The Evolution of DNA Uptake Sequences in *Neisseria* Genus from *Chromobacterium-violaceum.*

Cory Garnett

Introduction

In bacteria, transformation is conducted by the uptake of DNA followed by homologous recombination. The purpose of these sequences are to uptake exogenous DNA of closely related family members to keep the homology of alleles and protect themselves against hazardous DNA. (Frye et al. 2013) Not all bacteria have a DNA sequence that can be classified as a DNA uptake sequence, in order for a DUS to characterize as one then it must occur over 100 times its expected value (or ~1% of the genome). (Smith et al. 1999) There are only a few genus's that have been studied including *Neisseriaceae* and *Haemophilus* which, by bioinformatics analyses it is showed that there are eight variant DUS sequences present in *Neisseriaceae* which are known as DUS dialects which allows for communication between the species. (Frye et al. 2013)

N.meningitidis comes from the class *Proteobacteria* which has subdivisions of Alpha, Beta, Gamma, Delta and Epsilon: the Neisseriaceae family is within the *Betaproteobacteria* subunit. (Enright et al) In most bacteria there is no obvious preference of a specific uptake sequence, this is not the case for *Neisseria* and *Haemophilus.* (Maughan, Wilson, and Redfield 2010) These Gram-negative bacteria only interact with their genus to maintain homology; other organisms (Gram-positive bacteria) do not bind to the uptake system. (Mathis and Scocca 1981) Although a recent study shows they were able to find that there is a core connection between the different uptake sequences in the *Neisseriaceae* family, it contains the three nucleotide sequence of 5'-CTG-3' in the positions 6,7, and 8. (Frye et al. 2013) Many studies have shown the uptake sequence for *N.meningitidis* MC58 is the 10-bp sequence, 5'- GCCGTCTGAA – 3'. This sequence is found to be randomly distributed on both leading and lagging strands. (Smith et al. 1999)

It has been recently found that there are multiple uptake sequences within the family of *Neisseriaceae.* (Frye *et al* 2013) These sequences were taken to compare to the most represented sequences within *Chromobacterium-violaceum*. The question to be looked at here is to see rather there was a comparison of these sequences to the ones presented by Frye. Is there a core sequence that the present uptake sequences, in the different genus that comes from *C. violaceum.*

Methods

Phylogenetic Trees: With the use of BioBIKE and GenBank, the 16s rRNA was found by searching GENES-DESCRIBED-BY("SSU") IN NONCODING_GENES_OF(Neisseria-meningitidis-MC58). The sequence was confirmed by comparing the sequence of the *N.men* MC58 strain in the NCBI GenBank. By using the (GENES-SIMILAR-TO(SEQUENCE-OF(X)

(IN(Y)) where X was the sequence and Y represented the Family of *Neisseriaceae* the sequences were aligned to create the tree presented in (Figure 1).

DUS Alignment and Tree: With the information found by (Frye *et al.*) I was able to collect the DNA uptake sequences of different genus within the family of *Neisseriacea.* With the use of bioBIKE, the top five sequences in *C.violaceum-ATCC* were aligned with the sequences used in Frye. With this we are able to identify the similarities in the sequences (Figure 3) Once the alignment was complete, it was created into a phylogenetic tree by looking at the percent identity between the sequences to show the results of similarities. (Figure 2).

***Neisseriaceae* genus comparison:** Table 1 was created by creating a loop within bioBIKE. For Each of the *Neisseriaceae spp.* within bioBIKE the length of the different genomes were found as well as the number of genes. Within the genomes the DUS sequence 5' – GCCGTCTGAA was counted by finding the matches of pattern and with each match, it was collected and counted, as well as the times it occurs within the genes. The G-C fraction was calculated within the loop and multiplied by 100. Finally, the percentage of the DUS coverage was found by multiplying the amounts it was found in the genome by 10, and then this number was multiplied by 100 to get the final percentage.

Results and Discussion

The main idea of this paper is to look at the evolution of DUS within the *Neisseriacea* family. These uptake sequences match the requirements of the results presented by (Smith et al. 1999) where the DUS is represented around 1 % of the genome and is over represented by 100 fold of its expected value. By comparing to the DUS of different *Neisseria* to the most represented sequences of *Chromobacterium-violaceum,* I was able to locate some similarities that could have been the start to the uptake sequences occurring over long periods of evolutionary time. The consensus displayed in (Figure 3) shows a consistency in the 7 (cytosine) and 9 (guanine) positions, excluding CV 2-5 to only look at the most represented sequence of *C.violaceum*. These results rise potential as in the results given by (Frye et al 2013) show there is a consistent 5' CTG as a core sequence. This core sequence is shown to be the primary cause for transformation within these organisms and without it, DUSing would not happen. The first dialect to branch off after the beginning of the new genus is wadDUS, represented by *N. wadsworthii* 9715 and the sequence CCTGTCTGAA. This sequence showed the most diversity of all within the consensus. (Figure 3) (Table 2).

By looking at the consensus in Figure 3, we are able to look at the major differences of the dialects. The sequence "DUS" which is commonly found in *N. meningitides* only differs in three pairings, positions 6, 8 and 10. Which, by the results displayed by Frye, the first few nucleotides have a major impact in transformation efficacy where position 1 can decrease by 90%, when mutated into a different base pairing and a reduction of 60% in position 3 pairing. There are consistencies with the most represented pairings of *C.violaceum* and the different

dialects that have become over evolutionary time. It seems that through mutations the DUS system comes from the *C.violaceum.* Within the differences of the phylogenetic trees of Figures 1 (alignment of the 16s rRNA) and 2 (alignment of 10 –nt uptake sequences.), there is a clear breakoff to of *Neisseria* from *C. violaceum.*

Citations

Enright MC, Carter PE, Maclean IA, Mckenzie H. Phylogenetic relationships between some members of the genera Neisseria, Acinetobacter, Moraxella, and Kingella based on partial 16S ribosomal DNA sequence analysis. Int J Syst Bacteriol. 1994;44(3):387-91.

Frye SA, Nilsen M, Tønjum T, Ambur OH. Dialects of the DNA uptake sequence in Neisseriaceae. PLoS Genet. 2013;9(4):e1003458.

Harmsen D, Singer C, Rothgänger J, et al. Diagnostics of neisseriaceae and moraxellaceae by ribosomal DNA sequencing: ribosomal differentiation of medical microorganisms. J Clin Microbiol. 2001;39(3):936-42.

Mathis LS, Scocca JJ. Haemophilus influenzae and Neisseria gonorrhoeae recognize different specificity determinants in the DNA uptake step of genetic transformation. J Gen Microbiol. 1982;128(5):1159-61.

Maughan H, Wilson LA, Redfield RJ. Bacterial DNA uptake sequences can accumulate by molecular drive alone. Genetics. 2010;186(2):613-27.

Smith HO, Gwinn ML, Salzberg SL. DNA uptake signal sequences in naturally transformable bacteria. Res Microbiol. 1999;150(9-10):603-16.

Appendix

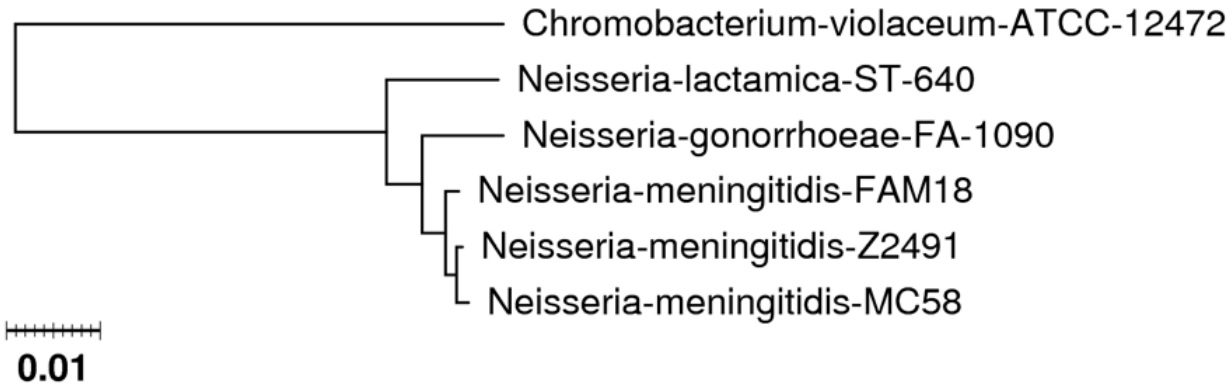

Figure 1: Shows a phylogenetic tree of the Neisseriacae Family. When comparing the 16s rRNA gene, it is found that the Neisseria genus spans from the *C. violaceum* organism. When looking at the tree you are able to see the very close relationships of *Neisseria* compared to the *C. violaceum*.

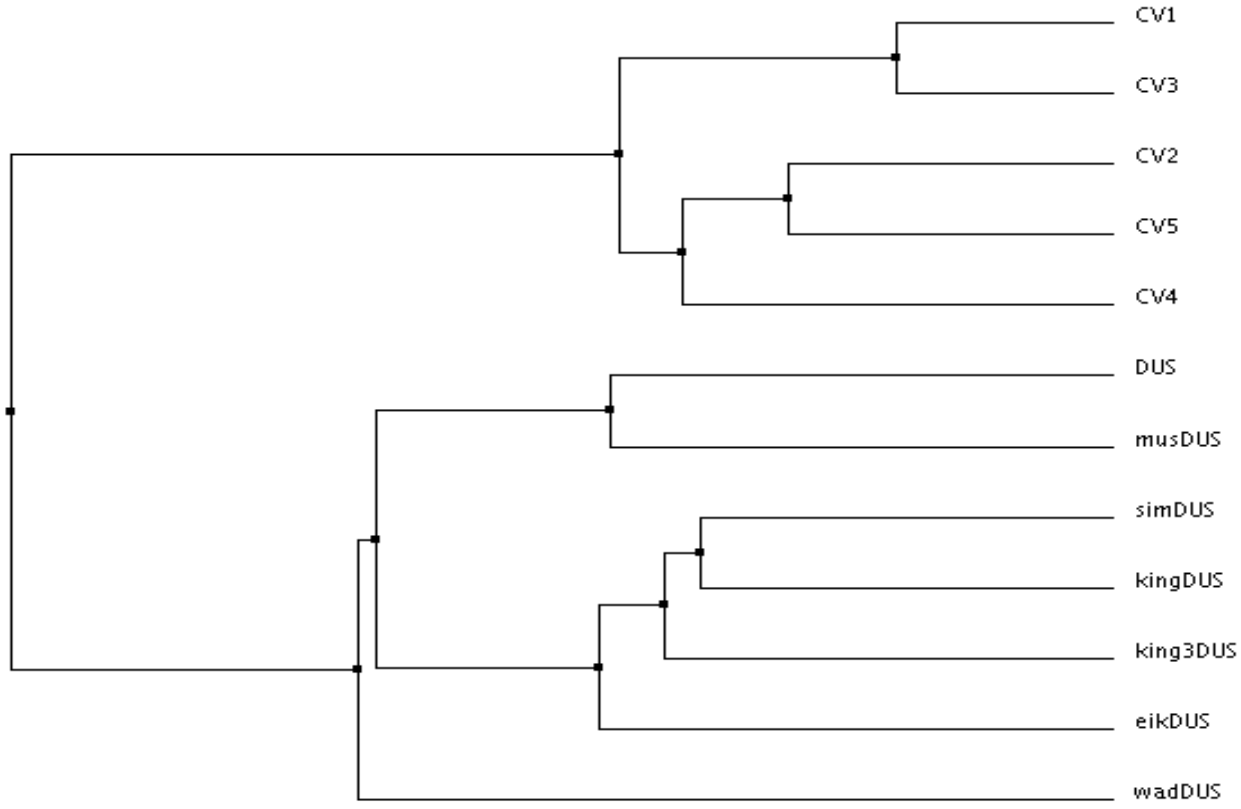

Figure 2: This shows the percent identity of the different DUS in *Neisseria* and the five most represented 10-nt sequences in *C.violaceum-ATCC*. The labels of CV1-5 represent the 5 most represented sequences in *C.violaceum-ATCC* while the different sequences of the different DUS labels represent the different sequences as shown in Table 2.

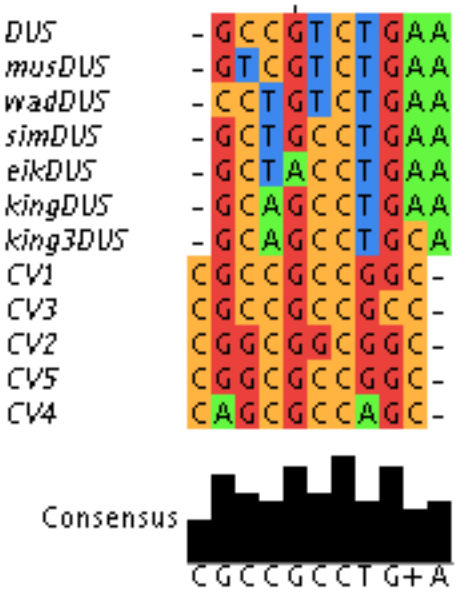

Figure 3: This alignment shows the different dialects found by (Frye *et al.* 2013) as well as the 5 most represented sequences in *C.violaceum-ATCC.* When looking at the sequences, positions 2, 5, 7, and 9 have peaks in the consensus sequence. Showing it is most likely that the 10 –bp sequence has evolved from this.

Table 1: Shows a comparison of the different *Neisseria sp.* and different statistics. By looking at the most common DUS sequence we see that the numbers of DUS within *C.violaceum-ATCC* are much lower then those represented by *Neisseria.*

Organism	Size in bp	# of Genes	# of DUS	#of DUS in Genes	G/C Fraction %	% of DUS in Genome
Neisseria-gonorrhoeae-FA-1090	2153922	2097	1965	764	52.68	.91
Neisseria-lactamica-ST-640	2181714	2496	2216	1275	52.15	1.02
Neisseria-meningitidis-MC58	2272351	2347	1935	741	51.53	.85
Neisseria-meningitidis-FAM18	2236675	2509	1917	862	51.61	.86
Neisseria-meningitidis-Z2491	2184406	2276	1892	742	51.81	.86
Chromobacterium-violaceum-ATCC-12472	4751080	4652	9	4	64.83	.0019

Table 2: Represents the Occurrences of the DUS sequences within *Neisseria* as presented in Frye.

Commonly Represented in:	DUS Sequence	Occurrences in *C.violaceum-ATCC*
N. meningitidis (DUS)	GCCGTCTGAA	9
N. wadsworthii 9715 (wadDUS)	CCTGTCTGAA	7
Neisseria mucosa (mucDUS)	GTCGTCTGAA	0
S. muelleri ATCC 29453 (simDUS)	GCTGCCTGAA	9
K. oralis ATCC 51147 (kingDUS)	GCAGCCTGAA	21
Kingella denitrificans (king3DUS)	GCAGCCTGCA	27
Eikenella corrodens (eikDUS)	GCTACCTGAA	14