each other and the *Escherichia coli* were grouped with each other (Tables 4.1-5.2). For *narGHJI* and *narZYVW*, the tree lengths were 1.625 and 3.5, respectively, showing low phylogenetic distances and higher homology among the nitrate-reducing operons across all examined bacteria (Tables 4.1 & 5.1).

napFDAGHBC was not found in either *Klebsiella pneumoniae* 43816 or IA565. *napA*, *napD*, and *napF* were not found in either of the *K. pneumoniae* strains. Therefore, no further analysis of the protein structure of the *napFDAGHBC* were performed given lack of the full operon present (Table 3.1).

Discussion

Based on the obtained data, the hypothesis was not supported. There was no evidence of any significant divergence in genetic or protein homologies of nitrate reductase loci between *Klebsiella pneumoniae* strains IA565 and 43816. This held true at both the nucleotide and protein level, and it held true for individual subunits as well as the entire operon, so any differences in virulence would likely not be found within the *narGHJI* or *narZYVW* operons.

Given that *napFDAGHBC* could not even be found in its entirety within either *Klebsiella pneumoniae* strain of interest, it was decided that, within the bounds of this investigation, the periplasmic nitrate reductase was also not the cause of virulence in IA565.

While this investigation did not yield data that could suggest the difference in IA565 virulence, further investigation into the biochemical composition of inflammatory cells in mice vs. humans could shed insight as to why IA565 may be virulent in inflammatory mice bowels.

Within the mice, it has been shown that reactive Oxygen and Nitrogen species in anaerobic environments can yield an antimicrobial known as peroxynitrite. However, peroxynitrite is rapidly converted to nitrate through reactivity with CO_2 in the gut, which serves as the terminal electron acceptor for nitrate reductase enzymes to function upon[11]. This mechanism can lead to bacterial bloom in the intestines or lungs as they are both mucosal surfaces that can undergo inflammation under certain conditions.

With the knowledge that mucosal environments that are anaerobic and inflamed are conducive to nitrate reduction in certain bacteria, the biochemical composition of the inflamed host microbiome itself can be an aspect to investigate so as to determine why the Klebsiella pneumoniae differences in virulence occur given the lack of significant genetic and protein data in the nitrate reductase loci.

(1237 words)

References

1. **Bertero M.** 2011. The Membrane-Bound Nitrate Reductase A from Escherichia Coli: NarGHI. Encyclopedia of Inorganic and Bioinorganic Chemistry.

2. **Blondelle SE.** 2006. Understanding Biology Using Peptides: Proceedings of the Nineteenth American Peptide Symposium. Springer Science Business Media, LLC. New York, NY.

3. **Bonnefoy V, Demoss JA.** 1994. Nitrate reductases in *Escherichia coli*. Antonie van Leeuwenhoek 66:47–56.

4. **Cortes G**. 2002. Molecular analysis of the contribution of the capsular polysaccharide and the lipopolysaccharide o side chain to the virulence of *Klebsiella pneumoniae* in a murine model of pneumonia. Infection and Immunity 70: 2583–2590.

5. **Dorman MJ, Feltwell T, Goulding DA, Parkhill J, Short FL.** 2018. The Capsule Regulatory Network of Klebsiella pneumoniae Defined by density-TraDISort. mBio 9.

6. Gene - NCBI. National Center for Biotechnology Information. U.S. National Library of Medicine.

7. **Lau, H. Y., G. B. Huffnagle, and T. A. Moore.** 2008. Host and microbiota factors that control *Klebsiella pneumoniae* mucosal colonization in mice. Microbes Infect 10: 1283-1290.

8. **Stewart V, Lu Y, Darwin AJ.** 2002. Periplasmic nitrate reductase (napABC enzyme) supports anaerobic respiration by *Escherichia coli* K-12. Journal of Bacteriology 184: 1314–1323.

9. **Vallve SG, Puigo P**. DendroUPGMA: Dendrogram construction using the UPGMA algorithm.

10. **Winter, S. E., and A. J. Baumler.** 2014. Dysbiosis in the inflamed intestine: chance favors the prepared microbe. Gut Microbes 5: 71-73.4.

11. **Winter, S. E., M. G. Winter, M. N. Xavier, P. Thiennimitr, V. Poon, A. M. Keestra, R. C. Laughlin, G. Gomez, J. Wu, S. D. Lawhon, I. E. Popova, S. J. Parikh, L. G. Adams, R. M. Tsolis, V. J. Stewart, and A. J. Baumler.** 2013. Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. Science 339: 708-711.

Reference Data and Models

Table 1.1: Nucleotide Homology (%) of *narGHJI* Associated Genes

narG(blastn)	***E. coli* K12**	***E. coli* Nissle**	***K. pneumoniae* IA565**	***K. pneumoniae* 43816**
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	97.22	100	x	x
***K. pneumoniae* IA565**	85.63	85.74	100	x
***K. pneumoniae* 43816**	85.63	85.74	99.73	100
narH(blastn)	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	97.34	100	x	x
***K. pneumoniae* IA565**	86.16	86.35	100	x
***K. pneumoniae* 43816**	86.16	86.35	99.67	100
narI(blastn)	***E. coli* K12**	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	96.9	100	x	x
***K. pneumoniae* IA565**	83.33	83.04	100	x
***K. pneumoniae* 43816**	83.63	82.89	99.41	100
narJ(blastn)	***E. coli* K12**	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	98.03	100	x	x
***K. pneumoniae* IA565**	81.77	80.93	100	x
***K. pneumoniae* 43816**	81.91	81.65	98.87	100
narGHJI(blastn)	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	97.17	100	x	x
***K. pneumoniae* IA565**	85	85.07	100	x
***K. pneumoniae* 43816**	84.94	85.13	99.59	100

(Table 1) narGHJI blastn Homology Matrix: This chart displays the percent identity/homology between bacterium using the nucleotide sequence of each gene and of the full narGHJI operon sequence.

(Table generated by ---------)

Table 1.2: Protein Homology (%) of *narGHJI* Associated Subunits

narG(tblastx)	***E. coli* K12**	***E. coli* Nissle**	***K. pneumoniae* IA565**	***K. pneumoniae* 43816**
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,99	100	x	x
***K. pneumoniae* IA565**	94,97	94,97	100	x
***K. pneumoniae* 43816**	94,97	94,97	100,100	100
narH(tblastx)	K12	Nissle	IA565	43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,100	100	x	x
***K. pneumoniae* IA565**	93,96	93,96	100	x
***K. pneumoniae* 43816**	93,96	93,96	100,100	100
narI(tblastx)	K12	Nissle	IA565	43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,100	100	x	x
***K. pneumoniae* IA565**	93,98	91,96	100	x
***K. pneumoniae* 43816**	91,96	91,96	100,100	100
narGHI(tblastx)	K12	Nissle	IA565	43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,99	100	x	x
***K. pneumoniae* IA565**	94,96	94,97	100	x
***K. pneumoniae* 43816**	94,97	94,97	99,99	100

(Table 2) narGHI tblastx Homology Matrix: This chart displays the percent identity/homology between bacterium using the translated protein sequence for each gene and for the narGHI operon. The numbers indicated are (percent identity of amino acids, positives/functionally equivalent amino acids). *narJ was not included as it encodes the chaperone of the complex.

(Table generated by --------)

Table 2.1: Nucleotide Homology (%) of *narZYWV* Associated Genes

narZ(blastn)	***E. coli* K12**	***E. coli* Nissle**	***K. pneumoniae* IA565**	***K. pneumoniae* 43816**
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	97.01	100	x	x
***K. pneumoniae* IA565**	80.19	80.15	100	x
***K. pneumoniae* 43816**	80.14	80.13	99.25	100
narY(blastn)	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	95.99	100	x	x
***K. pneumoniae* IA565**	80.8	80.87	100	x
***K. pneumoniae* 43816**	80.8	80.67	99.42	100
narV(blastn)	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	83.26	100	x	x
***K. pneumoniae* IA565**	74.45	77.4	100	x
***K. pneumoniae* 43816**	75.18	77.99	99.11	100
narW(blastn)	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	95.4	100	x	x
***K. pneumoniae* IA565**	73.93	73.35	100	x
***K. pneumoniae* 43816**	74.07	73.21	98.56	100
narZYWV(blastn)	E. coli K12	E. coli Nissle	K. pneumoniae IA565	K. pneumoniae 43816
E. coli K12	100	x	x	x
E. coli Nissle 1917	97.01	100	x	x
K. pneumoniae IA565	80.19	80.15	100	x
K. pneumoniae 43816	77.48	80.14	99.25	100

(Table 3) narZYVW blastn Homology Matrix: This chart displays the percent identity/homology between bacterium using the nucleotide sequence of each gene and of the full narZYWV operon.

(Table generated by --------)

Table 2.2: Protein Homology (%) of *narZYWV* Associated Subunits

narZ(tblastx)	***E. coli* K12**	***E. coli* Nissle**	***K. pneumoniae* IA565**	***K. pneumoniae* 43816**
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,99	100	x	x
***K. pneumoniae* IA565**	88,93	88,93	100	
***K. pneumoniae* 43816**	99,99	88,93	99,99	100
narY(tblastx)	K12	Nissle	IA565	43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	99,100	100	x	x
***K. pneumoniae* IA565**	88,95	89,95	100	
***K. pneumoniae* 43816**	88,95	89,95	99,100	100
narV(tblastx)	**K12**	**Nissle**	**IA565**	**43816**
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	85,94	100	x	x
***K. pneumoniae* IA565**	77,92	83,92	100	
***K. pneumoniae* 43816**	78,92	84,92	98,98	100
narZYV(tblastx)	K12	Nissle	IA565	43816
***E. coli* K12**	100	x	x	x
***E. coli* Nissle 1917**	98,99	100	x	x
***K. pneumoniae* IA565**	88,93	88,93	100	x
***K. pneumoniae* 43816**	88,93	88,93	99,99	100

(Table 4) narZYV tblastx Homology Matrix: This chart displays the percent identity/homology between bacterium using the protein sequences for each gene and for the narZYV operon. The numbers indicated are (percent identity of amino acids , positives/functionally equivalent amino acids). *narW was not included as it encodes the chaperone of the complex.

(Table generated by --------)

Table 3.1: Nucleotide Homology (%) of *napFDAGHBC* Associated Genes
(*Protein homology matrix not constructed due to lack of *nap*- subunits in *K. pneumoniae*)

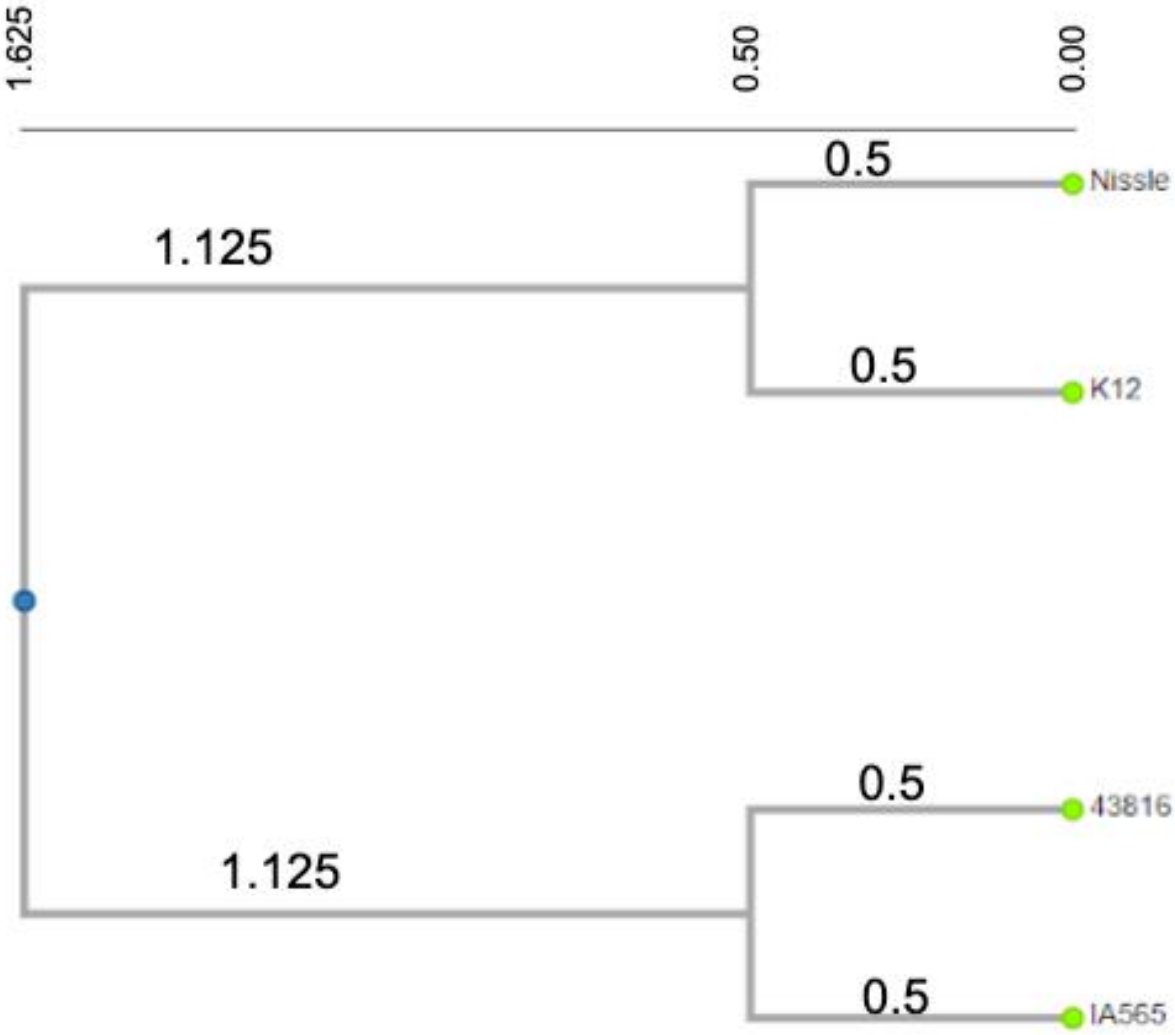
napA	*E. coli* K12	*E. coli* Nissle	*K. pneumoniae* IA565	*K. pneumoniae* 43816
E. coli K12	100	x	x	x
E. coli Nissle 1917	97.55	100	x	x
K. pneumoniae IA565	N/A	N/A	100	x
K. pneumoniae 43816	N/A	N/A	N/A	100
napD				
E. coli K12	100	x	x	x
E. coli Nissle 1917	99.62	100	x	x
K. pneumoniae IA565	N/A	N/A	N/A	x
K. pneumoniae 43816	N/A	N/A	N/A	N/A
napF				
E. coli K12	100	x	x	x
E. coli Nissle 1917	99.6	100	x	x
K. pneumoniae IA565	N/A	N/A	N/A	x
K. pneumoniae 43816	N/A	N/A	N/A	N/A
napH				
E. coli K12	100	x	x	x
E. coli Nissle 1917	97.8	100	x	x
K. pneumoniae IA565	N/A	N/A	100	x
K. pneumoniae 43816	N/A	N/A	N/A	100
napB				
E. coli K12	100	x	x	x
E. coli Nissle 1917	98.89	100	x	x
K. pneumoniae IA565	N/A	N/A	100	x
K. pneumoniae 43816	N/A	N/A	N/A	100
napC				
E. coli K12	100	x	x	x
E. coli Nissle 1917	99.17	100	x	x
K. pneumoniae IA565	N/A	N/A	100	x
K. pneumoniae 43816	N/A	N/A	N/A	100
napG				
E. coli K12	100	x	x	x
E. coli Nissle 1917	99.14	100	x	x
K. pneumoniae IA565	N/A	N/A	100	x
K. pneumoniae 43816	N/A	N/A	N/A	100
napFDAGHBC				
E. coli K12	100	x	x	x
E. coli Nissle	N/A	100	x	x
K.pneumoniae IA565	N/A	N/A	100	x
K.pneumoniae 43816	N/A	N/A	N/A	100

(Table 5) napFDAGHBC blastn Homology Matrix: This chart displays the percent identity/homology between bacterium using the nucleotide sequence of each gene and of the full napFDAGHBC operon.

Certain proteins in the nap operon(A,D,F) were not present in the *K.pnemoniae* strains, therefore the nap operon is not fully intact. *N/A values indicate no significant homology found and anything highlighted indicates a lower query coverage.

(Table generated by --------)

Figure 4.1: UPGMA Tree for *narGHI* Operon (blastn)



Description: Scale at top of figure, with tree length of 1.625.
Node Identities on at end of graph from top to bottom:

- Nissile = *E. coli* Nissile
- K12 = *E. coli* K12
- 43816 = *K. pneumoniae* 43816
- IA565 = *K. pneumoniae* IA565

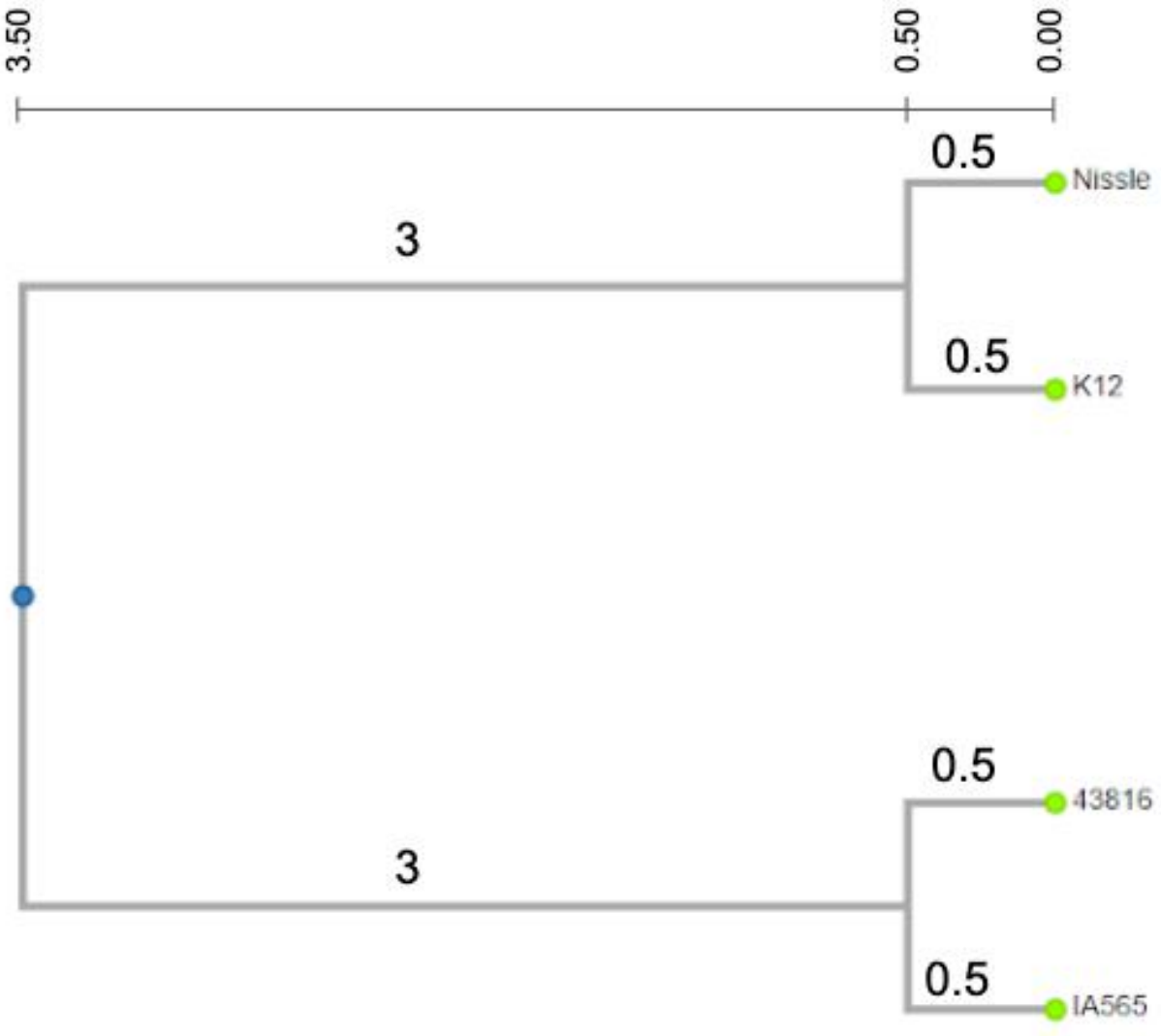
(Figure Generated by --------)

Figure 4.2: Dendrogram for NarGHI based on Protein Homology (tblastx)

***Scale bar represents 0.01 residue per 1 residue difference**

(Figure Generated by ----------)

Figure 5.1: UPGMA Tree for *narZYV* Operon (blastn)

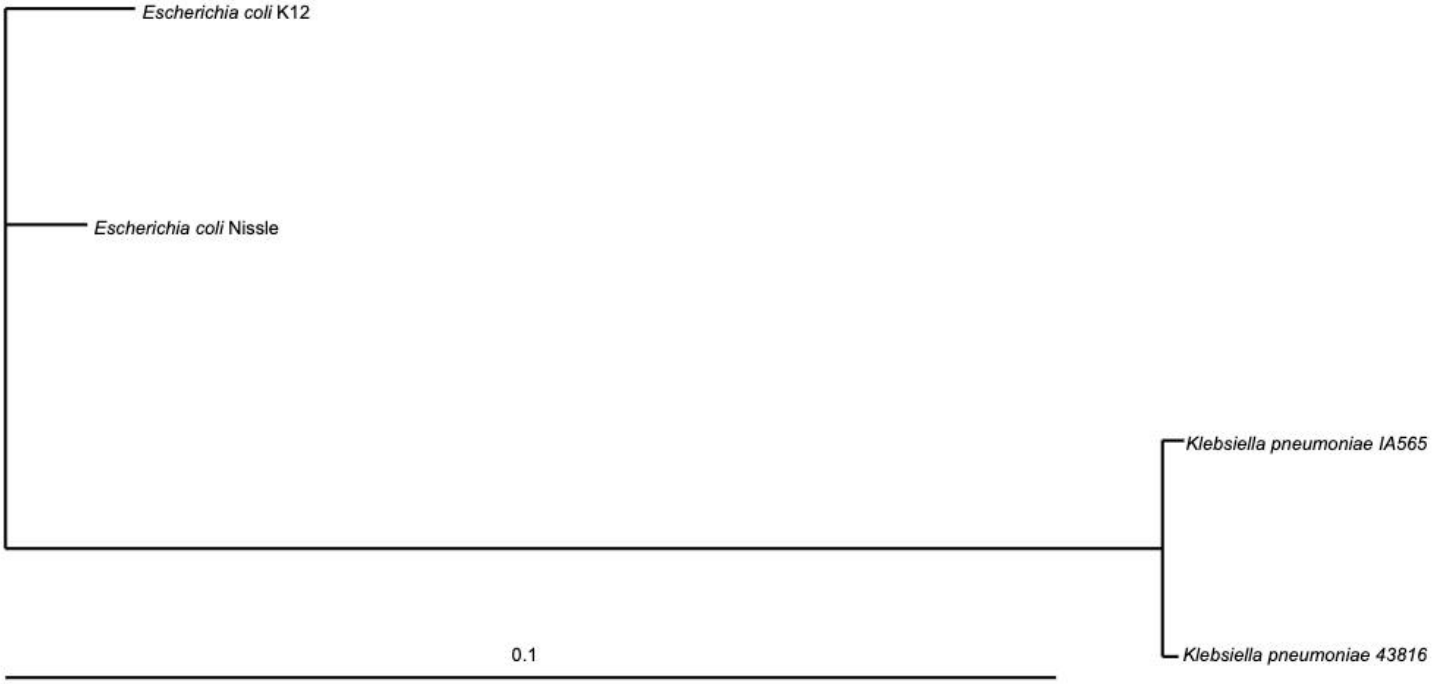


Description: Scale at top of figure, with tree length of 1.625.
Node Identities on at end of graph from top to bottom:

- Nissile = *E. coli* Nissile
- K12 = *E. coli* K12
- 43816 = *K. pneumoniae* 43816
- IA565 = *K. pneumoniae* IA565

(Figure Generated by ----------)

Figure 5.2: Dendrogram for NarZYV based on Protein Homology (tblastx)



***Scale bar represents 0.1 residue per 1 residue difference**

(Figure Generated by ----------)